RECEIVED
By USDA APHIS BRS Document Control Officer at 1:30 pm, Jan 28, 2020

Petition for a Determination of Nonregulated Status for Plant-Parasitic Nematode-Protected and Herbicide Tolerant GMB151 Soybean

OECD Unique Identifier BCS-GM151-6

The undersigned submits this petition under 7 CFR 340.6 to request that the Administrator, make a determination that the article should not be regulated under 7 CFR part 340.

Andrew Olson, Ph.D.
U.S. Regulatory Manager - Seeds

Submitted By
BASF Corporation
2 T.W. Alexander Drive
Research Triangle Park, NC 27709

Corresponding Author
Andrew Olson, Ph.D.
BASF Corporation
Telephone: (919) 323-0280
E-Mail: andrew.olson@basf.com

Contributing Authors

Michael Weeks

Submitted

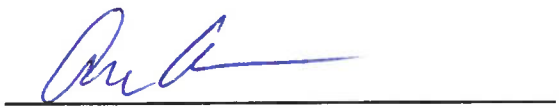
January 24, 2020

Release of information

As part of the regulatory requirement in the United States, BASF Corporation is submitting the information in this petition for review by the USDA. The submission of this information by BASF Corporation does not constitute an authorization to release any information contained within it to any third party. In the event the USDA receives a Freedom of Information Act request, pursuant to 5 U.S.C., § 552, and 7 CFR Part 1, covering all or some of this information, BASF Corporation expects that USDA will provide BASF Corporation with a copy of the material proposed to be released and the opportunity to object to the release of any information based on appropriate legal grounds, e.g., responsiveness, confidentiality, and/or competitive concerns in advance of the release of the document(s). BASF Corporation authorizes the posting of the petition on the USDA Biotechnology Regulatory Services website when it is also published in the Federal Register for public comment. BASF Corporation understands that a copy of this information may be made available to the public by individual request, as part of a public comment period. Except in accordance with the foregoing, BASF Corporation does not authorize the release, publication or other distribution of this information (including website posting) without prior notice and consent by BASF Corporation.

Certification

The undersigned certifies that, to the best knowledge of the undersigned, this petition includes all information and views on which to base a determination, and that it includes all relevant data and information known to the petitioner that are unfavorable to the petition.



Andrew Olson, Ph.D.
U.S. Regulatory Manager - Seeds

Address:

BASF Corporation
2 T.W. Alexander Dr.
Research Triangle Park, NC 27709

EXECUTIVE SUMMARY

GMB151 soybean was developed through disarmed *Agrobacterium*-mediated transformation using the vector pSZ8832 containing the *cry14Ab-1.b* and *hppdPf-4Pa* gene cassettes. GMB151 soybean produces the Cry14Ab-1 protein, a crystal protein derived from *Bacillus thuringiensis*, which confers resistance to the plant-parasitic nematode, soybean cyst nematode (*Heterodera glycines*). GMB151 also produces a modified 4-hydroxyphenylpyruvate dioxygenase (HPPD-4) derived from *Pseudomonas fluorescens* that confers tolerance to HPPD-inhibitor herbicides such as isoxaflutole. The OECD identifier of GMB151 soybean is BCS-GM151-6.

Molecular characterization studies using next generation sequencing (NGS) and junction sequence analysis (JSA) determined that GMB151 contains, at a single locus, a single complete copy of the *cry14Ab-1.b* gene cassette and a single *hppdPf-4Pa* gene cassette that lacks the 5′ part of the promoter. Generational stability analysis by NGS/JSA demonstrated that the transgenic locus of GMB151 is stably maintained across multiple generations. Segregation data were consistent with Mendelian principles, confirming that GMB151 has a single insert that is stably inherited over generations.

A thorough mammalian safety assessment was conducted for both the Cry14Ab-1 and HPPD-4 proteins as expressed in GMB151 soybean. No adverse effects were observed for either protein. The source organism of the Cry14Ab-1 protein, *B. thuringiensis* (*Bt*), is ubiquitous in the environment, is not known for allergenicity, and has a history of safe use as microbial *Bt*-derived biopesticides. Cry proteins have an established history of safe use and have been used for insect control in crops for over 50 years. The Cry14Ab-1 protein has no significant amino acid sequence similarity to known allergens or toxins, is rapidly degraded in simulated gastric fluid, and exhibited no effects in an acute oral mouse toxicity test. The source organism of the HPPD-4 protein, *P. fluorescens,* is a non-pathogenic bacterium that is ubiquitous in nature and has a history of safe use. HPPD proteins are ubiquitous in nature in nearly all aerobic organisms, including the kingdoms of bacteria, fungi, plants, and animals including mammals. The HPPD-4 protein has no significant amino acid sequence similarity to known allergens or toxins, is rapidly degraded in simulated gastric fluid, and exhibited no effects in an acute oral mouse toxicity test. Additionally, the U.S. Environmental Protection Agency (EPA) has granted a permanent exemption for the requirement of a tolerance for the HPPD-4 protein in all crops when used as a plant-incorporated protectant inert ingredient.

Composition analysis and a comparative assessment demonstrated that GMB151 soybean forage and grain are comparable to that of the non-GM counterpart and reference varieties. These results demonstrate that GMB151 soybean does not pose a plant pest risk and support the food and feed safety assessment.

Agronomic performance of GMB151 was evaluated at 11 field sites across U.S. soybean growing regions. Based on the comparative assessment, agronomic performance of GMB151 soybean is comparable to the non-genetically modified conventional counterpart and reference varieties. These data support the conclusion that GMB151 soybean lacks weediness potential and plant pest risk.

The environmental safety assessment conducted on GMB151 soybean included an evaluation of potential impacts to non-target organisms (NTOs). No adverse effects were observed from NTO species tested with Cry14Ab-1 protein including adult and larval honey bees, two soil-dwelling

organisms (Collembola and earthworms), three predator organisms (two species of ladybird beetle and green lacewing), one aquatic organism (water flea), one mammal (mouse), and one avian species (bobwhite quail). A field assessment also indicated that cultivation of GMB151 soybean was unlikely to have negative effects on non-target free-living nematodes. Based on the environmental safety assessment, the cultivation of GMB151 soybean is unlikely to pose any risk to NTOs at realistic field exposure levels.

Based on the information and data contained in this petition, it is concluded that GMB151 soybean does not have plant pest risk or weediness potential.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS

Abbreviation	Definition
a.i.	Active ingredient
ANOVA	Analysis of variance
APHIS	Animal and Plant Health Inspection Service
BBCH	A scale used to identify the phenological development stages of plants
BC_n	Back cross to the nth generation into conventional variety
BLAST	Basic local alignment search tool
bp	Base pair
Bt	*Bacillus thuringiensis*
Cry14Ab-1	A crystal protein derived from *Bacillus thuringiensis*, expressed by GMB151 soybean for protection against plant-parasitic nematodes
cry14Ab-1.b	Gene sequence coding for Cry14Ab-1 protein in GMB151 soybean
CV	Coefficient of variation
DW	Dry weight
DT_{50}	Degradation time for 50% of the measured substance
ED_{50}	Effective dose for achieving 50% of the measured effect
EEC	Expected environmental concentration
EED	Expected environmental dose
ELISA	Enzyme-linked immunosorbent assay
EPN	Entomopathogenic nematode
EUP	Experimental Use Permit
F_n	The nth filial generation, offspring from an open cross
FASTA	A DNA and protein sequence alignment software package
FDA	Food and Drug Administration
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
FLN	Free-living nematode
FW	Fresh weight
gDNA	Genomic DNA
GLP	Good Laboratory Practice
GM	Genetically modified
GMB151	GM event generated by transforming plasmid pSZ8832, containing the *cry14Ab-1.b* and *hppdPf-4Pa* gene cassettes for the expression of Cry14Ab-1 and HPPD-4 proteins for protection from plant-parasitic nematodes and tolerance to HPPD-inhibitor herbicides, respectively
ha	Hectare
HGA	Homogentisate. An aromatic precursor in the HPPD enzymatic pathway
HPPD	4-Hydroxyphenylpyruvate dioxygenase
HPPD-4	Modified HPPD protein conferring resistance to HPPD inhibitors

Abbreviation	Definition
hppdPf-4Pa	Gene sequence coding for HPPD-4 protein in GMB151 soybean
IFT	Isoxaflutole, HPPD inhibitor active ingredient in certain Group 27 herbicides
ILSI-CCDB	International Life Sciences Institute Crop Composition Database
IPM	Integrated pest management
JSA	Junction sequence analysis
L_n	Nth larval stage of nematode
LC-UV-MS	Liquid chromatography - ultra violet - mass spectroscopy
LLOQ	Lower limit of quantitation
LOQ	Limit of quantification
MDF	Minimum dilution factor
MOE	Margin of exposure
NASS	National Agricultural Statistics Service
NCBI	National Center for Biotechnology Information
NGS	Next generation sequencing
NOEC	No observed effect concentration
NOED	No observed effect dose
NTO	Non-target organism
OCSPP	Office of Chemical Safety and Pollution Prevention
OD	Optical density
OECD	Organisation for Economic Co-Operation and Development
ORF	Open reading frame
PCR	Polymerase chain reaction
PI	Plant Introduction (breeding line)
PPN	Plant-parasitic nematode
pSZ8832	Transforming plasmid containing the *cry14Ab-1.b* and *hppdPf-4Pa* gene cassettes
QC	Quality control
Q-Q	Quantile-quantile
R_n	Reproductive growth stage of soybean
RCBD	Randomized complete block design
RPM	Rotations per minute
SCN	Soybean cyst nematode
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	Simulated (human) gastric fluid
SIF	Simulated (human) intestinal fluid
T_n	The nth generation from in-breeding of the original transformant
TSD	Target site deletion

Abbreviation	Definition
UPLC-UV-MS	Ultra performance liquid chromatography - ultra violet - mass spectroscopy
U.S.	United States
USDA	United States Department of Agriculture
U.S. EPA	United States Environmental Protection Agency
V_n	Vegetative growth stage of soybean

1. RATIONALE FOR THE DEVELOPMENT OF PLANT-PARASITIC NEMATODE-PROTECTED AND HERBICIDE TOLERANT GMB151 SOYBEAN

1.1. Basis for the request for a determination of nonregulated status under 7 CFR § 340.6

The Animal and Plant Health Inspection Service (APHIS) of the U.S. Department of Agriculture (USDA) is responsible under the Plant Protection Act (7 U.S.C. § 7701-7772) for preventing the introduction or dissemination of plant pests into the United States (U.S.). Under APHIS regulation 7 CFR § 340.6, an applicant may petition APHIS to evaluate submitted data to determine that a particular regulated article does not present a plant pest risk and should no longer be regulated. If APHIS determines that the regulated article does not present a plant pest risk, the petition is granted and unrestricted introduction of the article is permitted.

BASF Corporation is submitting this request to APHIS for a determination of nonregulated status for the plant-parasitic nematode-protected and herbicide tolerant GMB151 soybean (OECD unique identifier BCS-GM151-6) and for any progeny derived from GMB151 soybean by traditional breeding methods. As part of this petition, BASF is submitting phenotypic and genotypic experimental data, agronomic data, and safety assessment data for GMB151 soybean as described in 7 CFR § 340.6.

1.2. Description of the trait and intended use of the product

GMB151 soybean was developed through disarmed *Agrobacterium*-mediated transformation using the vector pSZ8832 containing the *cry14Ab-1.b* and *hppdPf-4Pa* gene cassettes. GMB151 soybean produces the Cry14Ab-1 protein, a crystal protein derived from *Bacillus thuringiensis*, which confers resistance to the soybean cyst nematode (SCN) *Heterodera glycines* Ichinohe. GMB151 also produces a modified 4-hydroxyphenylpyruvate dioxygenase (HPPD-4) derived from *Pseudomonas fluorescens* that confers tolerance to HPPD-inhibitor herbicides such as isoxaflutole (IFT).

Soybeans and soybean products have been the leading U.S. agricultural export since 1962. By the early 1970s, soybeans passed corn to become the U.S.'s number one cash crop in terms of sales value, and in 1980, the sales value for soy was higher than other commodity crops (Shurtleff and Aoyagi, 2004).

Unlike the seeds of most other legumes (except the peanut), the soybean is rich in oil and is often called an "oilseed." The value of the soybean, which contains about 18% oil and 35% protein, lies in the fact that there is a strong demand for both of these ingredients, either directly or indirectly, in human foods. The great majority of the world's soybeans are processed by crushing to produce crude soybean oil and soybean meal. During the 1940s, the U.S. greatly increased its lead as the world's major soybean crushing country, and for the first time in history, the majority of the world's soybeans were crushed to yield oil and meal.

Soybean oil as a mostly edible product historically accounts for about 35% of the value of soybeans, with over 80% of the total fats and edible oil being consumed within the U.S. A breakdown of soybean oil uses includes baking and frying oils (46%), salad or cooking oils (43%), margarines (7%), other edible products (1%), and industrial products (3%). It is anticipated

that industrial oil uses will rapidly expand with greater use of biodiesel, polymers, and industrial chemicals developed from soybean oil (Smith, 2017).

Today, soybeans are an important part of U.S. agriculture. Soybeans were planted on more than 90 million acres in 2017, with total production of approximately 4.4 billion bushels (USDA-NASS, 2018). Soybean meal is the protein of choice for feed manufacturers, and the United Soybean Board estimates about 46% of the soybean meal produced is used by broiler chickens, layer chickens, and turkeys. Swine uses another 25% of annual soybean production, followed by beef (13%), dairy (8.5%), pet foods (2.5%), other feed (2%), and food and industrial uses 2.5%, respectively (Smith, 2017).

1.3. Description of the benefits and anticipated adoption of the product

SCN is one of the most yield-limiting pests of soybean in all cropping systems worldwide. SCN is an invasive pest of soybean occurring in most of the soybean growing regions of the U.S. (Niblack and Tylka, 2014) and globally (Riggs and Schmitt, 1989). Yield loss due to SCN can vary depending on location, genetic background of the variety, and cropping system. Losses of up to 30% have been reported (Niblack and Tylka, 2014), translating to economic loss in excess of U.S. $1 billion annually. Use of soybean varieties resistant to SCN coupled with rotations with non-host crops have been the most effective ways of managing SCN infestation.

To date, over 1000 SCN-resistant soybean varieties are available to farmers in the U.S. (Tylka and Mullaney, 2018). The majority of soybean varieties have SCN resistance derived from the breeding line named Plant Introduction (PI) 88788 (Niblack, 2005; Shannon et al., 2004). Two other commercially available sources of SCN resistance, Peking and PI 437654 (also known as Hartwig or CystX®), have limited but increasing availability. Extensive use of PI 88788 SCN-resistant varieties has resulted in decreasing effectiveness against the SCN pest in some parts of the Midwest U.S. (McCarville et al., 2017). Durable plant resistance is most likely to be achieved by combining multiple pest resistance traits, each of which contribute to pest population suppression through unique biological mechanisms (Parlevliet and Zadoks, 1977). GMB151 soybean expressing Cry14Ab-1 will be combined with commercially available SCN resistant varieties. The addition of GMB151 soybean to these SCN native resistant varieties will extend the durability of both GMB151 and native resistance against the SCN pest.

Furthermore, planting GMB151 soybean varieties will provide growers with an additional HPPD-inhibitor herbicide option for weed control.

Weed management in soybean is critical, especially during early vegetative growth and also during the early reproductive growth stages (Van Acker et al., 1993). Herbicides are essential for weed management in all cropping systems as they provide cost-effective means of weed control, thereby increasing profitability of the farming enterprise. The use of HPPD-inhibitor herbicides also offers an alternative mode of action for the control of herbicide resistant weeds.

1.4. Submissions to other regulatory agencies

Under the Coordinated Framework for Regulation of Biotechnology, the Office of Science and Technology Policy describes the U.S. Government agencies responsible for oversight of the products of modern agricultural biotechnology including USDA-APHIS, the U.S. Environmental

Protection Agency (U.S. EPA), and the U.S. Department of Health and Human Services' Food and Drug Administration (FDA).

As the FDA is responsible for ensuring the safety and proper labeling of all plant-derived food and feed, including those developed through modern biotechnology, BASF Corporation, in accordance with this policy, has initiated a consultation with the FDA that included molecular, composition, and nutrition data, as well as other food and feed safety assessment data related to plant-parasitic nematode-protected GMB151soybean.

The U.S. EPA, through a registration process, regulates the sale, distribution, and use of pesticides to protect human and animal health and the environment. The Cry14Ab-1 protein produced by GMB151 is a plant-incorporated protectant (PIP) and is regulated as a pesticide by the U.S. EPA. Currently, an Experimental Use Permit (EUP) as described under Section 5 of Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), is in effect for Cry14Ab-1 and the genetic material responsible for its production in GMB151 soybean. An application for a FIFRA Section 3 seed increase registration was submitted to the U.S. EPA on November 5, 2018 for Cry14Ab-1 and the genetic material responsible for its production in GMB151 soybean. The U.S. EPA also sets tolerance limits for residues of pesticides on and in food and feed or establishes an exemption from the requirement for a tolerance under the Federal Food, Drug, and Cosmetic Act (FFDCA). A temporary exemption from the requirement for a tolerance was granted to Cry14Ab-1 in conjunction with the previously mentioned EUP. In addition, the U.S. EPA established a permanent exemption from the requirement for a tolerance for the HPPD-4 protein expressed in all food commodities when used as an inert ingredient. A petition for a permanent exemption from the requirement for a tolerance for the Cry14Ab-1 protein when expressed in soybean was submitted to the U.S. EPA on November 5, 2018 in conjunction with the Section 3 registration submission.

Consistent with a commitment to Excellence Through Stewardship®, BASF will meet applicable regulatory requirements for GMB151 soybean in the country of intended production and for key import countries with functioning regulatory systems based on a market and trade assessment and the intended use of the product. This will assure regulatory compliance, maintain product integrity, and assist in minimizing the potential for trade disruptions.

2. THE BIOLOGY OF SOYBEAN (*Glycine max*)

The OECD consensus document on soybean biology (OECD, 2000) provides information pertaining to the following aspects of soybean biology:

- Taxonomy, morphology, and uses
- Centers of origin of the species and domestication
- Agronomic practices
- Reproductive biology and dispersal
- Genetics and hybridization
- Interactions with other organisms

3. DEVELOPMENT OF GMB151 SOYBEAN

3.1. Description of the transformation technology

3.1.1. Transformation methods

Explants from soybean variety Thorne were prepared and were transformed with the transformation vector pSZ8832 using a disarmed *Agrobacterium tumefaciens* transformation method.

Soybean variety Thorne explants were exposed to a culture of a disarmed *A. tumefaciens*, strain LBA4404 (Hoekema et al., 1983), harboring two plasmids: the helper Ti-plasmid pAL4404 and the T-DNA region-containing transformation vector pSZ8832, derived from plasmid pGSC1700 (Cornelissen and Vandewiele, 1989). After the exposure, the explants were transferred to selection media containing the selection agent tembotrione to select for transformed cells and ticarcillin to eliminate residual *A. tumefaciens*. Transformed shoots were transferred to fresh selection media for shoot elongation and then subsequently transferred to a rooting medium. Rooted shoots were transferred to the greenhouse where they were allowed to flower and set seed.

3.1.2. Description of the transformation vector and gene construct

The vector pSZ8832 is derived from pGSC1700 (Cornelissen and Vandewiele, 1989). The map of the vector pSZ8832 is presented in Figure 1, and the genetic elements are described in Table 1.

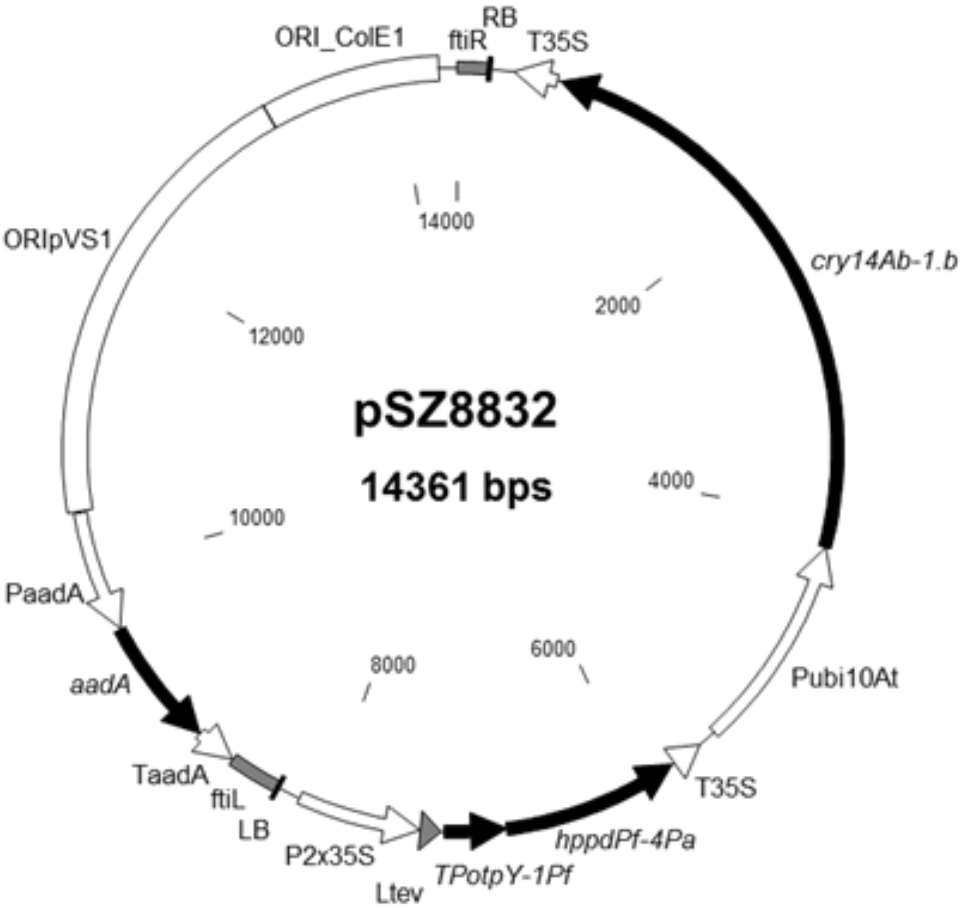


Figure 1. Map of vector pSZ8832

Table 1. Description of the genetic elements of pSZ8832

Nt Positions	Orientation	Origin
1 - 184		**ftiR:** Ti-plasmid sequence of pTiAch5 flanking the T-DNA right border region (Zhu et al., 2000)
185 - 189		Polylinker sequences: sequence used in cloning
190 - 214		**RB:** right border region of the T-DNA of *Agrobacterium tumefaciens* (Zambryski, 1988)
215 - 344		Polylinker sequences: sequence used in cloning
345 - 614	Counter clockwise	**T35S:** sequence including the 3´ untranslated region of the 35S transcript of the Cauliflower Mosaic Virus (Sanfaçon et al., 1991)
615 - 625		Polylinker sequences: sequence used in cloning
626 - 4183	Counter clockwise	***cry14Ab-1.b:*** coding sequence of the delta-endotoxin gene of *Bacillus thuringiensis* (GenBank accession number: AGU13817.1)
4184 - 5490	Counter clockwise	**Pubi10At:** sequence including the promoter region of ubiquitin-10 gene of *Arabidopsis thaliana* (Grefen et al., 2010)
5491 - 5595		Polylinker sequences: sequence used in cloning
5596 - 5790	Counter clockwise	**T35S:** sequence including the 3´ untranslated region of the 35S transcript of the Cauliflower Mosaic Virus (Sanfaçon et al., 1991)
5791 - 5802		Polylinker sequences: sequence used in cloning
5803 - 6879	Counter clockwise	***hppdPf-4Pa:*** coding sequence of a variant 4-hydroxyphenylpyruvate dioxygenase gene of *Pseudomonas fluorescens* (Porée et al., 2014)
6880 - 7251	Counter clockwise	***TPotpY-1Pf:*** coding sequence of an optimized transit peptide derivative (position 55 changed into Tyr), containing sequence of the RuBisCO small subunit genes of *Zea mays* and *Helianthus annuus* (Lebrun et al., 1996)
7252 - 7272		Polylinker sequences: sequence used in cloning
7273 - 7399	Counter clockwise	**Ltev:** sequence including the leader sequence of the Tobacco Etch Virus genomic RNA (Allison et al., 1985)
7400 - 7405		Polylinker sequences: sequence used in cloning
7406 - 8155	Counter clockwise	**P2x35S:** sequence including the double enhanced promoter region of the Cauliflower Mosaic Virus 35S genome transcript (Kay et al., 1987)
8156 - 8282		Polylinker sequences: sequence used in cloning
8283 - 8307		**LB:** left border region of the T-DNA of *Agrobacterium tumefaciens* (Zambryski, 1988)
8308 - 8612		**ftiL:** Ti- plasmid sequence of pTiAch5 flanking the T-DNA left border region (Zhu et al., 2000)
8613 - 8864	Counter clockwise	**TaadA:** sequence including the 3′ termination region of the aminoglycoside adenyltransferase gene of transposon Tn7 of *Escherichia coli* (Fling et al., 1985)
8865 - 9656	Counter clockwise	***aadA*:** the coding sequence of the aminoglycoside adenyltransferase gene (*aadA*) of transposon Tn7 of *Escherichia coli* (Fling et al., 1985)

Table 1. Description of the genetic elements of pSZ8832 (continued)

Nt Positions	Orientation	Origin
9657 - 10394	Counter clockwise	**PaadA:** sequence including the promoter region of the aminoglycoside adenyltransferase gene of transposon Tn7 of *Escherichia coli* (Fling et al., 1985)
10395 - 10400		Polylinker sequences: sequence used in cloning
10401 - 13187		**ORIpVS1:** fragment including the origin of replication of the plasmid pVS1 of *Pseudomonas aeruginosa* (Heeb et al., 2000)
13188 - 14251		**ORI_ColE1:** fragment including the origin of replication of the plasmid pBR322 for replication in *Escherichia coli* (Bolivar et al., 1977)
14252 - 14361		Polylinker sequences: sequence used in cloning

Nt = nucleotide

3.2. Description of the breeding process for the parent organism

Following *Agrobacterium*-mediated transformation of the conventional breeding line Thorne, the GMB151 T_0 plant was self-pollinated to generate T_1 seed. All subsequent T generations were produced through self-pollination. The T_3 generation was used for crossing into conventional soybean lines.

The breeding diagram is shown in Figure 2 below. Table 2 describes the GMB151 generations used for analysis and the associated reports describing these studies.

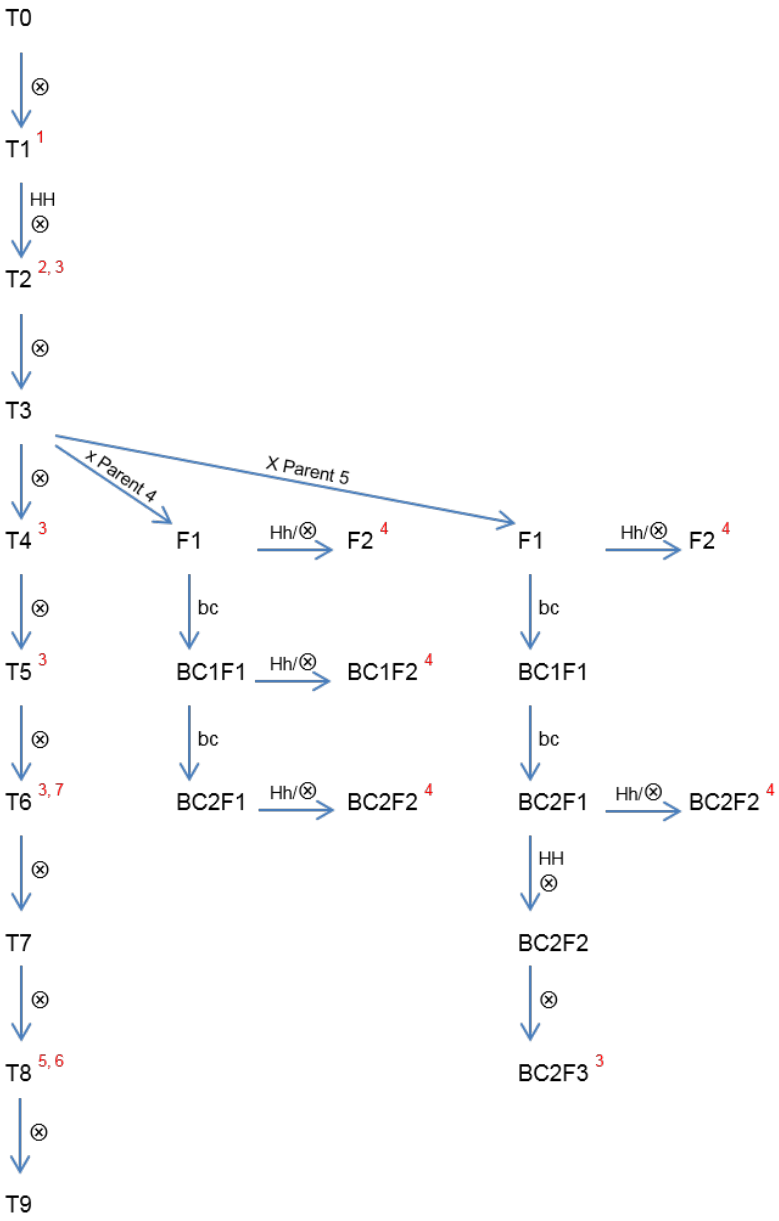


Figure 2. Breeding diagram of GMB151 soybean

Legend:

T0 - original transformant

⊗ - self pollination

bc - backcross

x - cross

HH - selection of homozygous plants

Hh - selection of hemizygous plants

Red numbers indicate generations used for particular experiments (see Table 2).

Table 2. Generations used for analysis of GMB151

No. in tree [a]	Experiment	Generation(s)	Comparator
1	DNA sequencing of insert and flanking region	T_1	Thorne
2	Molecular characterization by NGS/JSA	T_2	Thorne
3	Stable inheritance by NGS/JSA	T_2, T_4, T_5, T_6, BC_2F_3 (parent 5)	Thorne, parent 5
4	Inheritance of the insert	F_2 (parent 4) BC_1F_2 (parent 4) BC_2F_2 (parent 4) F_2 (parent 5) BC_2F_2 (parent 5)	None
5	Agronomic and phenotypic analysis	T_8	Thorne
6	Composition analysis [b]	T_8	Thorne
7	Protein expression analysis [b]	T_6	Thorne

[a] Numbers in this column correspond to the red numbers in Figure 2.
[b] The grain analyzed in the study is the next generation from that planted.

4. GENETIC MATERIAL USED FOR TRANSFORMATION OF GMB151 SOYBEAN

4.1. Description of the transferred genes and gene products

***Cry14Ab-1.b* gene and Cry14Ab-1 protein**

Cry14Ab-1 protein is a member of the Cry (crystal)-type protein family with a fully conserved three-domain structure. Cry proteins are produced by *B. thuringiensis* strains and demonstrate specific toxicity towards either insects or nematodes. Cry14Ab-1 belongs to the "nematicidal branch" of Cry proteins, with the closest homology to Cry14Aa1 (87% identity) (Sanahuja et al., 2011).

The *cry14Ab-1.b* gene encodes for the Cry14Ab-1 protein. The gene was identified with activity against the nematode *Caenorhabditis elegans* in a bioassay using *B. thuringiensis* organisms but also was found to be active against the plant-parasitic nematode SCN *in planta*. The amino acid sequence of Cry14Ab-1 is provided in Figure 3.

***hppdPf-4Pa* gene and HPPD-4 protein**

To create the *hppdPf-4Pa* gene, the HPPD coding sequence of *P. fluorescens* strain A32 was modified at positions 335, 336, 339, and 340 as follows: a glutamic acid residue was changed for a proline residue at position 335, a glycine residue for a tryptophan residue at position 336, a lysine residue for an alanine residue at position 339, and an alanine residue for a glutamine residue at position 340. The four introduced mutations lead to reduced HPPD-inhibitor herbicide binding efficacy. The modified protein is designated as HPPD-4. The amino acid sequence of HPPD-4 is provided in Figure 4.

```
1    MDCNLQSQQN IPYNVLAIPV SNVNSLTDTV GDLKKAWEEF QKTGSFSLTA LQQGFSASQG
61   GTFNYLTLLQ SGISLAGSFV PGGTFVAPII NMVIGWLWPH KNKNADTENL INLIDSEIQK
121  QLNKALLDAD RNEWSSYLES IFDSSNNLNG AIVDAQWSGT VNTTNRTLRN PTESDYTNVV
181  TNFIAADGDI ANNENHIMNG NFDVAAAPYF VIGATARFAA MQSYIKFCNA WIDKVGLSDA
241  QLTTQKANLD RTKQNMRNAI LNYTQQVMKV FKDSKNMPTI GTNKFSVDTY NVYIKGMTLN
301  VLDIVAIWPS LYPDDYTSQT ALEQTRVTFS NMVGQEEGTD GSLRIYNTFD SFSYQHSPIP
361  NNNVNLISYY NDELQNLELG VYTPPKKGSG YSYPYGFVLN YANSKYKYGD SNDPESLGGL
421  STLSAPIQQV NAATQNSKYL DGEILNGIGA SLPGYCTTGC SPTEPPFSCT STANGYKASC
481  NPSDTNQKIN ALYPFTQANV KGNTGKLGVL ASLVSYDLNP KNVFGELDSD TNNVILKGIP
541  AEKGYFPNNA RPTVVKEWIN GASAVPLDSG NTLFMTATNL TATQYRIRIR YANPNSNTQI
601  GVRITQNGSL ISSSNLTLYS TTDMNNTLPL NVYVIGENGN YTLQDLYNTT NVLSTGDITL
661  QITGGDQKIF IDRIEFVPTM PVPGNTNNNN GNNNGNNNPP HHVCAIAGTQ QSCSGPPKFE
721  QVSDLEKITT QVYMLFKSSP YEELALEVSS YQISQVALKV MALSDELFCE EKNVLRKLVN
781  KAKQLLEASN LLVGGNFETT QNWVLGTNAY INYDSFLFNG NYLSLQPASG FFTSYAYQKI
841  DESTLKPYTR YKVSGFIGQS NQVELIISRY GKEIDKILNV PYAGPLPITA DASITCCAPE
901  IGQCDGEQSD SHFFNYSIDV GALHPELNPG IEIGLKIVQS NGYITISNLE IIEERPLTEM
961  EIQAVNRKNQ KWEREKLLEC ASISELLQPI INQIDSLFKD GNWYNDILPH VTYQDLKNII
1021 IPELPKLKHW FIENLPGEYH EIEQKMKEAL KYAFTQLDEK NLIHNGHFTT NLIDWQVEGD
1081 AQMKVLENDA LALQLFNWDA SASQSINILE FDEDKAYKLR VYAQGSGTIQ FGNCEDEAIQ
1141 FNTNSFIYQE KIVYFDTPSV NLHIQSEGSE FIVSSIDLIE LSDDQ
```

Figure 3. Amino acid sequence of the Cry14Ab-1 protein

```
  1 MADLYENPMG LMGFEFIEFA SPTPGTLEPI FEIMGFTKVA THRSKNVHLY RQGEINLILN
 61 NEPNSIASYF AAEHGPSVCG MAFRVKDSQK AYNRALELGA QPIHIDTGPM ELNLPAIKGI
121 GGAPLYLIDR FGEGSSIYDI DFVYLEGVER NPVGAGLKVI DHLTHNVYRG RMVYWANFYE
181 KLFNFREARY FDIKGEYTGL TSKAMSAPDG MIRIPLNEES SKGAGQIEEF LMQFNGEGIQ
241 HVAFLTDDLV KTWDALKKIG MRFMTAPPDT YYEMLEGRLP DHGEPVDQLQ ARGILLDGSS
301 VEGDKRLLLQ IFSETLMGPV FFEFIQRKGD DGFGPWNFAQ LFESIERDQV RRGVLTAD
```

Figure 4. Amino acid sequence of the HPPD-4 protein

5. CHARACTERIZATION OF THE GENETIC MODIFICATIONS IN GMB151 SOYBEAN

Molecular analysis was conducted to characterize the transgenic locus of GMB151 soybean, including determination of the number, arrangement, and copy number of T-DNA insertions, the DNA sequence of the GMB151 transgenic locus and corresponding insertion locus, the absence/presence of vector backbone sequence, the stable inheritance and segregation ratios of the GMB151 locus across generations, and bioinformatics analysis of the GMB151 transgenic and insertion loci. Methods used for the characterization included next generation sequencing (NGS)/junction sequence analysis (JSA), Sanger sequencing, and bioinformatics analysis.

5.1. Number of insertion sites, and arrangement and copy number of transferred DNA

Insert characterization by next generation sequencing and junction sequence analysis

The inserted sequences of GMB151 soybean were characterized by means of NGS and JSA on genomic DNA (gDNA) prepared from seeds (Appendix 2).

To characterize the inserted DNA sequences in GMB151 soybean, a JSA was performed on whole-genome sequence data from GMB151 soybean (T_2 generation) as well as appropriate control samples. The non-genetically modified (non-GM) counterpart Thorne was used as negative control sample. As a positive control, the non-GM counterpart Thorne was supplemented with an equimolar amount of the transforming plasmid pSZ8832. The non-GM counterpart Thorne supplemented with 1/10th equimolar amount of the plasmid pSZ8832 was used as a sensitivity control sample.

Ready-to-sequence library pools with an average fragment size of 500–600 base pairs (bp) were prepared for each GMB151 soybean sample and for the negative, positive, and sensitivity control samples and used as templates for whole-genome sequencing. Whole genome sequencing was performed using the Illumina® HiSeq™ 2500 technology. For each fragment, paired-end reads of 125 bp each were obtained.

Low quality sequencing data were trimmed. The remaining sequencing reads were mapped to both the soybean genome and transforming plasmid reference sequences. Duplicate reads, which were the result of a polymerase chain reaction (PCR) amplification step during library preparation, were removed, and the effective median genome coverage was examined by the alignment of the reads to a known single copy locus of the soybean genome (*lectin* gene, Glyma.02G012600, Chr02:1123507–1125658). The median coverage depth of a known single copy locus represented the genome coverage at the insertion locus and was higher than 75-fold for each of the GMB151 soybean and control samples.

Alignment of the sequencing reads to the transforming plasmid reference sequence demonstrated that the median coverage depths of the transforming plasmid in the positive and sensitivity control samples were 96-fold and 7-fold, respectively, and that 100% of the transforming plasmid sequence was covered. This demonstrated that the coverage was adequate to perform a high quality JSA. A 100% nucleotide identity between the obtained transforming plasmid sequences and the transforming plasmid reference sequence was obtained, demonstrating that the entire transforming plasmid sequence was amenable to sequencing with the technology used.

Alignment of sequencing reads of the negative control sample to the transforming plasmid reference sequence demonstrated that the median coverage depth was 0-fold. The few reads that did map to the transformation plasmid sequence were the result of low quality or short sequence similarities between the transformation plasmid and host genome sequences.

An in-depth molecular characterization was performed through the identification of all novel sequence junctions created upon transformation. A novel junction created upon transformation is typically covered by sequence reads containing sequence from both pSZ8832 and the soybean genome (junction reads). Therefore, reads partially mapping to the transforming plasmid were selected and further analyzed using a bioinformatics JSA. Based on the results of the parameter optimization for the bioinformatics JSA pipeline, aiming for the optimal sensitivity and specificity, the minimum length of the mapped subsequence was set to 30 bp. Hence, insertions of plasmid sequences smaller than 30 bp would be below the detection limit.

The junction positions in the transforming plasmid that separated the mapped portion from the unmapped portion of the junction reads were collected. The selected junction reads were grouped based on their junction position(s) within the transforming plasmid. The junction position(s) determine the portion of the plasmid that is inserted in GMB151.

Due to the presence of a sequence repeat in the P2x35S promoter of the transforming plasmid, the correct position of some junction reads could not be defined, leading to two possible junction positions. This was resolved by broken pair analysis. Broken pairs are sequence reads coming from both ends of a DNA fragment that either map to a different target sequence or that map to the same target sequence but with an unexpected insert length or non-standard orientation. Analysis of the broken pairs defined the correct position of the junction.

The consensus sequences of all junction reads corresponding to the same junction position within the transforming plasmid were generated. The junctions were further identified as “plasmid/flank” or “plasmid/plasmid” junctions. An insertion site is defined as having two unique plasmid/flank junctions. A plasmid/plasmid junction would be the result of rearrangements within the inserted plasmid sequences.

Results

JSA of GMB151 identified two unique plasmid/flank junctions and no plasmid/plasmid junctions. This result demonstrated that GMB151 soybean contains one copy of the T-DNA insert, without rearrangements, at a single insertion site. The *cry14Ab-1.b* gene cassette is complete, and the *hppdPf-4Pa* gene cassette has a truncation at the 5’ end of the P2x35S promoter of 482 bp.

Alignment of junction reads and their mates to the transforming plasmid reference sequence demonstrated that the junctions in pSZ8832 are at positions 213 and 7673. Mapping of all reads to the transforming plasmid showed that the region between these two junction positions is sufficiently covered.

BLAST searches performed to identify the host genomic sequences in the flanking regions identified 39 bp of filler DNA between the GMB151 T-DNA sequence and the 3’ flanking genomic sequences. Twenty-one bp of this filler DNA showed sequence identity to the ORIpVS1 in the vector backbone region of the transformation plasmid. The short size of this fragment was below

the limit of detection (30 bp) in the NGS/JSA analysis. The filler DNA also contained 17 bp with sequence identity to the 3' flanking genomic region.

In conclusion, these results demonstrate that GMB151 soybean contains one copy of the T-DNA insert, without rearrangements, at a single insertion site in the soybean genome. The *cry14Ab-1.b* gene cassette is complete, and the 5' end of the P2x35S promoter of the *hppdPf-4Pa* gene cassette is truncated by 482 bp. The flanking region 3' to the T-DNA has 39 bp of filler DNA, consisting of 21 bp with sequence identity to the ORIpVS1 in the vector backbone region of the transformation plasmid and a 17 bp duplication of genomic sequence in the 3' flanking genomic region.

5.2. DNA sequence of the transgenic and insertion loci

The DNA sequence of the GMB151 soybean transgenic locus and the corresponding insertion locus was determined.

Two overlapping PCR fragments were prepared for the determination of the GMB151 transgenic locus, using GMB151 gDNA as a template. To determine the GMB151 insertion locus, one fragment was amplified from gDNA extracted from the non-GM counterpart Thorne (Table 3). Sanger sequencing was performed using the ABI PRISM® BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

The obtained consensus sequences of the transgenic and insertion loci were annotated by pairwise alignments using the Clone Manager software (Sci-Ed Central). An alignment between the GMB151 transgenic locus and the GMB151 insertion locus sequence was made to identify sequence regions of soybean origin within the GMB151 transgenic locus as well as the target site deletion (TSD) within the GMB151 insertion locus. The consensus sequence of the GMB151 transgenic locus was compared with the pSZ8832 sequence to identify the T-DNA region. Further sequence annotation within the T-DNA was performed by comparing the GMB151 transgenic locus sequence with each feature of the pSZ8832 T-DNA region.

By means of pairwise alignment, two large sections of sequence identity were found between the GMB151 soybean transgenic locus sequence and the GMB151 soybean insertion locus sequence. These regions are identified as the 5′ flanking genomic sequence (Region A of Table 4) and the 3′ flanking genomic sequence (Region B of Table 4) of the GMB151 soybean transgenic and insertion loci sequences. One base pair difference between the GMB151 soybean transgenic locus (position 8675) and the GMB151 soybean insertion locus (position 1240) was identified. Sequence differences between plants are commonly found and are due to natural genetic variation (Zhu et al., 2003).

In the GMB151 soybean insertion locus sequence, 63 bp of gDNA were observed that are not present in the GMB151 transgenic locus sequence. These base pairs were deleted during the transformation process and are referred to as the TSD.

Pairwise alignment between the GMB151 transgenic sequence and the pSZ8832 plasmid sequence identified one large region sharing 100% sequence identity (Table 5). Region D on the GMB151 soybean transgenic locus sequence is identical to the T-DNA region of pSZ8832. Two base pairs at the 5′ end of the T-DNA region (bp 999–1000) were identical to both the transforming plasmid pSZ8832 and the GMB151 soybean insertion locus sequence. These base pairs were

annotated as 5′ flanking genomic sequence. Base pairs 1001–8459 of the GMB151 soybean transgenic locus (7459 bp) were annotated as T-DNA.

The GMB151 soybean transgenic locus sequence corresponds to one copy of the T-DNA region of transforming vector pSZ8832, which includes the complete *cry14Ab-1.b* gene cassette and an incomplete *hppdPf-4Pa* gene cassette, lacking 482 bp of the P2x35S promoter.

Between the T-DNA sequences and the 3′ flanking genomic sequences, 39 bp of filler DNA were found. Twenty-one bp of this filler DNA show sequence identity to a region containing part of ORIpVS1 from the transforming plasmid pSZ8832 (Region E in Table 5), and 17 bp have sequence identity to the 3′ flanking genomic sequence (Region C in Table 4). Considering the small size of these separate regions showing sequence identity, the 39 bp sequence was annotated as filler DNA.

A schematic representation of the GMB151 soybean transgenic locus sequence in relation to the transforming plasmid pSZ8832 and the GMB151 soybean insertion locus is provided in Figure 5.

In summary, the results demonstrated that upon transformation, 63 bp from the non-GM counterpart were replaced by 7498 bp of inserted sequences, including 7459 bp of T-DNA sequence and 39 bp of filler DNA. Part of the filler DNA (21 bp) shows sequence identity to ORIpVS1 in the vector backbone region from pSZ8832, and another part (17 bp) shows sequence identity to the 3’ flanking genomic sequence. The flanking genomic sequences obtained at the transgenic locus were similar to the homologous sequence from the insertion locus, demonstrating the soybean origin. The GMB151 soybean transgenic locus contains one copy of the T-DNA region of transforming vector pSZ8832, comprising a complete *cry14Ab-1.b* gene cassette and an *hppdPf-4Pa* gene cassette lacking 482 bp from the 5’ end of the P2x35S promoter.

Table 3. Overview of the sequencing fragments prepared

	Fragment ID	Template DNA	Primer pair	Target	Length of the consensus sequence (bp)
GMB151 transgenic locus	FR-VLT137-A-01	GMB151 soybean	GLPA210 GLPB167	GMB151 - 5′ integration fragment	9498
	FR-VLT137-A-02		GLPB170 GLPA212	GMB151 - 3′ integration fragment	
GMB151 insertion locus	FR-VLT137-A-03	Non-GM counterpart	GLPA210 GLPA212	GMB151 - insertion locus	2063

Table 4. Results of the pairwise alignment between the sequence of the GMB151 soybean transgenic and insertion loci

Region showing sequence identity	% matches	Length (bp)	GMB151 soybean insertion locus sequence		GMB151 soybean transgenic locus sequence	
			Start	End	Start	End
Region A: 5′ flanking genomic sequence	100	1000	1 bp	1000 bp	1 bp	1000 bp
Region B: 3′ flanking genomic sequence	99*	1000	1064 bp	2063 bp	8499 bp	9498 bp
Region C: Part of filler DNA	100	17	1090 bp	1106 bp	8482 bp	8498 bp

* One base pair difference between the GMB151 soybean transgenic locus (3′ flanking genomic sequence position 8675) and the GMB151 soybean insertion locus (position 1240) was found. Sequence differences between plants are commonly found and are due to natural genetic variation (Zhu et al., 2003).

Table 5. Results of the pairwise alignment between the sequence of the GMB151 soybean transgenic locus and the transforming plasmid pSZ8832 sequence

Region showing sequence identity	% matches	Length (bp)	Transforming plasmid pSZ8832		GMB151 soybean transgenic locus sequence	
			Start	End	Start	End
Region D: T-DNA	100	7461 *	213 bp	7673 bp	999 bp	8459 bp
Region E: Part of filler DNA	100	21	11923 bp	11943 bp	8481bp	8461 bp

* Two base pairs at the 5′ end of the T-DNA region (bp 999–1000) were identical to both the transforming plasmid pSZ8832 and the GMB151 soybean insertion locus sequence. These base pairs were annotated as 5′ flanking genomic sequence. Base pairs 1001–8459 of the GMB151 soybean transgenic locus (7459 bp) were annotated as T-DNA.

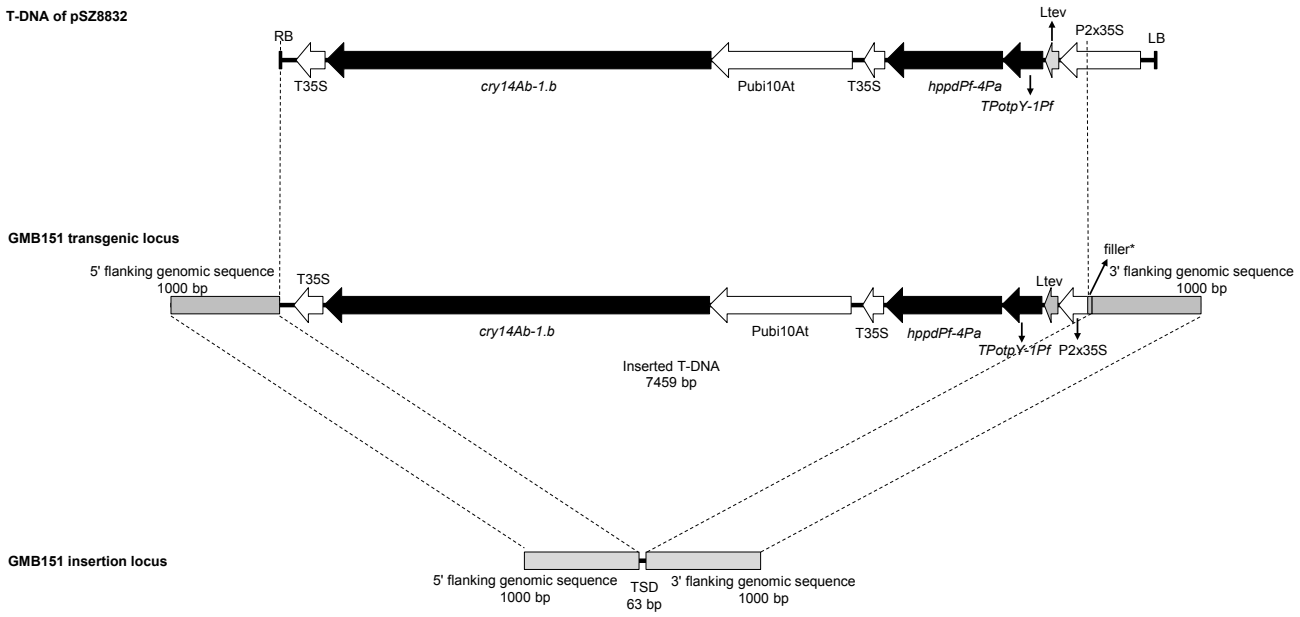


Figure 5. Schematic drawing of the GMB151 soybean transgenic locus in relation to the GMB151 insertion locus and the T-DNA of the transforming plasmid pSZ8832

TSD = target site deletion.
* 39 bp filler DNA contains 21 bp showing sequence identity to a region containing ORIpVS1 from the transforming plasmid pSZ8832 and 17 bp showing sequence identity to the 3′ flanking genomic sequence.

5.3. Presence of vector and/or other DNA of non-host origin

The presence of vector backbone sequences in GMB151 soybean was investigated by NGS/JSA on the T_2 generation as described in Section 5.1.

JSA of GMB151 identified only two unique plasmid/flank junctions, at positions 213 and 7673 of the transformation plasmid pSZ8832. Since these junction positions are in the T-DNA and not within the vector backbone of the transformation plasmid, the absence of vector backbone in GMB151 soybean was indicated. However, consensus sequences were generated from the junctions, and BLAST searches of these sequences, performed to identify host genomic sequences in the flanking regions, identified 39 bp of filler DNA between the GMB151 T-DNA sequence and the 3' flanking genomic sequences. Twenty-one bp of this filler DNA showed sequence identity to the ORIpVS1 in the vector backbone region of the transformation plasmid. The short size of this fragment was below the limit of detection (30 bp) in the NGS/JSA. In conclusion, the GMB151 transgenic locus has 21 bp containing sequence identity to ORIpVS1 in the pSZ8832 vector backbone region adjacent to the 3' end of the T-DNA insert. These 21 bp do not originate from an antibiotic resistance gene.

5.4. Stable inheritance of the insert over generations

The stable inheritance of the GMB151 insertion locus across generations was investigated by performing NGS/JSA on plants from five different breeding generations (T_2, T_4, T_5, T_6, and BC_2F_3) as described in Section 5.1 and Appendix 2.

For each of the GMB151 breeding generations, two plasmid/flank junctions were determined after NGS/JSA. A single copy, single insert in the genome has two unique plasmid/flank junctions. The positions of both junctions within the plasmid were the same in all generations (Table 6). Also, multiple sequence alignments between the obtained junction consensus sequences showed that both novel junctions are conserved between all breeding generations analyzed.

In summary, the GMB151 transgenic locus contains a single copy, single T-DNA insert that is stably maintained across different breeding generations.

Table 6. Junctions identified in GMB151 soybean over generations using JSA

Sample	Number of junctions	Junction positions in pSZ8832
GMB151 soybean - T_2	2	213 and 7673
GMB151 soybean - T_4	2	213 and 7673
GMB151 soybean - T_5	2	213 and 7673
GMB151 soybean - T_6	2	213 and 7673
GMB151 soybean - BC_2F_3	2	213 and 7673
Non-GM counterparts	0	-

Following background subtraction

5.5. Mendelian inheritance of the insert

The segregation ratios of the GMB151 insert over five segregating generations were determined to confirm inheritance of the insert in a predictable and stable manner according to Mendelian inheritance principles.

Genomic DNA from individual plants of five segregating GMB151 soybean generations was tested for the genotype of GMB151 by PCR. In addition, the presence or absence of the *cry14Ab-1.b* and *hppdPf-4Pa* genes was confirmed. The results from event-specific PCR analysis were used to calculate the segregation ratios of the GMB151 insert. Chi-square (χ^2) analysis of the segregation data was performed to test the hypothesis that the GMB151 soybean insert is inherited in a manner that is predictable according to Mendelian principles and is consistent with insertion into a single locus.

To determine the genotype of the GMB151 soybean insert, plant samples were analyzed using event-specific PCR. This included the amplification of the GMB151 event-specific sequence and the amplification of the pre-insertion locus sequence. Samples with fluorescence signal corresponding to the GMB151 event-specific sequence only were recorded as homozygous for the GMB151 insert. Samples with fluorescence signal corresponding to the pre-insertion locus sequence only were recorded as null segregant. Samples with fluorescence signal corresponding

to the GMB151 event-specific sequence and the pre-insertion locus sequence were recorded as hemizygous for the GMB151 insert.

To confirm the presence or absence of the *cry14Ab-1.b* and *hppdPf-4Pa* genes, plant samples were also analyzed using gene-specific PCR. This PCR analysis included the amplification of the gene-specific sequence and the amplification of an endogenous gene sequence. Samples with signal corresponding to the gene-specific sequence and the endogenous sequence were recorded as positive for the gene tested. Samples with signal corresponding to the endogenous sequence only were recorded as negative. The gene-specific PCR analysis confirmed that the *cry14Ab-1.b* and *hppdPf-4Pa* genes are present for samples positive for GMB151 and are absent for samples negative for GMB151.

Chi-square analysis was performed to confirm the segregation and stability of the GMB151 insert. The Chi-square analysis is based on testing the observed segregation ratio relative to the segregation ratio expected from Mendelian inheritance principles. The expected segregation ratio of homozygous, hemizygous, and null segregates is 1:2:1. The critical value to reject the hypothesis of a 1:2:1 ratio at the 5% significance level with two degrees of freedom is 5.99 (Strickberger, 1976). The Chi-square values were calculated using the following equation.

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

The results are summarized in Table 7.

Segregation ratios determined for plants from two F_2, two BC_2F_2, and one BC_1F_2 generations of GMB151 soybean confirmed that the GMB151 insert is inherited as expected for a single insert. These data are consistent with Mendelian principles and support the conclusion that GMB151 soybean consists of a single insert integrated at a single locus within the soybean genome that is stably inherited over generations. In addition, the presence or absence of the *cry14Ab-1.b* and *hppdPf-4Pa* genes was confirmed.

Table 7. Observed versus expected (1:2:1) identity for five generations of GMB151 soybean as determined by PCR analysis

GMB151 Insert	F_2 (14MRGM010348)		BC_2F_2 (16MRGM006629)		F_2 (15MRGM011089)		BC_1F_2 (16MRGM005436)		BC_2F_2 (16MRGM012834)	
	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected
Homozygous	56	61.5	57	61.5	54	54	62	55.5	56	57.25
Hemizygous	133	123	120	123	102	108	101	111	104	114.5
Wild-type	57	61.5	69	61.5	60	54	59	55.5	69	57.25
Total Number of Plant Samples	246		246		216		222		229	
χ^2 Value *	1.634		1.317		1.000		1.883		3.402	

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

5.6. Bioinformatics analyses of the insertion and transgenic loci

Bioinformatics analysis of the GMB151 soybean insertion locus

A bioinformatics analysis was performed on the GMB151 insertion locus sequence to identify the position of the insertion locus in the genome and to determine whether regulatory sequences or endogenous soybean genes were interrupted upon the insertion of T-DNA sequences.

The GMB151 insertion locus sequence was used as the query sequence. The GMB151 insertion locus consists of a 5′ and a 3′ flanking genomic sequence region and a TSD of 63 bp that is removed upon integration of the GMB151 T-DNA.

BLASTn compares a nucleotide query sequence against a nucleic sequence database. This tool was used to identify similarities between the GMB151 insertion locus and sequences within the nucleotide collection and the Expressed Sequence Tag (EST) databases available on the National Center for Biotechnology Information (NCBI) website. The MEGABLAST algorithm (Zhang et al., 2000), which uses parameters that efficiently find long alignments between very similar sequences, and the *G. max* dataset of the NCBI Genomic Reference Sequences database were used to locate the GMB151 insertion locus in the soybean genome.

BLASTx compares the six-frame theoretical translation products of the nucleotide query sequence (both strands) against a protein sequence database. A BLASTx search of the GMB151 insertion locus sequence against the NCBI non-redundant protein database was performed.

The bioinformatics analysis demonstrated that the GMB151 soybean insertion locus originates from *G. max* chromosome 7. The results indicate that the insertion of T-DNA sequences in GMB151 soybean is located in the 3′ untranslated region (UTR) of a putative endogenous gene annotated as a BON1-associated protein 1 (BAP1)-like protein. Based on literature, the BAP1 protein has a function in a signal transduction cascade in *Arabidopsis thaliana*. The biological function of BAP1-like protein in *G. max* is yet uncharacterized. The bioinformatics analysis of GMB151 soybean indicates the underrepresentation of expressed sequence tag (EST) sequences homologous to the putative BAP1-like locus in EST libraries sampled from standard soybean tissue. In the BLASTn search of the EST database, a very strong sequence identity between the GMB151 soybean insertion locus and only three soybean mRNA sequences, derived from dark-cultivated soybean callus tissue, was confirmed. This result is indicative of rather specific transcriptional expression conditions for this locus. The 3' UTR is involved in transcriptional expression regulation of a gene (Schoenberg and Maquat, 2012). An interruption could provoke differences in corresponding mRNA levels. However, there are no indications that such potential change in expression levels would have an impact on the safety assessment of GMB151 soybean. The putative BAP1-like protein has no homologies with any toxins or allergens. There are also no indications that the putative expression of this genetic locus has an impact on the phenotype of GMB151 soybean in relation to the non-GM counterpart. An agronomic performance and compositional analysis study was performed in which GMB151 soybean was compared to the non-GM counterpart. Based on the comparative assessment, there are no unexpected or unintended effects and no impact on GMB151 agronomic performance (Section 8) or the nutritional value of forage and grain (Section 7) from GMB151 soybean. Consequently, there are no reasons to assume an effect on plant pest risk due to the interruption of the putative BAP1-like locus in GMB151 soybean.

Identification of open reading frames and homology search of sequences with known allergens and toxins

A bioinformatics analysis was performed on the transgenic locus sequence of the GMB151 soybean to identify open reading frames (ORFs).

The GMB151 transgenic locus, containing the inserted DNA together with the 5′ and 3′ flanking sequences, was used as query sequence. The GetORF search program was used to identify all ORFs crossing a junction or overlapping the inserted DNA between two translation stop codons, with a minimum size coding for three amino acids. Six hundred one ORFs were identified. In the next step, the translated amino acid sequences from the identified ORFs with a minimum size of 30 amino acids were used as query sequences in homology searches to known allergens and toxins. After elimination of duplicates, they represented 115 unique sequences.

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

- An 8-mer search was carried out to identify any short sequences of eight amino acids or longer that share 100% identity to an allergenic protein. This search was performed using SeqMatchAll from the EMBOSS suite, which compared each ORF sequence with all known allergens present in the allergen database.

- An overall identity search was carried out by using the FASTA algorithm, which compared each complete query sequence with all protein sequences present in the allergen database. The scoring matrix was BLOSUM50. An E-value threshold of one was used. The criterion indicating potentially relevant identity to an allergen was ≥ 35% identity over at least 80 amino acids for sequences of ≥ 80 amino acids or ≥ 35% recalculated over a hypothetical 80-amino acid window for sequences of < 80 amino acids.

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

The overall search showed no biologically relevant identity between the query sequences and any known allergenic proteins. One ORF matched with a contiguous 8-amino acid sequence (SSPTTTTS) present in the Cas s 5 allergen. No match with this protein was found in the overall search for an allergen match, and an 8-mer match in isolation is unlikely to indicate any potential cross-reactivity. Also, because this putative ORF lacks both a translation start codon and an RNA splice site upstream of the 8-mer, translation of this 8-mer is not possible.

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

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

6. CHARACTERIZATION OF THE INTRODUCED PROTEINS

6.1. Identity and function of the HPPD-4 protein

The coding sequence of the HPPD protein was isolated from the *P. fluorescens* strain A32. Four amino acids were substituted (glutamic acid at position 335 with proline (Glu335Pro), glycine at position 336 with tryptophan (Gly336Trp), lysine at position 339 with alanine (Lys339Ala) and alanine at position 340 with glutamine (Ala340Gln)) to reduce HPPD-inhibitor binding efficacy. The modified protein is designated as HPPD-4 (Porée et al, 2014).

6.1.1. Background information and history of safe use of the HPPD-4 protein

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

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

In particular, HPPD proteins have been characterized in organisms present in human food, such as carrot (*Daucus carota*, Accession number O23920), barley (*Hordeum vulgare,* Accession number O48604), pig (*Sus scrofa*, Accession number Q02110), and cow (*Bos taurus*, Accession number Q5EA20).

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

6.1.2. Biochemistry and mode of action of the HPPD-4 protein

HPPD is an Fe(II)-dependent, non-heme oxygenase. HPPD is a key enzyme involved in the catabolism of tyrosine, which catalyzes the conversion of 4-hydroxyphenylpyruvate to homogentisate (HGA; 2,5-dihydroxyphenylacetate). In plants, HPPD is also involved in several anabolic pathways; its reaction product HGA is the aromatic precursor of tocopherol and plastoquinone, which are essential to the photosynthetic transport chain and antioxidative systems (Fritze et al., 2004). Figure 6 shows a diagram of the different metabolic pathways in which HPPD is involved in plants and non-photosynthetic organisms.

HPPD enzymes require an α-keto acid and molecular oxygen to oxidize or oxygenate a third molecule. The activity of HPPD is suppressed by HPPD-inhibiting herbicides (Figure 7). HPPD enzyme inhibition results in the disruption of the biosynthesis of carotenoids. Plants lacking carotenoids cannot protect themselves from the radicals generated by the light activation of chlorophyll, causing bleaching, necrosis, and death.

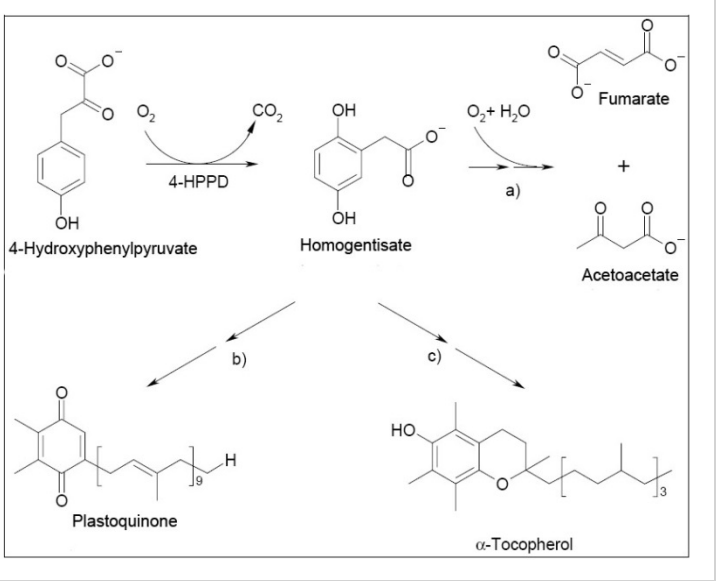
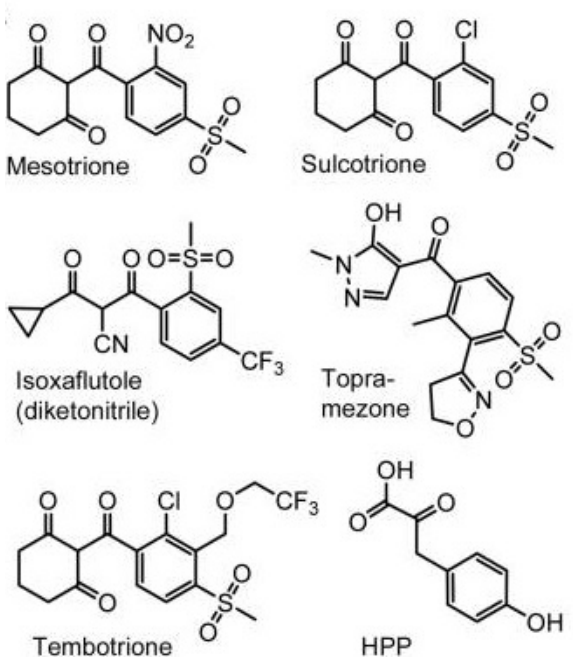


Figure 6. Biochemical pathways of HPPD proteins

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



From Siehl et al. (2014).

Figure 7. Structures of HPPD-inhibiting herbicides and substrate 4-hydroxyphenylpyruvate

6.2. Identity and function of the Cry14Ab-1 protein

Cry14Ab-1 protein is a member of the Cry-type protein family with a fully conserved three-domain structure. Cry proteins are produced by *B. thuringiensis* (*Bt*) strains and demonstrate specific toxicity towards either insects or nematodes. Cry14Ab-1 belongs to the "nematicidal branch" of Cry proteins with the closest homology to Cry14Aa1 (87% identity) (Sanahuja et al., 2011).

6.2.1. Background information and history of safe use of the Cry14Ab-1 protein

Cry proteins are synthesized by the Gram positive, spore-forming bacterium *Bt*. To date, more than 700 *cry* gene sequences that code for Cry proteins have been identified. Cry proteins typically have narrow target specificity that limits activity of individual Cry proteins to just a few genera. They have been successfully used as biopesticides against Lepidoptera, Coleoptera, and Diptera (Palma et al., 2014).

6.2.2. Biochemistry and mode of action of the Cry14Ab-1 protein

Many studies and reviews have described the mode of action of Cry proteins (WHO/IPCS, 1999; Betz et al., 2000; Siegel, 2001; Bravo et al., 2007; OECD, 2007; Federici and Siegel, 2008). Cry proteins are active only when ingested.

Cry14Ab-1 activity has been observed with nematodes but not other species (Appendix 3). Target organism panel assessments conducted as part of the discovery process of Cry14Ab-1 and non-target organism (NTO) testing conducted as part of the development of GMB151 and Cry14Ab-1 have revealed negligible effects against arthropods, annelids, and vertebrates.

With an insecticidal Cry protein, the crystals are solubilized in the midgut upon ingestion by an insect. The toxins are then proteolytically activated by midgut proteases and bind to specific receptors located in the cell membrane, leading to cell disruption and death (Palma et al., 2014). An essential component of the highly selective properties of most Cry proteins is the requirement that the toxin interact with one or more specific receptors. Many studies have demonstrated that the gastrointestinal tract epithelial surface of non-target animals, including humans, lack specific high-affinity Cry protein receptors (Koch et al., 2015).

Nematicidal Cry proteins have been observed to use a similar mode of action, damaging the intestines of *C. elegans* upon ingestion (see Appendix 3).

6.3. Expression levels of the introduced proteins

6.3.1. Expression levels of Cry14Ab-1 and HPPD-4 in field-grown soybean tissue

Protein expression levels of Cry14Ab-1 and HPPD-4 were determined by enzyme-linked immunosorbent assay (ELISA) in field-grown soybean tissue from GMB151 plants treated and not treated with trait-specific herbicide grown at three field trials in the U.S. in 2016.

Protein expression analysis was conducted on tissue samples harvested from plants grown in the U.S. in 2016. Four field sites were located in areas representative of the commercial production of soybean in the U.S. and sampled throughout the growing season for different tissues. The three randomly selected sites for expression analysis (Pennsylvania, Missouri, and Kansas) were

representative of the locations of commercial soybean production with respect to cultural practices, soil type, and climatic conditions.

Each site had two plots of GMB151 soybean. One plot was treated with trait-specific herbicide while the other plot was not treated. The IFT application to the treated GMB151 entry was made at a rate of 70.1 (69.2–71.1) g active ingredient per hectare (a.i./ha) before emergence (BBCH 00). All entries were of the Thorne background. The tissues analyzed are summarized in Table 8.

The quantitation of Cry14Ab-1 protein in leaf, root, flowers, forage, whole plant, and grain samples was conducted with a validated Cry14Ab-1-specific ELISA method using the EnviroLogix QualiPlate™ Kit for Cry14Ab-1. The quantitation of HPPD-4 protein in leaf, root, flowers, forage, whole plant, and grain samples was conducted with a validated HPPD-4-specific ELISA method using the EnviroLogix QuantiPlate™ Kit.

Expression of Cry14Ab-1 in Soybean Tissues

The level of Cry14Ab-1 protein in not treated and treated GMB151 soybean leaf, root, flower, forage, grain, and whole plant tissues ranged from 0.34–290.44 µg/g dry weight (DW) and 0.75–279.73 µg/g DW, respectively (Table 9).

Leaf at BBCH 16–17 demonstrated the highest mean Cry14Ab-1 protein expression levels (Table 9). Mean ± standard deviation (SD) Cry14Ab-1 expression levels in not treated and treated GMB151 soybean leaf at BBCH 16–17 were 191.99 ± 50.01 µg/g DW and 168.73 ± 64.98 µg/g DW, respectively.

Root at BBCH 60–66 demonstrated the lowest mean Cry14Ab-1 protein expression levels (Table 9). Mean ± SD Cry14Ab-1 expression levels in not treated and treated GMB151 soybean root at BBCH 60–66 were 8.78 ± 9.90 µg/g DW and 4.80 ± 4.26 µg/g DW, respectively.

Mean ± SD Cry14Ab-1 expression levels in not treated and treated GMB151 soybean grain were 95.91 ± 43.11 µg/g DW and 83.14 ± 37.69 µg/g DW, respectively.

Expression of HPPD-4 in Soybean Tissues

The level of HPPD-4 protein in not treated and treated GMB151 soybean leaf, root, flower, forage, grain, and whole plant tissues ranged from less than the lower limit of quantitation (< LLOQ) to 779.67 µg/g DW and 1.31–1028.76 µg/g DW, respectively (Table 10).

Leaf at BBCH 13–14 demonstrated the highest mean HPPD-4 protein expression levels (Table 10). Mean ± SD HPPD-4 expression levels in not treated and treated GMB151 soybean leaf at BBCH 13–14 were 430.04 ± 207.57 µg/g DW and 429.22 ± 271.72 µg/g DW, respectively.

Whole plant demonstrated the lowest mean HPPD-4 protein expression levels (Table 10). Mean ± SD HPPD-4 expression levels in not treated and treated GMB151 soybean whole plant were 2.44 ± 1.02 µg/g DW and 2.89 ± 1.33 µg/g DW, respectively.

Mean ± SD HPPD-4 expression levels in not treated and treated GMB151 soybean grain were 4.46 ± 2.90 µg/g DW and 4.45 ± 3.57 µg/g DW, respectively.

Table 8. Tissues analyzed for protein expression levels

Growth Stage [a]	Tissue	Sample Description
BBCH 13–14	Leaf	All true leaves from 1 plant
	Root	All roots from 1 plant
BBCH 16–17	Leaf	All true leaves from 1 plant
	Root	All roots from 1 plant
BBCH 60–66	Leaf	All true leaves from 1 plant
	Root	All roots from 1 plant
	Flowers	Composite sample of flowers collected from plants in a row
BBCH 76–79	Leaf	All true leaves from 1 plant
	Root	All roots from 1 plant
	Forage	Entire above-ground portion of 1 plant
BBCH 89–99	Grain	Grain from 1 plant
	Whole plant	1 whole plant (including roots)

[a] The BBCH scale is a system for a uniform coding of phenologically similar growth stages of mono- and dicotyledonous plant species. A conversion of BBCH stages to V/R stages of soybean is available online at https://www.soybase.org/OntologyConversion.php.

Table 9. Expression of Cry14Ab-1 in soybean tissues harvested from treated and not treated GMB151 grown at three sites

Tissue	BBCH Growth Stage	Entry	Cry14Ab-1 (µg/g DW)				Cry14Ab-1 (µg/g FW)			
			Mean	SD	Min	Max	Mean	SD	Min	Max
Leaf	13–14	C	156.36	49.91	98.54	280.32	36.07	17.01	18.92	75.59
		D	104.34	46.76	31.30	212.07	20.45	7.74	7.14	37.23
Root	13–14	C	28.28	16.19	9.86	58.56	4.99	2.69	1.49	10.54
		D	20.61	11.55	6.11	37.24	2.88	1.45	0.95	5.41
Leaf	16–17	C	191.99	50.01	129.60	290.44	44.48	15.08	27.01	72.95
		D	168.73	64.98	56.92	279.73	37.77	17.59	9.90	68.31
Root	16–17	C	21.59	13.48	3.65	48.61	4.01	3.02	0.64	10.60
		D	12.92	8.60	3.72	33.01	2.37	1.37	0.60	5.28
Leaf	60–66	C	79.62	39.18	42.87	186.45	19.90	12.09	10.80	53.68
		D	73.98	28.09	42.22	139.53	18.98	8.91	10.09	40.19
Root	60–66	C	8.78	9.90	0.61	29.90	1.99	2.09	0.17	6.28
		D	4.80	4.26	0.75	12.90	1.13	0.87	0.22	2.82
Flower	60–66	C	49.47	10.63	34.36	66.37	8.94	1.70	6.44	11.30
		D	48.03	17.08	30.71	72.10	8.52	3.10	4.62	12.71
Leaf	76–79	C	126.15	37.34	99.65	234.43	35.25	11.07	23.18	63.52
		D	107.77	42.86	46.41	214.47	29.86	12.49	14.00	59.99
Root	76–79	C	13.24	12.95	2.23	38.01	4.05	3.91	0.71	11.79
		D	13.24	12.65	3.57	40.91	4.09	4.05	1.06	13.30
Forage	76–79	C	51.34	9.25	36.92	64.05	12.19	3.02	7.56	17.01
		D	48.72	9.38	38.25	66.90	11.37	3.59	7.04	18.52
Grain	89–99	C	95.91	43.11	43.13	166.22	84.99	38.66	38.45	147.89
		D	83.14	37.69	15.17	144.81	73.93	33.86	13.50	129.82
Whole plant	89–99	C	40.84	18.13	0.34	69.19	33.66	15.06	0.28	55.05
		D	49.03	16.15	25.70	72.64	40.22	13.68	22.15	62.25

Entry C = GMB151 (not treated); Entry D = GMB151 (treated)
FW = fresh weight; DW = dry weight
Mean and standard deviation (SD) for each entry were based on the total sample population (n = 12).

Table 10. Expression of HPPD-4 in soybean tissues harvested from treated and not treated GMB151 grown at three sites

Tissue	BBCH Growth Stage	Entry [a]	HPPD-4 (µg/g DW)				HPPD-4 (µg/g FW)			
			Mean	SD	Min	Max	Mean	SD	Min	Max
Leaf	13–14	C	430.04	207.57	191.44	779.67	102.86	66.40	35.66	219.07
		D	429.22	271.72	191.29	1028.76	89.78	63.50	38.09	205.75
Root	13–14	C	57.71	29.15	26.43	107.86	10.49	5.46	3.78	20.54
		D	73.67	41.89	26.42	138.82	10.74	5.70	3.55	20.15
Leaf	16–17	C	416.62	158.43	175.60	688.68	97.30	44.48	37.33	174.06
		D	350.86	236.25	59.11	830.47	78.91	59.82	13.27	202.79
Root	16–17	C	61.90	16.24	40.71	91.93	11.30	4.81	7.04	20.83
		D	51.05	16.86	25.90	83.86	9.49	2.47	5.81	14.85
Leaf	60–66	C	201.65	100.90	82.72	372.05	50.53	29.72	20.84	101.80
		D	178.76	73.51	90.83	350.74	46.18	23.27	21.70	101.04
Root	60–66	C	24.60	12.98	9.70	49.41	5.94	2.69	2.48	11.33
		D	20.93	7.83	9.59	40.96	5.40	1.98	2.60	9.80
Flower	60–66	C	54.29	13.73	40.95	75.72	9.86	2.50	7.39	13.98
		D	70.87	26.45	39.59	121.60	12.45	4.62	7.53	20.67
Leaf	76–79	C	287.64	208.74	112.19	715.95	78.85	56.72	36.04	193.98
		D	225.73	141.37	34.57	449.58	61.38	39.21	11.34	125.74
Root	76–79	C	34.89	22.58	12.81	74.92	11.12	7.57	3.82	23.88
		D	24.69	9.71	9.58	40.41	7.45	2.94	2.88	12.76
Forage	76–79	C	120.18	42.47	78.96	203.83	29.09	12.99	15.55	54.92
		D	129.03	45.32	87.61	196.27	30.48	14.07	15.26	53.51
Grain	89–99	C	4.46	2.90	1.23	9.83	3.95	2.59	1.11	8.74
		D	4.45	3.57	1.31	12.71	3.97	3.20	1.18	11.39
Whole plant [a]	89–99	C	2.44	1.02	< LLOQ	4.12	2.03	0.88	< LLOQ	3.28
		D	2.89	1.33	1.33	4.59	2.40	1.15	1.01	3.93

Entry C = GMB151 (not treated); Entry D = GMB151 (treated)
FW = fresh weight; DW = dry weight
[a] Mean and standard deviation (SD) for each entry were based on the total sample population (n = 12), except for whole plant Entry C. Mean and SD for whole plant Entry C were calculated on a population of n = 11 since one of the samples yielded HPPD-4 values below the Lower Limit of Quantification (< LLOQ).

6.3.2. Comparison of Cry14Ab-1 protein expression in whole flower versus the plant reproductive structures

The Cry14Ab-1 expression levels in soybean pollen and nectar are an additional consideration when determining expected environmental concentrations and margins of exposure for NTO safety assessments of pollinator species, namely *Apis mellifera*. The anatomy of a soybean flower, specifically its small size and closed structure, creates practical limitations to the sampling of pollen and nectar at sufficient quantities to accurately measure protein expression levels. Therefore, for the purposes of safety assessments, inferences are made on the potential concentrations in pollen and nectar based on concentrations in the stamens and pistils on which pollen and nectar are accessed by pollinators. There are also practical limitations to dissecting field-grown flowers into stamens and pistils for direct analysis. Therefore, a greenhouse study

was conducted to correlate pistil and stamen concentrations to that in the whole flower so that whole flower expression data from field-grown plants could be applied to the NTO assessment for pollinator species.

Briefly, GMB151 soybean and non-GM comparator were grown in a greenhouse. Flowers were collected at BBCH 60–66 (approximately R_2). Flowers were pooled and a portion dissected to extract stamen and pistil material (Figure 8). Six pooled flower samples and six pooled stamen/pistil samples were analyzed for Cry14Ab-1 expression levels. The quantitation of Cry14Ab-1 protein in the samples was conducted with a validated Cry14Ab-1-specific ELISA method. The combined weight of a separate set of 30 whole flowers was measured. These flowers were dissected and the combined weight of the stamen/pistil materials was measured. The percent weight of the stamen/pistil materials was determined relative to the weight of the whole flowers. Mean concentrations and SD values from ELISA analysis across six samples for each tissue type were determined. The mean Cry14Ab-1 concentrations for the whole flowers and stamen/pistil tissues were compared using a two-tailed t-test (α=0.05).

Figure 9 summarizes the mean Cry14Ab-1 ELISA results obtained by analyzing six flower and six stamen/pistil samples. The mean concentration of Cry14Ab-1 in the flower samples was 17.24 µg/g fresh weight (FW; ranging from 10.27–31.41 µg/g FW), and the mean concentration in stamen/pistil samples was 12.77 µg/g FW (ranging from 7.44–23.10 µg/g FW). A t-test analysis of these results indicated that there was no significant difference between the concentration of Cry14Ab-1 in the whole flower samples relative to the stamen/pistil plant reproductive parts from GMB151 soybean plants ($p = 0.28$). The total fresh weight of 30 whole flowers was 357.1 mg, and the combined stamen/pistil materials from these flowers weighed 48.0 mg, indicating that the stamen/pistil parts constitute approximately 13% of the mass of whole flowers.

Statistical analysis of the ELISA data indicated that there was no significant difference between the concentration of Cry14Ab-1 expressed in the GMB151 soybean flower and stamen/pistil tissues. The stamen/pistil reproductive parts of these soybean flowers composed approximately 13% of the mass of whole flowers. Therefore, it is concluded that expression levels in field-grown flowers of GMB151 are acceptable for calculation of expected environmental concentrations and margins of exposure for pollinator species.

Figure 8. Illustrative images of soybean flower and stamen/pistil tissues
[a] A whole soybean flower.
[b] A flower with front sepal removed and petals spread to reveal the stamen and pistil.
[c] The stamen/pistil flower reproductive structures dissected from a flower.

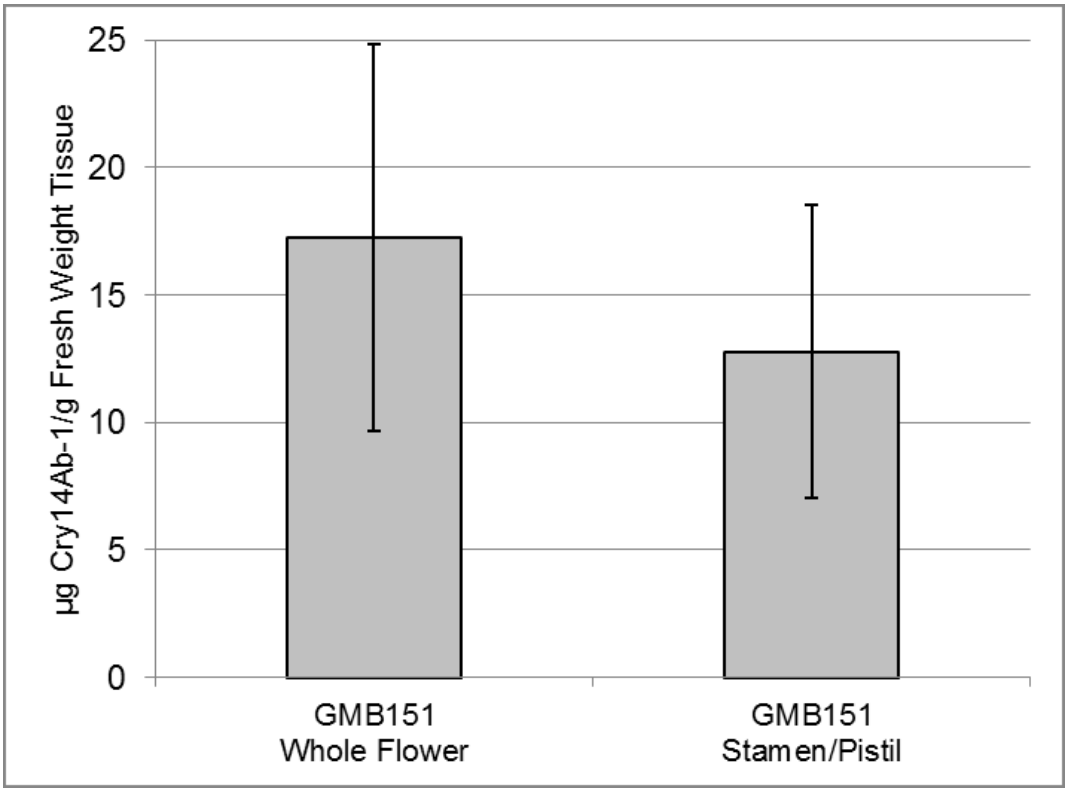


Figure 9. Cry14Ab-1 in GMB151 whole flower and stamen/pistil tissues
The mean Cry14Ab-1 concentration from six samples of GMB151 soybean whole flower and stamen/pistil tissues are presented. Error bars represent standard deviation. (p = 0.28).

6.4. Comparability of *in planta* and *in vivo* proteins

For the safety assessment of GM crops, certain safety tests and studies require large amounts of protein. The expression levels of Cry14Ab-1 and HPPD-4 in GMB151 were too low to allow for purification of sufficient quantities of the two proteins directly from GMB151 for use in the safety assessment studies. Therefore, the Cry14Ab-1 and HPPD-4 proteins were produced in high-expressing recombinant host organisms, *Bt* and *Escherichia coli*, respectively. The recombinant proteins were engineered to match the amino acid sequences of their counterparts expressed in GMB151. The equivalence of GMB151 soybean-produced and bacterially-produced proteins was examined to ensure that the proteins from the two host sources were equivalent so that the bacterially-produced proteins could be used as a surrogate in the studies.

6.4.1. The equivalence of GMB151-purified and microbially-produced Cry14Ab-1 proteins

A purification of Cry14Ab-1 protein was performed from the GMB151 soybean leaf tissue using affinity chromatography. GMB151 soybean-purified Cry14Ab-1 protein was characterized and evaluated for equivalence with bacterially-produced Cry14Ab-1 protein based on a panel of analytical tests and assays, including analysis of a Coomassie-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot analysis, glycostaining analysis, mass spectroscopy, N-terminal sequence analysis, and *C. elegans* bioassay.

Assessment and comparison of the apparent molecular mass

The GMB151 soybean-purified Cry14Ab-1 protein and the bacterially-produced Cry14Ab-1 protein were compared side-by-side by means of an SDS-PAGE analysis (Figure 10). Additionally, the bacterially-produced Cry14Ab-1 protein was spiked into the sample resulting from non-GM soybean subjected to the same affinity purification procedure as the plant-purified Cry14Ab-1 protein sample (i.e. non-GM control) to allow comparison in a similar soybean plant tissue.

A specific, predominant band was observed for both the plant-purified and the bacterially-produced Cry14Ab-1 that corresponds to the expected molecular mass of the Cry14Ab-1 protein (131.1 kDa). This demonstrated that the apparent molecular masses of the GMB151 soybean-purified and the bacterially-produced Cry14Ab-1 protein are comparable. No signal was observed for the negative control sample.

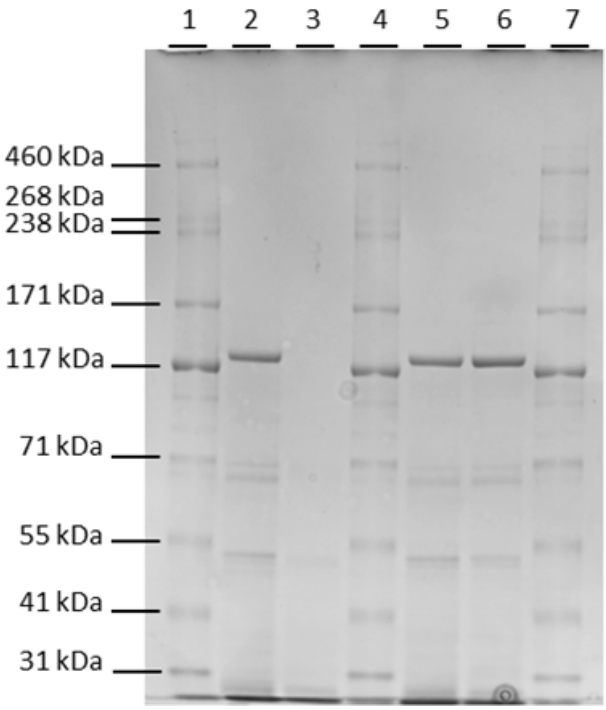


Figure 10. Assessment and comparison of the apparent molecular mass of the GMB151 soybean-purified Cry14Ab-1 protein and the bacterially-produced Cry14Ab-1 protein

Both plant-purified and bacterially-produced Cry14Ab-1 protein samples were loaded on a Criterion XT Tris-acetate 3–8% gel, and SDS-PAGE gel electrophoresis was performed in 1X Tris-Acetate running buffer, followed by Coomassie staining.

Loading order:

Lane	Sample
1	5 µL of the HiMark™ Pre-Stained High Molecular Mass Protein Standard (Life Technologies)
2	1 µg of plant-purified Cry14Ab-1 protein from GMB151 soybean
3	10 µL (= corresponding volume) of the non-GM control
4	5 µL of the HiMark™ Pre-Stained High Molecular Mass Protein Standard (Life Technologies)
5	1 µg of plant-purified Cry14Ab-1 protein from GMB151 soybean
6	1 µg of bacterially-produced Cry14Ab-1 protein spiked in 10 µL non-GM control
7	5 µL of the HiMark™ Pre-Stained High Molecular Mass Protein Standard (Life Technologies)

Assessment and comparison of the immunoreactivity

The GMB151-purified Cry14Ab-1 protein and the bacterially-produced Cry14Ab-1 protein were compared side-by-side by means of western blot analysis (Figure 11).

Using a Cry14Ab-1-specific polyclonal antibody, a signal corresponding to the expected molecular mass of the Cry14Ab-1 protein was detected for all plant-purified and bacterially-produced Cry14Ab-1 samples. No signal was observed for the negative control samples.

The obtained results confirmed the immunoreactivity of the GMB151-purified Cry14Ab-1 protein and the comparability to the bacterially-produced Cry14Ab-1 protein.

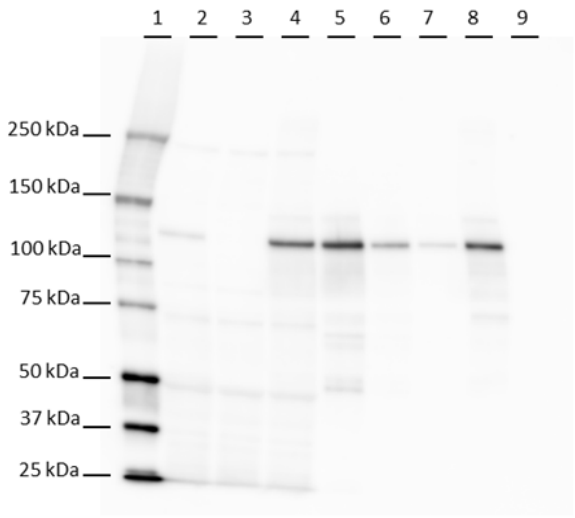


Figure 11. Assessment and comparison of immunoreactivity of GMB151 soybean-purified Cry14Ab-1 protein and bacterially-produced Cry14Ab-1 protein

Both plant-purified and bacterially-produced Cry14Ab-1 protein samples were loaded on a Criterion XT Tris-Acetate gel 3–8% gel, and SDS-PAGE gel electrophoresis was performed in 1X Tris-Acetate running buffer. After semi-dry blotting, the proteins were visualized by enhanced chemiluminescence (ECL) detection using the rabbit anti-Cry14Ab-1 antibody as primary antibody, the goat anti-rabbit HRP conjugated as secondary antibody, and the Precision Protein™ StrepTactin-HRP Conjugate to detect the ECL molecular mass marker.

Loading order:

Lane	Sample
1	5 µL of Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
2	10 µL of crude protein extract from GMB151 soybean
3	10 µL of crude protein extract from the non-GM counterpart
4	40 ng bacterially-produced Cry14Ab-1 protein spiked in 10 µL of the crude protein extract from the non-GM counterpart
5	40 ng of plant-purified Cry14Ab-1 protein from GMB151 soybean
6	10 ng of plant-purified Cry14Ab-1 protein from GMB151 soybean
7	4 ng of plant-purified Cry14Ab-1 protein from GMB151 soybean
8	40 ng of bacterially-produced Cry14Ab-1 protein
9	4 µl of a 1/10 diluted non-GM control

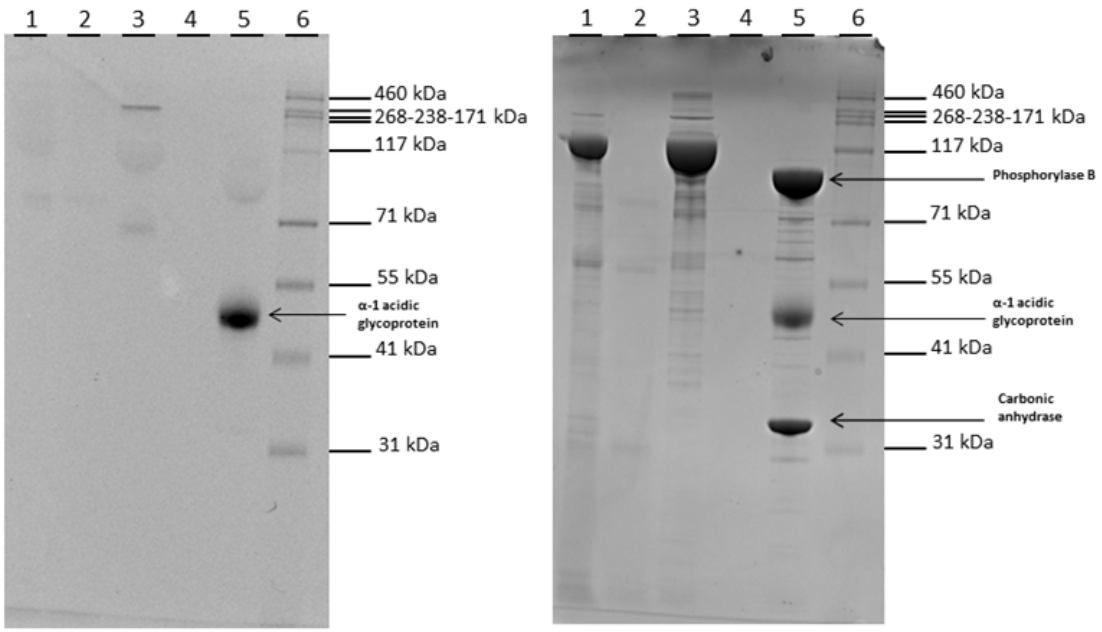
Assessment and comparison of the glycosylation status

The glycosylation statuses of plant-purified Cry14Ab-1 protein and the bacterially-produced Cry14Ab-1 protein were assessed side-by-side. The results of the glycostaining analysis are shown in Figure 12.

The α-1 acidic glycoprotein of the glycoprotein mix was visualized as a bright band. For the Cry14Ab-1 protein samples, weak background signals similar to the background signal of the non-glycosylated proteins in the glycoprotein mix were observed. In the bacterially-produced Cry14Ab-1 sample, a weak band was observed between 171 kDa and 238 kDa, which is most probably an impurity consisting of a potential glycosylated or glycosylation-like structure from the *Bt* spores (Figure 12, panel A).

The presence of sufficient Cry14Ab-1 protein on the gel was demonstrated by staining the gels with Coomassie after the glycostaining procedure (Figure 12, panel B).

The absence of glycosylation was demonstrated for both the GMB151-purified Cry14Ab-1 protein and the bacterially-produced Cry14Ab-1 protein. Consequently, both Cry14Ab-1 protein samples have a comparable glycosylation status.



Panel A Panel B

Figure 12. Assessment and comparison of the glycosylation status of plant-purified Cry14Ab-1 protein and bacterially-produced Cry14Ab-1 protein

One Criterion XT Bis-Tris 12% SDS-PAGE gel was prepared and cut into two pieces, each part containing 7 µg Cry14Ab-1 protein of both the plant-purified Cry14Ab-1 protein and bacterially-produced Cry14Ab-1 protein together with the appropriate positive and negative controls to assess the glycosylation status.

Panel A shows the result of the staining using the Glycoprotein Detection Kit (Sigma) to demonstrate the absence of glycosylation of both the Cry14Ab-1 protein samples.

Panel B shows a Coomassie staining of the SDS-PAGE gel to demonstrate the presence of the proteins on the gel.

Loading order of Panel A and B:

Lane	Sample
1	53.21 pmoles of plant-purified Cry14Ab-1 protein from GMB151 soybean
2	9.46 (corresponding volume) µL of the non-GM control
3	53.21 pmoles of bacterially-produced Cry14Ab-1 protein
4	empty
5	45.45 pmoles of Glycoprotein mix
6	5 µL of the HiMark™ Pre-Stained High Molecular Mass Protein Standard (Life Technologies)

Assessment of the peptide mapping

The peptide mapping was established using ultra performance liquid chromatography-ultra violet-mass spectrometry (UPLC-UV-MS) analysis for the GMB151-purified and the bacterially-produced Cry14Ab-1 protein.

Peptides resulting from a trypsin digest of the GMB151-purified Cry14Ab-1 protein were analyzed using UPLC-UV-MS. Figure 13 provides an overview of the mapped peptides against the theoretical amino acid sequence of the Cry14Ab-1 protein. A coverage of 70.8% was determined, which confirms the identity of Cry14Ab-1 protein. The observation of a 42 atomic mass unit (amu) mass increment at the N-terminal peptide indicates acetylation of the N-terminus. Cleavage of the N-terminal methionine and N-terminal acetylation are by far the most common modifications of eukaryotic proteins (Bradshaw et al., 1998; Polevoda and Sherman, 2000, 2003). These post-translational modifications affect the majority of soluble eukaryotic proteins. Over 70% of the analyzed proteins contain a modification on the N-terminus. Examples of N-terminal acetylation of plant proteins are described for soybean (Peariso et al., 2008).

Peptides resulting from a trypsin digest of bacterially-produced Cry14Ab-1 protein were analyzed using UPLC-UV-MS. Figure 14 provides an overview of the mapped peptides against the theoretical amino acid sequence of the Cry14Ab-1 protein. A coverage of 86.8% was determined, which confirms the identity of Cry14Ab-1 protein.

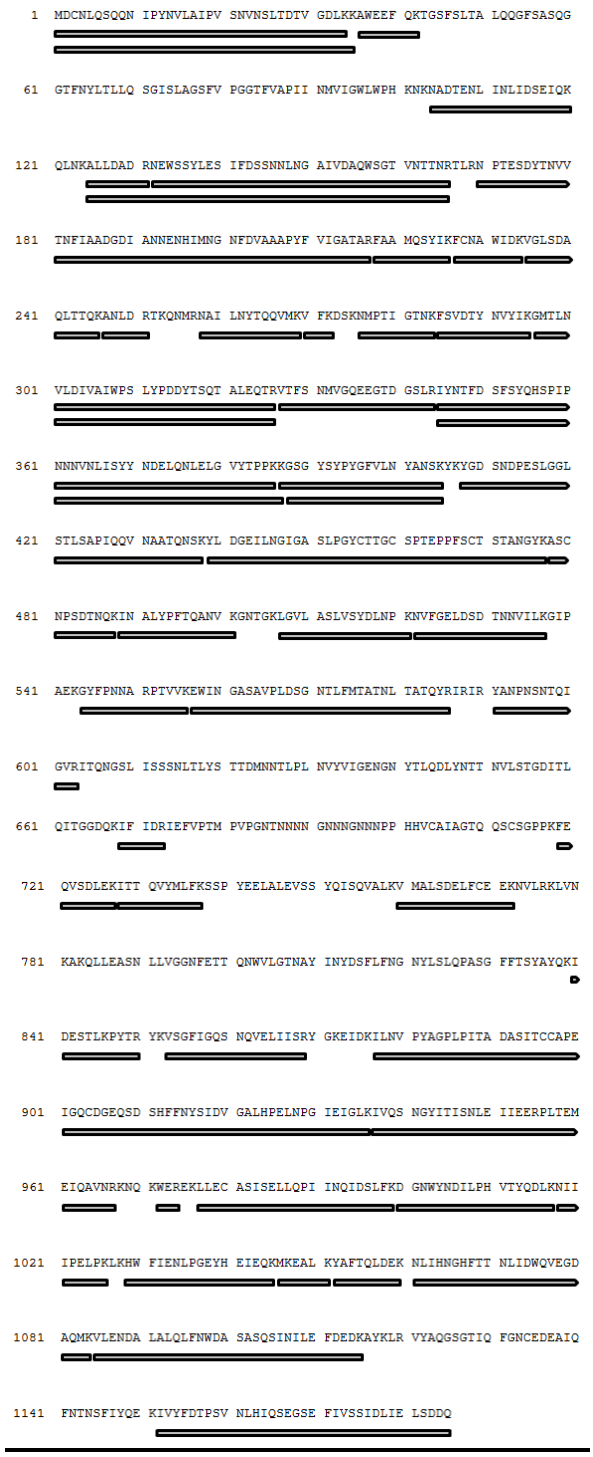


Figure 13. Schematic overview of the coverage of the theoretical Cry14Ab-1 sequence by the tryptic peptides from the GMB151 soybean-purified Cry14Ab-1 protein

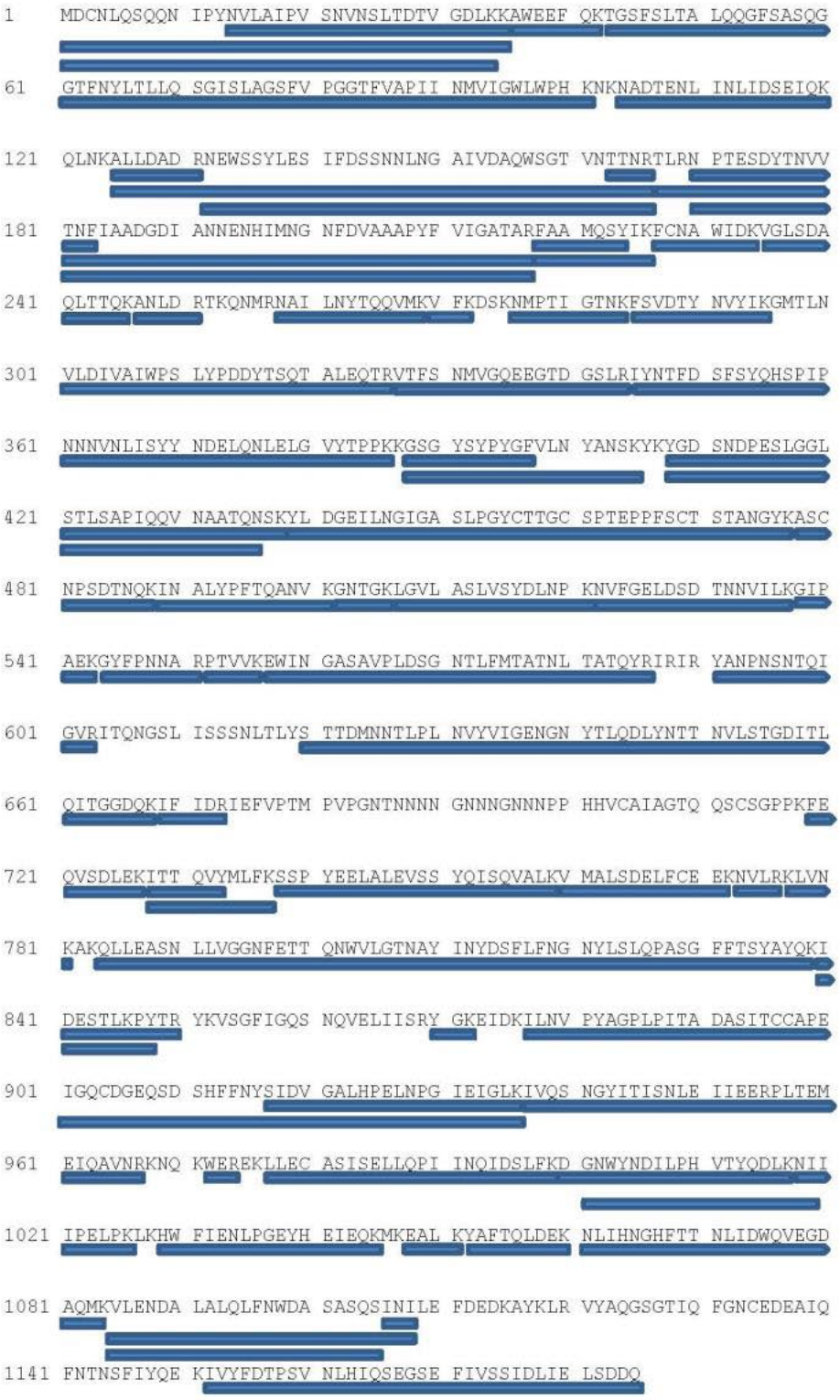


Figure 14. Schematic overview of the detected peptides derived from tryptic digestion of bacterially-produced Cry14Ab-1

Assessment of the N-terminal sequence

The N-terminal sequence of the GMB151 soybean-purified Cry14Ab-1 protein was determined by UPLC-UV-MS^E analysis of a chymotrypsin-digested sample.

The obtained mass of the N-terminal peptide, containing amino acid residues 1–13, agreed with the theoretical monoisotopic mass of the acetylated N-terminal peptide.

The N-terminal sequence of bacterially-produced Cry14Ab-1 protein was determined by Edman degradation. The obtained N-terminal sequence confirmed the theoretical amino acid sequence (MDCNLQSQQN) with the exception of cysteine. Cysteines are likely to form disulfide bridges that cannot be detected by Edman degradation.

Assessment and comparison of the Cry14Ab-1 functional activity

The functional activity of the GMB151 soybean-purified and the bacterially-produced Cry14Ab-1 proteins was assessed by means of a *C. elegans* bioassay (Appendix 6).

The mean ED_{50} value (the effective dose required to inhibit *C. elegans* growth by 50%) for the GMB151 soybean-purified Cry14Ab-1 protein was 11.39 µg/mL. The mean ED_{50} value for the bacterially-produced Cry14Ab-1 protein was 2.87 µg/mL.

When proteins are produced in different systems and purified using different methods, it is not unusual to have variability in the functional activity (Raybould et al., 2013). The bacterially-produced protein was used as a surrogate for the soybean-purified Cry14Ab-1 protein to assess the risks to human and animal health and to the environment. The apparent molecular mass, immunoreactivity, glycosylation status, functional activity, N-terminal sequence, and peptide mapping of the two batch proteins were compared side-by-side and found to be comparable, both biochemically and functionally. Since the mean ED_{50} of the plant-purified Cry14Ab-1 protein was within 10-fold of the mean ED_{50} of the bacterially-produced Cry14Ab-1 protein, plant-purified and bacterially-produced Cry14Ab-1 proteins were considered functionally equivalent.

Conclusion

The equivalence of the GMB151 soybean-purified Cry14Ab-1 protein with bacterially-produced Cry14Ab-1 protein was demonstrated based on a panel of analytical tests and assays, including analysis of a Coomassie-stained SDS-PAGE, western blot analysis, glycostaining analysis, mass spectroscopy, N-terminal sequence analysis, and *C. elegans* bioassay.

6.4.2. The equivalence of GMB151-purified and microbially-produced HPPD-4 proteins

A purification of HPPD-4 protein was performed from the GMB151 soybean leaf tissue using affinity chromatography. GMB151 soybean-purified HPPD-4 protein was characterized, and the equivalence with bacterially-produced HPPD-4 protein was evaluated based on a panel of analytical tests and assays, including analysis of a Coomassie-stained SDS-PAGE, western blot analysis, glycostaining analysis, mass spectroscopy, and N-terminal sequence analysis.

Assessment and comparison of the apparent molecular mass

The GMB151 soybean-purified HPPD-4 protein and the bacterially-produced HPPD-4 protein were compared side-by-side by means of an SDS-PAGE analysis (Figure 15). The bacterially-produced HPPD-4 protein was spiked into non-GM soybean tissue subjected to the same affinity purification procedure as the plant-purified HPPD-4 protein sample (i.e. non-GM control) to allow comparison in a similar soybean plant tissue.

A specific, predominant band that corresponds to the expected molecular mass of the HPPD-4 protein (40.3 kDa) was observed for both the plant-purified HPPD-4 protein and the bacterially-produced HPPD-4 protein samples. This demonstrated that the apparent molecular masses of the GMB151 soybean-purified and the bacterially-produced HPPD-4 proteins are comparable. The non-GM control showed some non-specific background staining derived from the plant tissue. These background fragments are visible in all soybean samples.

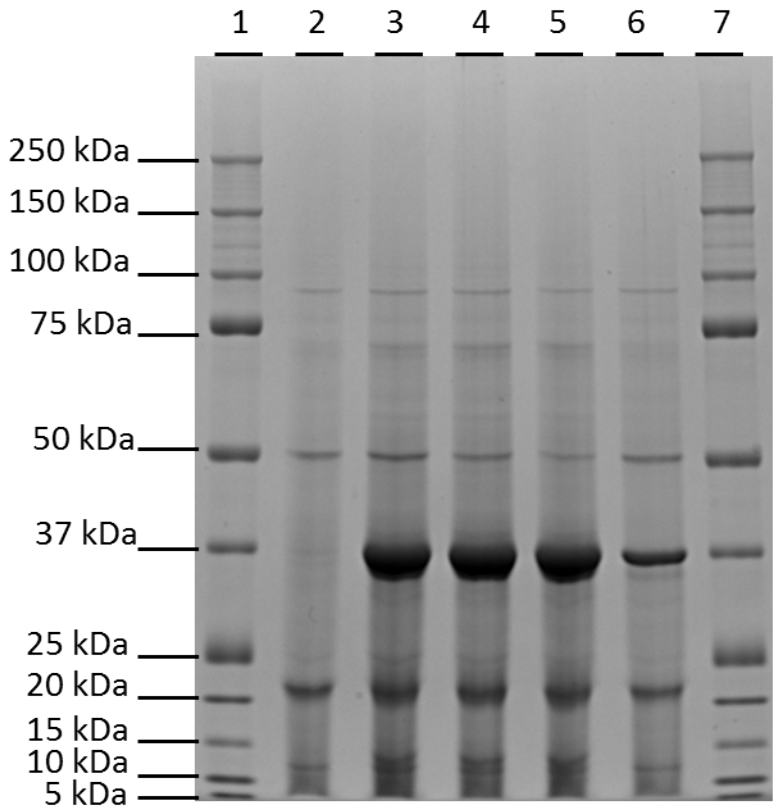


Figure 15. Assessment and comparison of the apparent molecular mass of GMB151 soybean-purified HPPD-4 protein and bacterially-produced HPPD-4 protein

Both plant-purified and bacterially-produced HPPD-4 protein samples were loaded on a NuPAGE Bis-Tris 4–12% gel, and SDS-PAGE gel electrophoresis was performed in 1X MOPS running buffer, followed by Coomassie staining.

Loading order:

Lane	Sample
1	2 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
2	10 µL of the non-GM control
3	10 µl of plant-purified HPPD-4 protein sample GM/VLN028-A/05 from GMB151 soybean
4	10 µl of plant-purified HPPD-4 protein sample GM/VLN028-A/06 from GMB151 soybean
5	10 µl of plant-purified HPPD-4 protein sample GM/VLN028-A/07 from GMB151 soybean
6	1 µg of bacterially-produced HPPD-4 protein spiked in 10 µL non-GM control
7	2 µL of Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)

Assessment and comparison of the immunoreactivity

The GMB151 soybean-purified HPPD-4 protein and the bacterially-produced HPPD-4 protein were compared side-by-side by means of a western blot (Figure 16).

Using an HPPD-4-specific polyclonal antibody, a signal corresponding to the expected molecular mass of the HPPD-4 protein was detected for all plant-purified and bacterially-produced HPPD-4 samples.

The obtained results confirmed the immunoreactivity of the GMB151 soybean-purified HPPD-4 protein and the comparability to the bacterially-produced HPPD-4 protein.

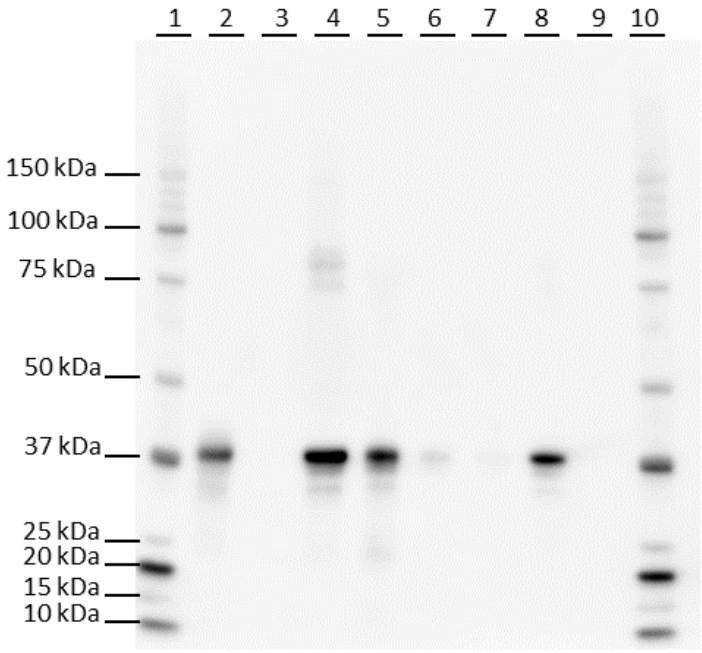


Figure 16. Assessment and comparison of immunoreactivity of GMB151 soybean-purified HPPD-4 protein and bacterially-produced HPPD-4 protein

Both plant-purified and bacterially-produced HPPD-4 protein samples were loaded on a NuPAGE Bis-Tris 4–12% gel, and SDS-PAGE gel electrophoresis was performed in 1X MOPS running buffer. After semi-dry blotting, the proteins were visualized by enhanced chemiluminescence detection (ECL) using a 1:5000 dilution of the rabbit anti-HPPD-4 antibody as primary antibody, a 1:10000 dilution of the goat anti-rabbit HRP conjugated as secondary antibody, and a 1:10000 dilution of the Precision Protein™ StrepTactin-HRP Conjugate to detect the ECL molecular mass marker.

Loading order:

Lane	Sample
1	5 µL of Precision Plus Protein™ WesternC Standards (Bio-Rad)
2	5 µL of crude protein extract from GMB151 soybean
3	5 µL of crude protein extract from the non-GM counterpart
4	40 ng bacterially-produced HPPD-4 protein spiked in 5 µL of the crude protein extract from the non-GM counterpart
5	40 ng of plant-purified HPPD-4 protein from GMB151 soybean
6	10 ng of plant-purified HPPD-4 protein from GMB151 soybean
7	4 ng of plant-purified HPPD-4 protein from GMB151 soybean
8	40 ng of bacterially-produced HPPD-4 protein
9	6.56 µl of a 1/100 diluted non-GM control
10	5 µL of Precision Plus Protein™ WesternC Standards (Bio-Rad)

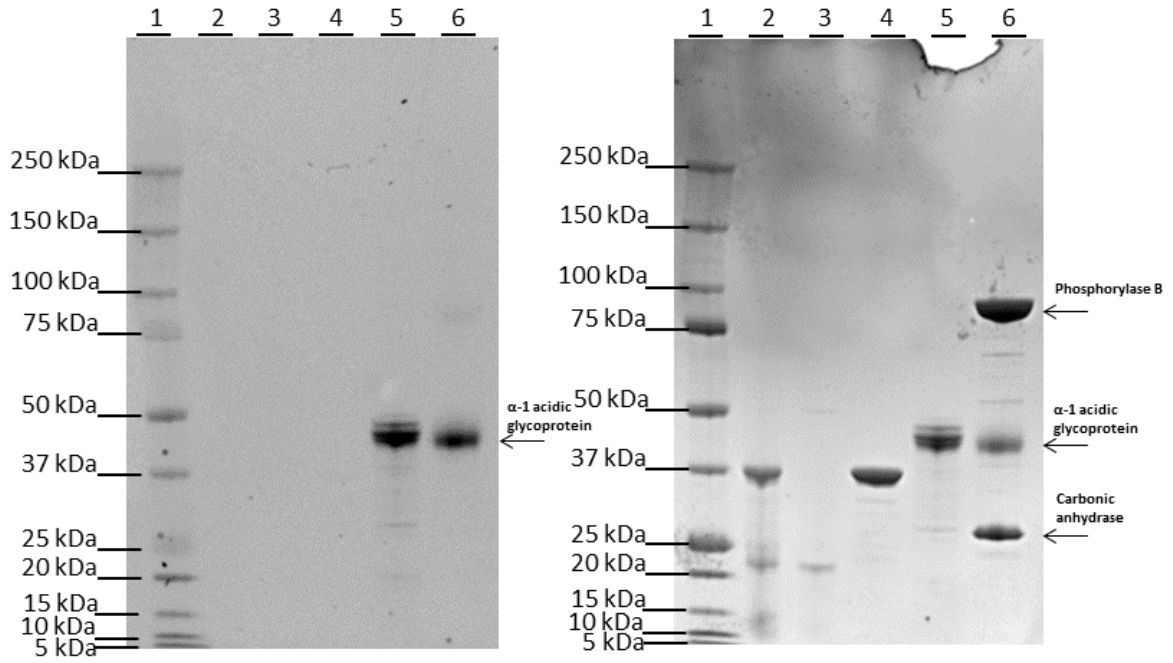
Assessment and comparison of the glycosylation status

The results of the glycostaining analysis are shown in Figure 17.

The glycosylated proteins of the horseradish peroxidase positive control and the α-1 acidic glycoprotein of the glycoprotein mix were visualized as bright bands on the gel while for the HPPD-4 protein samples, no signal was observed (Figure 17, panel A).

The presence of sufficient HPPD-4 protein on the gel was demonstrated by staining the gels with Coomassie after the glycostaining procedure (Figure 17, panel B).

The absence of glycosylation was demonstrated for both the GMB151 soybean-purified HPPD-4 protein and the bacterially-produced HPPD-4 protein. Consequently, both HPPD-4 protein samples have a comparable glycosylation status.



Panel A **Panel B**

Figure 17. Assessment and comparison of the glycosylation status of GMB151 soybean-purified HPPD-4 protein and bacterially-produced HPPD-4 protein

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

Panel A shows the result of the staining using the Glycoprotein Detection Kit (Sigma) to demonstrate the absence of glycosylation of both the HPPD-4 protein samples.

Panel B shows a Coomassie staining of the SDS-PAGE gel to demonstrate the presence of the proteins on the gel.

Loading order of Panel A and B:

Lane	Sample
1	5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
2	2 µg of plant-purified HPPD-4 protein from GMB151 soybean
3	3.28 (corresponding volume) µL of the non-GM control
4	2 µg of bacterially-produced HPPD-4 protein
5	45.45 pmoles Horseradish Peroxidase (positive control)
6	45.45 pmoles of Glycoprotein mix

Assessment of the intact molecular mass

The intact molecular mass of the GMB151 soybean-purified HPPD-4 protein was determined using UPLC-UV-MS. The determined intact molecular mass allowed the identification of three major molecular masses. The molecular mass of 39748.3 Da corresponds to the mature form of the HPPD-4 protein minus the initial methionine residue, which lacks the four C-terminal residues. A second molecular mass (40014.7 Da) corresponds to the mature HPPD-4 protein with an N-terminal cysteinic sulfinic acid, lacking the four C-terminal residues. The third observed molecular mass (40415.1 Da) corresponds to the mature HPPD-4 protein with an N-terminal

cysteinic sulfinic acid. An overview of the different forms mapped against the theoretical form is shown in Figure 18.

Degradation of the plant-purified HPPD-4 protein sample was noticed in several experiments. The forms lacking four residues at the C-terminus are considered as partially degraded HPPD-4 protein forms derived from the two major forms present in the plant.

Cleavage of the N-terminal methionine is mediated by methionine aminopeptidase (Bradshaw et al., 1998; Polevoda and Sherman, 2000) and is a common process in many organisms. The additional cysteinic sulfinic acid at the N-terminus is a modified cysteine residue derived from incomplete cleavage of the transit peptide.

The intact molecular mass of the bacterially-produced HPPD-4 protein was determined using liquid chromatography - ultra violet - mass spectroscopy (LC-UV-MS) analysis. The determined intact molecular mass (40148.1 Da) confirms the theoretical molecular mass of the protein corresponding with the mass of the HPPD-4 protein lacking the N-terminal methionine residue (40149 Da).

Both bacterially-produced and GMB151 soybean-purified HPPD-4 proteins consist of the mature HPPD-4 protein lacking the N-terminal methionine. The plant also contains a second form, the mature HPPD-4 protein carrying a modified cysteine residue on the N-terminus. This additional modified cysteine residue is derived from incomplete cleavage of the transit peptide.

Theoretical HPPD-4 sequence			**MADLYENPMG**	…	**RRGVLTAD**
Bacterially-produced			ADLY	…	RRGVLTAD
Plant-purified	Form 1	39748.3 Da	ADLY	…	RRGV
	Form 2	40014.7 Da	**X**MADLYE	…	RRGV
	Form 3	40415.1 Da	**X**MADLYE	…	RRGVLTAD

Figure 18. Schematic overview of the results of the intact molecular mass determination

Mapping of the three major forms of the plant-purified HPPD-4 protein and the bacterial-produced HPPD-4 against the N- and C termini of the theoretical amino acid sequence of the mature HPPD-4 protein. **X** is cysteinic sulfinic acid, a modified cysteine residue derived from the transit peptide.

Assessment of the peptide mapping

The peptide mapping for the GMB151 soybean-purified HPPD-4 protein was established using UPLC-UV-MS^E analysis. Peptides resulting from a trypsin digest of the GMB151 soybean-purified HPPD-4 protein were analyzed. Figure 19 provides an overview of the mapped peptides against the theoretical amino acid sequence of the HPPD-4 protein. A coverage of 72.9% was determined, which confirmed the identity of the HPPD-4 protein. Numerous peptides resulted from protein cleavage at sites that are not trypsin-specific, which could possibly have been caused by protein degradation.

The peptide mapping for bacterially-produced HPPD-4 protein was established using LC-UV-MS analysis. Peptides resulting from a trypsin digest of bacterially-produced HPPD-4 protein were analyzed. Figure 20 provides an overview of the mapped peptides against the theoretical amino acid sequence of the HPPD-4 protein. A coverage of 96.6% was determined, which confirms the identity of HPPD-4 protein.

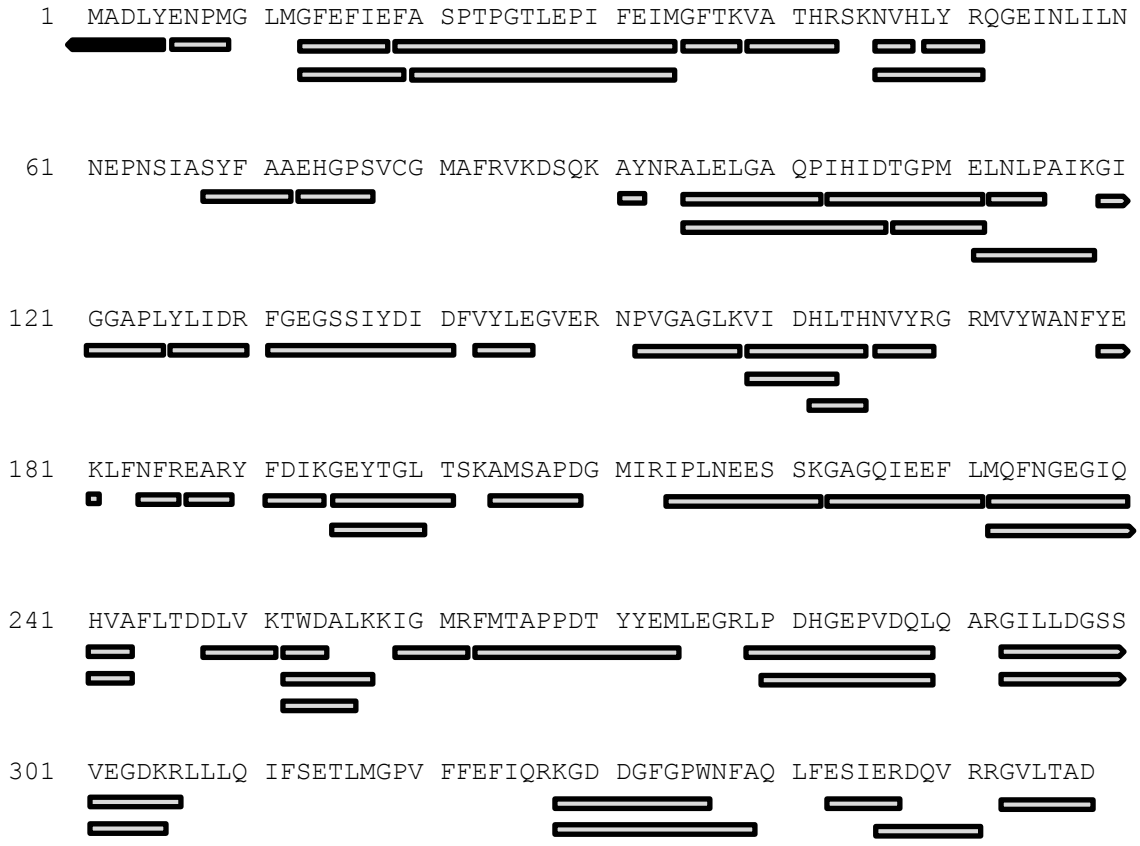


Figure 19. Schematic overview of the coverage of the theoretical HPPD-4 sequence by the tryptic peptides from the GMB151 soybean-purified HPPD-4 protein detected by UPLC-UV-MS[E]

Mapping of the peptides derived from the trypsin-digested plant-purified HPPD-4 protein sample against the theoretical amino acid sequence of the HPPD-4 protein. The black bar represents the N-terminal peptide corresponding to the HPPD-4 protein with an N-terminal cysteinic sulfinic acid, derived from the transit peptide.

Figure 20. Schematic overview of the coverage of the theoretical HPPD-4 sequence by the tryptic peptides from the bacterially-purified HPPD-4 protein detected by UPLC-UV-MS^E

Mapping of the peptides derived from the trypsin-digested bacterially-produced HPPD-4 protein sample against the theoretical amino acid sequence of the HPPD-4 protein.

Assessment of the N-terminal sequence

The N-terminal sequence of the GMB151 soybean-purified HPPD-4 protein and the bacterially-produced HPPD-4 protein was determined by Edman degradation.

The obtained data for the GMB151 soybean-purified sample suggested the presence of two N-termini. The obtained sequence result could be resolved as ADLYENPMGL, corresponding to the N-terminus of the mature HPPD-4 lacking the initial methionine, and sequence XMADLYENPM, corresponding to incomplete cleavage of the transit peptide of the HPPD-4 protein. These results support the data obtained within the intact molecular mass determination.

The N-terminal sequence of bacterially-produced HPPD-4 protein was determined by Edman degradation. The obtained N-terminal sequence (ADLYENPMGL) corresponds to the N-terminus of the mature HPPD-4 lacking the initial methionine. This confirms the results observed with the intact molecular mass determination.

Assessment and comparison of the HPPD-4 functional activity

The functional activity of the HPPD-4 protein in GMB151 soybean and the bacterially-produced HPPD-4 protein was assessed by means of a quantitative enzyme assay.

A protein extraction was performed on GMB151 soybean leaf samples, and the concentration of HPPD-4 was determined by means of ELISA. The activity of the HPPD-4 was determined in this crude extract by measuring the HGA formed over time using liquid chromatography tandem-mass spectrometry (LC-MS/MS).

As the soybean background negatively impacts the enzymatic assay, the measurement of the activity of the bacterially-produced HPPD-4 protein had to be performed in a similar matrix. Therefore, the bacterially-produced HPPD-4 protein was added during protein extraction from non-GM soybean, and the determination of its activity was performed in the same way.

The GMB151 soybean-extracted HPPD-4 protein was shown to have a specific activity of $1.88E^{-02}$ nmol/min/µg enzyme. This activity is comparable with the bacterially-produced HPPD-4 protein as measured in the leaf extracts ($2.77E^{-02}$ nmol/min/µg). As the two values are comparable, the HPPD-4 in GMB151 soybean and the bacterially-produced HPPD-4 protein can be considered as functionally equivalent.

Conclusion

The equivalence of the GMB151 soybean-purified HPPD-4 protein with bacterially-produced HPPD-4 protein was demonstrated based on a panel of analytical tests and assays, including analysis of a Coomassie-stained SDS-PAGE, western blot analysis, glycostaining analysis, mass spectroscopy, and N-terminal sequence analysis. Both bacterially-produced and GMB151 soybean-purified HPPD-4 proteins consist of the mature HPPD-4 protein lacking the N-terminal methionine. *In planta*, a second form is present: the mature HPPD-4 protein carrying a modified cysteine residue at the N-terminus. This additional modified cysteine residue is derived from incomplete cleavage of the transit peptide. Overall, the comparison of the plant-purified HPPD-4 protein from GMB151 soybean with the bacterially-produced HPPD-4 protein demonstrated that both HPPD-4 proteins are comparable for all assessed structural characteristics.

6.5. Digestibility and stability of the introduced proteins

6.5.1. Digestibility and stability of the Cry14Ab-1 protein

***In vitro* digestibility in human simulated gastric fluid**

The bacterially-produced Cry14Ab-1 protein was tested for digestibility in human simulated gastric fluid (SGF) containing pepsin at pH 1.2 for incubation times from 0.5 to 60 minutes.

The Cry14Ab-1 protein was incubated at 37°C in SGF, and samples were taken for analysis at time points of 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. The resulting protein solution was analyzed for presence of the test proteins and potential stable protein fragments by SDS-PAGE followed by Coomassie blue staining (Figure 21) and by western blot (Figure 22).

The Cry14Ab-1 protein was digested very rapidly, within 30 seconds of incubation in SGF in the presence of pepsin, at pH 1.2.

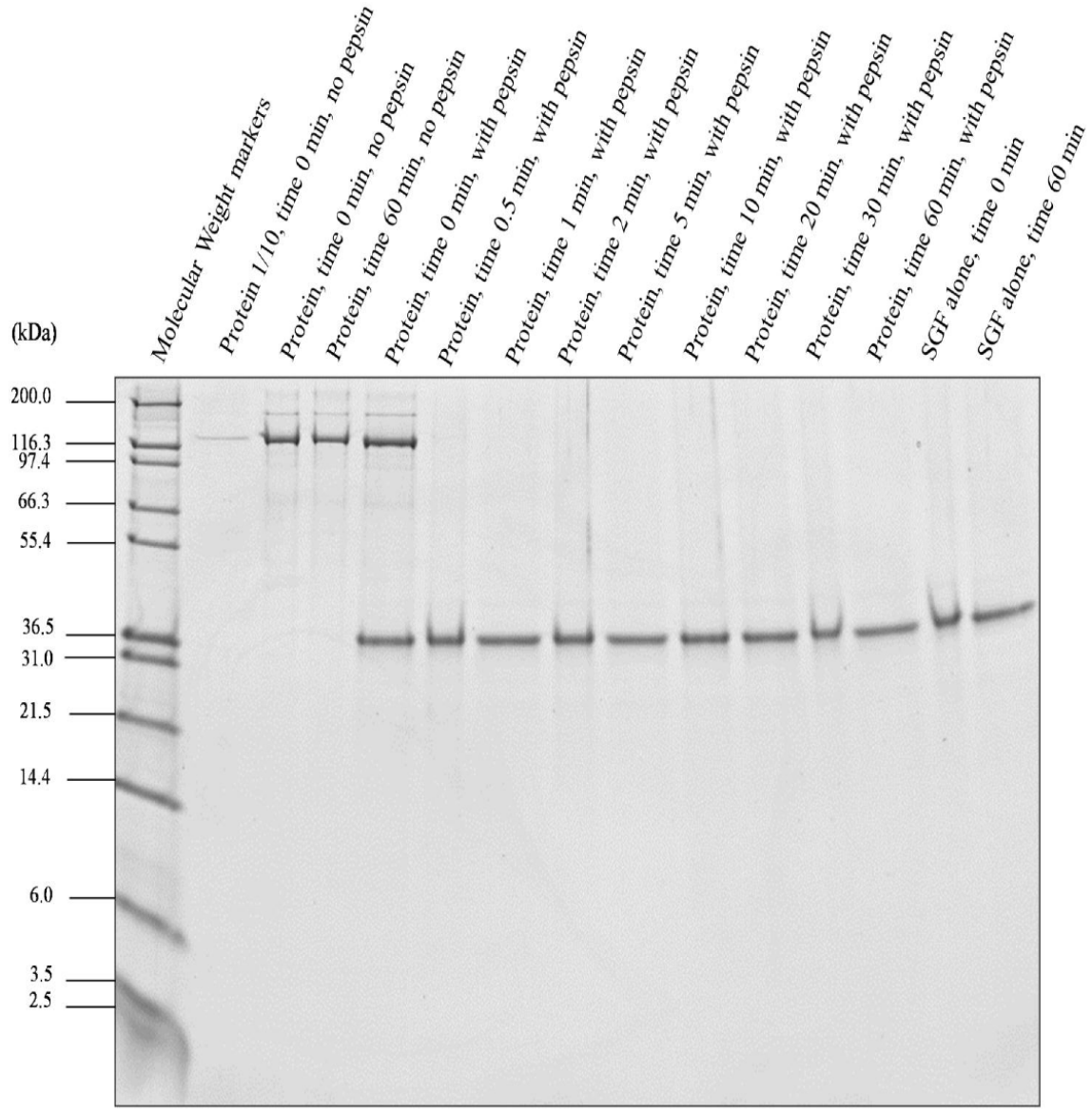


Figure 21. Coomassie blue stained SDS-PAGE of Cry14Ab-1 protein after incubation in human simulated gastric fluid

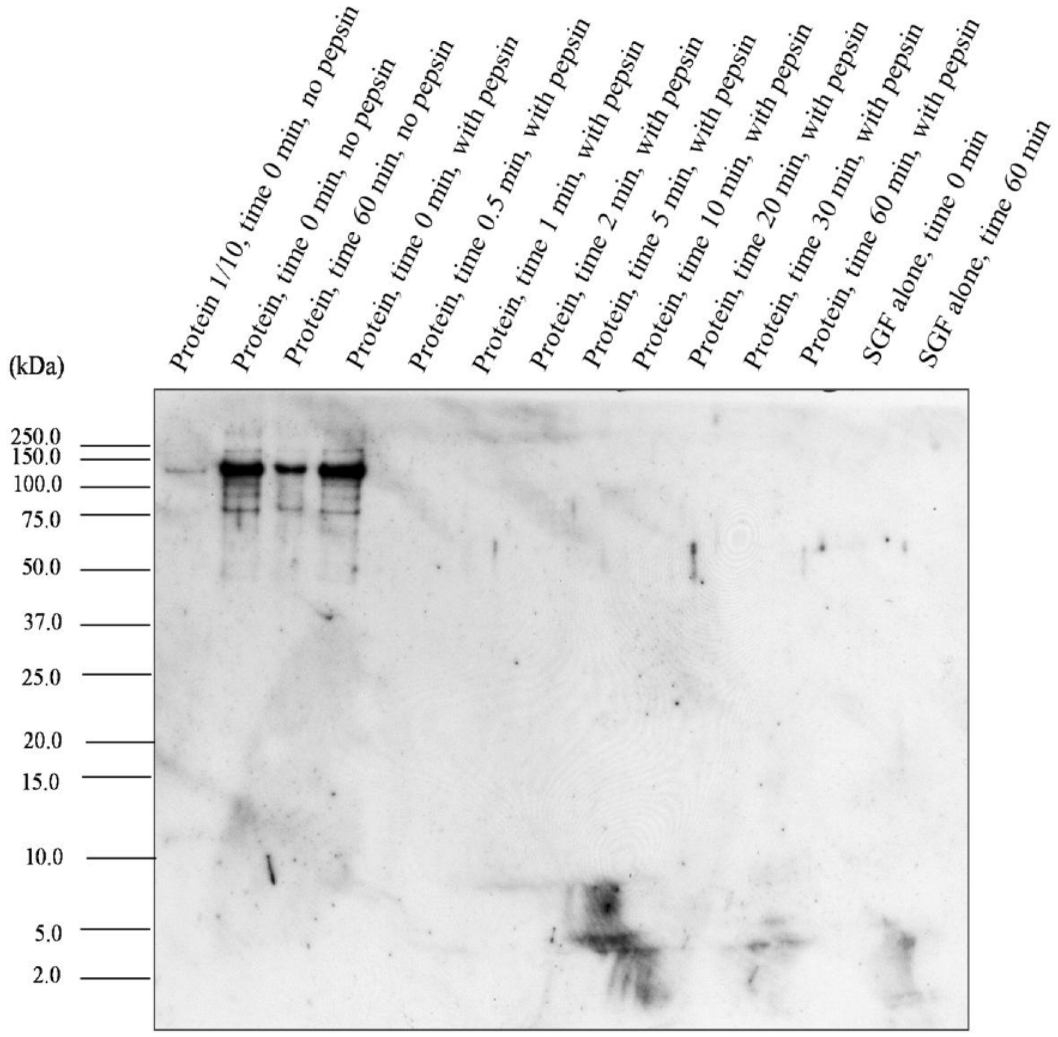


Figure 22. Western blot of Cry14Ab-1 protein after incubation in human simulated gastric fluid

Note: The observed signals below the 10 kDa molecular weight marker correspond to unspecific background that was generated during the chemiluminescence detection step. As no degradation band was visible at these molecular weights in the 0.5 to 5 minute lanes, this background was considered not to compromise interpretation of the results.

The effect of temperature as assessed by SDS-PAGE and western blot

The effect of temperature was assessed on bacterially-produced Cry14Ab-1 using SDS-PAGE and western blot analysis.

Samples of Cry14Ab-1 were incubated for 30 minutes at 4, 25, 37, 55, 75, and 95°C, followed by SDS-PAGE and western blot analysis. The sample treated at 4°C was used for comparison of the other temperature-treated samples.

The SDS-PAGE (Figure 23) and western blot analysis (Figure 24) produced similar results. After temperature treatments at 25, 37, 55, and 75°C, the majority of Cry14Ab-1 remained in the

supernatant as soluble protein. After temperature treatments at 75°C, the intensity of the Cry14Ab-1 band decreased and a smearing pattern became more pronounced in the higher molecular weight region of the lanes. After temperature treatment at 95°C, the Cry14Ab-1 band was no longer visible, and only a faint smearing pattern remained visible. The SDS-PAGE and western blot results indicate that Cry14Ab-1 loses its stability after temperature treatment of 75°C.

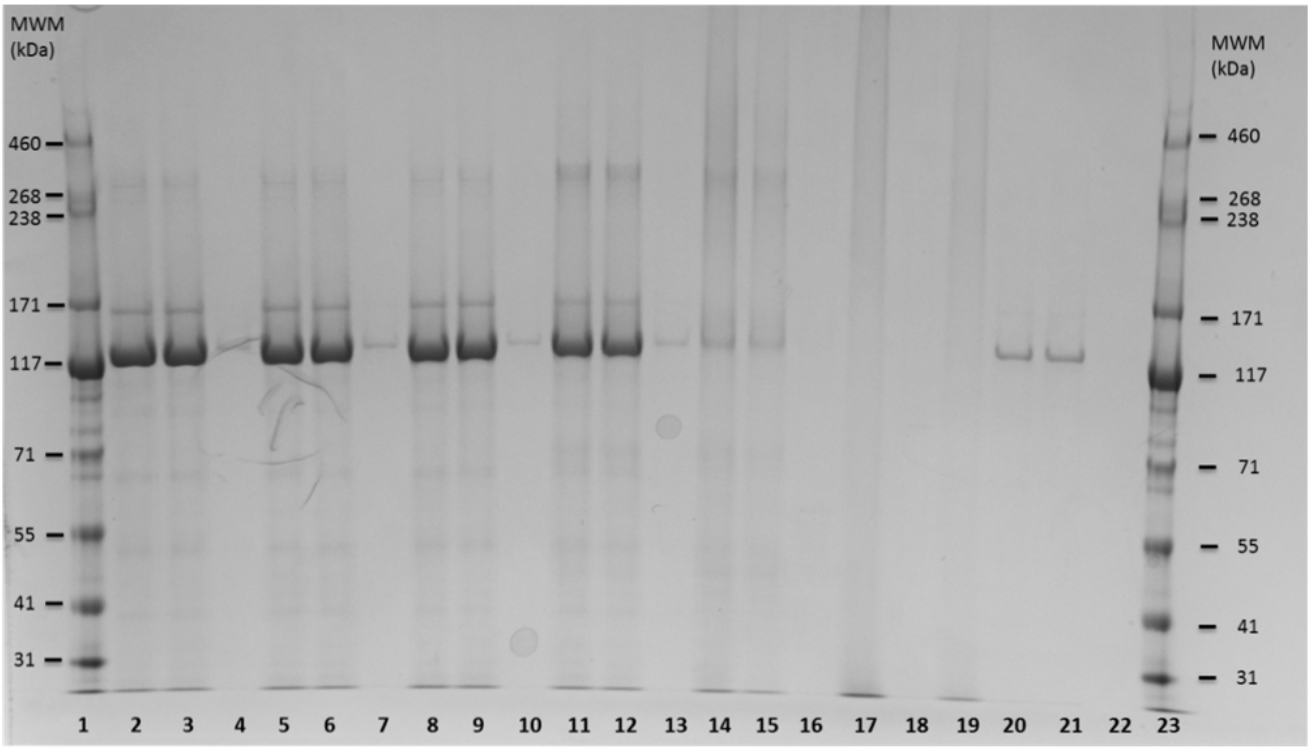


Figure 23. Coomassie blue-stained SDS-PAGE analysis of Cry14Ab-1 protein after temperature treatment

Loading Order:

Lane	Sample
1	HiMark Pre-Stained Protein Standard Molecular Weight Marker (MWM)
2	2 µg Cry14Ab-1, 4°C - Uncentrifuged (UC)
3	2 µg Cry14Ab-1, 4°C - Supernatant (S)
4	2 µg Cry14Ab-1, 4°C - Pellet (P)
5	2 µg Cry14Ab-1, 25°C - UC
6	2 µg Cry14Ab-1, 25°C - S
7	2 µg Cry14Ab-1, 25°C - P
8	2 µg Cry14Ab-1, 37°C - UC
9	2 µg Cry14Ab-1, 37°C - S
10	2 µg Cry14Ab-1, 37°C - P
11	2 µg Cry14Ab-1, 55°C - UC
12	2 µg Cry14Ab-1, 55°C - S
13	2 µg Cry14Ab-1, 55°C - P
14	2 µg Cry14Ab-1, 75°C - UC
15	2 µg Cry14Ab-1, 75°C - S
16	2 µg Cry14Ab-1, 75°C - P
17	2 µg Cry14Ab-1, 95°C - UC
18	2 µg Cry14Ab-1, 95°C - P
19	2 µg Cry14Ab-1, 95°C - S
20	0.2 µg Cry14Ab-1, 4°C (10% Control) - UC
21	0.2 µg Cry14Ab-1, 4°C (10% Control) - S
22	0.2 µg Cry14Ab-1, 4°C (10% Control) - P
23	HiMark Pre-Stained Protein Standard Molecular Weight Marker

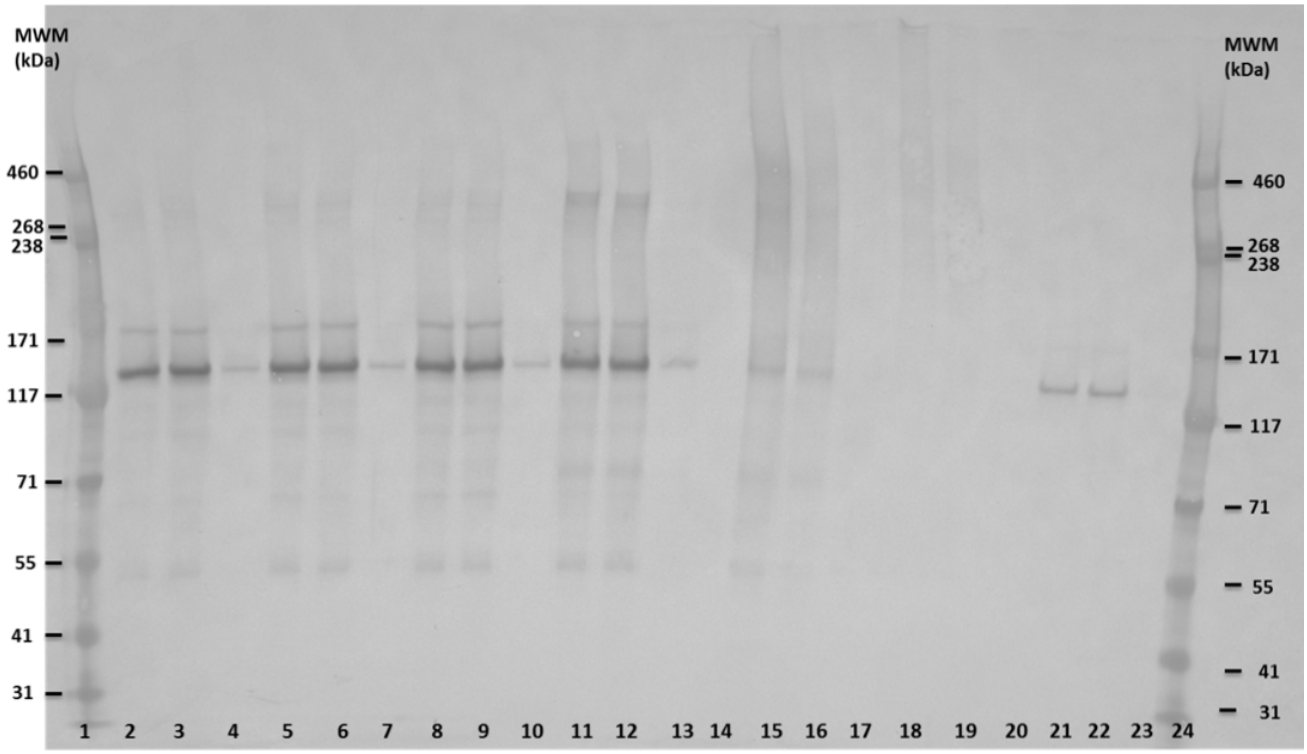


Figure 24. Western blot analysis of Cry14Ab-1 protein after temperature treatment

Loading Order:

Lane	Sample
1	HiMark Pre-Stained Protein Standard Molecular Weight Marker (MWM)
2	50 ng Cry14Ab-1, 4°C - Uncentrifuged (UC)
3	50 ng Cry14Ab-1, 4°C - Supernatant (S)
4	50 ng Cry14Ab-1, 4°C - Pellet (P)
5	50 ng Cry14Ab-1, 25°C - UC
6	50 ng Cry14Ab-1, 25°C - S
7	50 ng Cry14Ab-1, 25°C - P
8	50 ng Cry14Ab-1, 37°C - UC
9	50 ng Cry14Ab-1, 37°C - S
10	50 ng Cry14Ab-1, 37°C - P
11	50 ng Cry14Ab-1, 55°C - UC
12	50 ng Cry14Ab-1, 55°C - S
13	50 ng Cry14Ab-1, 55°C - P
14	Blank lane
15	50 ng Cry14Ab-1, 75°C - UC
16	50 ng Cry14Ab-1, 75°C - S
17	50 ng Cry14Ab-1, 75°C - P
18	50 ng Cry14Ab-1, 95°C - UC
19	50 ng Cry14Ab-1, 95°C - S
20	50 ng Cry14Ab-1, 95°C - P
21	5 ng Cry14Ab-1, 4°C (10% Control) - UC
22	5 ng Cry14Ab-1, 4°C (10% Control) - S
23	5 ng Cry14Ab-1, 4°C (10% Control) - P
24	HiMark Pre-Stained Protein Standard Molecular Weight Marker

6.5.2. Digestibility and stability of the HPPD-4 protein

***In vitro* digestibility in human simulated gastric fluid**

The HPPD-4 protein was tested for digestibility in human SGF containing pepsin at pH 1.2 for incubation times from 0.5 to 60 minutes.

The HPPD-4 protein was incubated at 37°C in SGF, and samples were taken for analysis at time points of 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. The resulting protein solution was analyzed for presence of the test proteins and potential stable protein fragments by SDS-PAGE followed by Coomassie blue staining (Figure 25) and by western blot (Figure 26).

The HPPD-4 protein was digested very rapidly, within 30 seconds of incubation in SGF in the presence of pepsin at pH 1.2.

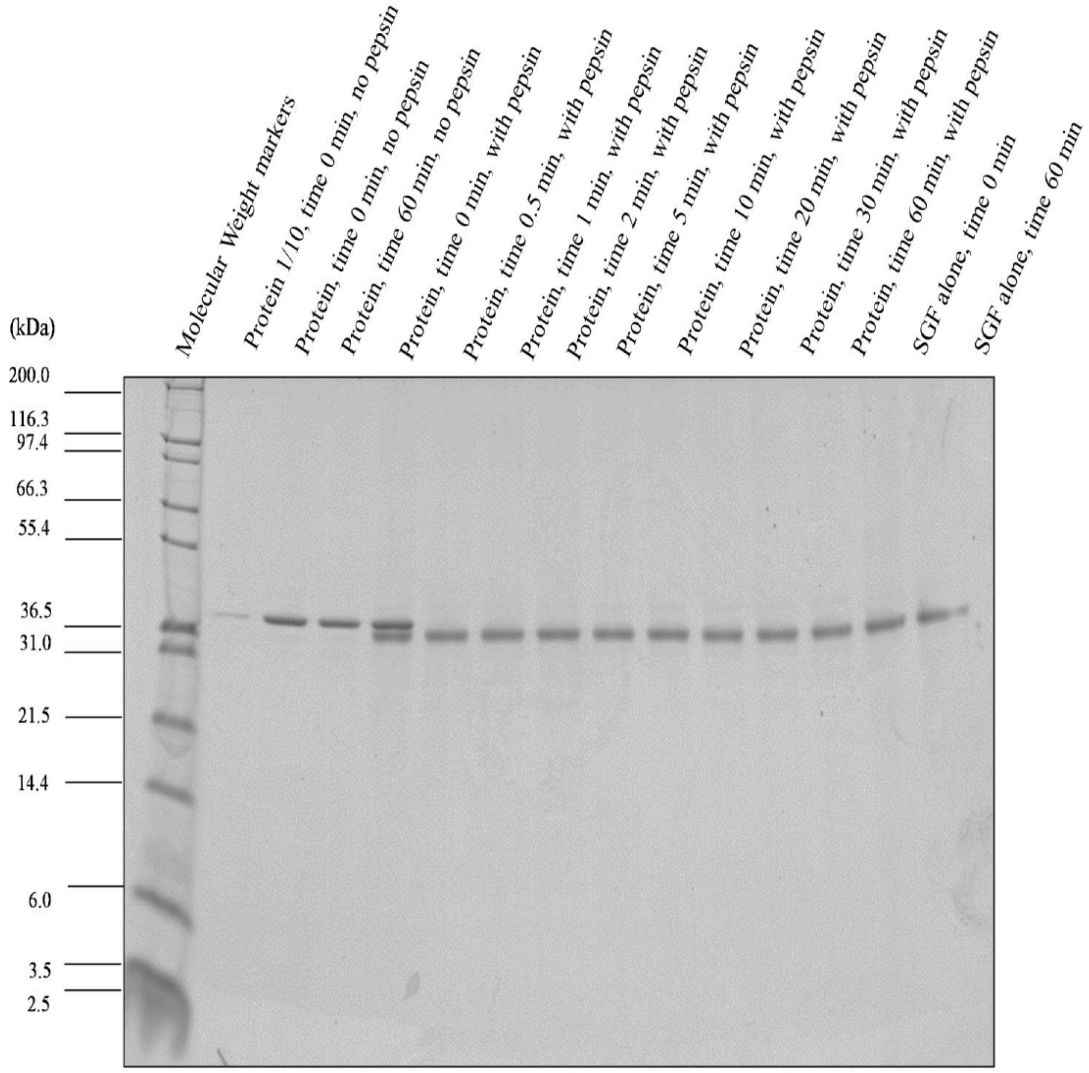


Figure 25. Coomassie blue-stained SDS-PAGE of HPPD-4 protein after incubation in human simulated gastric fluid

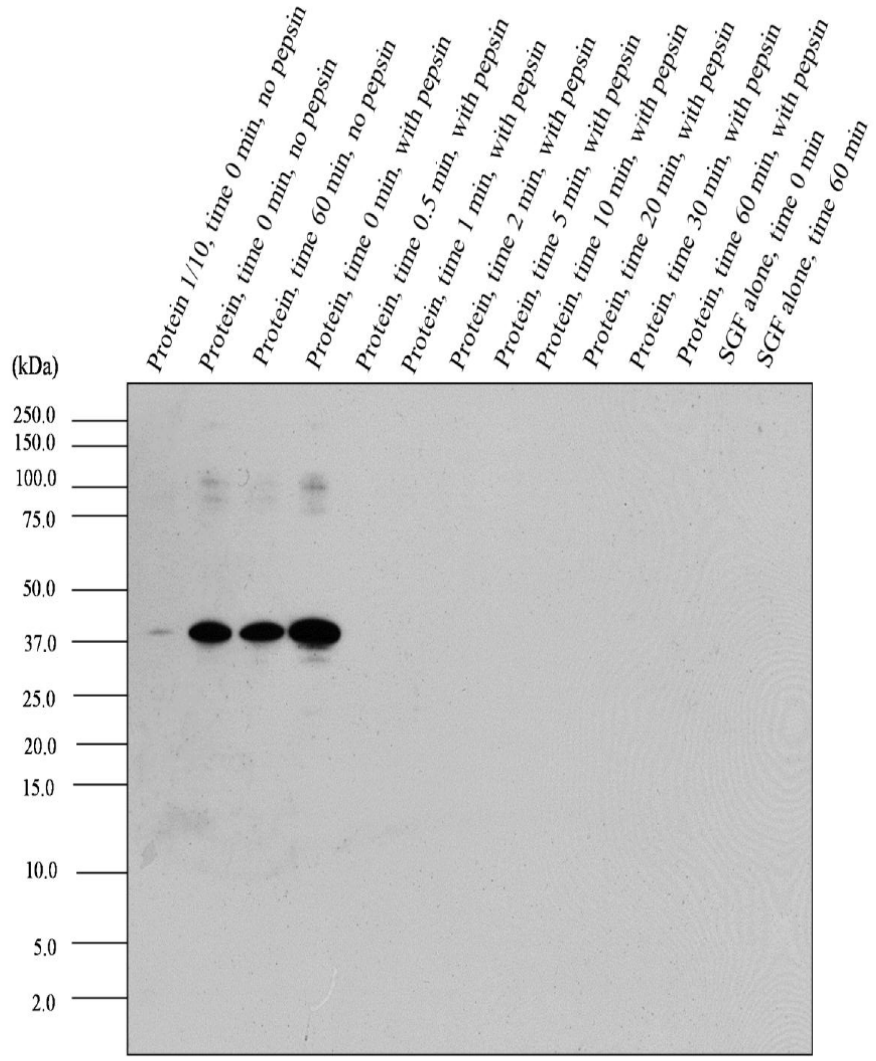


Figure 26. Western blot of HPPD-4 protein after incubation in human simulated gastric fluid

The effect of temperature as assessed by SDS-page and western blot

The effect of temperature was assessed on the HPPD-4 protein using SDS-PAGE and western blot analysis.

The HPPD-4 protein was incubated at 4, 25, 37, 55, 75, and 95°C for 30 minutes. The sample treated at 4°C was used for comparison of the other temperature-treated samples.

The SDS-PAGE (Figure 27) and the western blot (Figure 28) produced similar results. After temperature treatments at 25 and 37°C, the majority of HPPD-4 remained in the supernatant as soluble protein. After temperature treatments at 55°C and above, HPPD-4 began to appear as insoluble protein. The SDS-PAGE and western blot results indicate that HPPD-4 begins to lose its solubility upon temperature treatments at 55°C and greater.

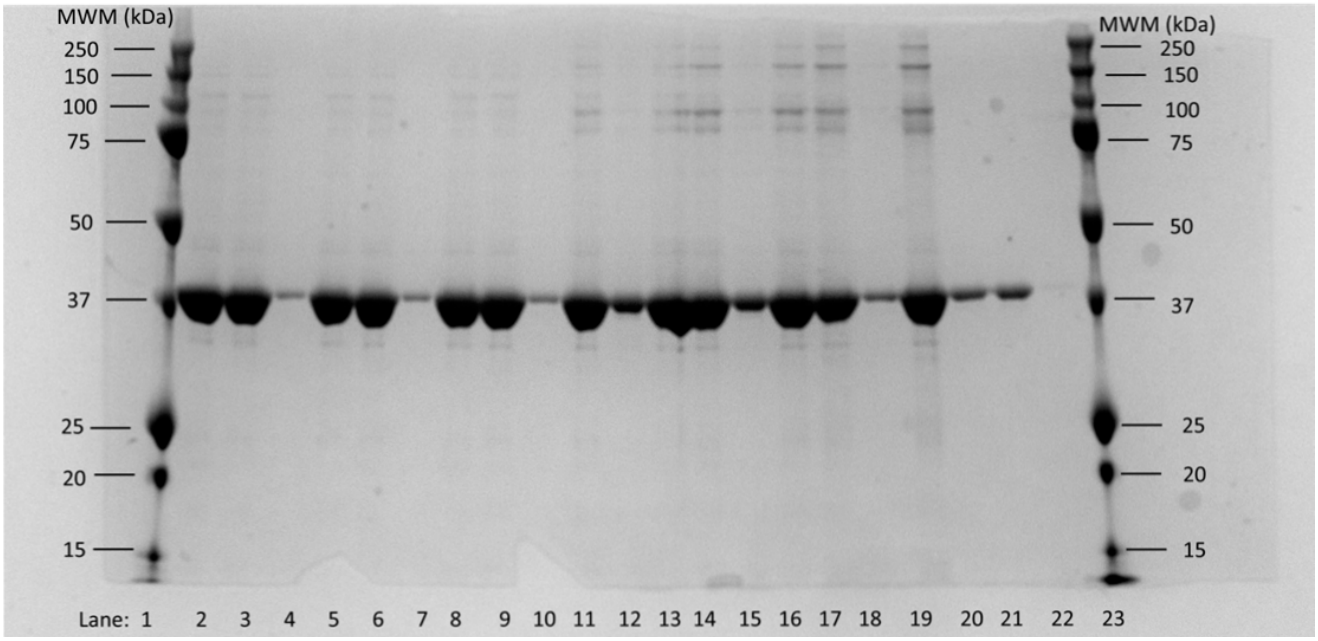


Figure 27. Coomassie blue-stained SDS-PAGE analysis of HPPD-4 protein after temperature treatment

Loading Order:

Lane	Sample
1	Precision Plus Protein™ Dual Xtra Standards Molecular Weight Marker
2	5 µg HPPD-4, 4°C - Uncentrifuged (UC)
3	5 µg HPPD-4, 4°C - Supernatant (S)
4	5 µg HPPD-4, 4°C - Pellet (P)
5	5 µg HPPD-4, 25°C - UC
6	5 µg HPPD-4, 25°C - S
7	5 µg HPPD-4, 25°C - P
8	5 µg HPPD-4, 37°C - UC
9	5 µg HPPD-4, 37°C - S
10	5 µg HPPD-4, 37°C - P
11	5 µg HPPD-4, 55°C - UC
12	5 µg HPPD-4, 55°C - S
13	5 µg HPPD-4, 55°C - P
14	5 µg HPPD-4, 75°C - UC
15	5 µg HPPD-4, 75°C - S
16	5 µg HPPD-4, 75°C - P
17	5 µg HPPD-4, 95°C - UC
18	5 µg HPPD-4, 95°C - S
19	5 µg HPPD-4, 95°C - P
20	0.5 µg HPPD-4, 4°C (10% Control) - UC
21	0.5 µg HPPD-4, 4°C (10% Control) - S
22	0.5 µg HPPD-4, 4°C (10% Control) - P
23	Precision Plus Protein™ Dual Xtra Standards Molecular Weight Marker

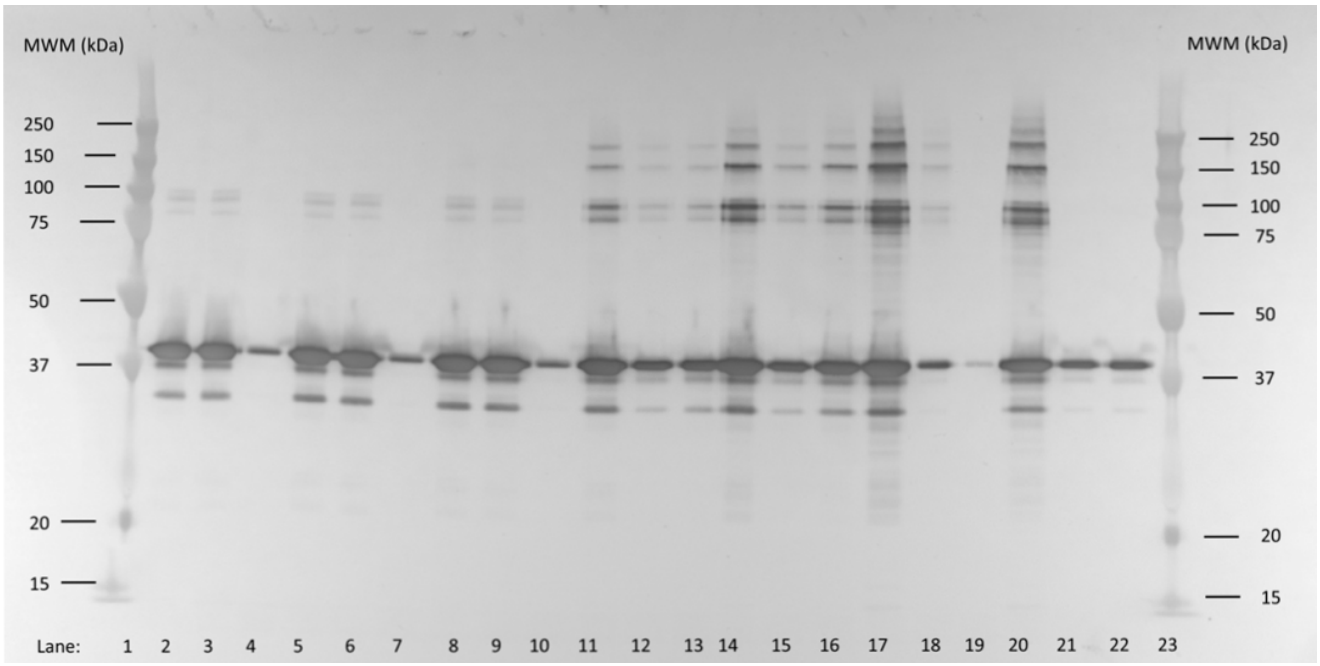


Figure 28. Western blot analysis of HPPD-4 protein after temperature treatment

Loading Order:

Lane	Sample
1	Precision Plus Protein™ Dual Xtra Standards Molecular Weight Marker
2	250 ng HPPD-4, 4°C - Uncentrifuged (UC)
3	250 ng HPPD-4, 4°C - Supernatant (S)
4	250 ng HPPD-4, 4°C - Pellet (P)
5	250 ng HPPD-4, 25°C - UC
6	250 ng HPPD-4, 25°C - S
7	250 ng HPPD-4, 25°C - P
8	250 ng HPPD-4, 37°C - UC
9	250 ng HPPD-4, 37°C - S
10	250 ng HPPD-4, 37°C - P
11	250 ng HPPD-4, 55°C - UC
12	250 ng HPPD-4, 55°C - S
13	250 ng HPPD-4, 55°C - P
14	250 ng HPPD-4, 75°C - UC
15	250 ng HPPD-4, 75°C - S
16	250 ng HPPD-4, 75°C - P
17	250 ng HPPD-4, 95°C - UC
18	250 ng HPPD-4, 95°C - S
19	0.25 ng HPPD-4, 4°C (10% Control) - P
20	250 ng HPPD-4, 95°C - P
21	0.25 ng HPPD-4, 4°C (10% Control) - UC
22	0.25 ng HPPD-4, 4°C (10% Control) - S
23	Precision Plus Protein™ Dual Xtra Standards Molecular Weight Marker

6.6. Aerobic degradation of Cry14Ab-1 in soil

A study was conducted to estimate the aerobic degradation rate of the Cry14Ab-1 protein in representative agricultural soils. The time required for 50% degradation of the initial protein concentration (DT_{50}) was determined for each type of soil. U.S. EPA Guideline OPPTS 835.4100

Aerobic Soil Metabolism (EPA 712-C-08-016) was followed in soil handling and study design. Detailed materials and methods are provided in Appendix 5.

Briefly, the aerobic soil degradation study was performed with Cry14Ab-1 protein at a concentration of 80 µg/g in four U.S. soils. Soils used in the study were collected from the top layer (6 inches) of the soil from Sanger, CA, Bagley, IA, Stilwell, KS, and Louisville, NE. Sampling site information is summarized in Table 11. These soils were selected because they provided a diversity of physiochemical soil characteristics, as well as diverse geographical and climatic conditions representative of the agricultural soil types for soybean cultivation in the U.S. Soil properties such as soil type, texture, and organic matter content were characterized by Agvise Laboratories (Northwood, ND) and are summarized in Table 12.

Received soils were cleaned by removing the stones and plant parts and sieving through a 2-mm mesh sieve. The cleaned soils were stored at 4.5–6.0°C, pre-incubated (acclimated) for five days to restore the microbial activity, and used in the study within a month of collection. A maximum of three months when stored at 4 ± 2°C is allowed under EPA guidelines. Cry14Ab-1 protein solution was applied to acclimated soil at 80 µg Cry14Ab-1 protein per gram DW of soil. The approach was to provide a worst-case scenario of protein concentration in the soil, greater than what would be expected in root exudate or sloughed-off plant tissue. Treated soil samples were incubated at approximately 20°C in the dark under aerobic conditions for 0, 0.1, 0.3, 1, 1.3, 2, 5, and 7 days. At each sampling interval, samples were analyzed by quantitative Cry14Ab-1 ELISA. The samples at Day 0 and 7 were also analyzed by the *C. elegans* bioassay (Appendix 6).

Adsorption of Cry14Ab-1 to soil particles was not a factor in measurement of the degradation of Cry14Ab-1 as at least 77% of the Cry14Ab-1 applied was extracted out and quantified at time 0 (Table 13). DT_{50} values were calculated using Cry14Ab-1 protein concentration detected by ELISA and KinGUI 2 software. The degradation of Cry14Ab-1 protein was best described by Double First Order in Parallel kinetics in all soil types. The DT_{50} values were 0.3 day for the CA soil, 0.1 day for the IA soil, 0.2 day for the KS soil, and 0.2 day for the NE soil (Table 14).

Bioassay results indicate that the Cry14Ab-1 protein was active in Day 0 soil samples (Table 15). There was no quantifiable active Cry14Ab-1 protein in Day 7 soil samples (Table 16).

Table 11. Soil sample site information

Designation	CA soil	IA soil	KS soil	NE soil
US Textural Class	Sandy Loam	Loam	Silt Loam	Silt Loam
Geographic Location	17915 E. Annadale Ave. Sanger, CA	2699 Hwy 141, Bagley, IA	17745 S. Metcalf Ave. Stilwell, KS	Louisville, Sarpy, NE
Site Description	Agricultural research field station	Field	Grass covered field	Fallow area of alfalfa/grass mix
Pesticide Use History at Site	None applied since 2014	Glyphosate	None	None applied since 2004
Soil Collected: Depth from Surface of Soil	0–6 inches	0–6 inches	0–6 inches	0–6 inches
Collection Date	10/13/2016	10/12/2016	10/11/2016	10/10/2016

Table 12. Soil properties for soil aerobic degradation study

Designation	CA soil	IA soil	KS soil	NE soil
Soil ID	101416-S	101316-S	101216-S	101116-S
Source	Sanger, CA	Bagley, IA	Stilwell, KS	Louisville, NE
US Textural Class	Sandy Loam	Loam	Silt Loam	Silt Loam
Percent Sand	62.7	46.9	5.1	15.6
Percent Silt	29.7	31.6	80.6	60.3
Percent Clay	7.6	21.5	14.3	24.1
Bulk Density (disturbed) gm/cc	1.37	0.98	0.90	0.99
Cation Exchange Capacity (meq/100 g)	8.4	17.5	9.4	16.3
Max. Water Hold Capacity (MWHC) (gm/100 gm)	26.7	50.8	47.4	55.4
% Moisture at 1/10 Bar (pF* 2.0)	24.7	36.7	37.0	54.0
% Moisture at 1/3 Bar (pF* 2.5)	19.3	23.7	24.7	26.6
%Organic Carbon – Walkley Black	0.46	2.3	1.3	1.9
%Organic Matter – Walkley Black	0.80	4.0	2.3	3.3
pH in 1:1 soil: water ratio	7.0	6.5	7.7	7.0
Total Nitrogen (Analyzer) (%)	0.05	0.20	0.12	0.15
Soluble Salts (mmhos/cm)	0.57	0.13	0.15	0.19
Total Phosphorus (ppm)	554	371	457	497
Calcium (ppm)	987	2300	1510	2200
Magnesium (ppm)	187	300	56	236
Sodium (ppm)	20	15	14	16
Potassium (ppm)	207	108	61	289
Hydrogen (ppm)	13	31	11	25

*: pF value represents the logarithm of the height of water column (cm) needed to provide suction

Table 13. Cry14Ab-1 detected in soil extracts (µg/g dry soil)

Soil samples		Cry14Ab-1 concentration at each sampling interval (day)							
		0.0	**0.1**	**0.3**	**1.0**	**1.3**	**2.0**	**5.0**	**7.0**
CA soil	Sample 1	80.4	53.3	36.8	18.7	11.8	6.3	1.4	0.8
	Sample 2	76.1	44.4	40.6	17.6	9.2	5.6	1.8	0.8
	Sample 3	69.7	50.1	39.3	17.4	14.9	7.4	1.6	0.8
	Mean	**75.4**	**49.3**	**38.9**	**17.9**	**11.9**	**6.5**	**1.6**	**0.8**
	SD	5.4	4.5	1.9	0.7	2.9	0.9	0.2	0.0
	% Protein recovered (Time 0)	94.3	--	--	--	--	--	--	--
IA soil	Sample 1	59.4	34.4	25.9	10.2	7.1	4.5	1.8	1.0
	Sample 2	65.5	32.0	22.4	9.8	7.5	4.2	1.6	1.0
	Sample 3	61.3	31.8	26.9	9.7	7.2	4.2	1.4	1.0
	Mean	**62.1**	**32.7**	**25.0**	**9.9**	**7.3**	**4.3**	**1.6**	**1.0**
	SD	3.1	1.5	2.4	0.3	0.2	0.1	0.2	0.0
	% Protein recovered (Time 0)	77.6	--	--	--	--	--	--	--
KS soil	Sample 1	68.1	42.9	27.4	8.2	4.5	2.3	0.9	0.6
	Sample 2	66.0	42.7	27.9	8.0	4.6	2.3	0.9	1.2
	Sample 3	69.6	46.3	30.9	7.5	4.5	2.4	1.0	0.6
	Mean	**67.9**	**44.0**	**28.7**	**7.9**	**4.5**	**2.4**	**0.9**	**0.8**
	SD	1.8	2.0	1.9	0.3	0.1	0.1	0.0	0.4
	% Protein recovered (Time 0)	84.9	--	--	--	--	--	--	--
NE soil	Sample 1	66.5	36.9	29.1	10.1	5.9	2.7	0.8	0.6
	Sample 2	70.3	36.0	30.1	10.3	5.2	2.5	0.8	0.6
	Sample 3	73.2	38.0	29.5	9.8	5.3	2.6	0.8	0.6
	Mean	**70.0**	**36.9**	**29.6**	**10.1**	**5.5**	**2.6**	**0.8**	**0.6**
	SD	3.4	1.0	0.5	0.3	0.4	0.1	0.0	0.0
	% Protein recovered (Time 0)	87.5	--	--	--	--	--	--	--

Table 14. DT_{50} values

Soil	Kinetic model	DT_{50} (Days)	Chi^2 error (%)	Visual assessment
CA soil	Double First Order in Parallel (DFOP)	0.3263	2.406	Good
IA soil	DFOP	0.1281	4.106	Good
KS soil	DFOP	0.2110	2.650	Good
NE soil	DFOP	0.1525	3.603	Good

Table 15. Bioassay result of Day 0 samples

Soil extract of treatment solution	Mean %Inhibition	Dose-response curve		
		LLOQ (%)	R^2	No. of non-zero dose
CA soil	42.4	23.8	0.99	7
IA soil	37.9	17.1	0.98	7
KS soil	43.5	21.9	0.98	7
NE soil	45.3	18.4	0.99	6
2 mg/mL Cry14Ab-1 treatment solution	61.7	19.7	0.99	7

Table 16. Bioassay result of Day 7 samples

Soil extract of treatment solution	Mean %Inhibition	Dose-response Curve		
		LLOQ (%)	R^2	No. of non-zero dose
CA soil	6.4	8.8	0.98	5
IA soil	5.7	8.5	0.96	5
KS soil	6.3	9.2	0.95	6
NE soil	9.5	16.0	0.97	5

Conclusions

The results of this study demonstrate that the Cry14Ab-1 protein degrades in CA, IA, KS, and NE soils in less than 0.5 days at 20°C (68°F). This temperature regime and study setup were selected because of recommendations in the U.S. EPA guidance document EPA 712-C-08-016. Temperature is an important factor regulating microbial activity and shaping the soil microbial community. As the temperature increases (to a point), the degradation rate of the protein in the soil would increase. Conversely, at lower temperatures, the rate would most likely decrease. The absolute minimum soil temperature for sowing soybean is 10°C (50°F) at a depth of 5 cm (2 inches), and preferred soil temperatures for sowing are 25°C (77°F) (ISU Extension and Outreach, 2019). Over a growing season, 20°C is a realistic soil temperature within the primary rooting zone of soybean in the primary soybean production regions of the U.S. Using Iowa in 2018 as an example, the 10 cm (4 inch) year-long soil temperature average (average of the daily averages of all reporting weather stations) was 52°F (11°C). Over the growing season (defined May 1 to October 15 for the purpose of this calculation), the average was 71°F (22°C) (ISU Department of Agronomy, 2019). This demonstrates the appropriateness of 20°C as the temperature regime for assessments of aerobic soil degradation.

6.7. Homology of the introduced proteins with toxins and allergens

The amino acid sequence of the Cry14Ab-1 and HPPD-4 proteins were evaluated for potential homology with known allergens and known toxins by using several *in silico* approaches. This search evaluated the potential amino acid sequence identity of the query proteins with known allergens by using two *in silico* approaches:

- An overall identity search was carried out to compare the complete query sequence with all protein sequences present in the public allergen database COMPARE. The FASTA algorithm was used, with the BLOSUM50 scoring matrix and an E-value threshold of 10. The criterion indicating potential allergenicity was ≥ 35% identity over at least 80 consecutive amino acids with an allergenic protein.

- An 8-mer search was carried out to identify any short sequences of eight amino acids or longer that share 100% identity to an allergenic protein. This search was performed using SeqMatchAll from the EMBOSS suite, which compared the query sequence with all known allergens present in the allergen database.

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

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

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

For Cry14Ab-1, the overall identity search showed no biologically relevant identity between the query protein and any known allergenic proteins. In addition, the 8-mer search showed no 100% identity with known allergenic proteins. As expected, the overall identity search against the general protein database showed that in most cases, the Cry14Ab-1 protein matched with Cry family sequences from *Bt* that are considered safe. Furthermore, no biologically relevant identities were found with any toxic proteins from the BASF toxin database. In conclusion, there are neither allergenic nor toxicological *in silico* findings associated with the Cry14Ab-1 protein.

For HPPD-4, the overall identity search showed no biologically relevant identity between the query protein and any known allergenic proteins. In addition, the 8-mer search showed no 100% identity with known allergenic proteins. As expected, the overall identity search against the general protein database showed that in most cases, the HPPD-4 protein matched with other HPPD proteins from various origins. In addition, no significant similarities were found with any toxic protein from the BASF toxin database. In conclusion, there are neither allergenic nor toxicological *in silico* findings associated with the HPPD-4 protein.

6.8. Acute toxicity study in the mouse

Cry14Ab-1

The Cry14Ab-1 protein was evaluated for acute oral toxicity in male and female C57BL/6J mice.

Groups of six male and six female C57BL/6J mice were administered the Cry14Ab-1 protein by oral gavage at the limit dose level of 2000 mg/kg body weight. Similarly constituted groups of six male and six female mice received vehicle alone, administered in the same manner, and acted as controls. All animals were observed for clinical signs daily for 15 days while their body weights and food consumption were measured weekly. At the termination of the study period, the animals were weighed and subjected to a necropsy, including a macroscopic examination. The tissues were retained for possible microscopic examinations.

There were no mortalities, no treatment-related clinical signs, no effects on the body weight and food consumption parameters, and no macroscopic changes at necropsy in C57BL/6J mice after an acute oral administration of the Cry14Ab-1 protein at 2000 mg/kg body weight.

In conclusion, the treatment with the Cry14Ab-1 protein at 2000 mg/kg body weight via the oral route did not produce any signs of systemic toxicity in the male and female C57BL/6J mice.

HPPD-4

The HPPD-4 protein was evaluated for acute oral toxicity in male and female C57BL/6J mice.

Groups of 10 male and 10 female C57BL/6J mice were administered the HPPD-4 protein by oral gavage at the limit dose level of 2000 mg/kg body weight. Similarly constituted groups of 10 male and 10 female mice received the vehicle alone, administered in the same manner, and acted as controls. All animals were observed for clinical signs daily for 15 days while their body weights and food consumption were measured weekly. At the termination of the study period, the animals were weighed and subjected to a necropsy, including a macroscopic examination. The tissues were retained for possible microscopic examinations.

There were no mortalities, no treatment-related clinical signs, no effects on the body weight and food consumption parameters, and no macroscopic changes at necropsy in C57BL/6J mice after an acute oral administration of the HPPD-4 protein at 2000 mg/kg body weight.

In conclusion, the treatment with the HPPD-4 protein at 2000 mg/kg body weight via the oral route did not produce any signs of systemic toxicity in the male and female C57BL/6J mice.

7. COMPOSITIONAL ANALYSIS OF KEY COMPONENTS

Composition analyses were conducted to determine the levels of nutrients and anti-nutrients in soybean grain and select nutrients in soybean forage samples from multiple field trials as part of the comparative assessment between GMB151 soybean (not treated and treated with trait-specific herbicide), the non-GM counterpart Thorne, and non-GM reference varieties.

7.1. Field productions

Composition analysis was conducted on forage and grain samples from eight field trials conducted in the 2017 continental U.S. season. Selected sites are representative of likely environments where GMB151 soybean will be commercially grown. The field trial site information for the eight sites selected is presented in Table 17 below.

In addition to GMB151 soybean and its non-GM counterpart Thorne, nine reference varieties that represent the natural variability existing in soybean were included in this study to provide reference ranges for the composition assessment. The non-GM reference varieties were provided by Schillinger Genetics Inc. (Des Moines, IA) and NuPride Genetics Network, LLC (Lincoln, NE). Each field trial site planted three of the nine varieties. The entries included are presented in Table 18.

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

Conventional herbicide management was applied to all entries. The HPPD-inhibitor herbicide IFT was applied to Entry E at a target rate of 70.1 g a.i./ha at growth stage BBCH 00–03 (pre-emergence).

Forage samples were harvested at BBCH 71–78 (pod formation) and grain samples were harvested at BBCH 89–99 (grain maturity) from all plots for subsequent composition analysis. Forage samples were shipped frozen and grain samples were shipped at ambient temperatures to EPL Bio Analytical Services (EPL Bio Analytical Services, 9095 West Harristown Blvd, Niantic, IL, U.S). All samples were stored frozen upon receipt at -20°C.

Table 17. Field trial sites for compositional analysis sample production

Site Code	Nearest Town or City	Site County	State
01	Richland	Jefferson	Iowa
02	York	York	Nebraska
03	Elk Horn	Shelby	Iowa
04	Stewardson	Shelby	Illinois
05	Germansville	Lehigh	Pennsylvania
06	Fisk	Butler	Missouri
07	Larned	Pawnee	Kansas
11	Carlyle	Clinton	Illinois

Table 18. Description of entries in production trials for composition analysis

Entry ID	Material Name	Other Specification	Site Code
A	Thorne	Non-GM Counterpart	All
D	GMB151	GM not treated	All
E	GMB151	GM treated	All
F	E3494	Non-GM Reference	06, 07, 11
G	NGN 3121STS	Non-GM Reference	06, 07, 11
H	NGN 3347C	Non-GM Reference	06, 07, 11
I	E2282	Non-GM Reference	01, 03, 05
J	E2692	Non-GM Reference	01, 03, 05
K	NGN 3292C	Non-GM Reference	01, 03, 05
L	E3066	Non-GM Reference	02, 04
M	E2993	Non-GM Reference	02, 04
N	E3192	Non-GM Reference	02, 04

7.2. Composition analysis

Composition analysis of the soybean forage and grain samples was conducted at EPL Bio Analytical Services. Soybean forage and grain samples were homogenized in the presence of dry ice using a Robot Coupe R10 and an Ultra Centrifugal Mill for forage and grain samples, respectively. Samples were randomized prior to analysis by EPL Bio Analytical Services in order to control bias.

Forage samples were analyzed for proximates, fiber, calcium, and phosphorus. Grain samples were analyzed for proximates, fiber, amino acids, fatty acids, minerals, vitamins, and anti-nutrients. Table 19 provides detailed descriptions of all analytes measured.

Table 19. Composition analytes, units, and methods for soybean forage and grain

Tissue	Parameter	Units
	Proximates and Fiber	
Grain, Forage	Moisture	% FW
	Ash	% FW, DW
	Carbohydrates	
	Crude Fat	
	Crude Protein	
	Acid Detergent Fiber	
	Neutral Detergent Fiber	
Grain	Total Dietary Fiber	
	Amino Acids	
Grain	Alanine, Arginine, Aspartic Acid, Glutamic Acid, Glycine, Histidine, Isoleucine, Leucine, Lysine, Phenylalanine, Proline, Serine, Threonine, Tyrosine, Valine	% FW, DW
	Cystine, Methionine	
	Tryptophan	
	Fatty Acids	
Grain	C8:0, C10:0, C12:0, C14:0, C14:1, C15:0, C15:1, C16:0, C16:1, C17:0, C17:1, C18:0, C18:1, C18:2, C18:3, C18:4, C19:0, C20:0, C20:1, C20:2, C20:3, C20:4, C20:5, C22:0, C22:1, C22:5 n-3, C22:5 n-6, C22:6, C24:0	% FW, DW, % Total Fatty Acids
	Minerals	
Grain, Forage	Calcium, Phosphorus	mg/kg FW, DW
Grain	Copper, Iron, Magnesium, Manganese, Potassium, Sodium, Zinc	
	Vitamins	
Grain	Alpha Tocopherol (Vitamin E)	mg/kg FW, DW
	Vitamin B1(Thiamine)	
	Vitamin B2 (Riboflavin)	
	Vitamin B3 (Niacin)	
	Vitamin B5 (Pantothenic Acid)	
	Vitamin B6 (Pyridoxine)	
	Vitamin B9 (Folic Acid)	
	Vitamin K1	
	Anti-Nutrients	
	Isoflavones	
	Total Daidzein	mg/kg FW, DW
	Total Genistein	
	Total Glycitein	
	Total Isoflavones	

Table 19. Composition analytes, units, and methods for soybean forage and grain (continued)

Tissue	Parameter	Units
	Lectins	mg/g FW, DW
	Phytic Acid	% FW, DW
	Raffinose	% FW, DW
	Stachyose	
	Trypsin Inhibitor	TIU/mg FW, DW

FW = fresh weight; DW = dry weight; TIU = Trypsin Inhibitor Unit.

7.3. Statistical analysis

Composition data for 192 samples collected from eight sites were statistically analyzed using SAS version 9.4 (SAS, 2002–2012).

Analytes with more than one third of sample values below the limit of quantification (LOQ) were excluded from further statistical evaluation (i.e. analysis of variance (ANOVA) and mean comparisons) and discussion. These analytes are presented in Table 20. Minimum and maximum values for analytes where some values were above LOQ but there were insufficient data for statistical analysis are presented in Table 26 and Table 28.

For several grain analytes (C17:1 heptadecenoic acid, C24:0 lignoceric acid, and vitamin B1), some samples had values less than the LOQ but there were sufficient (more than two thirds) sample values above LOQ for statistical evaluation (Table 20). In these cases, the below LOQ value was substituted by a value equal to half the LOQ in order to be included in the statistical analysis.

Soybean forage and grain samples from eight sites generated a total of 192 observations for each analyte for each soybean tissue. There were 32 observations for Entry A, 32 observations each for Entry D and Entry E, and 96 observations for the reference variety entries (Entries F, G, H, I, J, K, L, M, and N), which were statistically analyzed collectively. Observations for the non-GM counterpart Thorne (Entry A) were compared to GMB151 soybean not treated (Entry D) and GMB151 soybean treated with trait-specific herbicide (Entry E).

Combined-site analysis

Composition data were statistically analyzed for each analyte combined over all sites with a mixed model ANOVA. ANOVA was selected as the statistical method because the field trials contained multiple sites, and within each site, the experimental design structure was an RCBD with multiple entries. A linear mixed model was chosen to appropriately account for Block and Site effects while increasing the precision of Entry comparisons. Block effect is random since it is a source of replication and since its levels are not of interest. The selected sites represent a random sample from many potential sites, and it is desirable to make broad inference to all field sites in the crop growing areas. Therefore, Site is also included in the model as a random effect. Entry effect is fixed because its levels are of specific interest, and if the study were repeated, the same levels would be used again. Interaction between Site and Entry is a random effect since Site is random. This analysis is also referred to as a mixed-model ANOVA and is the most standard model for this type of experimental design.

The additive model for the design is:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \delta_{k(j)} + \varepsilon_{ijk}$$

where Y_{ijk} is the individual value measurement, μ the overall mean, α_i the fixed effect associated with entry, β_j the random effect associated with site, $\alpha\beta_{ij}$ the random effect for the interaction of entry by site, $\delta_{k(j)}$ the random effect associated with block nested within site, and ε_{ijk} the random error. The degrees of freedom were estimated using the method specified by Kenward and Roger (1997).

Studentized residuals and quantile-quantile (Q-Q) plots based on the mixed model were visually checked for model assumptions, including normality and equality of variances. Overall, no serious departure from model assumptions was observed for the majority of the parameters. A small proportion of the parameters showed a moderate departure from normality, in which situation the use of the mixed model was still valid based on the robustness of ANOVA to moderate departures from normality of the error distribution (Schabenberger and Pierce, 2002; Jacqmin-Gadda et al., 2007). Departures from normality were detected on forage parameter “Ash” and grain parameter “Iron.” Log transformation was applied on these two parameters, and normality was improved. A sensitivity analysis of before/after log transformation was then conducted, and the results are shown in Table 21. T-test comparison p-values are insignificant both before and after log transformation for Ash and Iron, so no transformation was applied to keep consistency with other parameters.

Based on the mixed model, entry differences (Entry A vs Entry D and Entry A vs Entry E) were estimated and presented with 95% confidence intervals, along with the p-values (t-test) for the entry differences. Statistical significance was evaluated at $\alpha < 0.05$.

Descriptive statistics

Descriptive statistics for each of the nine forage analytes and 64 quantifiable grain analytes were calculated for combined-sites, including mean, standard deviation, minimum, and maximum for Entry A, Entry D, Entry E, and the collective reference varieties. For reference varieties, a range of means was also calculated for comparison to the means for Entries D and E. Additionally, tolerance intervals for each analyte, specified to contain 99% of the population with 95% confidence (NIST, 2012), were calculated based on the reference varieties. Tolerance interval values that were negative were set to zero. The comparison of means was conducted in a stepwise manner to determine the biological relevance. First, the GMB151 mean for an analyte was compared to the mean range of the reference varieties, second to the tolerance intervals, and finally, in situations where the means were outside that of the mean range of the reference variety or tolerance intervals, to the International Life Science Institute Crop Composition Database (ILSI-CCDB).

Table 20. Soybean grain analytes with samples below the limit of quantitation (LOQ)

Parameter	Number of Values ≥ LOQ	Number of Values < LOQ	Excluded from analysis
C8:0 Caprylic Acid	0	192	Yes
C10:0 Capric Acid	0	192	Yes
C12:0 Lauric Acid	0	192	Yes
C14:1 Myristoleic Acid	0	192	Yes
C15:0 Pentadecanoic Acid	0	192	Yes
C15:1 Pentadecenoic Acid	0	192	Yes
C17:1 Heptadecenoic Acid	191	1	No
C18:4 Stearidonic Acid	2	190	Yes
C19:0 Nonadecanoic Acid	18	174	Yes
C20:2 Eicosadienoic Acid	106	86	Yes
C20:3 Eicosatrienoic Acid	0	192	Yes
C20:4 Arachidonic Acid	0	192	Yes
C20:5 Eicosapentaenoic Acid	66	126	Yes
C22:1 Erucic Acid	0	192	Yes
C22:5 n-3 Docosapentaenoic Acid	0	192	Yes
C22:5 n-6 Docosapentaenoic Acid	3	189	Yes
C22:6 Docosahexaenoic Acid	15	177	Yes
C24:0 Lignoceric Acid	155	37	No
Sodium	112	80	Yes
Vitamin B1 (Thiamine)	149	43	No

Table 21. Sensitivity analysis for normality transformations of parameters "Forage-Ash" and "Grain-Iron"

	Parameter	Transformation	Entry or Comparison	t-test p-value (Comparison)
Forage	Ash	Original Scale	A vs D	0.285
			A vs E	0.826
		Log Transformation	A vs D	0.106
			A vs E	0.745
Grain	Iron	Original Scale	A vs D	0.959
			A vs E	0.786
		Log Transformation	A vs D	0.888
			A vs E	0.542

7.4. Results and discussion

Composition forage

No significant differences were observed between the non-GM counterpart (Entry A) and GMB151 soybean not treated or treated with trait-specific herbicide (Entry D and Entry E, respectively) for any of the proximates, fiber, or select minerals (Table 22 and Table 23).

Composition grain – Proximates and fiber

A statistically significant difference ($p < 0.05$) was observed between the non-GM counterpart (Entry A) and GMB151 soybean treated with trait-specific herbicide (Entry E) for neutral detergent fiber. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GMB151 soybean not treated or treated with trait-specific herbicide (Entry D and Entry E, respectively) for moisture, carbohydrates, and crude protein (Table 24).

However, the GMB151 soybean mean values for neutral detergent fiber, moisture, carbohydrates, and crude protein were within the mean range of the reference varieties and the tolerance intervals (Table 24). Therefore, the statistically significant differences are not considered biologically relevant.

Composition grain – Amino acids

No significant differences were observed between the non-GM counterpart (Entry A) and GMB151 soybean not treated or treated with trait-specific herbicide (Entry D and Entry E, respectively) for any of the amino acids (Table 25).

Composition grain – Fatty acids

A statistically significant difference ($p < 0.05$) was observed between the non-GM counterpart (Entry A) and GMB151 treated with trait-specific herbicide (Entry E) for C24:0 lignoceric acid (Table 26).

Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GMB151 soybean not treated or treated with trait-specific herbicide (Entry D and Entry E, respectively) for C16:0 palmitic acid, C17:1 heptadecenoic acid, C18:1 oleic acid, C18:2 linoleic acid, C20:1 eicosenoic acid, and C22:0 behenic acid (Table 26 and Table 27).

However, the GMB151 soybean mean values for C18:1 oleic acid, C18:2 linoleic acid, C20:1 eicosenoic acid, and C24:0 lignoceric acid were within the mean range of the reference varieties and the tolerance intervals. The GMB151 soybean mean values for C16:0 palmitic acid, C17:1 heptadecenoic acid, and C22:0 behenic acid were slightly outside the mean range of the reference varieties but within the range of the tolerance intervals and the ILSI-CCDB ranges (ILSI, 2016). Therefore, the statistically significant differences are not considered biologically relevant.

Composition grain – Minerals

Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GMB151 soybean treated with trait-specific herbicide (Entry E) for calcium, copper, and zinc (Table 28).

However, the GMB151 soybean mean values for calcium, copper, and zinc were within the mean range of the reference varieties and tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Composition grain – Vitamins

Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GMB151 soybean treated with trait-specific herbicide (Entry E) for vitamin B1 (thiamine), vitamin B5 (pantothenic acid), and vitamin B9 (folic acid) (Table 29).

However, the GMB151 soybean mean values for vitamin B1 (thiamine), vitamin B5 (pantothenic acid), and vitamin B9 (folic acid) were within the mean range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Composition grain – Anti-nutrients

A statistically significant difference ($p < 0.05$) was observed between the non-GM counterpart (Entry A) and GMB151 soybean treated with trait-specific herbicide (Entry E) for phytic acid (Table 30).

Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GMB151 soybean not treated or treated with trait-specific herbicide (Entry D and Entry E, respectively) for total daidzein, total genistein, total glycitein, and total isoflavones (Table 30).

However, the GMB151 soybean mean values for phytic acid, total daidzein, total genistein, total glycitein, and total isoflavones were within the mean range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Table 22. Comparison of proximates and fiber in forage of GMB151 soybean with its non-GM counterpart

	Non-GM counterpart (Entry A)	GMB151 not treated (Entry D)	GMB151 treated (Entry E)	Non-GM reference varieties range (Entries F–N) [a]	Non-GM reference varieties mean range (Entries F–N) [a]	Tolerance interval non-GM reference varieties (Entries F–N) [b]	Comparison t-test A vs D [c]	Comparison t-test A vs E [c]
Parameter	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**(Min–Max)**	**(Min–Max)**	**(Lower–Upper)**	**p-value**	**p-value**
Moisture (% FW)	79.2 ± 3.95	79.2 ± 3.41	79.1 ± 4.04	70.9–85.6	75.8–79.2	65.9–88.7	0.960	0.906
Ash (% DW)	8.02 ± 2.91	8.60 ± 3.48	8.14 ± 2.82	5.15–29.10	6.65–11.52	0–21.43	0.285	0.826
Carbohydrates (% DW)	68.7 ± 3.96	66.3 ± 4.94	67.1 ± 4.64	48.2–76.3	65.0–69.6	51.5–82.7	0.053	0.182
Crude Fat (% DW)	3.78 ± 1.58	4.18 ± 1.82	4.31 ± 1.54	0.26–11.40	3.99–5.32	0–11.37	0.410	0.279
Crude Protein (% DW)	19.6 ± 2.50	20.9 ± 3.71	20.5 ± 2.86	15.2–30.6	17.7–21.7	11.5–28.3	0.099	0.246
Acid Detergent Fiber (% DW)	32.4 ± 5.19	33.7 ± 7.27	30.3 ± 5.59	20.7–58.1	29.4–38.2	11.9–54.4	0.428	0.170
Neutral Detergent Fiber (% DW)	37.4 ± 6.05	37.1 ± 8.56	34.0 ± 5.21	25.2–58.0	33.3–43.1	17.0–55.8	0.889	0.093

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
FW = fresh weight; DW = dry weight
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 23. Comparison of select minerals in forage of GMB151 soybean with its non-GM counterpart (mg/kg DW)

	Non-GM counterpart (Entry A)	GMB151 not treated (Entry D)	GMB151 treated (Entry E)	Non-GM reference varieties range (Entries F–N) [a]	Non-GM reference varieties mean range (Entries F–N) [a]	Tolerance interval non-GM reference varieties (Entries F–N) [b]	Comparison t-test A vs D [c]	Comparison t-test A vs E [c]
Parameter	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**(Min–Max)**	**(Min–Max)**	**(Lower–Upper)**	**p-value**	**p-value**
Calcium	1054 ± 297	1115 ± 277	1068 ± 261	522–1843	835–1268	201.3–1870.9	0.255	0.799
Phosphorus	319 ± 77.1	336 ± 75.3	309 ± 73.2	164–463	265–348	122.0–479.7	0.416	0.606

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
DW = dry weight
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 24. Comparison of proximates and fiber in grain of GMB151 soybean with its non-GM counterpart

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
Moisture (% FW)	12.2 ± 2.70	11.5 ± 2.33	11.3 ± 2.30	8.9–19.4	9.6–14.8	4.0–19.7	<0.001	<0.001
Ash (% DW)	4.93 ± 0.23	4.91 ± 0.19	4.94 ± 0.24	4.09–5.63	4.68–5.17	4.27–5.60	0.705	0.814
Carbohydrates (% DW)	35.2 ± 0.85	34.4 ± 0.97	34.5 ± 1.14	32.3–39.1	33.4–37.7	32.1–40.0	0.005	0.013
Crude Fat (% DW)	20.2 ± 0.77	20.3 ± 0.99	20.1 ± 0.92	17.5–23.8	18.9–22.1	15.8–24.2	0.624	0.721
Crude Protein (% DW)	39.7 ± 1.07	40.4 ± 1.37	40.4 ± 1.37	33.6–44.1	36.2–42.9	32.6–45.5	0.013	0.009
Acid Detergent Fiber (% DW)	16.2 ± 2.10	15.5 ± 1.63	15.7 ± 1.41	10.6–21.2	13.9–16.6	9.8–20.9	0.168	0.271
Neutral Detergent Fiber (% DW)	16.3 ± 1.52	16.1 ± 0.97	15.7 ± 1.00	13.1–19.1	14.6–17.0	11.7–20.1	0.439	0.019
Total Dietary Fiber (% DW)	17.5 ± 1.64	17.4 ± 1.68	17.4 ± 2.07	12.1–21.9	16.3–19.0	12.2–23.6	0.846	0.766

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
FW = fresh weight; DW = dry weight
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 25. Comparison of amino acids in grain of GMB151 soybean with its non-GM counterpart (% DW)

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
Alanine	1.74 ± 0.066	1.75 ± 0.071	1.76 ± 0.083	1.48–2.00	1.60–1.85	1.43–1.98	0.379	0.230
Arginine	2.72 ± 0.182	2.77 ± 0.217	2.78 ± 0.245	2.05–3.48	2.40–3.12	1.90–3.50	0.167	0.111
Aspartic Acid	4.75 ± 0.149	4.84 ± 0.283	4.82 ± 0.286	3.93–5.66	4.36–5.31	3.68–5.80	0.148	0.243
Cystine	0.514 ± 0.053	0.529 ± 0.051	0.527 ± 0.055	0.314–0.645	0.448–0.601	0.285–0.692	0.278	0.368
Glutamic Acid	7.49 ± 0.25	7.68 ± 0.43	7.64 ± 0.46	6.15–8.81	6.83–8.31	5.85–9.21	0.054	0.121
Glycine	1.76 ± 0.09	1.79 ± 0.10	1.79 ± 0.11	1.41–1.96	1.60–1.88	1.40–2.06	0.209	0.137
Histidine	1.08 ± 0.067	1.09 ± 0.076	1.08 ± 0.091	0.845–1.25	0.98–1.18	0.815–1.32	0.553	0.741
Isoleucine	1.88 ± 0.081	1.90 ± 0.088	1.90 ± 0.100	1.52–2.13	1.72–2.02	1.50–2.19	0.207	0.213
Leucine	3.08 ± 0.12	3.12 ± 0.14	3.11 ± 0.16	2.53–3.51	2.81–3.33	2.46–3.59	0.195	0.235
Lysine	2.89 ± 0.24	2.92 ± 0.23	2.92 ± 0.19	2.33–3.41	2.67–3.10	2.11–3.56	0.632	0.635
Methionine	0.513 ± 0.039	0.516 ± 0.039	0.528 ± 0.035	0.376–0.583	0.466–0.524	0.361–0.625	0.706	0.105
Phenylalanine	2.06 ± 0.15	2.09 ± 0.16	2.09 ± 0.20	1.58–2.56	1.85–2.29	1.50–2.61	0.386	0.517
Proline	2.07 ± 0.088	2.11 ± 0.100	2.11 ± 0.118	1.66–2.46	1.87–2.30	1.61–2.46	0.119	0.109
Serine	2.11 ± 0.11	2.13 ± 0.12	2.13 ± 0.13	1.71–2.35	1.91–2.25	1.67–2.46	0.330	0.350
Threonine	1.63 ± 0.071	1.64 ± 0.085	1.64 ± 0.095	1.33–1.79	1.48–1.71	1.30–1.88	0.347	0.487
Tryptophan	0.542 ± 0.025	0.527 ± 0.041	0.531 ± 0.022	0.370–0.603	0.470–0.559	0.398–0.624	0.083	0.199
Tyrosine	1.24 ± 0.10	1.26 ± 0.10	1.25 ± 0.12	0.92–1.46	1.12–1.34	0.88–1.56	0.310	0.613
Valine	1.90 ± 0.076	1.92 ± 0.084	1.93 ± 0.104	1.55–2.21	1.75–2.08	1.52–2.23	0.342	0.208

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
DW = dry weight
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 26. Comparison of fatty acids in grain of GMB151 soybean with its non-GM counterpart (% total fatty acids)

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
C14:0 Myristic Acid	0.0701 ± 0.0072	0.0690 ± 0.0072	0.0702 ± 0.0084	0.0574–0.0873	0.0643–0.0779	0.0520–0.0895	0.578	0.934
C16:0 Palmitic Acid	11.0 ± 0.24	10.6 ± 0.27	10.7 ± 0.27	10.7–12.4	11.0–12.2	10.2–12.8	<0.001	0.002
C16:1 Palmitoleic Acid	0.0833 ± 0.0054	0.0812 ± 0.0044	0.0847 ± 0.0047	0.0648–0.1010	0.0728–0.0936	0.0593–0.1075	0.128	0.300
C17:0 Heptadecanoic Acid	0.0913 ± 0.0053	0.0904 ± 0.0070	0.0902 ± 0.0062	0.0725–0.1080	0.0781–0.1011	0.0643–0.1160	0.498	0.368
C17:1 Heptadecenoic Acid	0.0618 ± 0.0038	0.0653 ± 0.0041	0.0652 ± 0.0050	0.0257–0.0673	0.0515–0.0631	0.0433–0.0724	0.005	0.006
C18:0 Stearic Acid	4.38 ± 0.32	4.34 ± 0.32	4.38 ± 0.34	3.79–5.24	4.04–4.79	3.41–5.36	0.540	0.996
C18:1 Oleic Acid	21.8 ± 1.37	23.6 ± 2.39	23.4 ± 2.05	17.2–28.0	20.0–23.8	15.3–28.8	0.003	0.006
C18:2 Linoleic Acid	53.7 ± 1.18	52.6 ± 1.80	52.6 ± 1.51	49.4–56.9	50.8–54.8	47.2–58.3	0.007	0.009
C18:3 Linolenic Acid	7.81 ± 0.43	7.49 ± 0.52	7.52 ± 0.51	6.26–10.40	6.95–8.77	5.00–10.98	0.064	0.096
C18:4 Stearidonic Acid	<LOQ	<LOQ–0.113	<LOQ	<LOQ–0.0859	<LOQ–0.0107	NA	NA	NA
C19:0 Nonadecanoic Acid	<LOQ–0.0943	<LOQ–0.449	<LOQ–0.0653	<LOQ–0.307	<LOQ–0.0486	NA	NA	NA
C20:0 Arachidic Acid	0.315 ± 0.018	0.320 ± 0.015	0.324 ± 0.019	0.277–0.393	0.293–0.365	0.250–0.403	0.289	0.062

Table 26. Comparison of fatty acids in grain of GMB151 soybean with its non-GM counterpart (% total fatty acids) (continued)

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
C20:1 Eicosenoic Acid	0.182 ± 0.019	0.189 ± 0.022	0.191 ± 0.019	0.135–0.256	0.165–0.216	0.115–0.260	0.010	0.003
C20:2 Eicosadienoic Acid	<LOQ–0.0546	<LOQ–0.0516	<LOQ–0.0579	<LOQ–0.0579	0.0161–0.0366	NA	NA	NA
C20:5 Eicosapentaenoic Acid	<LOQ–0.104	<LOQ–0.0534	<LOQ–0.0657	<LOQ–0.0856	0.0055–0.0306	NA	NA	NA
C22:0 Behenic Acid	0.331 ± 0.010	0.378 ± 0.013	0.382 ± 0.012	0.320–0.390	0.328–0.377	0.294–0.401	<0.001	<0.001
C22:5 n-6 Docosapentaenoic Acid	<LOQ–0.0515	<LOQ	<LOQ	<LOQ–0.081	<LOQ–0.0114	NA	NA	NA
C22:6 Docosahexaenoic Acid	<LOQ–0.0567	<LOQ–0.0415	<LOQ–0.109	<LOQ–0.0654	<LOQ–0.0123	NA	NA	NA
C24:0 Lignoceric Acid	0.111 ± 0.021	0.101 ± 0.024	0.098 ± 0.032	0.038–0.180	0.0076–0.1370	0–0.200	0.080	0.024

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
NA = Not Applicable because more than 1/3 of the values are < LOQ. Minimum and maximum are reported instead of mean and standard deviation.
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 27. Comparison of fatty acids in grain of GMB151 soybean with its non-GM counterpart, including ILSI ranges (% total fatty acids)

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	ILSI Crop Composition Database (Version 6.0) (Min–Max)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
C16:0 Palmitic Acid	11.0 ± 0.24	10.6 ± 0.27	10.7 ± 0.27	10.7–12.4	11.0–12.2	10.2–12.8	8.03–15.99	<0.001	0.002
C17:1 Heptadecenoic Acid	0.0618 ± 0.0038	0.0653 ± 0.0041	0.0652 ± 0.0050	0.0257–0.0673	0.0515–0.0631	0.0433–0.0724	<LOQ–0.088	0.005	0.006
C22:0 Behenic Acid	0.331 ± 0.010	0.378 ± 0.013	0.382 ± 0.012	0.320–0.390	0.328–0.377	0.294–0.401	0.181–0.723	<0.001	<0.001

Composition samples were derived from eight field trials conducted in the U.S. in 2017.

[a] Range of results from nine reference varieties (Entries F–N).

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

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

Table 28. Comparison of minerals in grain of GMB151 soybean with its non-GM counterpart (mg/kg DW)

	Non-GM counterpart (Entry A)	GMB151 not treated (Entry D)	GMB151 treated (Entry E)	Non-GM reference varieties range (Entries F–N) [a]	Non-GM reference varieties mean range (Entries F–N) [a]	Tolerance interval non-GM reference varieties (Entries F–N) [b]	Comparison t-test A vs D [c]	Comparison t-test A vs E [c]
Parameter	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**(Min–Max)**	**(Min–Max)**	**(Lower–Upper)**	**p-value**	**p-value**
Calcium	2547 ± 362	2435 ± 345	2343 ± 334	1480–3689	1952–3186	1004–3987	0.172	0.017
Copper	14.8 ± 4.50	14.8 ± 3.44	13.3 ± 2.61	6.8–23.2	10.4–13.9	3.2–20.7	0.939	0.040
Iron	118 ± 29.2	118 ± 41.1	115 ± 45.3	59–364	83.9–134.7	0–233	0.959	0.786
Magnesium	2580 ± 201	2497 ± 204	2498 ± 268	1954–3177	2348–2814	1758–3278	0.163	0.164
Manganese	39.2 ± 9.72	40.4 ± 9.79	38.2 ± 10.46	23.8–74.7	28.6–47.5	3.7–68.1	0.389	0.520
Phosphorus	6135 ± 762	6170 ± 673	6090 ± 730	4789–7773	5608–6716	4088–8109	0.788	0.723
Potassium	19545 ± 1214	19937 ± 1601	19862 ± 1425	16973–22831	18666–20630	15691–24021	0.292	0.392
Sodium	<LOQ–45.7	<LOQ–51.4	<LOQ–51.2	<LOQ–55.5	6.55–17.73	NA	NA	NA
Zinc	58.4 ± 11.47	57.7 ± 13.47	54.2 ± 9.90	33.0–85.6	38.8–60.6	18.0–80.6	0.692	0.024

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
NA = Not Applicable because more than 1/3 of the values are < LOQ. Minimum and maximum are reported instead of mean and standard deviation.
DW = dry weight
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 29. Comparison of vitamins in grain of GMB151 soybean with its non-GM counterpart (mg/kg DW)

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
Alpha Tocopherol (Vitamin E)	27.6 ± 9.59	27.4 ± 9.67	28.1 ± 9.36	12.3–53.2	16.6–42.2	0–57.2	0.835	0.717
Vitamin B1 (Thiamine)	1.80 ± 0.72	1.67 ± 0.75	1.46 ± 0.74	0.41–2.54	0.08–1.72	0–2.93	0.128	<0.001
Vitamin B2 (Riboflavin)	4.44 ± 0.74	4.51 ± 0.70	4.33 ± 0.88	2.26–5.92	3.84–4.49	2.32–6.03	0.675	0.507
Vitamin B3 (Niacin)	26.7 ± 6.57	29.2 ± 6.34	29.3 ± 6.58	13.7–46.9	19.5–30.8	5.7–47.0	0.084	0.075
Vitamin B5 (Pantothenic Acid)	8.64 ± 1.76	9.27 ± 2.23	9.50 ± 2.22	5.42–14.60	8.5–12.0	4.01–16.05	0.098	0.026
Vitamin B6 (Pyridoxine)	4.59 ± 1.18	4.63 ± 1.13	4.81 ± 1.20	2.12–7.69	3.70–5.25	0.85–8.43	0.878	0.461
Vitamin B9 (Folic Acid)	5.50 ± 1.64	6.43 ± 2.02	6.77 ± 2.50	2.19–13.30	4.65–8.62	0–12.57	0.084	0.021
Vitamin K1	0.703 ± 0.330	0.700 ± 0.334	0.707 ± 0.335	0.344–1.440	0.50–1.17	0–1.592	0.937	0.895

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
DW = dry weight
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 30. Comparison of anti-nutrients in grain of GMB151 soybean with its non-GM counterpart

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [a] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [b] (Lower–Upper)	Comparison t-test A vs D [c] p-value	Comparison t-test A vs E [c] p-value
Isoflavones (mg/kg DW):								
Total Daidzein	850 ± 217	712 ± 199	709 ± 203	380–1696	491–1333	0–1788	<0.001	<0.001
Total Genistein	808 ± 211	690 ± 192	703 ± 211	428–1498	546–1140	236–1582	<0.001	0.002
Total Glycitein	160 ± 29.0	136 ± 32.2	134 ± 18.2	102–350	121–250	24–320	0.004	0.002
Total Isoflavones	1818 ± 436	1538 ± 396	1546 ± 421	947–3455	1159–2662	364–3506	<0.001	<0.001
Lectins (mg/g DW)	2.54 ± 0.71	2.44 ± 0.67	2.43 ± 0.76	0.78–4.64	1.66–2.73	0.23–4.41	0.587	0.579
Phytic Acid (% DW)	1.45 ± 0.25	1.53 ± 0.23	1.60 ± 0.25	0.88–2.11	1.32–1.74	0.77–2.26	0.079	0.001
Raffinose (% DW)	1.02 ± 0.23	1.03 ± 0.26	0.99 ± 0.21	0.40–1.38	0.73–1.04	0.39–1.40	0.798	0.583
Stachyose (% DW)	3.18 ± 0.18	3.18 ± 0.18	3.19 ± 0.25	2.59–4.50	3.01–3.85	2.40–4.53	0.970	0.890
Trypsin Inhibitor (TIU/mg DW)	35.8 ± 5.66	35.9 ± 4.74	36.3 ± 5.88	18.1–41.8	26.6–33.8	15.1–44.4	0.949	0.657

Composition samples were derived from eight field trials conducted in the U.S. in 2017.
DW = dry weight; TIU = Trypsin Inhibitor Unit
[a] Range of results from nine reference varieties (Entries F–N).
[b] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[c] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

8. AGRONOMIC AND PHENOTYPIC EVALUATION

8.1. History of field activities

GMB151 soybean has been field tested in adapted growing regions of the U.S., Argentina, and Brazil. The field trials and associated notification numbers in the U.S. are summarized in Appendix 1. Field trials were conducted for the development of regulatory data, product development and characterization, and production of seed for the latter activities.

8.2. Field study of agronomic and phenotypic characteristics

To evaluate the agronomic performance of GMB151 soybean relative to its non-GM counterpart under field conditions representative of commercial cultivation, a multi-site field evaluation was undertaken during the 2017 growing season. The agronomic assessment included 11 locations representative of diverse soybean growing regions of the U.S. (Table 31). The goal of this study was a comparative assessment of GMB151 relative to its conventional counterpart and reference varieties. This study did not seek to characterize the agronomic performance of GMB151 under all possible conditions or agronomic systems.

The selected field trial sites (Figure 29) and management systems used for the agronomic and phenotypic assessment of GMB151 soybean represented geographically diverse regions, cropping practices (e.g., irrigation, pest management, etc.), and diverse soil types in areas of soybean production. All trials were managed to replicate typical practices for grain production in the region of the trial. Therefore, the sites captured enough variability within the set of possible receiving environments where the GMB151 soybean can be grown to enable comparative assessment of potential differences between GMB151 and the conventional counterpart. Details regarding irrigation (Table 32), soil characteristics (Table 33), tillage (Table 34), and pest management (Table 35) for each trial site are provided.

Entries relevant for the agronomic assessment of GMB151 soybean included at each field trial site are presented in Table 36. Entries were replicated four times in an RCBD. Nine total non-GM reference varieties were used, and each field trial contained three reference varieties. The non-GM reference varieties were provided by Schillinger Genetics Inc. and NuPride Genetics Network, LLC.

All plots within a field site were subjected to the same growing conditions and management (i.e. cultivation, irrigation, fertilizer, and maintenance pesticide treatments). Each plot within each field trial was identically sized and separated by fallow area to achieve adequate separation of each material.

The GMB151 soybean plots treated with the trait-specific herbicide received one application of IFT (Balance Pro) at a target rate of 70.1 g a.i./ha at the growth stage BBCH 00–03 (pre-emergence).

Table 31. Field trial locations that completed field production

Site code	Nearest town or city	Site county	State
01	Richland	Jefferson	Iowa
02	York	York	Nebraska
03	Elk Horn	Shelby	Iowa
04	Stewardson	Shelby	Illinois
05	Germansville	Lehigh	Pennsylvania
06	Fisk	Butler	Missouri
07	Larned	Pawnee	Kansas
09	Manilla	Rush	Indiana
10	Stilwell	Johnson	Kansas
11	Carlyle	Clinton	Illinois
12	Bangor	Marshall	Iowa

Site 08 was removed from the study due to flood damage.

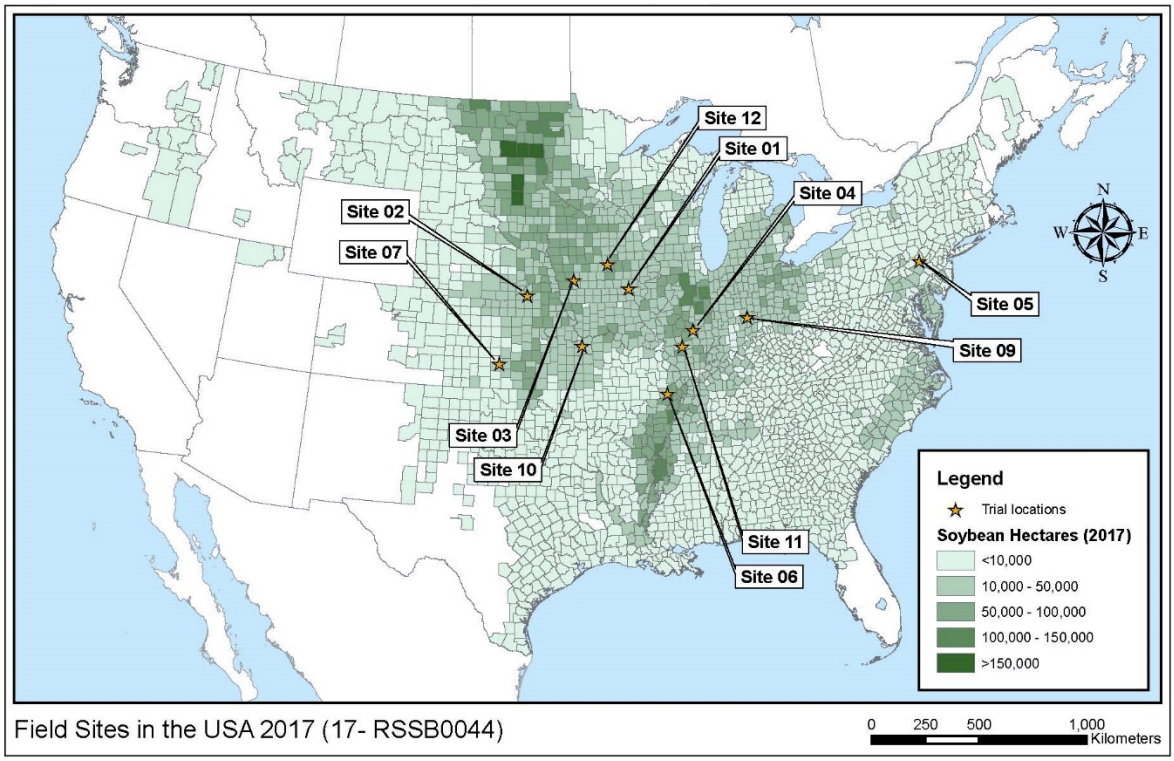


Figure 29. Trial site locations and soybean growing regions in the U.S.

Table 32. USDA agricultural region, soil type, cropping practices, seasonal temperatures, and rainfall

Site	USDA ag. region	Soil	Planting date 2017	Irrigation	Rotation	Growing season average temp (°C)	Growing season rainfall (mm)
01	Upper Midwest	Clay Loam	25-May	Overhead	Corn/Soybean/ Corn	12.9–25.8	584
02	Northern Plains	Silty Clay Loam	30-May	Overhead	Sorghum/Soybean/ Soybean	13.3–26.2	679
03	Upper Midwest	Silt Loam	5-Jun	Overhead	Corn/Soybean/ Corn	14.5–25.3	571
04	Heartland	Silt Loam	31-May	None	Soybean/Corn/ Corn	15.5–28.7	308
05	Northeastern	Loam	13-Jun	None	Teffgrass and Annual Rye/Soybean/Canola	14.7–25.2	849
06	Heartland	Loam	25-May	Overhead	Rice/Soybean/ Corn	18.4–30.6	351
07	Northern Plains	Sandy Loam	2-Jun	Drip	Wheat/Soybean/ Sorghum	15.2–29.5	226
09	Central High Plains	Loam	7-Jun	None	Corn/Soybean/ Soybean and Wheat	14.9–29.1	697
10	Southern High Plains	Silt Loam	8-Jun	Overhead	Buckwheat and Phacelia/Ornamental Flowers, Cabbage, Corn, Leaf Lettuce, Soybean, Sunflower, and Watermelon/ Buckwheat	14.7–28.4	797
11	Heartland	Silt Loam	8-Jun	None	Soybean/Wheat/ Corn	15.8–28.5	303
12	Corn Belt	Silty Clay Loam	31-May	None	Soybean/Corn/ Soybean	13.9–26.1	407

Rotation column indicates the crops produced in the prior three seasons (2014/2015/2016).

Table 33. Soil characterization summary

Site code	Sample soil type	Average % sand	Average % silt	Average % clay	Average soil pH	Average cation exchange capacity (CEC; meq/100g)	Average % organic matter
01	Clay Loam	22	48	31	6.1	17.4	3.0
02	Silty Clay Loam	19	50	32	6.3	17.4	2.5
03	Silt Loam	20	60	21	6.2	18.0	3.4
04	Silt Loam	30	58	13	6.5	8.3	1.8
05	Loam	38	34	29	6.0	8.4	2.5
06	Loam	51	30	20	5.1	10.4	1.3
07	Sandy Loam	61	22	18	7.0	12.6	1.9
09	Loam	43	37	21	6.4	11.1	2.3
10	Silt Loam	19	61	21	6.2	14.0	2.4
11	Silt Loam	24	60	16	5.7	6.5	2.1
12	Silty Clay Loam	20	48	32	6.7	19.4	4.2

Table 34. Tillage practices and dates performed

Site	Date performed	Average crop growth stage (BBCH)	Type	Depth (cm)
01	16-May-2017	Pre-plant	Seedbed preparation	10.00
	07-Jun-2017	10	Hoeing	2.00
	14-Jun-2017	13	Hoeing	1.00
02	12-Apr-2017	Pre-plant	Disc	15.00
	30-May-2017	Pre-plant	Rototill	12.50
	27-Jun-2017	13	Hoeing	2.50
03	27-Apr-2017	Pre-plant	Disc	20.00
	07-May-2017	Pre-plant	Seedbed preparation	15.00
	07-Jul-2017	13	Hoeing	0.00
04	30-Apr-2017	00	Disc	4.00
	30-May-2017	00	Seedbed preparation	3.00
	11-Jul-2017	17	Hoeing	1.00
	30-Sep-2017	99	Crop destruction	4.00
05	02-May-2017	Pre-plant	Chisel	25.40
	02-May-2017	Pre-plant	Disc	17.78
	02-Jun-2017	Pre-plant	Hoeing	12.70
	13-Jun-2017	Pre-plant	Harrow	12.70
	03-Jul-2017	12	Hoeing	0.00
	19-Jul-2017	15	Hoeing	0.00
	14-Aug-2017	65	Hoeing	0.00
06	25-Apr-2017	Pre-plant	Seedbed preparation	10.00
	26-Apr-2017	Pre-plant	Furrow	15.00
	26-Jun-2017	17	Cultivation between rows	5.00
	25-Sep-2017	99	Crop destruct	15.00
07	15-Apr-2017	Pre-plant	Disc	10.00
	29-May-2017	Pre-plant	Disc	10.00
	23-Jun-2017	08	Hoeing	4.00
	27-Jun-2017	09	Harrow	15.00
	21-Jul-2017	61	Hoeing	4.00
	01-Aug-2017	69	Hoeing	4.00
	11-Aug-2017	76	Harrow	15.00
	11-Aug-2017	76	Hoeing	4.00
	05-Sep-2017	86	Harrow	15.00
	15-Sep-2017	89	Hoeing	4.00
09	24-May-2017	Pre-plant	Rototill	10.16

Table 34. Tillage practices and dates performed (continued)

Site	Date performed	Average crop growth stage (BBCH)	Type	Depth (cm)
	25-May-2017	Pre-plant	Rototill	10.16
	06-Jun-2017	Pre-plant	Rototill	10.16
	10-Jul-2017	14	Hoeing	5.08
	10-Jul-2017	14	Rototill	10.16
	25-Jul-2017	18	Mechanical	0.00
	31-Jul-2017	65	Rototill	10.16
	26-Nov-2017	99	Crop destruct	10.16
	07-Mar-2017	Pre-plant	Inversion	20.00
	13-Apr-2017	Pre-plant	Disc	10.00
	24-Apr-2017	Pre-plant	Disc	10.00
	16-May-2017	Pre-plant	Disc	10.00
10	06-Jun-2017	Pre-plant	Disc	10.00
	08-Jun-2017	Pre-plant	Disc	10.00
	28-Jun-2017	15	Hoeing	NA
	24-Jul-2017	62	Hoeing	NA
	31-Oct-2017	99	Crop destruct	10.00
11	19-May-2017	Pre-plant	Seedbed preparation	18.00
	28-Apr-2017	Pre-plant	Disc	8.00
	15-May-2017	Pre-plant	Rototill	8.00
12	14-Jun-2017	03	Mechanical	2.00
	01-Jul-2017	17	Rototill	4.00
	12-Jul-2017	19	Cultivation between rows	6.00

Table 35. Maintenance pesticides and target pests

Application date (2017)	Average crop growth stage (BBCH)	Trade name	Formulation type	Active ingredient	Type	Target(s)	Application method	Volume (L/ha)	Rate (g a.i/.ha)
Site 01									
25-May	00	Dimetric	DF	Metribuzin	Herbicide	Weeds	Broadcast on soil	149.3	420.52
25-May	00	Framework	EC	Pendimethalin				149.3	1269.18
25-May	00	Roundup WeatherMax	AS	Glyphosate				149.3	1541.96
25-May	00	Sonic	DF	Cloransulam-methyl				149.3	27.71
25-May	00	Sonic	DF	Sulfentrazone				149.3	217.81
Site 02									
31-May	00	Blanket	L	Sulfentrazone	Herbicide	Broadleaf weeds	Broadcast on soil	186.6	39.48
31-May	00	Boundary	L	Metribuzin		Grasses and broadleaf weeds		186.6	350.52
31-May	00	Boundary	L	S-metolachlor				186.6	1469.85
Site 03									
08-Jun	00	Anthem	L	Fluthiacet-methyl	Herbicide	Annual weeds	Broadcast on soil	153.9	4.09
08-Jun	00	Anthem	L	Pyroxasulfone				153.9	146.24
08-Jun	00	Prowl H2O	AS	Pendimethalin				153.9	1064.63
Site 04									
02-Jun	01	Dual II Magnum	EC	S-metolachlor	Herbicide	Grass weeds	Broadcast on soil	93.3	1603.63
02-Jun	01	Zidua	WDG	Pyroxasulfone		Broadleaf weeds		93.3	178.88
03-Jul	15	Fusilade DX	EC	Fluazifop-p-butyl		Grass weeds	Foliar	112.0	105.29
Site 05									
15-Jun	00	Canopy 75 DF	DF	Chlorimuron ethyl	Herbicide	Broadleaf weeds and grass weeds	Broadcast on soil	279.9	45.04
15-Jun	00	Canopy 75 DF	DF	Metribuzin				279.9	270.63
15-Jun	00	Dual II Magnum	L	S-metolachlor				279.9	1389.81

Table 35. Maintenance pesticides and target pests (continued)

Application date (2017)	Average crop growth stage (BBCH)	Trade name	Formulation type	Active ingredient	Type	Target(s)	Application method	Volume (L/ha)	Rate (g a.i/.ha)
Site 06									
23-May	Pre-plant	Liberty	SL	Glufosinate	Herbicide	Broadleaf and grass weeds	Broadcast on soil	186.6	655.16
27-May	01	Zidua	WDG	Pyroxasulfone				186.6	119.25
14-Jun	15	Prefix	EC	Fomesafen			Foliar	186.6	266.74
14-Jun	15	Prefix	EC	Metolachlor				186.6	1216.72
28-Jun	18	Warrant	EC	Acetochlor				186.6	1260.01
21-Jul	67	Section	EC	Clethodim		Grass weeds		186.6	140.39
Site 07									
03-Jun	01	Authority Elite	L	S-metolachlor	Herbicide	Winter/summer annual weeds	Broadcast on soil	140.0	358.29
03-Jun	01	Authority Elite	SC	Sulfentrazone				140.0	36.85
24-Jun	09	Acifin	SC	Acifluorfen		Palmer amaranth		186.6	80.62
23-Aug	82	Hero	XX	Bifenthrin	Insecticide	Fall armyworm, insects	Foliar	28.0	1.83
23-Aug	82	Hero	XX	Zeta cypermethrin				28.0	0.37
Site 09									
27-Jun	12	Arrow	L	Clethodim	Herbicide	Weeds	Foliar	140.0	140.39
27-Jun	12	Classic	WDG	Chlorimuron ethyl				140.0	8.77

Table 35. Maintenance pesticides and target pests (continued)

Application date (2017)	Average crop growth stage (BBCH)	Trade name	Formulation type	Active ingredient	Type	Target(s)	Application method	Volume (L/ha)	Rate (g a.i/.ha)
Site 10									
08-Jun	Pre-plant	Authority MTZ	DF	Metribuzin (27%)	Herbicide	Broadleaf weeds (whole plot area)	Foliar	205.3	227.28
08-Jun	Pre-plant	Authority MTZ	DF	Sulfentrazone (18%)				205.3	151.52
08-Jun	Pre-plant	Pavillion	EC	Pendimethalin		Broadleaf weeds and grass weeds (whole plot area)		205.3	1107.64
08-Jun	Pre-plant	Dual II Magnum	L	S-metolachlor				205.3	1421.89
06-Jul	NA	RoundUp PowerMax	L	Glyphosate		Broadleaf weeds and grass weeds (buffers and fallow zone)		186.6	1542.29
05-Sep	NA	RoundUp PowerMax	L	Glyphosate				186.6	1542.29
Site 11									
08-Jun	Pre-plant	Authority MTZ	DF	Metribuzin (27%)	Herbicide	Broadleaf weeds (whole plot area)	Foliar	205.3	227.28
08-Jun	Pre-plant	Authority MTZ	DF	Sulfentrazone (18%)				205.3	151.52
08-Jun	Pre-plant	Pavillion	EC	Pendimethalin		Broadleaf weeds and grass weeds (whole plot area)		205.3	1107.64
08-Jun	Pre-plant	Dual II Magnum	L	S-metolachlor				205.3	1421.89
06-Jul	NA	RoundUp PowerMax	L	Glyphosate		Broadleaf weeds and grass weeds (buffers and fallow zone)		186.6	1542.29
05-Sep	NA	RoundUp PowerMax	L	Glyphosate				186.6	1542.29
Site 12									
16-Jun	11	Prefix	L	Fomesafen	Herbicide	Broadleaf	Foliar	130.6	266.74
16-Jun	11	Prefix	L	Metalochlor		Broadleaf and grasses		130.6	1385.19

Formulation types: AS = aqueous solution, DF = dry flowable, EC = emulsifiable concentrate, L = liquid, SL = soluble (liquid) concentrate, WDG = water dispersible granule, XX = other formulation types

Table 36. Description of entries in field study of agronomic and phenotypic characteristics

Entry ID	Material name	Other specification	Site code
A	Thorne	Non-GM counterpart	All
D	GMB151 soybean	GM (not treated)	All
E	GMB151 soybean	GM (treated)	All
F	E3494	Non-GM reference	06, 07, 11
G	NGN 3121STS	Non-GM reference	06, 07, 11
H	NGN 3347C	Non-GM reference	06, 07, 11
I	E2282	Non-GM reference	01, 03, 05, 12
J	E2692	Non-GM reference	01, 03, 05, 12
K	NGN 3292C	Non-GM reference	01, 03, 05, 12
L	E3066	Non-GM reference	02, 04, 09, 10
M	E2993	Non-GM reference	02, 04, 09, 10
N	E3192	Non-GM reference	02, 04, 09, 10

8.2.1. Agronomic data

Agronomic data were collected during the growing season. Agronomic parameters/observations, crop growth stage, and units are presented in Table 37. In addition to continuous agronomic parameters (e.g., stand count, days to flowering, plant height), evaluations of severity of incidence and plant response to biotic (arthropods and disease) and abiotic stress were made by visual observation. Three stressors from each category were selected by the local collaborator before plot ratings were assigned. Stressors were selected based on biological and economic importance and diversity in their local area.

Data were collected at four different growth phases: 1) early vegetative stage (BBCH 11–15), 2) late vegetative/early reproductive stage (BBCH 17–19), early reproductive/mid-reproductive stage (BBCH 63–71), and late reproductive stage (BBCH 75–80). Data collected in this way allowed the comparative assessment of GMB151 relative to its conventional counterpart and reference varieties. Although there was no need to characterize the agronomic performance of GMB151 under all possible conditions or agronomic systems, any stressors other than the three pre-selected were recorded. If the severity of any of the stressors was likely to have an impact on agronomic performance if no control measures were taken, the entire trial was treated equally. In addition, trait-specific herbicide (IFT) injury at BBCH 11–15 (only on Entry E, GMB151 treated) and symptoms of SCN injury at time points BBCH 17–19 and BBCH 63–71) were rated. These categorical evaluation parameters are described in Table 37 and further elaborated upon in Table 38.

Table 37. Agronomic parameters for soybean

Growth stage (BBCH) [a]	Agronomic parameter/observation	Units
11–15	**Early stand count:** The number of seedlings in one row of each plot	plants/m^2
	IFT injury: Any symptoms of IFT herbicide injury (bleaching, etc.) were observed and recorded.	None, Slight, Moderate, Severe [b]
	Stressor rating: Prevalence and severity of arthropod, disease, and abiotic stressors (three of each) were recorded.	
17–19	**Stressor rating:** Prevalence and severity of arthropod, disease, and abiotic stressors (three of each) were recorded.	
	SCN evaluation: Foliar symptoms of infestation (e.g., patchy stunting, yellowing of leaves, early senescence) were monitored.	
51–59	**Crop development:** Visual estimate of the percent soybean ground cover (row closure) from an over-head perspective within the plot area. Higher percentages indicate more rapid crop development relative to lower percentages.	%
61	**Days to flowering:** The date when 10% of plants were flowering was recorded and the number of days to flowering calculated.	days
69	**Flowering duration:** The date at which 90% of plants had completed flowering was recorded and the duration of flowering calculated.	days
63–71	**Stressor rating:** Prevalence and severity of arthropod, disease, and abiotic stressors (three of each) were recorded.	None, Slight, Moderate, Severe [b]
	SCN evaluation: Foliar symptoms of infestation (e.g., patchy stunting, yellowing of leaves, early senescence) were monitored	
75–80	**Stressor rating:** Prevalence and severity of arthropod, disease, and abiotic stressors (three of each) were recorded.	
71–89	**Plant height:** The height of five representative and competitive plants per plot was measured.	cm
	Lodging: The number of plants lodged by visually estimating to the nearest 10% of the plot	%
	Final stand count: The number of plants in one row of each plot	plants/m^2

Table 37. Agronomic parameters for soybean (continued)

Growth stage (BBCH) [a]	Agronomic parameter/observation	Units
86–89	**Days to maturity:** The date when 90% of plants have physiologically mature pods was used to calculate the days to maturity.	days
	Seed loss: The number of shattered pods on five competitive and representative plants in every plot	count
	Fruit count: The number of pods on five competitive and representative plants per plot	count
89–99	**Seed moisture [c]:** The moisture content of a representative sample immediately after harvest	%
	Seed weight: The weight of 100 seeds (and seed moisture if seed weight was not determined at time of yield measurement)	g
	Yield: The grain yield from rows 4 and 5	kg/ha at 13% moisture

[a] Target growth stage.
[b] Scale interpreted in Table 38.
[c] Seed moisture was used for normalizing yield and not as an agronomic endpoint.

Table 38. Categorical evaluation parameters

Observation/parameter	Rating scale	
	Category	**Description**
Trait-specific herbicide injury [a] Diseases Abiotic stress SCN injury Insect pests	None	No symptoms observed
	Slight	Symptoms not damaging to plant development. Mitigation likely not required
	Moderate	Intermediate between Slight and Severe; likely requires mitigation
	Severe	Symptoms damaging to plant development (e.g., stunting or death); mitigation unlikely to be effective

[a] Trait-specific herbicide injury only applies to evaluations of Entry E (GMB151 treated). The requirement for mitigation for “trait-specific herbicide injury” at moderate does not apply since the herbicide tolerance trait allows plants to overcome moderate damage from IFT application that may occur.

8.2.2. Statistical analysis of agronomic data

Agronomic data from 11 sites were statistically analyzed using a mixed model ANOVA (SAS version 9.4). Observations for the non-GM counterpart Thorne (Entry A) were compared to GMB151 soybean not treated (Entry D) and GMB151 treated with trait-specific herbicide (Entry E).

For characteristics where "0" values accounted for a large proportion of the data (i.e. lodging and seed loss), statistical analysis was not performed and tolerance intervals were not calculated due to sparse data. Only descriptive statistics were provided.

Combined-site analysis

The continuous data were analyzed for each characteristic with a mixed model ANOVA. ANOVA was selected as the statistical method because the field trials contained multiple sites, and within each site, the experimental design structure was an RCBD with multiple entries. A linear mixed model was chosen to appropriately account for Block and Site effects while increasing the precision of Entry comparisons. Block effect is random since it is a source of replication and since its levels are not of interest. The selected sites represent a random sample from many potential sites, and it is desirable to make broad inference to all field sites in the crop growing areas. Therefore, Site is also included in the model as a random effect. Entry effect is fixed because its levels are of specific interest, and if the study were repeated, the same levels would be used again. Interaction between Site and Entry is a random effect since Site is random. This analysis is also referred to as a mixed-model ANOVA and is the most standard model for this type of experimental design.

The additive model for the design is:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \delta_{k(j)} + \varepsilon_{ijk}$$

where Y_{ijk} is the individual value measurement, μ the overall mean, α_i the fixed effect associated with entry, β_j the random effect associated with site, $\alpha\beta_{ij}$ the random effect for the interaction of entry by site, $\delta_{k(j)}$ the random effect associated with block nested within site, and ε_{ijk} the random error. The degrees of freedom were estimated using the method specified by Kenward and Roger (1997).

Studentized residuals and Q-Q plots based on the mixed model were visually checked for model assumptions, including normality and equality of variances. Overall, no serious departure from model assumptions was observed for majority of the parameters. Some deviations from normality were found for agronomic parameters "Flowering Duration (Days)" and "Days to Flowering" due to moderate data discreteness. No transformation can significantly improve data discreteness in this case. Since most of the data have good continuity and because of the robustness of the mixed model (Schabenberger et al., 2002; Jacqmin-Gadda et al., 2007), the linear mixed model is still the most appropriate model based on study design. No transformation was applied to these two parameters for the consistency of statistical analysis and interpretation with other parameters.

Based on the ANOVA model, entry differences (A versus D and A versus E) were estimated and presented with 95% confidence intervals, along with the p-values (t-test) for the entry differences. Statistical significance was evaluated at $p < 0.05$.

Descriptive statistics

For the continuous characteristics, mean values, standard deviations, minimums, and maximums were calculated for combined sites. In addition, tolerance intervals, specified to contain 99% of the population with 95% confidence, were calculated for the reference varieties over all sites (NIST, 2012). The comparison of means was conducted in a stepwise manner to determine the biological relevance. First, the GMB151 mean for a parameter was compared to the mean range of the reference varieties, and second, to the tolerance intervals.

For categorical data of IFT injury, no comparisons can be made among all entries since IFT was only applied on Entry E. For the assessments of other categorical data, the GM crop and conventional control are considered different in susceptibility or tolerance for a particular observation date if the range of injury severity to the GM crop across all four replications did not overlap with the range of injury severity to the conventional control. Any observed differences between the GM crop and conventional control are assessed for biological significance in the context of the range of the commercial reference varieties and for consistency in other observation times and sites.

8.2.3. Results and discussion of agronomic evaluation

The combined-site summary of the statistical analysis of all continuous agronomic parameters is presented in Table 39.

No significant differences were observed between the non-GM counterpart (Entry A) and GMB151 soybean not treated (Entry D) or treated (Entry E) with trait-specific herbicide for early stand count, days to flowering, flowering duration, final stand count, days to maturity, fruit count, seed weight, and yield.

Statistically significant differences were detected for the continuous parameters crop development and plant height between the non-GM counterpart (Entry A) and GMB151 soybean treated (Entry E). However, the mean values of GMB151 for crop development and plant height were within the mean range of the reference varieties. Thus, statistically significant differences were considered not biologically relevant.

Categorical agronomic parameters were analyzed by-site as described in Section 8.2.2. Similar responses to stressor injury were observed for all sites, varying from none to moderate, except for site 09. Site 09 reported severe injury to several plots due to excessive moisture and flooding that resulted in soil crusting during the first two evaluation time points. IFT injury to GMB151 soybean treated with trait-specific herbicide (Entry E) varied from slight to moderate for sites 01, 05, and 11. This level of injury would not be expected to impact agricultural performance as the plants quickly recovered out of the injury and yield was not affected. Aside from IFT injury, which was only expected for Entry E because only Entry E received IFT applications, no other abiotic or biotic stressors were different across replications from that of the conventional comparator. An overview of categorical data comparisons is presented in Table 40. Detailed comparisons for abiotic stressors (Table 41), pests (Table 42), and disease (Table 43) are also provided. Stressors evaluated by cooperators other than those pre-selected as described in Section 8.2.1 were similar across all entries.

Based on the agronomic assessment, GMB151 soybean demonstrated no biologically relevant differences from the non-GM counterpart and showed equivalent agronomic performance to non-GM reference varieties.

Table 39. Summary of continuous agronomic parameters

Parameter	Non-GM counterpart (Entry A) Mean ± SD	GMB151 not treated (Entry D) Mean ± SD	GMB151 treated (Entry E) Mean ± SD	Non-GM reference varieties range (Entries F–N) [a] (Min–Max)	Non-GM reference varieties mean range (Entries F–N) [b] (Min–Max)	Tolerance interval non-GM reference varieties (Entries F–N) [c] (Lower–Upper)	Comparison t-test A vs D [d] p-value	Comparison t-test A vs E [d] p-value
Early stand count (plants/m²)	24.79 ± 6.50	23.77 ± 7.13	23.90 ± 6.94	3.51–33.33	18.4–26.8	1.47–45.54	0.306	0.370
Crop development (%)	56.9 ± 17.6	54.5 ± 19.4	49.4 ± 18.5	20–80	45.0–59.7	0.8–100.0	0.211	<0.001
Days to flowering	46.5 ± 4.8	46.5 ± 4.7	46.5 ± 5.2	35–54	38.6–46.1	28.5–57.7	0.969	0.969
Flowering duration (days)	65.6 ± 8.5	65.3 ± 8.6	64.4 ± 8.2	50–78	58.3–65.4	40.1–84.2	0.741	0.199
Plant height (cm)	80.8 ± 17.3	81.6 ± 15.1	77.2 ± 17.3	42.8–112.2	61.6–92.0	34.5–123.5	0.640	0.033
Lodging (%)	13.9 ± 20.7	17.7 ± 25.0	15.7 ± 24.4	0–50	0–9.4	NA	NA	NA
Final stand count (plants/m²)	22.63 ± 6.08	21.69 ± 6.58	21.43 ± 6.65	5.78–32.89	17.7–24.8	1.61–42.79	0.298	0.186
Days to maturity	118.8 ± 12.8	118.9 ± 12.7	119.1 ± 12.9	96–138	104–123	84.8–147.6	0.870	0.625
Seed loss (number of pods shattered per plant)	0.32 ± 0.71	0.32 ± 0.77	0.30 ± 0.70	0–3	0–0.563	NA	NA	NA
Fruit count	47.1 ± 16.6	51.4 ± 18.1	48.5 ± 19.0	21–117	39.2–56.6	0–106.0	0.129	0.607
Seed weight (g)	18.2 ± 2.0	18.6 ± 2.2	18.3 ± 2.2	14.2–25.6	16.1–21.9	11.1–25.0	0.204	0.729
Yield (kg/ha at 13% moisture)	3574.6 ± 997.8	3632.0 ± 820.4	3455.2 ± 838.3	1382.8–6024.3	2838–4538	982.2–6896.3	0.700	0.425

NA = Not Applicable; analysis was not run due to sparse data.
[a] Range of results from nine reference lines (Entries F–N).
[b] Range of means for results from nine reference lines (Entries F–N).
[c] 99% Tolerance Interval: Range of reference varieties based on tolerance intervals specified to contain 99% of the population with 95% confidence.
[d] t-test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 40. Overview of categorical data

Stressor	n° of observations	n° of observations with no differences between GMB151 and the conventional counterpart
Abiotic stressors	132	132
Disease	132	132
Pest	132	132
SCN	22	22
Total	418	418

Table 41. Overview of the abiotic stressors investigated

Abiotic stressors investigated	n° of observations	n° of observations with no differences between GMB151 and the conventional counterpart
Cloudy / low light	1	1
Drought	19	19
Excess moisture in the soil	11	11
Flooding [a]	8	8
Hail injury	7	7
Heat stress	17	17
Mineral toxicity	2	2
Nutrient deficiency	21	21
Soil compaction	5	5
Soil crusting	7	7
Sun scald	10	10
Wind damage	24	24
Total	132	132

[a] Standing water above ground

Table 42. Overview of pests investigated

Pests investigated	n° of observations	n° of observations with no differences between GMB151 and the conventional counterpart
Alfalfa caterpillars	1	1
Aphids	11	11
Armyworms	2	2
Bean leaf beetles	27	27
Caterpillars, woolybear	3	3
Cutworms	3	3
Grape colaspis	1	1
Grasshoppers	22	22
Green cloverworms	11	11
Japanese beetles	24	24
Leafminer	1	1
Loopers	4	4
Mexican bean beetles	2	2
Soybean skipper	1	1
Spider mites	2	2
Stem borers	1	1
Stink bugs	8	8
Thistle caterpillar	5	5
Thrips	1	1
Whitefly	2	2
Total	**132**	**132**

Table 43. Overview of the diseases investigated

Diseases investigated	n° of observations	n° of observations with no differences between GMB151 and the conventional counterpart
Bacterial blight	13	13
Bacterial pustule	5	5
Bean pod mottle virus	5	5
Brown spot	13	13
Brown stem rot	3	3
Cercospora leaf spot	8	8
Charcoal rot	2	2
Downy mildew	9	9
Frogeye leaf spot	22	22
Phytophthora blight	2	2
Phytophthora root rot	7	7
Powdery mildew	4	4
Rhizoctonia foliar blight	1	1
Rhizoctonia rot	3	3
Rust	5	5
Septoria	9	9
Soybean mosaic virus	4	4
Soybean vein necrosis virus	4	4
Sudden death syndrome	9	9
Target spot	2	2
White mold	2	2
Total	**132**	**132**

8.3. Seed germination evaluation

To compare the germination potential of GMB151 soybean to that of the non-GM counterpart Thorne, warm and cold germination tests were conducted.

For warm germination testing, eight replicates of 50 seeds per replicate were incubated using germination paper in a plant growth chamber at 25 ± 5°C without light for approximately eight days. For cold germination testing, four replicates of 50 seeds each were incubated using germination paper in a plant growth chamber at 10 ± 5°C without light for seven days. Following the seven days of incubation, the cold germination test seeds were then incubated at 25 ± 5°C without light for an additional six days.

At the end of the specified incubation, normal seedlings (seedlings with well-developed essential structures), abnormal seedlings (seedlings with any of the essential structures missing or badly damaged), and ungerminated seeds were scored based on the Association of Official Seed Analysts guidelines (AOSA, 2013). Ungerminated seeds included seeds that had not germinated and remained hard (Hard Seeds) or seeds that were neither hard nor had produced any part of the seedling at the end of the test period (Dead Seeds). Chi-square analysis for the warm and cold germination studies is presented in Table 44 and Table 45, respectively. Descriptive statistics for the warm and cold germination studies are provided in Table 46 and Table 47, respectively. Analysis of the warm germination data showed that GMB151 and Thorne soybean had germination percentages of 78.50% and 79.75%, respectively. In the cold germination test, GMB151 and Thorne soybean had germination percentages of 84.00% and 81.00%, respectively. These data demonstrate that GMB151 soybean and the non-GM counterpart Thorne were similar.

Table 44. Comparison of germination categories between GMB151 and Thorne in the warm germination test

Genotype	Category	Count	Percent	p-value [a]
GMB151	Abnormal/ungerminated	86	21.50	
GMB151	Normal	314	78.50	
Thorne	Abnormal/ungerminated	81	20.25	
Thorne	Normal	319	79.75	
				0.664

[a] Chi-square test p-value for comparison of germination categories between GMB151 and Thorne in the warm germination test.

Table 45. Comparison of germination categories between GMB151 and Thorne in the cold germination test

Genotype	Category	Count	Percent	p-value [a]
GMB151	Abnormal/ungerminated	32	16.00	
GMB151	Normal	168	84.00	
Thorne	Abnormal/ungerminated	38	19.00	
Thorne	Normal	162	81.00	
				0.430

[a] Chi-square test p-value for comparison of germination categories between GMB151 and Thorne in the cold germination test.

Table 46. Descriptive statistics of the percentages for different germination categories in the warm germination test

Genotype	Replicate	Category	Percent mean	Percent standard deviation	Percent minimum	Percent maximum
		Abnormal	15.25	6.92	6	26
GMB151	8	Normal	78.50	7.23	70	88
		Ungerminated	6.25	2.92	2	10
		Abnormal	19.75	4.95	10	26
Thorne	8	Normal	79.75	4.33	74	88
		Ungerminated	0.50	0.93	0	2

Table 47. Descriptive statistics of the percentages for different germination categories in the cold germination test

Genotype	Replicate	Category	Percent mean	Percent standard deviation	Percent minimum	Percent maximum
		Abnormal	9.50	2.52	6	12
GMB151	4	Normal	84.00	7.66	78	94
		Ungerminated	6.50	5.51	0	12
		Abnormal	17.50	7.00	10	26
Thorne	4	Normal	81.00	9.59	68	90
		Ungerminated	1.50	3.00	0	6

9. ENVIRONMENTAL SAFETY AND IMPACT ON AGRONOMIC PRACTICES

9.1. Environmental safety

Persistence, weediness and invasiveness

Commercial soybean varieties in the U.S. are neither problematic volunteer weeds in other cropping systems nor are they found as feral populations on unmanaged lands (OECD, 2000). The potential fate of soybean seeds remaining in the field after harvest includes rot, predation, herbicides from rotational crops, and winter weather. Soybeans generally do not survive over the winter season. When climatic conditions are permissive, volunteer soybeans provide minimal interference in the rotational crop and are not recognized as an economic problem in soybean production. Volunteer soybeans are not competitive and can be managed by existing agronomic practices.

GMB151 soybean is tolerant to HPPD-inhibitor herbicides such as IFT and remains sensitive to herbicides registered for pre-plant, pre-emergence, and post-emergence use for weed control in soybean and other crops commonly rotated with soybean. Volunteer GMB151 soybeans can be controlled with pre-emergence or post-emergence herbicides such as 2,4-D, atrazine, glyphosate, glufosinate, acetochlor, and dicamba. These products are also widely used for weed control in the rotational crops of soybean.

As soybeans are not difficult to control as volunteers in a subsequent crop, and as GMB151 soybean has been shown to be no different from cultivated soybean in any of the traits that might impact weediness, the current practice to control volunteers will be effective.

Gene flow and its consequences

Soybean is a self-pollinating crop. Anthers mature in the floral buds and directly pollinate the stigma of the same flower (cleistogamy). Natural cross-pollination with nearby soybean plants is reported to be less than 1% (OECD, 2000). The extent of outcrossing can be influenced by the distance between individual plants, floral characteristics of different varieties, environmental conditions, and insect activities. In seed production fields, the occurrence of cross-pollination is so low that the standards of certified seed production require isolation distances to prevent mechanical mixture (7 CFR § 201.76) based upon the width of the harvest machinery.

Soybean is a non-native crop of the Americas. The origin of the *Glycine* genus is Asia, and there are no wild or native soybean relatives in the Americas that could be considered potential targets for gene flow.

Potential selective advantage to wild relatives

Sexually compatible wild relatives of soybean do not occur in the U.S. or its territories (Mammadov et al., 2018). Therefore, there would be no outcrossing of GMB151 soybean to wild relatives or potential for selective advantage.

Potential for horizontal gene transfer

Scientific evidence to date supports that horizontal gene transfer is an extremely rare event in the environment under natural conditions (Nielsen et al., 1998; Keese, 2008).

The factors affecting the potential for horizontal gene transfer between GM plants and microorganisms in the environment or in the gastrointestinal system have been extensively studied. Successful gene transfer via transformation requires many factors that are highly unlikely to be encountered under field conditions, including available free DNA, the development of competent bacteria, gene sequence homology between plant and bacterial DNA, stable integration of DNA from plant cells into the bacterial genome, and a strong selective pressure (Keese, 2008). Transformation frequencies (the frequency of foreign DNA incorporation into the microbial genome) likely to be encountered in the field are exceedingly low, representing environmental significance only on an evolutionary time scale.

Altered disease/pest susceptibility

Protecting soybean against nematodes such as SCN can help the plants better cope with other stresses. GMB151 soybean will have decreased susceptibility to SCN. As a result of reduced susceptibility to SCN, there can be reduced occurrence or severity of soybean diseases that interact with SCN, such as sudden death syndrome caused by the soil borne pathogen *Fusarium virguliforme* in North America (Back et al., 2002) and iron deficiency chlorosis.

There is no known mechanism that would result in increased susceptibility to diseases or pests by the introduction of GMB151 soybean.

Potential interaction with the abiotic environment

Decreased susceptibility to SCN should allow GMB151 soybean plants to better cope with other environmental stresses such as weather conditions.

There are no indications from the agronomic assessment or observations made during regulated field trial monitoring that GMB151 soybean would otherwise affect the interaction of soybean plants with the abiotic environment.

Survival

To compare the germination potential of GMB151 soybean to that of the non-GM counterpart Thorne, the warm-cold germination test was conducted (Section 8.3).

Analysis of the warm germination data showed that GMB151 and Thorne soybean had germination percentages of 78.50 and 79.75%, respectively. In the cold germination test, GMB151 and Thorne soybean had germination percentages of 84.00 and 81.00%, respectively.

Evaluation of the germination potential of GMB151 soybean compared to the non-GM counterpart Thorne demonstrated that there was no significant difference in the germination potential of the two genotypes under the two germination conditions.

9.2. Potential effects on non-target organisms, including beneficial organisms and threatened and endangered species

Tier 1 testing of standard non-target organism panel with purified Cry14Ab-1 protein

For the environmental risk assessment of pesticidal Cry proteins expressed in GM plants, studies are performed with NTOs that are directly or indirectly exposed to these proteins but are not the

target pests. The detailed NTO assessment was provided to the U.S. EPA as part of the Section 3 registration submitted in November 2018.

The safety assessment starts with the selection of NTOs exposed to the Cry protein, taking into account their importance and beneficial function in the environment and their potential sensitivity to the toxin based on their taxonomic relatedness to the target species (e.g., beneficial pollinators, endangered species, predators or parasites of the target pests, and species with a specific cultural value). Additional selection criteria consider that the test species can be reared under artificial conditions, an international established and validated Good Laboratory Practices (GLP) study protocol is available, and a well-characterized, specific toxic reference substance is available to demonstrate sensitivity of the test system (ILSI-CERA, 2013; Koch et al., 2015).

Therefore, the test organisms selected for Cry14Ab-1 for these NTO studies are:

1. Adult and larval honey bees (*A. mellifera*) as a representative pollinator,
2. Collembola (*Folsomia candida*) and earthworm (*Eisenia fetida*) as representative soil-dwelling organisms,
3. Ladybird beetles (*Coleomegilla maculata* and *Coccinella septempunctata*) and green lacewings (*Chrysoperla carnea*) as representative predator organisms,
4. Water flea (*Daphnia magna*) as a representative aquatic organism, and
5. Bobwhite quail (*Colinus virginianus*).

NTO testing demonstrated no adverse effects on any species tested. For these Tier 1 NTO studies to translate to the field, the expected environmental concentration (EEC) of Cry14Ab-1 in relevant exposure routes must be calculated and compared to the No Observed Effect Concentration (NOEC) or Dose (NOED) determined in the Tier 1 studies. This generates a margin of exposure (MOE), or safety factor, between the concentrations studied in the Tier 1 and expected real world exposures. The MOE for all NTOs tested was determined to be greater than 10 for all species except *C. septempunctata.* Therefore, it is concluded that cultivation of GMB151 soybean is unlikely to have effects on non-target invertebrates, vertebrates, or annelids including endangered species. Calculations of the EEC and resulting NOEC or NOED for each NTO study are provided in Appendix 8. Details regarding each NTO test, including test species, concentration, test conditions, and results, are provided in Appendix 7.

Potential effects on non-target, free-living nematodes

An assessment of the potential risk of the cultivation of GMB151 soybean on free-living nematodes (FLNs) was conducted (Appendix 9). Briefly, GMB151 soybean and its non-GM isoline were cultivated for two consecutive seasons in field trials in Fruitlant, IA. Each year, the GMB151 soybean and isoline were sown into the same plots and, as near as possible, in the same furrow. This maximized the potential impact of GMB151 soybean cultivation on potentially susceptible FLNs through continuous cultivation.

In August 2019 during podfill (R5/R6), soil samples were taken from each plot. Bulk soil between soybean rows in each plot was sampled. In addition, soybean plants were dug from each plot and soil sampled as rhizosphere soil. Nematodes were extracted from soil samples, and population densities were estimated of all genera of FLNs present in the samples. These populations were

categorized into their respective functional niches, and calculations of appropriate ecological indices were made for each plot type (GMB151 or non-GM isoline) and sampling location (bulk soil or rhizosphere). The presence of GMB151 did not affect the population densities of any FLN trophic group or any of the ecological indices calculated.

This study indicates that the cultivation of GMB151 soybean is unlikely to have negative effects on non-target FLNs or the ecological services which they may provide as a dimension of soil health, and ultimately plant health.

9.3. Impact on agronomic practices

GMB151 soybean is genetically modified to confer resistance to the plant-parasitic SCN. In addition, GMB151 soybean is modified for tolerance to HPPD-inhibitor herbicides. Petitions for HPPD-inhibitor tolerant soybean varieties have previously been analyzed by the USDA. The potential impact of GMB151 soybean introduction into U.S. agriculture will be similar to these previously reviewed events. Therefore, this section will focus on the impacts of GMB151 soybean introduction as it relates to SCN resistance.

9.3.1. Importance of soybean cyst nematode as a pest

SCN is considered to be the most damaging pathogen of soybeans in the U.S. and Canada (Wrather et al. 2010; Allen et al. 2017). In the U.S., *H. glycines* was discovered for the first time in North Carolina in 1954 and subsequently has spread to nearly every soybean producing state in the U.S. (Figure 30). Yield loss due to SCN can be variable depending on location, genetic background of the variety, and cropping system. Losses of up to 30% have been reported (Niblack and Tylka, 2014). Use of soybean varieties resistant to SCN coupled with rotations with non-host crops have been the most effective ways of managing SCN infestation.

SCN is a microscopic, soil-dwelling animal that reproduces sexually and must feed on host plants to complete its developmental cycle. It has a very high rate of increase on susceptible host plants and can complete between three and five generations per year, depending on temperature. Mated females produce 50–100 eggs outside the body, and then their body fills internally with 200 or more eggs each. Upon dying, the body wall toughens into a protective cyst around the eggs, some of which will remain dormant in soil for multiple years (Tylka, 2012). Normal plant growth and development is hampered by SCN feeding within the vascular system and developing within roots. Female nematodes become so large that they break through the fine roots they infest, making them visible to the unaided eye on exhumed roots. Nonetheless, yield reductions caused by SCN can sometimes be without obvious visible symptoms in the above-ground portion of plants. SCN damage is commonly misdiagnosed as the result of other plant stresses such as drought, compaction, or nutrient deficiencies (Tylka, 2012).

To date, there are over 1000 SCN-resistant soybean varieties (Tylka and Mullaney, 2018) available to farmers in the U.S. However, most of them have been derived from two prominent parental lines, PI 88788 and Peking (Niblack, 2005; Shannon et al., 2004). Extensive use of PI 88788 SCN-resistant varieties has resulted in decreasing effectiveness against the SCN pest in some parts of the Midwest U.S. (McCarville et al., 2017). There is a need to provide farmers with additional protection against the SCN pest.

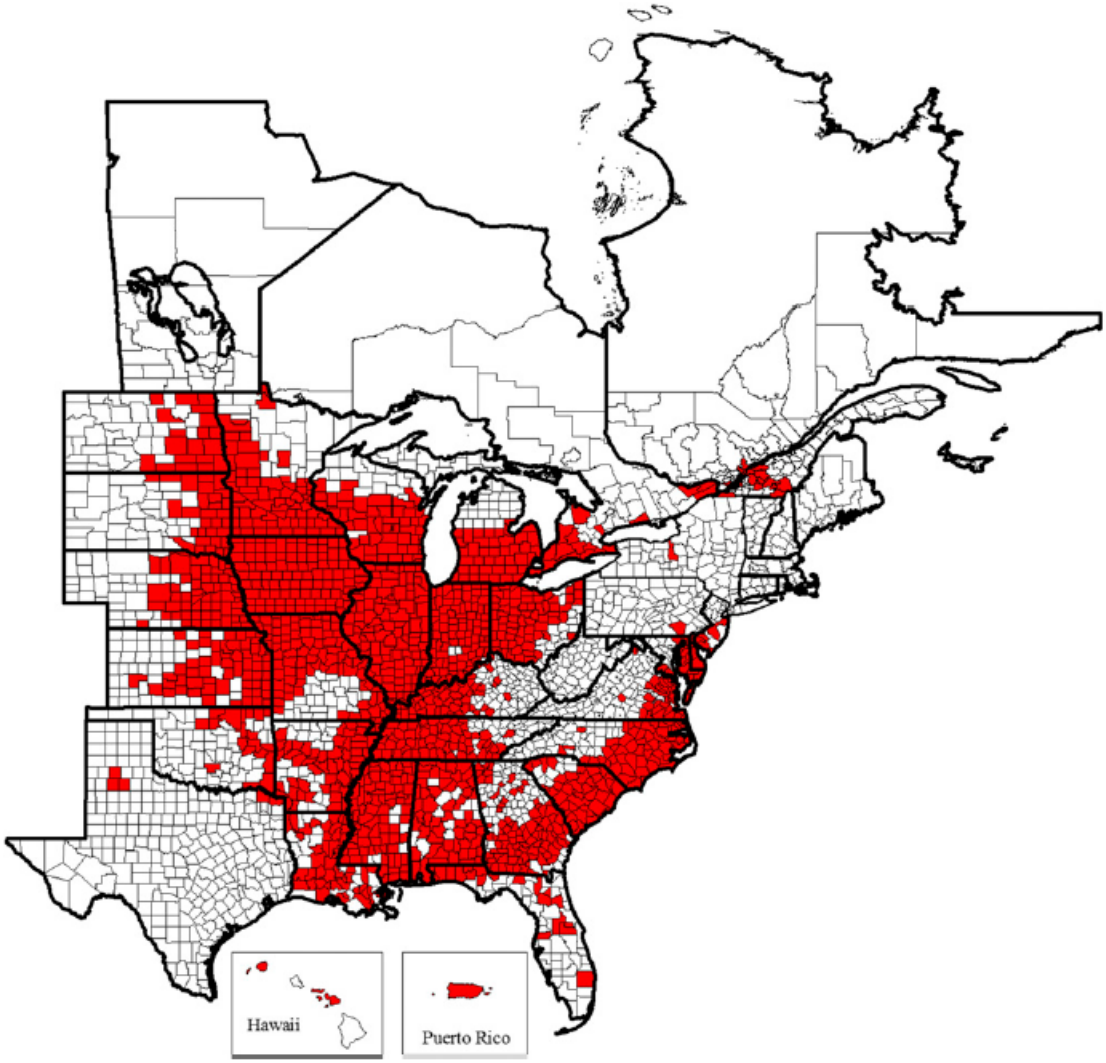


Figure 30. Map of the known distribution of soybean cyst nematode, *Heterodera glycines*, in the United States and Canada from 1954 to 2017

Known infested counties are indicated in red.
Map © C. C. Marett and G. L. Tylka, Iowa State University, 2017 (Tylka and Marett, 2017).

9.3.2. Plans for commercial development of GMB151

GMB151 soybean will be sold in the U.S. only in combination with other “native” SCN resistance trait(s) and as breeding stacks with other events. The most common source of SCN resistance in current soybean cultivars in the Midwest U.S. and Canada is PI 88788. However, two other commercially-available sources of SCN resistance, Peking and PI 437654 (also known as Hartwig or CystX®), have limited but increasing availability.

9.3.3. Resistance risk and ramifications

Georghiou and Taylor (1977) identified the major genetic, biological, and operational factors influencing resistance development. These provided a framework for the resistance risk assessments required by the U.S. EPA for *Bt* crops (Glaser and Matten, 2003; U.S. EPA, 2017, 2018). Pest resistance is an expectation in biological systems with all selective agents that have high adoption rates. Resistance risk assessment seeks to estimate the time-to-resistance under relevant biological assumptions. BASF concludes that the risk of SCN developing resistance to the Cry14Ab-1 protein in GMB151 is moderate or low in the soybean production systems of U.S. and Canada. This is due to multiple factors—most significantly, the uniqueness of the mode of action of Cry14Ab-1, the low/moderate intensity of selection imposed on SCN by GMB151, and GMB151 being deployed only in combination with other sources of native resistance to SCN.

A novel *Bt* trait

GMB151 soybean expresses Cry14Ab-1, the first nematicidal *Bacillus thuringiensis* protein to be commercialized. There are no other known *Bt* proteins in commercial agricultural use, whether sprayable or expressed in transformed crops, that control nematodes (Wei et al., 2003; ISAAA, 2018).

Low/moderate intensity of selection

GMB151 causes moderate suppression of SCN. Population-level selection of SCN by GMB151 is further limited by the mobility of these microscopic animals. Only the proportion of the nematode population in close proximity to GMB151 roots will experience selection with Cry14Ab-1. This leaves significant portions of each generation of SCN in soil that will not encounter the *Bt* protein expressed by GMB151. Beginning with the first seedling roots and continuing through mature soybean root system development, the Cry14Ab-1 protein is conveyed to the pest via growth of the root system. The net result is low or moderate selection for resistance.

GMB151 combined with PI 88788 or other native resistance traits

There is general agreement that durable plant resistance is most likely to be achieved by combining multiple pest resistance traits, each of which contribute to pest population suppression through unique biological mechanisms (Parlevliet and Zadoks, 1977). Yet, there are remarkable examples of single-gene resistance to nematodes remaining effective for long periods of time. Roberts (1992) cited three major examples of simply-inherited (i.e. single-gene) nematode resistance having durability more than 30 years—when used in an integrated management program. GMB151 soybean will be combined with other commercially-available varieties possessing SCN resistance. PI 88788 has a history of successful SCN suppression spanning over 30 years. This long durability contrasts dramatically with examples of single-gene resistance to insect and fungal pests that have failed in less than 10 years (see Briggs and Knowles, 1967). The long durability of resistance of PI 88788 to SCN (Niblack and Tylka, 2014) supports the conclusion of low-to-moderate resistance risk of SCN. However, there is clear evidence of SCN populations with the ability to develop on resistant soybean lines (Niblack et al., 2008). Adding GMB151 to varieties possessing native resistance will extend the durability of both GMB151 and native resistance.

9.3.4. IPM best management practices

Recent resistance issues with western corn rootworm (U.S. EPA, 2016) and earworm/bollworm (U.S. EPA, 2018) in the U.S. have prompted a renewed emphasis on integrated pest management (IPM), and especially pest monitoring and pest avoidance activities that are essential for success of resistance management. BASF will take specific measures to promote IPM of SCN with the commercialization of products containing GMB151 (see below). Specific measures, whether mandatory requirements or recommendations (e.g., Best Management Practice (BMP) recommendations), should be practical and effective and have the imperatives of the farmer foremost in mind.

As detailed by Tylka (2012), the most effective SCN management program integrates as many different management strategies as possible to maintain high soybean yields while preventing increases in SCN population densities. BASF will promote the active management of SCN populations, as detailed by the SCN Coalition[1]. This will include testing of fields to estimate SCN numbers and, where possible, rotation of native resistant varieties. Similarly, BASF will promote rotating to non-host crops and use of a nematode-protectant seed treatment. BASF will continue to support the SCN Coalition and will conduct grower educational programs built around their IPM messages. In addition, IPM messages will be included in the product technology guides and on labels, as appropriate.

[1] The SCN Coalition, composed of University researchers, Extension specialists, and ag company representatives, is concerned about the evolving threat from soybean cyst nematode (https://www.thescncoalition.com/resources).

10. STATEMENT OF GROUNDS UNFAVORABLE

BASF Corporation knows of no study data and/or observations associated with GMB151 soybean that will result in adverse environmental consequences for its introduction. The only biologically relevant phenotypic difference between GMB151 soybean and conventional soybean is the expression of the Cry14Ab-1 and HPPD-4 proteins, which provide protection against plant-parasitic nematodes and tolerance to HPPD-inhibitor herbicides such as isoxaflutole, respectively. Planting GMB151 soybean varieties will provide growers with a new option for genetic resistance to the plant-parasitic soybean cyst nematode. In addition, growers will have another option for HPPD-inhibitor based weed control in soybean.

11. REFERENCES

Allen TW, Bradley CA, Sisson AJ, Byamukama E, Chilvers MI, Coker CM, Collins AA, Damicone JP, Dorrance AE, Dufault NS, Esker PD, Faske TR, Giesler LJ, Grybauskas AP, Hershman DE, Hollier CA, Isakeit T, Jardine DJ, Kelly HM, Kemerait RC, Kleczewski NM, Koenning SR, Kurle JE, Malvick DK, Markell SG, Mehl HL, Mueller DS, Mueller JD, Mulrooney RP, Nelson BD, Newman MA, Osborne L, Overstreet C, Padgett GB, Phipps PM, Price PP, Sikora EJ, Smith DL, Spurlock TN, Tande CA, Tenuta AU, Wise KA and Wrather JA, 2017. Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. Plant Health Progress 18, 19–27.

Allison RF, Dougherty WG, Parks TD, Willis L, Johnston RE, Kelly M and Armstrong FB, 1985. Biochemical analysis of the capsid protein gene and capsid protein of tobacco etch virus: N-terminal amino acids are located on the virion's surface. Virology 147, 309–316.

AOSA, 2013. Rules for Testing Seeds, Volume 4, Seedling Evaluation. Association of Official Seed Analysts, Washington, D.C.

Back MA, Haydock PPJ and Jenkinson P, 2002. Disease complexes involving plant parasitic nematodes and soilborne pathogens. Plant Pathology 51, 683–697.

Betz FS, Hammond BG and Fuchs RL, 2000. Safety and advantages of *Bacillus thuringiensis*-protected plants to control insect pests. Regulatory Toxicology and Pharmacology 32, 156–173.

Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL and Boyer HW, 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2, 95–113.

Bradshaw RA, Brickey WW and Walker KW, 1998. N-terminal processing: The methionine aminopeptidase and N^{α}-acetyl transferase families. Trends in Biochemical Science 23, 263–267.

Bravo A, Gill SS and Soberón M, 2007. Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. Toxicon 49, 423–435.

Briggs FN and Knowles PF, 1967. Stability of resistance. In: Introduction to Plant Breeding. Reinhold Publishing Corporation, USA, 365–366.

Cornelissen M and Vandewiele M, 1989. Nuclear transcriptional activity of the tobacco plastid psbA promoter. Nucleic Acids Research 17, 19–25.

Federici BA and Siegel JP, 2008. Safety assessment of *Bacillus thuringiensis* and Bt crops used in insect control. In: Food Safety of Proteins in Agricultural Biotechnology. Ed Hammond BG, CRC Press, Boca Raton, FL, 45–102.

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

Fritze IM, Linden L, Freigang J, Auerbach G, Huber R and Steinbacher S, 2004. The crystal structures of *Zea mays* and Arabidopsis 4-hydroxyphenylpyruvate dioxygenase. Plant Physiology 134, 1388–1400.

Georghiou GP and Taylor CE, 1977. Genetic and biological influences in the evolution of insecticide resistance. Journal of Economic Entomology 70, 319–323.

Glaser JA and Matten SR, 2003. Sustainability of insect resistance management strategies for transgenic Bt corn. Biotechnology Advances 22, 45–69.

Grefen C, Donald N, Hashimoto K, Kudla J, Schumacher K and Blatt MR, 2010. A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. Plant Journal 64, 355–365.

Heeb S, Itoh Y, Nishijyo T, Schnider U, Keel C, Wade J, Walsh U, O'Gara G and Haas D, 2000. Small, stable shuttle vectors based on the minimal pVS1 replicon for use in Gram-negative, plant-associated bacteria. Molecular Plant-Microbe Interactions 13, 232–237.

Hoekema A, Hirsch PR, Hooykaas PJJ and Schilperoort RA, 1983. A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303, 179–180.

ILSI, 2016, online. International Life Sciences Institute Crop Composition Database. Version 6. Accessed on August 15, 2018. available at www.cropcomposition.org.

ILSI-CERA, 2013. Surrogate species selection for assessing potential adverse environmental impacts of genetically engineered plants on non-target organisms. Proceedings of the Center for Environmental Risk Assessment, ILSI Research Foundation, Washington, D.C., June 26–28, 2012.

ISAAA, 2018, online. Commercial GM traits list. available at www.isaaa.org/gmapprovaldatabase/commercialtraitlist/default.asp.

ISU Department of Agronomy, 2019, online. Iowa Environmental Mesonet - Iowa Soil Moisture Network / Daily Download. Iowa State University of Science and Technology. Accessed on October 17, 2019. available at https://mesonet.agron.iastate.edu/agclimate/hist/daily.php.

ISU Extension and Outreach, 2019, online. Soybean planting date can have a significant impact on yield. Iowa State University of Science and Technology. Accessed on October 17, 2019. available at https://crops.extension.iastate.edu/soybean-planting-date-can-have-significant-impact-yield.

Jacqmin-Gadda H, Sibillot S, Proust C, Molina J-M and Thiébaut R, 2007. Robustness of the linear mixed model to misspecified error distribution. Computational Statistics & Data Analysis 51, 5142–5154.

Kay R, Chan A, Daly M and McPherson J, 1987. Duplication of CaMV 35*S* promoter sequences creates a strong enhancer for plant genes. Science 236, 1299–1302.

Keese P, 2008. Risks from GMOs due to horizontal gene transfer. Environmental Biosafety Research 7, 123–149.

Kenward MG and Roger JH, 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53, 983–997.

Koch MS, Ward JM, Levine SL, Baum JA, Vicini JL and Hammond BG, 2015. The food and environmental safety of *Bt* crops. Frontiers in Plant Science 6, 283.

Lebrun M, Leroux B and Sailland A, 1996. Chimeric gene for the transformation of plants. United States Patent Number 5,510,471. 5 pp.

Mammadov J, Buyyarapu R, Guttikonda SK, Parliament K, Abdurakhmonov IY and Kumpatla SP, 2018. Wild relatives of maize, rice, cotton, and soybean: Treasure troves for tolerance to biotic and abiotic stresses. Frontiers in Plant Science 9, 886.

McCarville MT, Marett CC, Mullaney MP, Gebhart GD and Tylka GL, 2017. Increase in soybean cyst nematode virulence and reproduction on resistant soybean varieties in Iowa from 2001 to 2015 and the effects on soybean yields. Plant Health Progress 18, 146–155.

Niblack TL, 2005. Soybean cyst nematode management reconsidered. Plant Disease 89, 1020–1026.

Niblack TL, Colgrove AL, Colgrove K and Bond JP, 2008. Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788. Online. Plant Health Progress doi:10.1094/PHP-2008-0118-01-RS.

Niblack TL and Tylka GL, 2014, online. Soybean cyst nematode management guide. 5th edition. Plant Health Initiative, North Central Soybean Research Program. available at http://soybeanresearchinfo.com/pdf_docs/SCNGuide_5thEd.pdf.

Nielsen KM, Bones AM, Smalla K and van Elsas JD, 1998. Horizontal gene transfer from transgenic plants to terrestrial bacteria — a rare event? FEMS Microbiology Reviews 22, 79–103.

NIST, 2012. Tolerance intervals for a normal distribution. In: NIST/SEMATECH e-Handbook of Statistical Methods. U.S. Department of Commerce, Technology Administration. available at www.itl.nist.gov/div898/handbook/prc/section2/prc263.htm.

OECD, 1997. Series on Harmonization of Regulatory Oversight in Biotechnology No. 6. Consensus document on information used in the assessment of environmental applications involving *Pseudomonas*. Organisation for Economic Co-Operation and Development, Paris, France, 110 pp.

OECD, 2000. Series on Harmonization of Regulatory Oversight in Biotechnology No. 15. Consensus document on the biology of *Glycine max* (L.) Merr. (soybean). Organisation for Economic Co-Operation and Development, Paris, France, 22 pp.

OECD, 2007. Series on Harmonization of Regulatory Oversight in Biotechnology No. 42. Consensus document on safety information on transgenic plants expressing *Bacillus thuringiensis* - derived insect control proteins. Organisation for Economic Co-Operation and Development, Paris, France, 109 pp.

Palma L, Muñoz D, Berry C, Murillo J and Caballero P, 2014. *Bacillus thuringiensis* toxins: An overview of their biocidal activity. Toxins 6, 3296–3325.

Parlevliet JE and Zadoks JC, 1977. The integrated concept of disease resistance: A new view including horizontal and vertical resistance in plants. Euphytica 26, 5–21.

Peariso AM, Nicholson KM, Jones RB, Green-Church KB and Funk Jr. MO, 2008. Electrospray ionization mass spectrometry of soybean lipoxygenases: N-terminal acetylation, chemical modification, and solution conformation. Proteins 70, 650–658.

Polevoda B and Sherman F, 2000. N^{α}-terminal acetylation of eukaryotic proteins. Journal of Biological Chemistry 275, 36479–36482.

Polevoda B and Sherman F, 2003. Composition and function of the eukaryotic N-terminal acetyltransferase subunits. Biochemical and Biophysical Research Communications 308, 1–11.

Porée F, Heinrichs V, Lange G, Laber B, Peters C and Schouten L, 2014. International Patent Application WO 2014/043435 A1. HPPD variants and methods of use. 20 March 2014, 104 pp.

Raybould A, Kilby P and Graser G, 2013. Characterising microbial protein test substances and establishing their equivalence with plant-produced proteins for use in risk assessments of transgenic crops. Transgenic Research 22, 445–460.

Riggs RD and Schmitt DP, 1989. Soybean cyst nematode. In: Compendium of Soybean Diseases. Eds Sinclair JB and Backman PA. 3rd edition, APS Press, St. Paul, MN, 65–67.

Roberts PA, 1992. Current status of the availability, development, and use of host plant resistance to nematodes. Journal of Nematology 24, 213–227.

Sanahuja G, Banakar R, Twyman RM, Capell T and Christou P, 2011. *Bacillus thuringiensis*: A century of research, development and commercial applications. Plant Biotechnology Journal 9, 283–300.

Sanfaçon H, Brodmann P and Hohn T, 1991. A dissection of the cauliflower mosaic virus polyadenylation signal. Genes & Development 5, 141–149.

Schabenberger O and Pierce FJ, 2002. Recursive and linearly recovered errors. In: Contemporary Statistical Models for the Plant and Soil Sciences. CRC Press, New York, NY, 135–139.

Schoenberg DR and Maquat LE, 2012. Regulation of cytoplasmic mRNA decay. Nature Reviews Genetics 13, 246–259.

Shannon JG, Arelli PR and Young LD, 2004. Breeding for resistance and tolerance. In: Biology and Management of Soybean Cyst Nematode. Eds Schmitt DP, Wrather JA and Riggs RD. 2nd edition, Schmitt & Associates of Marceline, Marceline, MO, 155–180.

Shurtleff W and Aoyagi A, 2004, online. History of world soybean production and trade. A chapter from the unpublished manuscript History of Soybeans and Soyfoods, 1100 B.C. to the 1980s. Soyfoods Center, Lafayette, CA. Accessed on October 17, 2019. available at http://www.soyinfocenter.com/HSS/production_and_trade1.php.

Siegel JP, 2001. The mammalian safety of *Bacillus thuringiensis*-based insecticides. Journal of Invertebrate Pathology 77, 13–21.

Siehl DL, Tao Y, Albert H, Dong Y, Heckert M, Madrigal A, Lincoln-Cabutu B, Lu J, Fenwick T, Bermudez E, Sandoval M, Horn C, Green JM, Hale T, Pagano P, Clark J, Udranszky IA, Rizzo N, Bourett T, Howard RJ, Johnson DH, Vogt M, Akinsola G and Castle LA, 2014. Broad 4-hydroxyphenylpyruvate dioxygenase inhibitor herbicide tolerance in soybean with an optimized enzyme and expression cassette. Plant Physiology 166, 1162–1176.

Smith K, 2017, online. Soybeans - history and future. Soybean Meal Info Center Fact Sheet. United Soybean Board, 5 pp. available at www.soymeal.org/soy-meal-articles/soybeans-history-and-future/.

Strickberger MW, 1976. Probability and statistical testing. In: Genetics. 2nd edition, Macmillan, New York, NY, 140–163.

Tylka GL, 2012, online. Soybean cyst nematode field guide. Extension and Outreach Publications 223. available at https://lib.dr.iastate.edu/extension_pubs/223.

Tylka GL and Marett CC, 2017. Known distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada, 1954 to 2017. Plant Health Progress 18, 167–168.

Tylka GL and Mullaney MP, 2018, online. Soybean cyst nematode-resistant soybean varieties for Iowa. Extension and Outreach Publications 496. available at https://lib.dr.iastate.edu/extension_pubs/496.

U.S. EPA, 2016, online. Registration terms for CRW IRM – final version 1-26-16. available at www.regulations.gov/docket?D=EPA-HQ-OPP-2014-0805.

U.S. EPA, 2017, online. Insect resistance management for Bt plant-incorporated protectants. Accessed on October 17, 2019. available at www.epa.gov/regulation-biotechnology-under-tsca-and-fifra/insect-resistance-management-bt-plant-incorporated.

U.S. EPA, 2018. White paper on resistance in Lepidopteran pests of *Bacillus thuringiensis* (Bt) plant-incorporated protectants in the United States. 82 pp.

USDA-NASS, 2018, online. Quick Stats. Accessed on July 6, 2018. available at http://quickstats.nass.usda.gov/.

Van Acker RC, Swanton CJ and Weise SF, 1993. The critical period of weed control in soybean (*Glycine max* (L.) Merr.). Weed Science 41, 194–200.

Wei J-Z, Hale K, Carta L, Platzer E, Wong C, Fang S-C and Aroian RV, 2003. *Bacillus thuringiensis* crystal proteins that target nematodes. Proceedings of the National Academy of Sciences of the United States of America 100, 2760–2765.

WHO/IPCS, 1999. Environmental Health Criteria 217. *Bacillus thuringiensis*. World Health Organization, Geneva.

Wrather A, Shannon G, Balardin R, Carregal L, Escobar R, Gupta GK, Ma Z, Morel W, Ploper D and Tenuta A, 2010. Effect of diseases on soybean yield in the top eight producing countries in 2006. Online. Plant Health Progress doi:10.1094/PHP-2010-0125-01-RS.

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

Zhang Z, Schwartz S, Wagner L and Miller W, 2000. A greedy algorithm for aligning DNA sequences. Journal of Computational Biology 7, 203–214.

Zhu J, Oger PM, Schrammeijer B, Hooykaas PJJ, Farrand SK and Winans SC, 2000. The bases of crown gall tumorigenesis. Journal of Bacteriology 182, 3885–3895.

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND and Cregan PB, 2003. Single-nucleotide polymorphisms in soybean. Genetics 163, 1123–1134.

Appendix 1. History of field activities under USDA notification or permit

List of Tables

Field trials of GMB151 soybean have been conducted in the U.S. since 2013 under USDA notification. Trials have been conducted for research, development, regulatory, breeding, and seed multiplication. Monthly planting reports, field trial termination reports and volunteer monitoring reports have been submitted to USDA as these trials and associated monitoring are completed. No reports of unusual or unexpected phenotypes, effects on NTOs, susceptibility to plant pests, or other unexpected interactions with the biotic or abiotic environment have been reported. Detailed location information by notification is provided in Table A1.1.

Table A1.1. Summary of field activities under USDA notification for GMB151 soybean

Year	Permit Number	State - # of locations
2013	13-063-108n	IA - 2, IL - 4, IN - 1
2013	13-066-102n	IN - 1
2013	13-280-105n	PR - 1
2014	14-070-108n	IA - 2, IL - 6, IN - 3, MN - 3, MS - 1, NC - 1
2014	14-077-101n	MO - 1
2014	14-241-102n	PR - 3
2015	15-063-103n	MO - 2
2015	15-064-103n	AR - 1, IA - 3, IL - 5, IN - 3, MN - 1, OH - 1, TN - 1
2015	15-247-105n	PR - 2
2016	16-075-101n	AR - 1, IA - 4, IL - 11, IN - 3, MN - 2, OH - 1, PR - 1, TN - 1, WI - 1
2016	16-078-106n	MO - 1
2016	16-082-104n	IA - 3, IL - 2, IN - 2, KS - 2, MO - 3, NE - 2, OK - 1, PA - 1
2016	16-116-105n	IA - 1, IL - 1, ND - 1
2016	16-235-102n	PR - 2
2017	17-062-103n	AR - 1, CA - 1, GA - 3, IA - 15, IL - 11, IN - 7, KS - 3, MN - 4, MO - 4, MS - 1, NC - 4, ND - 2, NE - 3, OH - 1, PA - 1, PR - 3, SD - 1, TN - 1, WI - 2
2017	17-067-104n	MS - 2
2017	17-250-101n	PR - 2
2018	18-016-101n	PR - 1
2018	18-064-101n	IA - 2, IL - 1, IN - 1, ND - 1, NE - 1
2018	18-068-105n	GA - 1, IA - 3, IL - 1, IN - 3, MN - 2, MO - 2, MS - 1, NC - 1, NE - 3, PR - 1, SC - 1, TN - 1, WI - 1
2018	18-276-102n	PR - 2
2019	19-072-101n	IA - 5, IL - 4, IN - 4, KS - 2, MN - 3, NE- 4, ND - 1, PR - 2, TN - 1, WI - 1

Appendix 2. Methods and detailed results for molecular characterization of GMB151 soybean

List of Tables

List of Figures

1. Objective of study

The objectives of this study were (1) the characterization of inserted T-DNA sequences in GMB151 soybean, (2) the assessment of the presence or absence of vector backbone sequences, and (3) the demonstration of the stable inheritance of the inserted sequences across different breeding generations. This was undertaken by means of a bioinformatics JSA of whole-genome NGS data.

2. Identification of test item

The test item in this study was GMB151 soybean. Seeds containing GMB151 soybean were used to characterize the test item.

Five GMB151 soybean generations were investigated in this study: four generations in the genetic background Thorne (T_2, T_4, T_5, and T_6) and one additional generation that was backcrossed in Parent 5 (BC_2F_3). All seed lots used in this study were non-segregating. For each generation, gDNA was extracted from a pool of ten GMB151 soybean seeds. The identity of the seed material was confirmed.

Generation	Seed lot number	Identification document of the seed material
T_2	13MRGM003079	STSa-IdConf-01-v01
T_4	14PRGM620044	STSa-IdConf-02-v01
T_5	14ILGM000146	STSa-IdConf-03-v01
T_6	15PGGM130007	STSa-IdConf-04-v01
BC_2F_3	16LGGM000006	STSa-IdConf-05-v01

3. Identification of reference item

No reference item was used in this study.

4. Control items

To ensure the correct interpretation of the generated results, negative and positive controls were included in the study.

4.1. Negative control items

Non-GM soybean variety Thorne (non-GM counterpart Thorne) was used as a negative control for the GMB151 soybean generations T_2 to T_6. Genomic DNA was extracted from a pool of ten non-GM counterpart Thorne seeds from seed lot 16LGGM000075.

Parent 5 (non-GM counterpart Parent 5) was used as a negative control for the backcross generation GMB151 soybean BC_2F_3. Genomic DNA was extracted from a pool of ten Parent 5 seeds from seed lot 16LGGM000070. The identity of the seed material was confirmed.

Negative controls: non-GM counterparts of GMB151 soybean		
Genetic background	**Seed lot number**	**Identification document of the seed material**
Thorne	16LGGM000075	STSa-IdConf-06-v01
Parent 5	16LGGM000070	STSa-IdConf-07-v01

4.2. Positive control item

The positive control in this study was the transforming plasmid pSZ8832 used to generate GMB151 soybean (Figure 1 of this Petition). Plasmid DNA was prepared from *E. coli* cell cultures containing pSZ8832.

5. Characterization of the test system

The test system is identified as the combination of analytical procedures used for the molecular characterization of the test item. The molecular characterization of the test item is done by means of the analytical procedure of whole-genome NGS.

Sequencing was done according to standard methods and in compliance with CLSI Guideline MM09-A2 (CLSI, 2014).

6. Method synopsis of NGS/JSA

A visual representation of the method synopsis is shown in Figure A2.1. Shotgun libraries were prepared from gDNA from GMB151 soybean seeds as well as from their non-GM counterparts. The libraries were used to generate NGS data that extensively covered the entire genome. To identify their origin, all sequencing reads were subsequently mapped to the transforming plasmid and host genome reference sequences. By visual representation of the mapping results to the transforming plasmid, the inserted transforming plasmid-derived sequences could be readily estimated based on their read coverage. The absence or presence of adequate coverage of vector backbone sequences demonstrates the absence or presence of vector backbone-derived sequences.

An in-depth molecular characterization was performed through the identification of all novel junctions that were created upon transformation. Novel junctions define which T-DNA sequences are inserted. Junctions containing both plasmid and flanking genomic sequences define the number of transgenic loci since each insertion of T-DNA sequences is characterized by two plasmid/flank junctions. Rearrangements within the inserted T-DNA sequences are further characterized by plasmid/plasmid junctions. The number of plasmid/plasmid junctions define the number of rearrangements as compared to the plasmid reference sequence. The alignment positions of the plasmid/flank and plasmid/plasmid junctions relative to the plasmid reference sequence determine the inserted portion of the T-DNA sequence as well as the positions of the rearrangements within the inserted T-DNA sequences.

For a transgenic locus containing a single, non-rearranged copy of the T-DNA, two different plasmid/flank junctions and no plasmid/plasmid junctions are expected. The alignment positions of the corresponding junction reads relative to the transforming plasmid sequence demonstrate that each of the plasmid/flank junctions originates from either the 5’ or 3’ end of the T-DNA sequence.

The transgenic locus is evaluated as stably maintained if the number and the identity of the junctions are conserved across different breeding generations.

A novel junction created upon transformation is typically covered by reads that only partially map to the transforming plasmid sequence. Therefore, reads partially mapping to the plasmid reference sequence were selected and further analyzed using a bioinformatics JSA.

Based on the positions that separated the mapped portion from the unmapped portion of the junction reads, the corresponding breakpoints within the transforming plasmid were collected. The selected junction reads were grouped based on their breakpoint position(s) within the transforming plasmid.

Breakpoint positions within the transforming plasmid were further supported by the presence of broken pairs following mapping of paired-end reads to the transforming plasmid. Broken pairs are mate-pairs for which only one mate maps to the transforming plasmid or for which both mates map to the transforming plasmid with unexpected insert lengths or non-standard orientation.

The consensus sequences of all junction reads corresponding to the same breakpoint position within the transforming plasmid were generated. The junctions were further identified as plasmid/flank or plasmid/plasmid junctions. The alignment positions of the junction reads relative to the transforming plasmid sequence demonstrated the origin of the identified junctions.

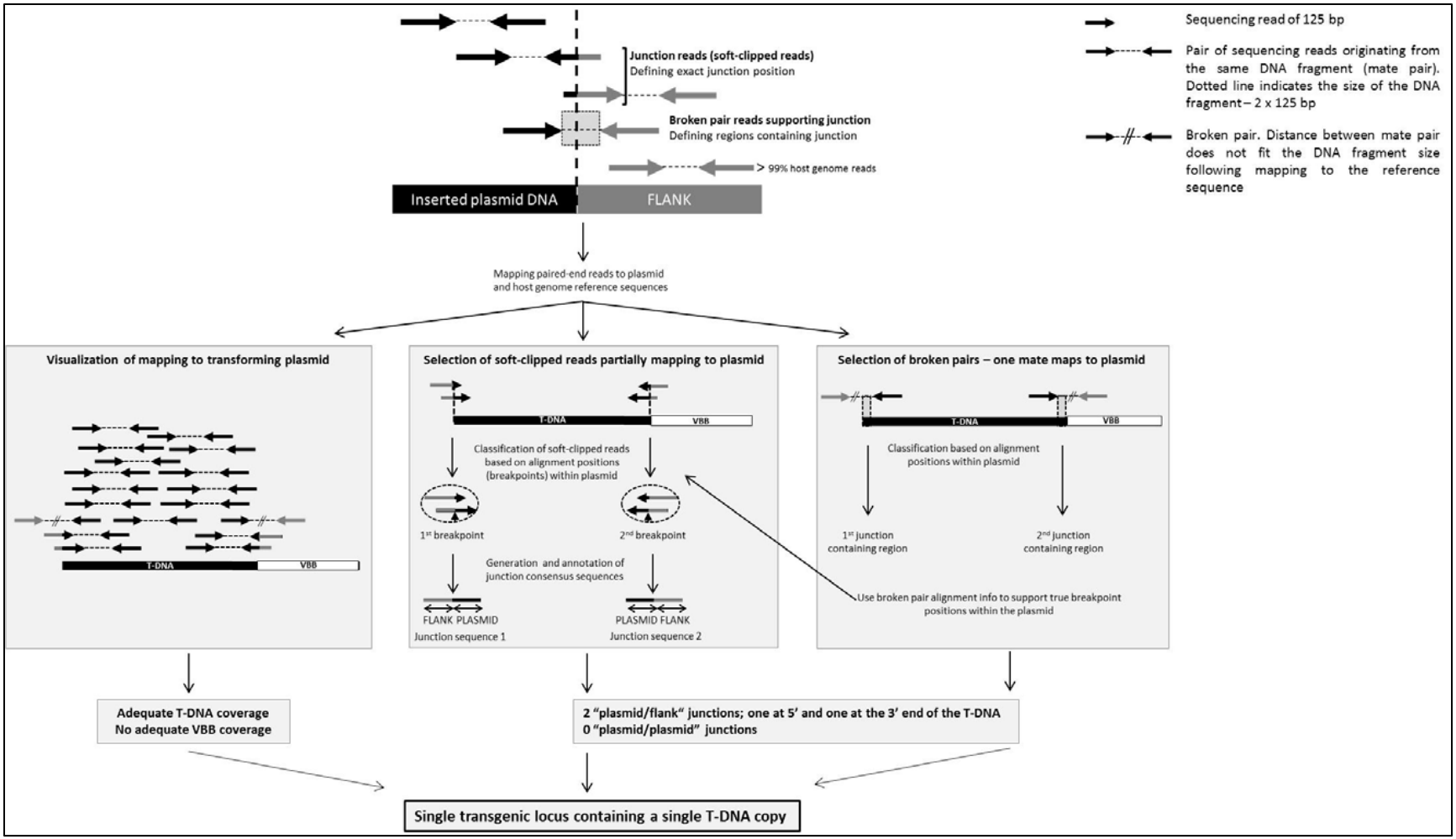


Figure A2.1. Method synopsis of whole-genome NGS/JSA

7. Complete description of experimental design

7.1. Preparation of genomic and plasmid DNA

Seeds from different generations of GMB151 soybean and from their non-GM counterparts were consecutively washed in 0.05% Tween, 0.5% NaClO, and 1% HCl to remove surface contaminants. Following each washing step, the seeds were rinsed with ultrapure water. Afterwards, seeds were left overnight at room temperature under a laminar flow hood to dry. Conventional manual coffee mills were used to grind the seeds. Ten seeds were used for each sample. Genomic DNA was isolated using the Qiagen DNeasy Plant Mini Kit (Qiagen). Genomic DNA preparation and initial quality control were performed for one sample at a time; preparation was started with seeds from the non-GM counterpart Parent 5, followed by seeds from the non-GM counterpart Thorne. Subsequently, gDNA was prepared from GMB151 soybean seeds and subjected to an initial quality control (QC). The transforming plasmid pSZ8832 was delivered as an already prepared DNA sample and could therefore directly be transferred into initial QC.

7.2. Quality control and concentration determination of genomic and plasmid DNA

QC of the genomic and plasmid DNA was done by optical density (OD) measurements at 260 nm and 280 nm using a BioPhotometer (Eppendorf) and agarose gel electrophoresis. The concentration of the DNA samples was measured using the fluorometric-based Quant-iT™ PicoGreen (Thermo Fisher Scientific). Only samples for which the following QC criteria were met were used for further analysis: the isolated DNA loaded on an agarose gel showed a clear high molecular weight band and no traces of degradation, and both the quality and quantity of the DNA were sufficient according to the OD and PicoGreen measurements.

7.3. DNA library generation

Four DNA libraries were prepared for each of the five different GMB151 soybean generations (T_2, T_4, T_5, T_6, and BC_2F_3), the non-GM counterpart Parent 5 (negative control), and the transforming plasmid pSZ8832. Libraries from the transforming plasmid pSZ8832 were used to generate the positive and sensitivity control samples used in the study setup. Twelve libraries were prepared for the non-GM counterpart Thorne since libraries from the non-GM counterpart Thorne were used both as a negative control sample and to generate the positive and sensitivity control samples. Hence, in total 40 indexed libraries were prepared using the TruSeq Nano DNA Library Prep Kit (Illumina) according to the Reference Guide Part No. 15041110 Rev. D June 2015. Approximately 0.1 µg DNA was used for each individual library preparation. Shearing of the DNA was performed on the Covaris™ E220 Systems using the SonoLAB™ 7.3 software (Covaris). The following settings were used: peak incident power: 175 W, duty cycle: 10%, cycles per burst: 200, time: 45 s, temperature: 5°C.

7.4. Quality control and initial concentration determination of the DNA libraries

The Bioanalyzer 2100 instrument (Agilent Genomics) and the Agilent DNA 1000 kit (Agilent Genomics) were used to determine the fragment size distribution and initial concentration of the individual DNA libraries. The measurement for the T_6 libraries was repeated once with the High Sensitivity kit (Agilent Genomics). Only libraries with expected size distribution (between 500–600 bp including adapters) and no traces of degradation were used for further analysis.

7.5. Pooling and spiking of the DNA libraries

To reduce potential bias that may occur during library preparation, equimolar amounts of four individual gDNA libraries corresponding to the same gDNA sample were pooled. Equimolar amounts of two individual plasmid DNA libraries were pooled as well.

The Bioanalyzer 2100 instrument and the Agilent DNA 1000 kit or High Sensitivity kit were used to check the fragment size distribution of the gDNA and plasmid DNA library pools and to determine the concentration of two library pools from the non-GM counterpart Thorne and two library pools from the transforming plasmid pSZ8832 that were used to generate the positive and sensitivity control samples. The positive control sample was prepared by spiking a first gDNA library pool from the non-GM counterpart Thorne with an equimolar amount from a first plasmid DNA library pool. Similarly, the sensitivity control sample was prepared by spiking a second gDNA library pool from the non-GM counterpart Thorne with $1/10_{th}$ equimolar amount from a second plasmid DNA library pool. The concentration determination using the Bioanalyzer 2100 instrument was used to calculate the equimolar and 1/10th equimolar spiking.

Determination of an equimolar amount of transforming plasmid pSZ8832:

- Size of *G. max* genome: 1.1 x 10^9 bp (Schmutz et al., 2010)
- Size of plasmid pSZ8832: 14361 bp; hence X µg gDNA corresponds to (X µg gDNA x 14361 bp / 1.1 x 10^9 bp) µg pSZ8832 plasmid DNA.

7.6. Final concentration determination of the DNA libraries

The final concentration of the ready-to-sequence DNA library pools was determined based on qPCR on prediluted DNA Standards and appropriately diluted gDNA library pools using the KAPA SYBR FAST qPCR Master Mix and primers targeting the Illumina P5 and P7 flow cell oligo sequences (KAPA Library Quantification Kit for Illumina platforms, KAPA Biosystems). qPCR was performed on the Roche LightCycler 480 II instrument (Roche). The concentration determination based on qPCR was used to adjust the quantity of the gDNA library pools for sequencing on an Illumina HiSeq instrument.

7.7. Sequencing

Following pooling, spiking, and concentration determination, the ready-to-sequence DNA library pools from five different generations of GMB151 soybean (T_2, T_4, T_5, T_6, and BC_2F_3), two different negative control samples (non-GM counterparts Thorne and Parent 5), the positive control sample (non-GM counterpart Thorne + equimolar amount of pSZ8832), and the sensitivity control sample (non-GM counterpart Thorne + $1/10_{th}$ equimolar amount of pSZ8832) were sequenced on an Illumina HiSeq 2500 instrument using the high-output run mode with v4 chemistry and the Illumina cluster generation and sequencing reagent kits (Illumina; HiSeq PE Cluster Kit v4, HiSeq SBS Kit v4). The pools were sequenced using the 2 x 125 bp paired end run mode. In total, three full flow cells with eight separate channels each were sequenced. To reduce potential bias during NGS sequencing, three different channels were used to sequence the library pools from the negative, positive, and sensitivity control samples, as well as from the GMB151 soybean samples of the T_2 and BC_2F_3 generations. Two different channels were used to sequence the library pools from the GMB151 soybean samples of the T_4, T_5, and T_6 generations.

7.8. Bioinformatics

Primary data processing and QC

The initial data analysis started directly on the HiSeq 2500 System during the run. The HiSeq control software and RTA software performed image processing and base calling. The software bcl2fastq generated and reported run statistics and the final FASTQ files comprising the sequence information, which were used in all subsequent bioinformatics analysis. Data quality controls were done before subsequent analysis. The Q-Score is a prediction of the probability of a wrong base call. The % Q30 value represents the percentage of bases with a quality score of at least 30. % Q30 scores of > 80 are characteristic for a good sequencing run, and the corresponding base calls are highly valid. % Q30 scores of > 80 were obtained for all sequencing runs used in this study.

Bioinformatic analysis of sequencing reads

Read data were analyzed with the Eurofins Genomics Junction-Sequence-Analysis software pipeline (JSA pipeline) and by the Eurofins Genomics Variant Calling software. The JSA pipeline is an integrated and automated system that consists of different computational analysis steps. The JSA pipeline was subjected to a user acceptance test using a comprehensive validation data set and was demonstrated robust for the molecular characterization of genetically modified plants. To maximize both specificity and sensitivity of the JSA, critical parameters were identified, and the optimal parameter settings were defined. Prior to use in this study, the JSA pipeline was validated according to the GAMP5 principles. A scheme of the bioinformatics analysis of the sequencing reads in the JSA pipeline and the output of each step in the analysis are provided in Figure A2.2.

Adaptor and quality trimming

Low quality read ends with Phred scores of 12 or lower were trimmed. Illumina sequencing adapters in reads were removed using software Trimmomatic. Trimmed reads of 30 bp or longer and only paired reads were accepted for further analysis.

Mapping of paired reads to the transforming plasmid and the soybean genome reference sequences

Paired reads were mapped to the transforming plasmid pSZ8832 and to the soybean reference genome (*G. max* Wm82.a2.v1) using software BWA-MEM. Although the soybean varieties used in this study are of another genetic background than the soybean reference genome used in this study, the latter one is the most complete and accurate reference genome currently available. In addition, validation of the JSA pipeline demonstrated that the soybean reference genome used was suitable for a comprehensive JSA.

Removal of duplicate reads

Duplicate reads that were the result of a PCR amplification step during library preparation were removed using software PICARD. The resulting de-replicated reads were used for further coverage, soft-clipped read, and broken pair analysis.

Sequence depth analysis

The effective sequencing depth of the soybean *lectin* gene (Glyma.02G012600, Chr02:1123507–1125658) was calculated to represent the genome coverage at the insertion locus. Validation of the pipeline showed that an effective genomic sequencing depth of 75-fold (75X) resulted in a highly robust JSA. Plasmid coverage was assessed by calculation of the effective sequencing depth of the transforming plasmid.

Soft-clipped read analysis

Soft-clipped reads are sequencing reads of which only a subsequence maps to the target sequence (partially mapping reads). The soft-clipping position separates the mapped subsequence from the non-mapped subsequence and corresponds to a breakpoint in the target sequence (Wang et al., 2011). Sequence reads that were identified as soft-clipped reads following mapping to the transforming plasmid reference sequence were used to identify breakpoints in the transforming plasmid sequence with base pair resolution.

- Soft-clipped reads partially mapping to the transformation plasmid sequence, and their mates, were selected and processed using customized Perl scripts as part of the JSA pipeline. Based on the results of the parameter optimization of the JSA pipeline, the minimum length of the mapped subsequence was set to 30 bp. Hence, insertions of plasmid sequences smaller than 30 bp were below the detection limit.

- The breakpoints in the transforming plasmid were identified using customized Perl scripts as part of the JSA pipeline.

- Soft-clipped reads that aligned to the same breakpoint in the transforming plasmid (“supporting reads”) were grouped, aligned, and assembled to a consensus junction sequence. Validation of the JSA pipeline showed that low support junctions with less than ten supporting reads were false positive junctions. Hence, low support junctions were rejected.

Broken pair analysis

Mate-pairs for which only one mate mapped to the target sequence or for which both mates mapped to the target sequence with unexpected insert lengths or non-standard orientation were considered as broken pairs. Mate-pairs that were identified as broken pairs following mapping to the transforming plasmid were used to identify transforming plasmid regions in which a breakpoint was expected. While soft-clipped read analysis may identify breakpoint positions with base pair resolution, ambiguous alignments of soft-clipped reads in homologous regions may result in ambiguities in breakpoint positions. Broken pair alignment information was used to resolve ambiguities in breakpoint positions since the breakpoint positions of the true junctions are located within the junction regions defined by the broken pair analysis. Hence, soft-clipped read and broken pair analysis are complementary components of the JSA pipeline.

Background subtraction

Plasmid regions that were equally covered in both GMB151 soybean samples and in the non-GM counterpart samples were considered as background, and corresponding junctions were not taken into account to build the insert model of GMB151 soybean.

Identification and annotation of junctions

The junction consensus sequences were further annotated using BLAST analysis against the transforming plasmid and soybean genome reference sequences:

- Regions showing sequence identity to the soybean genome were annotated as flanking genomic sequence.
- Regions showing sequence identity to the plasmid pSZ8832 were annotated as plasmid (e.g., T-DNA).
- Bases in between plasmid and flanking genomic sequences were annotated as filler DNA.
- Bases at the plasmid/flank junction showing sequence identity to both plasmid and soybean genome were annotated as flanking genomic sequences.

Based on the BLAST results, junctions were identified as plasmid/flank or plasmid/plasmid.

Junction consensus sequence comparison

The consensus sequences corresponding to the same junction and obtained from different breeding generations were compared to each other using multiple alignments created by the Muscle software.

Validation of the sequencing technology

The reads that mapped to the transforming plasmid reference sequence in the positive and sensitivity controls were further subjected to a variant calling using the Eurofins Genomics Variant Calling software to check if the entire transforming plasmid sequence was amenable to sequencing with the technology used. If there was 100% sequence identity between the obtained transforming plasmid sequence in the positive and sensitivity controls and the transforming plasmid reference sequence, the sequencing technology was validated.

7.9. Statistics

The obtained data were qualitative. Therefore, statistical analyses were not required.

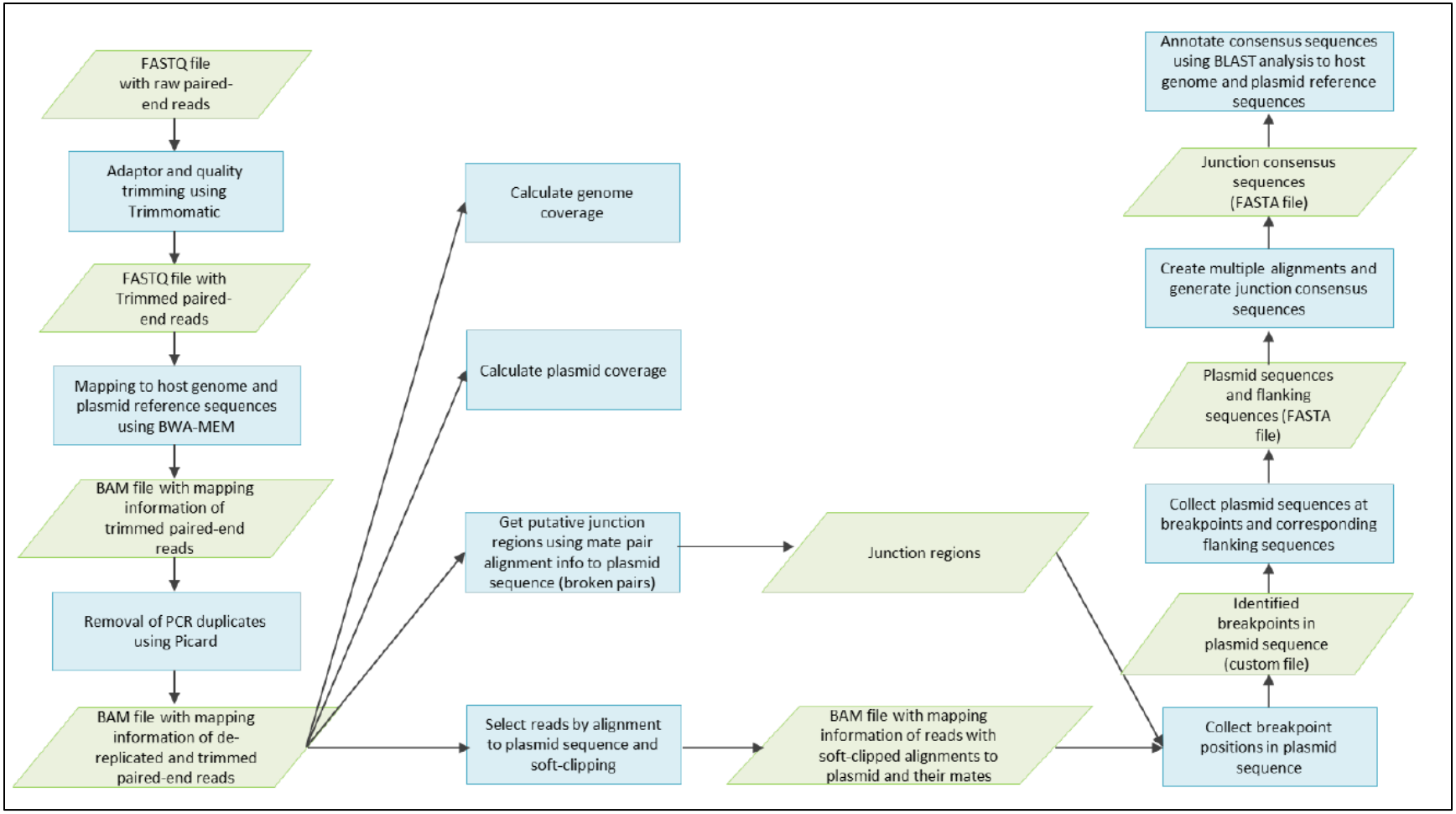


Figure A2.2. Scheme of the genomics junction sequence analysis

Each computational step of the analysis is indicated by rectangles. The output file of each step is indicated by parallelograms.

8. Results

The sequencing yield obtained from whole-genome NGS for each of the aforementioned samples is shown in Table A2.1.

To identify their original location, the whole-genome sequence data were mapped to both the soybean genome and transforming plasmid reference sequences. PCR duplicate reads were removed, and the effective median genome coverage was examined by the alignment of the paired-end reads to a known single copy locus of the soybean genome (*lectin* gene, Glyma.02G012600, Chr02:1123507–1125658). The effective median coverage depth of a known single copy locus represented the genome coverage at the insertion locus and was observed higher than 75-fold for each of the GMB151 soybean and control samples (Table A2.1).

Visual representation of the alignment of paired-end reads to the transforming plasmid reference sequence demonstrated that the effective median coverage depth of the transforming plasmid in the positive and sensitivity control samples were 96-fold and 7-fold, respectively, and that 100% of the transforming plasmid sequence was covered. Sequence depth analysis of the transforming plasmid in the positive and sensitivity control samples demonstrated that the coverage was adequate to perform a high quality JSA (Figure A2.3 and Figure A2.4). Variant calling on the combined sequence reads of positive and sensitivity controls indicated a 100% nucleotide identity between the obtained transforming plasmid sequences and the transforming plasmid reference sequence. This demonstrated that the entire transforming plasmid sequence was amenable to sequencing with the technology used.

Visual representation of the alignment of paired-end reads to the transforming plasmid reference sequence demonstrated that the effective median coverage depth of the negative control samples (non-GM counterparts Thorne and Parent 5) was 0-fold (Figure A2.5).

Table A2.1. Sequencing yield and coverage depth of the soybean genome

Sample	Number of channels sequenced	Total number of read-pairs sequenced	Total sequencing yield (Mb)	Effective median coverage depth (x-fold)
GMB151 soybean – T_2	3	546,852,393	136,713.09	115
GMB151 soybean – T_4	2	414,522,683	103,630.67	77
GMB151 soybean – T_5	2	428,545,152	107,136.28	87
GMB151 soybean – T_6	2	440,746,078	110,186.51	88
GMB151 soybean – BC_2F_3	3	689,557,786	172,389.44	127
Non-GM counterpart Parent 5	3	660,867,494	165,216.87	123
Non-GM counterpart Thorne	3	695,723,139	173,930.78	146
Non-GM counterpart Thorne + equimolar amount of pSZ8832	3	615,298,015	153,824.50	123
Non-GM counterpart Thorne + 1/10$_{th}$ equimolar amount of pSZ8832	3	534,045,912	133,511.47	111

Effective median coverage depth is examined by alignment and de-replication of quality-clipped paired-end reads to a known single copy locus within the soybean genome (*lectin* gene, Glyma.02G012600, Chr02:1123507–1125658).

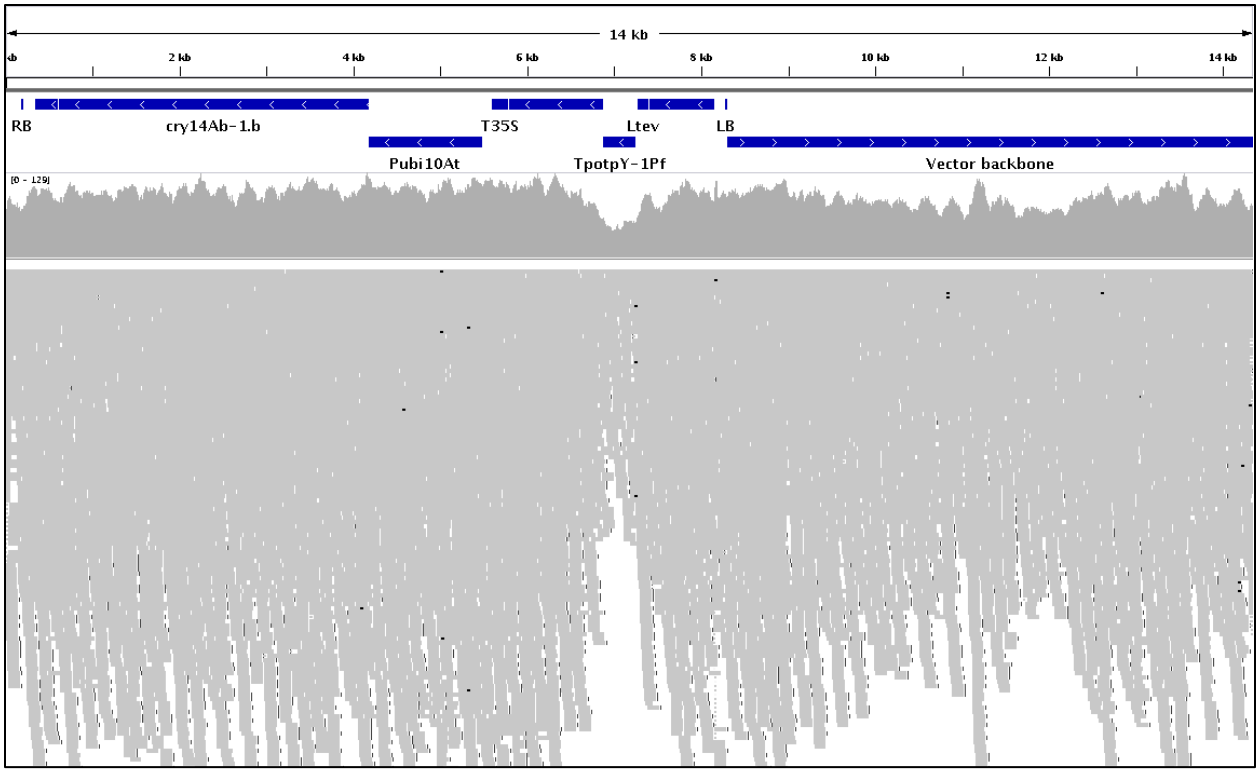


Figure A2.3. Mapping of trimmed and de-replicated paired reads of the positive control sample of pSZ8832

Top panel: pSZ8832 (range 1–14361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–129x) and aligned reads.

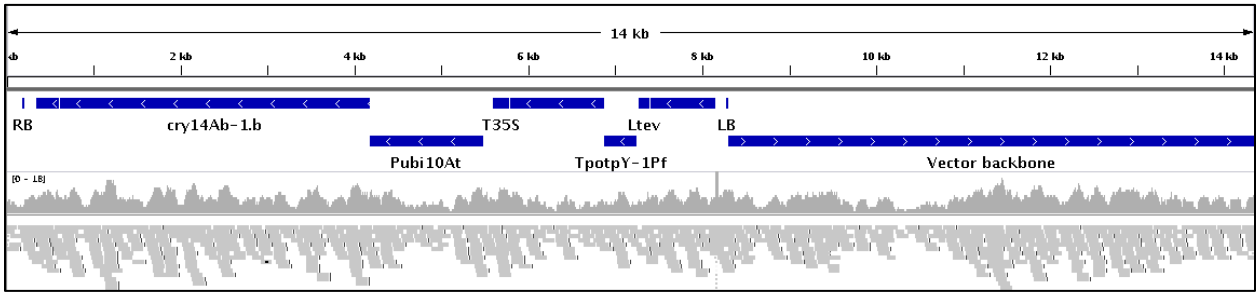


Figure A2.4. Mapping of trimmed and de-replicated paired reads of the sensitivity control sample to pSZ8832

Top panel: pSZ8832 (range 1–14361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–15x) and aligned reads.

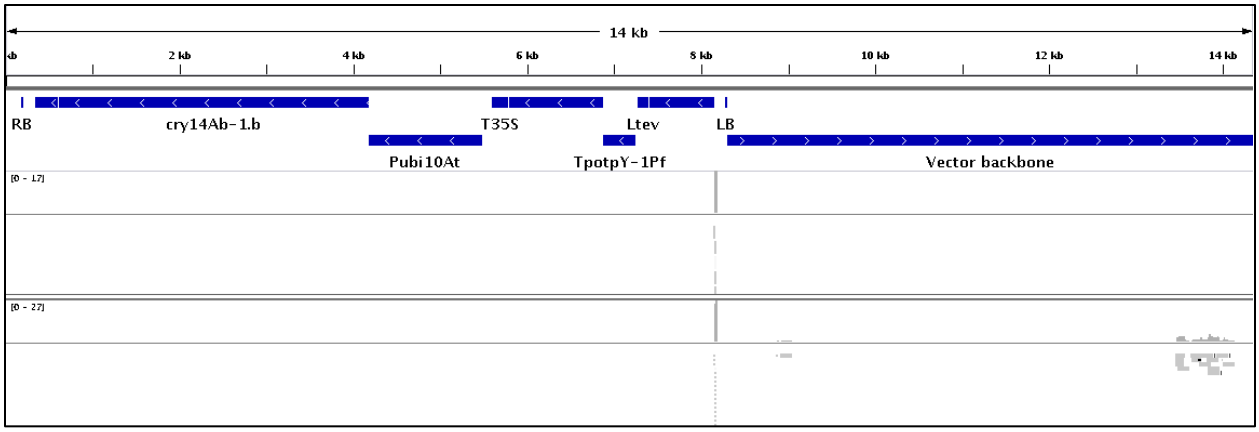


Figure A2.5. Mapping of trimmed and de-replicated paired reads of the negative control samples to pSZ8832

Top panel: pSZ8832 (range 1–14,361 bp). Some annotations were omitted to increase readability.
Center panel: coverage plot (range 0–17x) and aligned reads for the non-GM counterpart Thorne.
Bottom panel: coverage plot (range 0–27x) and aligned reads for the non-GM counterpart Parent 5.
The few reads that did map were partial or low quality alignments and the result of sequence similarities between the transforming plasmid and the soybean genome.

8.1. GMB151 soybean contains a single transgenic locus that consists of a single T-DNA copy

Mapping of paired-end reads to the transforming plasmid reference sequence demonstrated that the entire T-DNA, except for the 5' end of the P2x35S promoter, was inserted and comprehensively covered. The absence of vector backbone in GMB151 soybean was indicated since none of the vector backbone sequences were adequately covered (Figure A2.6 to Figure A2.10).

JSA on the GMB151 soybean samples (T_2, T_4, T_5, T_6, and BC_2F_3) identified two plasmid/flank junctions (NJ-A and NJ-B) and no plasmid/plasmid junctions (Table A2.2). As none of the plasmid/flank junctions were found in the non-GM counterparts, the specificity of the analysis was demonstrated.

Alignment of soft-clipped reads and mates in relation to the transforming plasmid sequence demonstrated that NJ-A (breakpoint position pSZ8832: 213) originated from the 5’ end of the T-DNA sequence and that NJ-B (breakpoint position pSZ8832: 7673) originated from the P2x35S promoter near the 3’ end of the T-DNA sequence (Figure A2.11 to Figure A2.15). Since none of the soft-clipped reads aligned to the vector backbone region of the transforming plasmid reference sequence, the absence of vector backbone in GMB151 soybean was confirmed.

The presence of a repetitive region within the P2x35S promoter of pSZ8832 resulted in mapping ambiguities of some soft-clipped reads, visible as an additional small peak in the coverage plot (pSZ8832: +/- 8kb, see footnote 1 in Figure A2.6 to Figure A2.15). Mapping ambiguities of these soft-clipped reads resulted in two putative breakpoint positions (pSZ8832: 7673 and pSZ8832: 7999). The true breakpoint position (pSZ8832: 7673) was determined using broken pair analysis as part of the JSA.

BLAST analysis of the junction consensus sequences against the transforming plasmid and host genome reference sequences revealed the presence of 39 bp filler DNA between the 3' end of the inserted T-DNA region and the 3' flanking genomic region. Filler DNA showed 21 bp sequence identity to the ORIpVS1 of the transforming plasmid pSZ8832 and 17 bp sequence identity to the 3' flanking genomic region.

These results demonstrated that GMB151 soybean contains a single transgenic locus that consists of a single T-DNA copy lacking the 5' part of the P2x35S promoter (Figure A2.16).

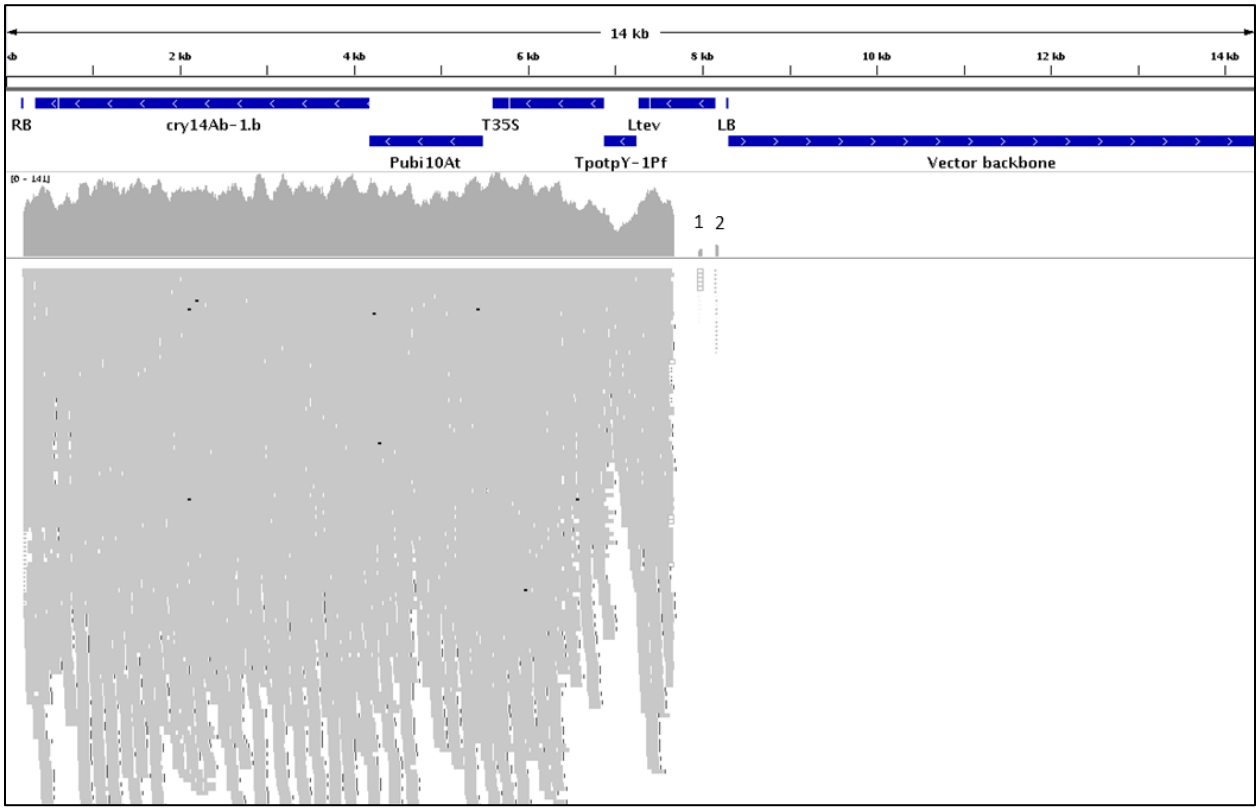


Figure A2.6. Mapping of trimmed and de-replicated paired reads of GMB151 soybean (T_2 generation) to pSZ8832

Top panel: pSZ8832 (range 1–14,361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–141x) and aligned reads (soft-clipped reads and their mates are shown in Figure A2.11).
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

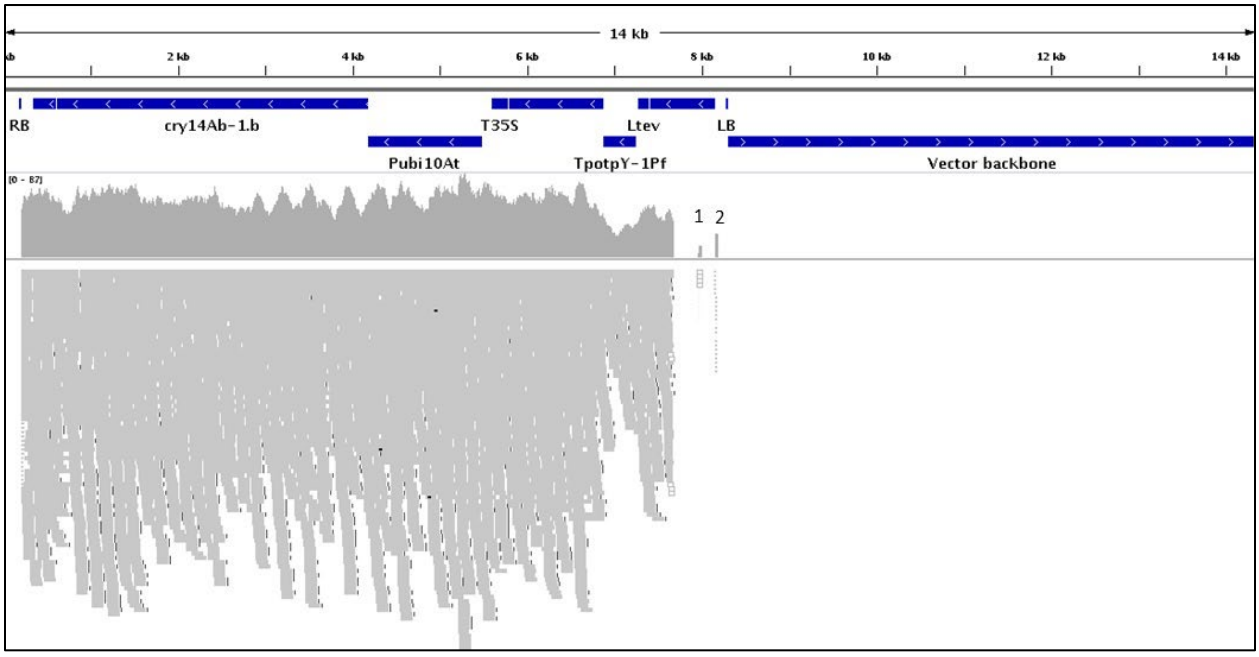


Figure A2.7. Mapping of trimmed and de-replicated paired reads of GMB151 soybean (T_4 generation) to pSZ8832

Top panel: pSZ8832 (range 1–14,361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–87x) and aligned reads (soft-clipped reads and their mates are shown in Figure A2.12).
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

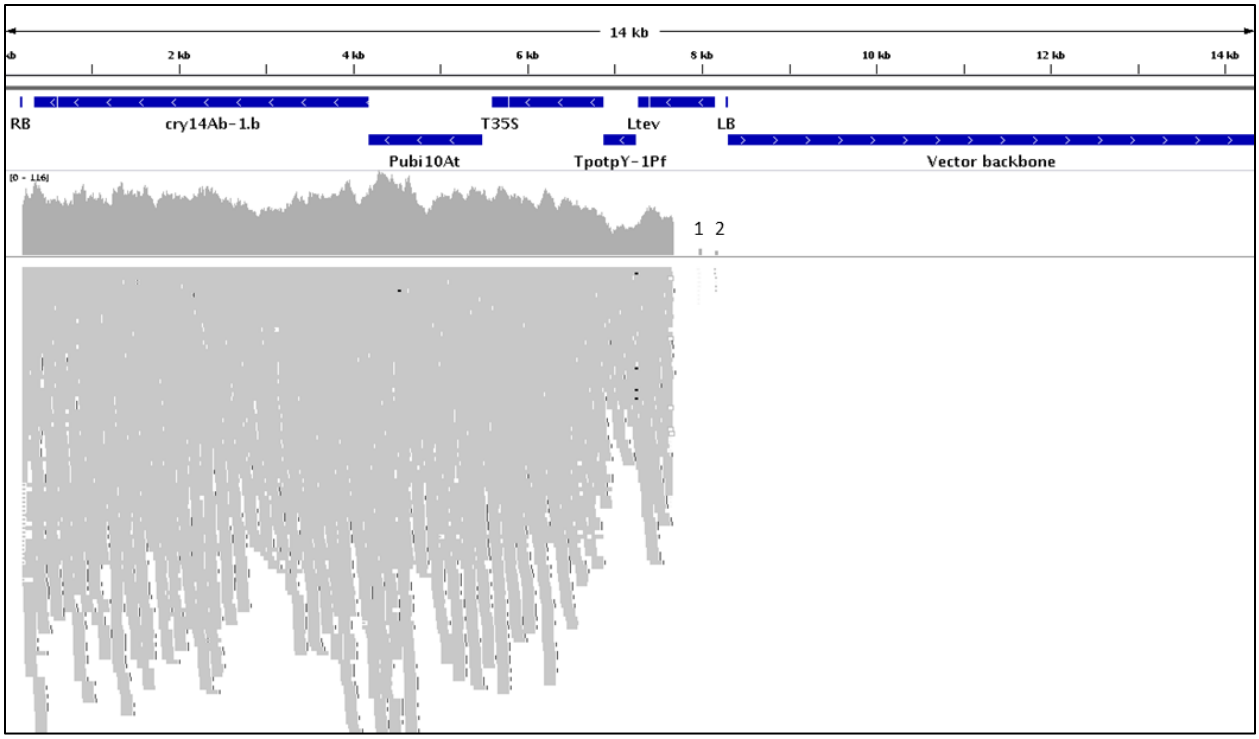


Figure A2.8. Mapping of trimmed and de-replicated paired reads of GMB151 soybean (T_5 generation) to pSZ8832

Top panel: pSZ8832 (range 1–14,361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–116x) and aligned reads (soft-clipped reads and their mates are shown in Figure A2.13).
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

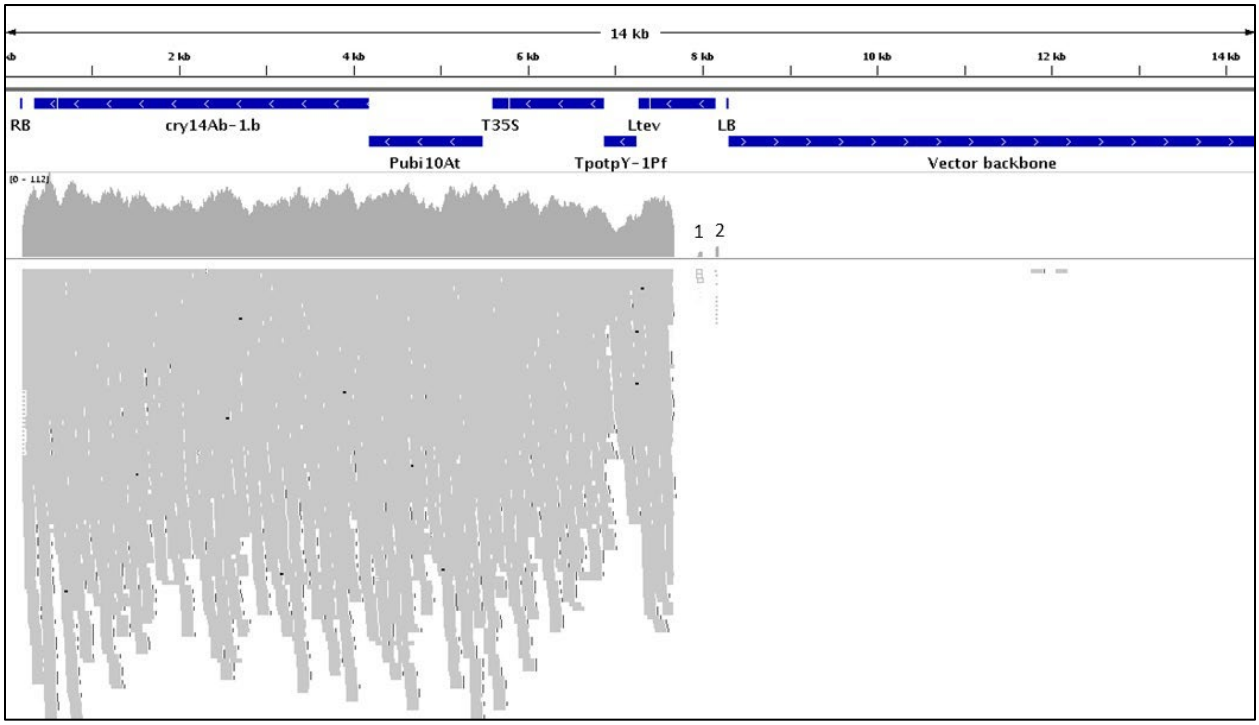


Figure A2.9. Mapping of trimmed and de-replicated paired reads of GMB151 soybean (T_6 generation) to pSZ8832

Top panel: pSZ8832 (range 1–14,361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–112x) and aligned reads (soft-clipped reads and their mates are shown in Figure A2.14).
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

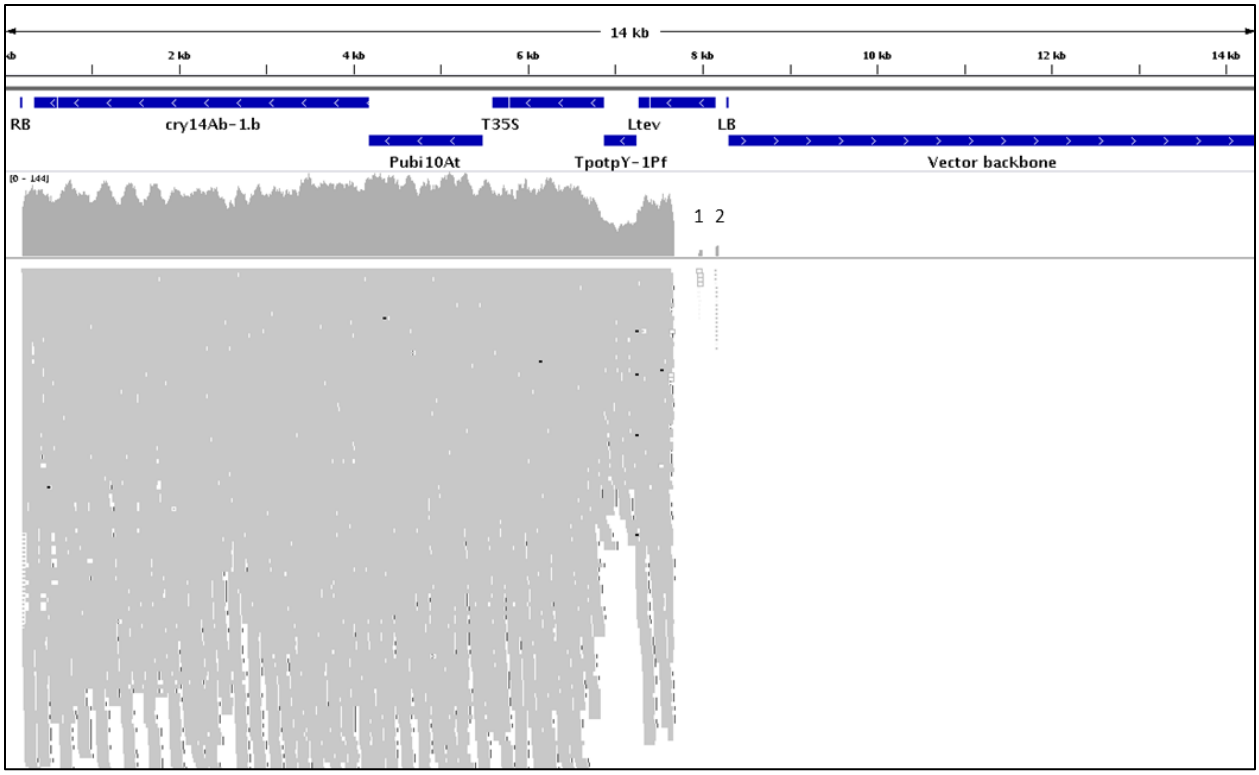


Figure A2.10. Mapping of trimmed and de-replicated paired reads of GMB151 soybean (BC_2F_3 generation) to pSZ8832

Top panel: pSZ8832 (range 1–14,361 bp). Some annotations were omitted to increase readability.
Bottom panel: coverage plot (range 0–144x) and aligned reads (soft-clipped reads and their mates are shown in Figure A2.15).
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

Table A2.2. The novel plasmid/flank and plasmid/plasmid junctions identified in GMB151 soybeans and the non-GM counterparts using JSA following background subtraction

Sample	Junction ID	Breakpoint position	Number of supporting soft-clipped reads	Number of supporting broken pairs	Junction type
GMB151 soybean – T_2	NJ-A	pSZ8832: 213	64	145	plasmid/flank
	NJ-B	pSZ8832: 7673	64	98	plasmid/flank
GMB151 soybean – T_4	NJ-A	pSZ8832: 213	45	101	plasmid/flank
	NJ-B	pSZ8832: 7673	36	45	plasmid/flank
GMB151 soybean – T_5	NJ-A	pSZ8832: 213	69	124	plasmid/flank
	NJ-B	pSZ8832: 7673	40	68	plasmid/flank
GMB151 soybean – T_6	NJ-A	pSZ8832: 213	41	91	plasmid/flank
	NJ-B	pSZ8832: 7673	37	59	plasmid/flank
GMB151 soybean – BC_2F_3	NJ-A	pSZ8832: 213	79	161	plasmid/flank
	NJ-B	pSZ8832: 7673	60	103	plasmid/flank
Non-GM counterpart Thorne	-	-	-	-	-
Non-GM counterpart Parent 5	-	-	-	-	-

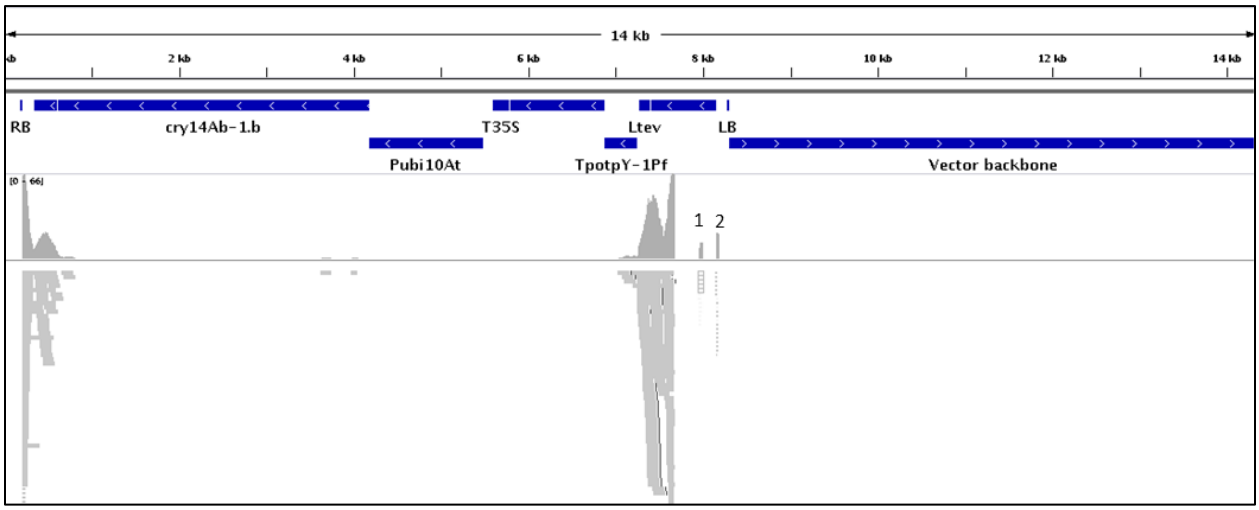


Figure A2.11. Alignment of soft-clipped reads and mates to pSZ8832 for GMB151 soybean (T_2 generation)

Left: NJ-A; Center: NJ-B
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

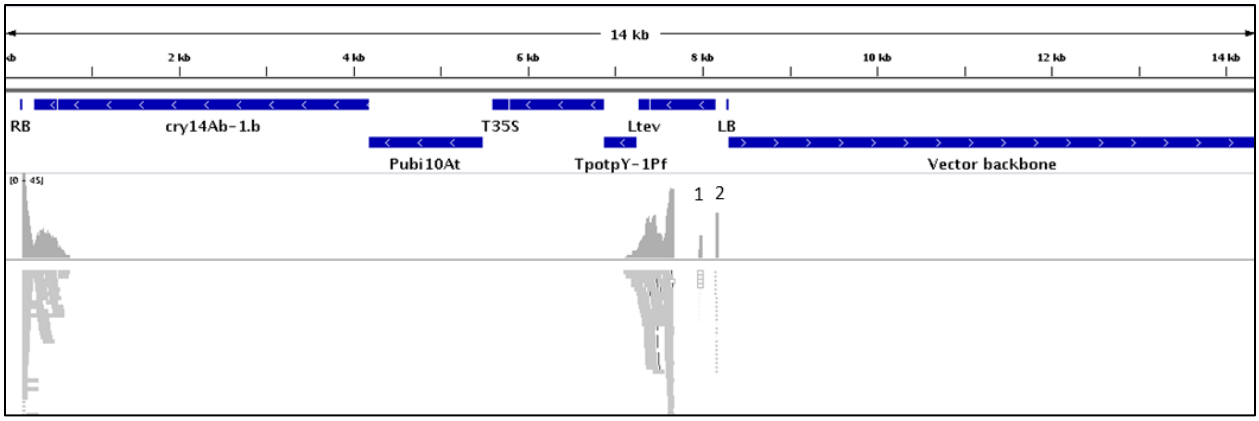


Figure A2.12. Alignment of soft-clipped reads and mates to pSZ8832 for GMB151 soybean (T_4 generation)

Left: NJ-A; Center: NJ-B
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

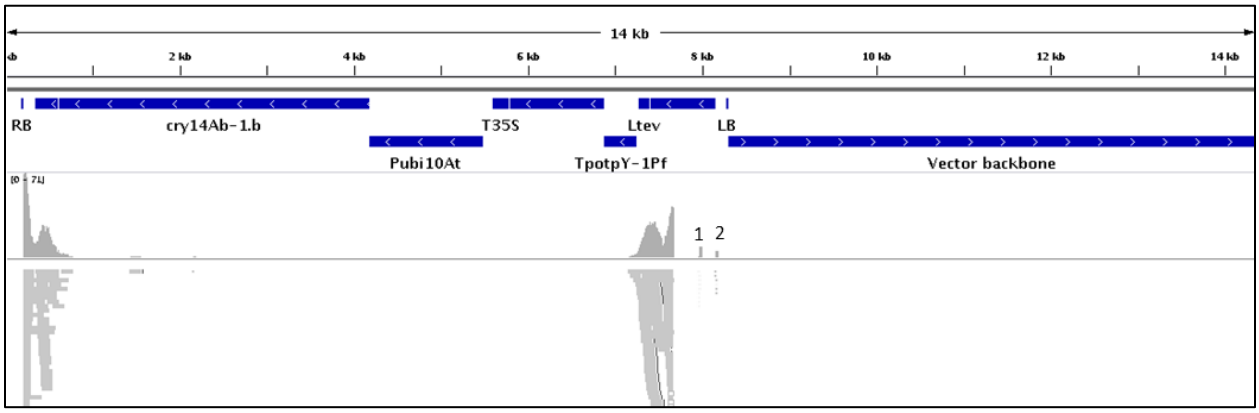


Figure A2.13. Alignment of soft-clipped reads and mates to pSZ8832 for GMB151 soybean (T_5 generation)

Left: NJ-A; Center: NJ-B
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

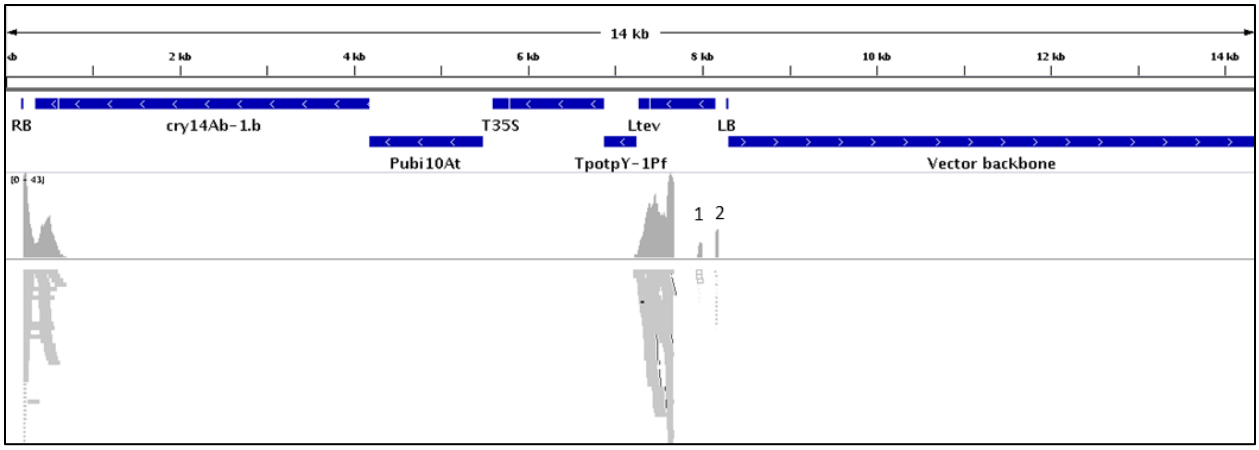


Figure A2.14. Alignment of soft-clipped reads and mates to pSZ8832 for GMB151 soybean (T_6 generation)

Left: NJ-A; Center: NJ-B
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

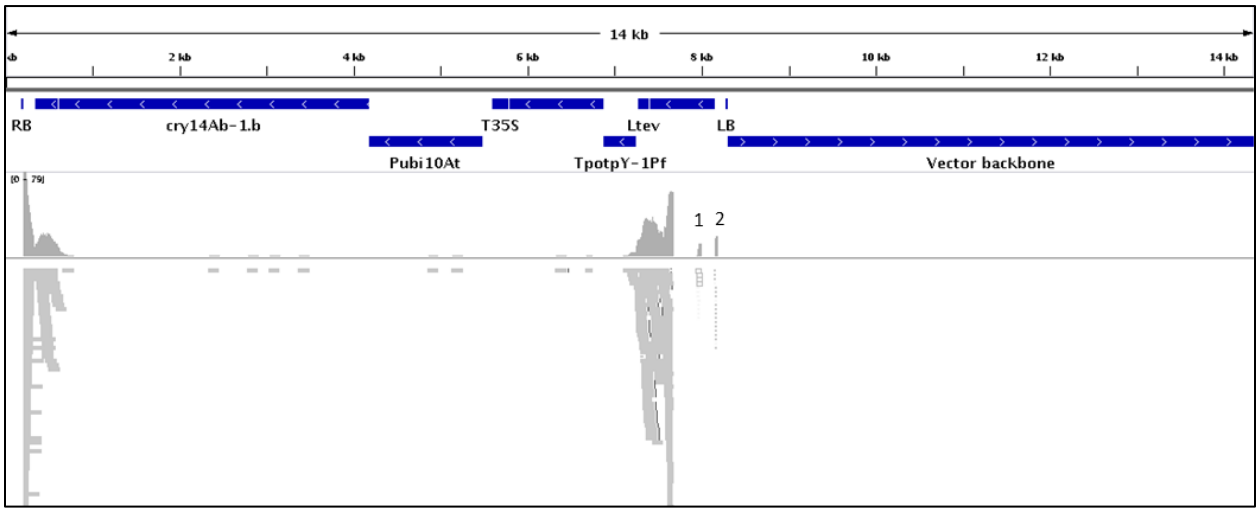


Figure A2.15. Alignment of soft-clipped reads and mates to pSZ8832 for GMB151 soybean (BC_2F_3 generation)

Left: NJ-A; Center: NJ-B
[1] Mapping ambiguities of soft-clipped reads
[2] Sequence similarities with endogenous soybean

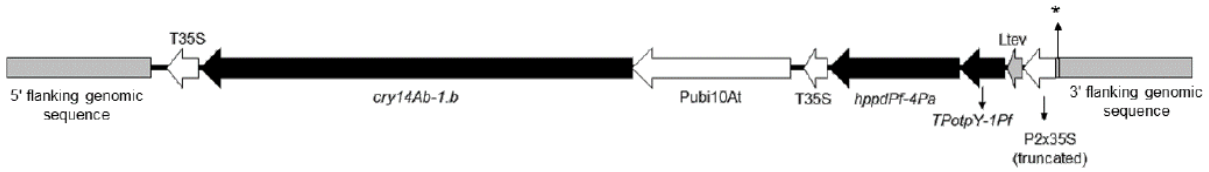


Figure A2.16. Schematic drawing of the GMB151 soybean insert model

* GMB151 soybean contains 39 bp filler DNA between the 3' end of the inserted T-DNA region and the 3' flanking genomic region. Filler DNA showed 21 bp sequence identity to the ORIpVS1 of the transforming plasmid pSZ8832 and 17 bp sequence identity to the 3' flanking genomic region.

8.2. The transgenic locus of GMB151 soybean is stably inherited across different breeding generations

Multiple sequence alignment between the obtained junction consensus sequences showed that NJ-A and NJ-B were conserved across all breeding generations analyzed (Table A2.2). This demonstrated that the transgenic locus of GMB151 soybean is stably maintained across different breeding generations.

Appendix 2 References

CLSI, 2014. MM09-A2. Nucleic acid sequencing methods in diagnostic laboratory medicine; Approved guideline—Second edition. Clinical and Laboratory Standards Institute. Vol. 34 No. 4. 15 pp.

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC and Jackson SA, 2010. Genome sequence of the palaeopolyploid soybean. Nature 463, 178–183.

Wang J, Mullighan CG, Easton J, Roberts S, Heatley SL, Ma J, Rusch MC, Chen K, Harris CC, Ding L, Holmfeldt L, Payne-Turner D, Fan X, Wei L, Zhao D, Obenauer JC, Naeve C, Mardis ER, Wilson RK, Downing JR and Zhang J, 2011. CREST maps somatic structural variation in cancer genomes with base-pair resolution. Nature Methods 8, 652–654.

Appendix 3. Background information on activity spectrum of Cry14Ab-1

List of Tables

List of Figures

1. Introduction

Plant-parasitic nematodes cause serious crop losses worldwide and are among the most important agricultural pests (Koenning et al., 1999). The control of nematodes is more difficult than other pests because nematodes mostly inhabit the soil and usually attack the underground parts of the plants (Stirling, 1991).

Bacteria are numerically the most abundant organisms in soil, and some of them, for example members of the genus *Bacillus* (Emmert and Handelsman, 1999), have shown potential for the biological control of nematodes.

This appendix summarizes the discovery and spectrum of activity against agricultural pests of Cry14Ab-1.

1.1. The gene and its product, history of discovery or isolation

Bt Cry proteins are pore-forming proteins used as biological pest control around the world. The activity of the bacterium was attributed to the proteins residing in the crystal that is produced upon sporulation.

Widespread screening of *Bt* strains and Cry protein gene sequencing has led to the identification of more than 700 gene sequences (Crickmore et al., 1998). These sequences were categorized based on their amino acid sequence similarity in at least 70 different cry gene groups, among which Cry5, Cry6, Cry12, Cry13, Cry14, and Cry21 displayed nematicidal activity (Bravo et al., 2013). The Cry proteins require specific conditions (e.g., specific proteases and receptors), resulting in a narrow host range. Several of the receptors for Cry proteins, including nematicidal Cry, have been characterized (Griffitts et al*.,* 2003; Griffitts and Aroian, 2005).

The Cry14Ab-1 protein was identified as having activity against the nematode *C. elegans* in a bioassay.

1.2. The use of the gene and its product in the target species

The c*ry14Ab-1.b* gene coding for the *Bt* native protein was optimized for plant expression and introduced in plants by *A. tumefaciens*-mediated transformation. The subsequent transformed plant, GMB151 soybean, expressing Cry14Ab-1 was found to be more resistant to the sedentary parasitic SCN than non-transformed plants.

2. Functional characterization of the Cry14Ab-1 protein

2.1. Description of the nematicidal activity

Cry proteins are recognized as pore-forming proteins that, upon ingestion by the susceptible organism, destroy epithelium midgut cells by causing an osmotic shock, leading to cell lysis.

The nematode-active Cry5B protein has been shown to damage the *C. elegans* intestine. Nematodes fed the Cry5B toxin showed vacuole-like structures, pitting, constriction, and degeneration of the intestine after two days (Marroquin et al., 2000). The Cry14A protein has also been shown to damage the gut of *C. elegans* (Wei et al., 2003).

2.2. Activity spectrum of the Cry14Ab-1 protein

Cry14Ab-1 was tested across several agriculturally relevant pest species and shown to be only active against nematodes. Cry14Ab-1 was tested against the indicated pests in their respective standardized *in vitro* pest screening assays. As summarized in Table A3.1, Cry14Ab-1 was highly active against the nematode *C. elegans*, but no activity was detected against any of the other pests.

Cry14Ab-1 has been shown to be active against nematodes using a *C. elegans* standard *in vitro* screen. No assay is currently available to test Cry14Ab-1 against plant-parasitic nematodes (PPNs) *in vitro*. However, activity against PPNs was shown *in planta*. Soybean events expressing Cry14Ab-1 protein show activity against the sedentary PPN *H. glycines* (SCN) (Figure A3.1 and Figure A3.2) and the migratory lesion nematode *Pratylenchus brachyurus* (Figure A3.3 and Figure A3.4). Both nematodes are of high agricultural importance in soybean, with SCN causing the highest losses to growers in the U.S. Midwest while both SCN and lesion nematodes cause significant crop losses in Brazil.

Efficacy against other PPNs has also been explored as part of product development. A field evaluation of GMB151 soybean efficacy against spiral nematode (*Helicotylenchus* spp.) was conducted. GMB151 soybean did not provide efficacy against spiral nematodes relative to the nullizygous comparator (Figure A3.5).

Experimental methods for activity spectrum of Cry14Ab-1

The research group at BASF conducted laboratory tests designed to evaluate the continuous exposure of each test species to the Cry14Ab-1 protein. Each species followed a different bioassay designed to evaluate activity for its pest group. Each taxonomic group had a different end point, depending on the assay. Species were exposed to the protein through feeding on treated diet or exposed in the media, depending on bioassay protocol.

Order Lepidoptera

Experimental Design: Neonates (less than 24 h old) were exposed to Cry14Ab-1 protein applied topically to diet in a well plate format. Assays were terminated after 3–5 days, depending on the species.

Insects: Eggs were obtained from a BASF internal insectary and a commercial vendor. Both groups utilize well established rearing methods to culture adults for egg production. After the eggs were received, the different species were kept in different environmental conditions to optimize hatching for the assays.

Assay: Cry14Ab-1 was topically applied to the diet and allowed to dry. Eggs or neonates were then transferred to each well. Eggs hatched within 24 h. Each experiment was repeated at least four times. Positive (commercial protein that controls Lepidoptera) and negative (buffer alone) controls were also run at the same time. Assays were kept in an environmental chamber between 3–5 days at 25°C, 85% relative humidity, and in darkness. Insects were evaluated for stunting and mortality.

Western corn rootworm (Diabrotica virgifera)

Experimental Design: First instar larvae were exposed to Cry14Ab-1 protein applied topically to diet in a well plate format. Assays were terminated after 5 days.

Insects: Eggs were obtained commercially. This company utilizes well established rearing methods to culture adults for egg production. After the eggs were received, they were kept in environmental chambers at 25°C, 85% relative humidity, and in darkness.

Assay: Cry14Ab-1 was applied topically to the diet and allowed to dry. Eggs were then transferred to each well. Eggs hatched within 24 h. Each experiment was repeated at least four times. Positive (commercial protein that controls western corn rootworm) and negative (buffer alone) controls were also run at the same time. Assays were kept in an environmental chamber for 5 days at 25°C, 85% relative humidity, and in darkness. Insects were evaluated for stunting and mortality.

Southern green stink bug (Nezara viridula)

Experimental Design: Second instar nymphs were exposed to Cry14Ab-1 protein mixed with a liquid diet into hand-made sachets. Assays were terminated after 7 days.

Insects: Eggs were obtained from a BASF internal insectary. The insectary utilizes well established rearing methods to culture adults for egg production. After the eggs were received for assays, they were kept in environmental chambers at 23°C, 65% relative humidity, and a photoperiod cycle of 14 h light: 10 h darkness.

Assay: Cry14Ab-1 was mixed with a liquid diet in hand-made sachets. After that, second instar nymphs were carefully transferred with a paint brush. Each experiment was repeated at least four times. Positive (internal protein that controls southern green stink bug) and negative (buffer alone) controls were also run at the same time. Assays were kept in environmental chamber for 7 days at 23°C, 65% relative humidity, and a photoperiod cycle of 14 h light: 10 h darkness. Development to third instar and mortality were evaluated.

Fungi

Experimental Design: Fungi inoculum broths were exposed to Cry14Ab-1 protein mixed with a liquid media in a 96-well plate format. Assays were terminated after 3–7 days.

Fungi: Fungi species were obtained internally, and well established methods to culture fungi were utilized.

Assay: Cry14Ab-1 was mixed with fungi inoculum broth and media in 96-well plates. Each experiment was repeated at least two times. Positive (commercial fungicide) and negative (buffer alone) controls were also run at the same time. Assays were kept in an environmental chamber for 3–7 days at 20°C, > 85% relative humidity, and a photoperiod of 24 h darkness. OD levels were measured, and percent inhibition of fungal growth was calculated and compared to the media-alone negative control.

Table A3.1. Cry14Ab-1 qualitative activity on a panel of *in vitro* pest assays

Common name	Scientific name	Taxonomy	Concentration Tested	Activity
Beet armyworm	*Spodoptera exigua*	Lepidoptera	100 µg/ml diet	No
Black cutworm	*Agrotis ipsilon*	Lepidoptera	100 µg/ml diet	No
Diamondback moth	*Plutella xylostella*	Lepidoptera	100 µg/ml diet	No
European corn borer	*Ostrinia nubilalis*	Lepidoptera	100 µg/ml diet	No
Fall armyworm	*Spodoptera frugiperda*	Lepidoptera	100 µg/ml diet	No
Tobacco budworm	*Heliothis virescens*	Lepidoptera	100 µg/ml diet	No
Corn earworm	*Helicoverpa zea*	Lepidoptera	100 µg/ml diet	No
Soybean looper	*Pseudoplusia includens*	Lepidoptera	100 µg/ml diet	No
Sugarcane borer	*Diatraea saccharalis*	Lepidoptera	100 µg/ml diet	No
Southwestern corn borer	*Diatraea grandiosella*	Lepidoptera	100 µg/ml diet	No
Velvetbean caterpillar	*Anticarsia gemmatalis*	Lepidoptera	100 µg/ml diet	No
Western corn rootworm	*Diabrotica virgifera*	Coleoptera	100 µg/ml diet	No
Southern green stink bug	*Nezara viridula*	Hemiptera	400 µg/ml diet	No
Roundworm	*Caenorhabditis elegans*	Nematoda	10–15 µg/ml	Yes
Rhizoctonia	*Rhizoctonia solani*	Fungi	250 µg/ml	No
Botrytis	*Botrytis cinerea*	Fungi	250 µg/ml	No
Alternaria	*Alternaria alternata*	Fungi	250 µg/ml	No

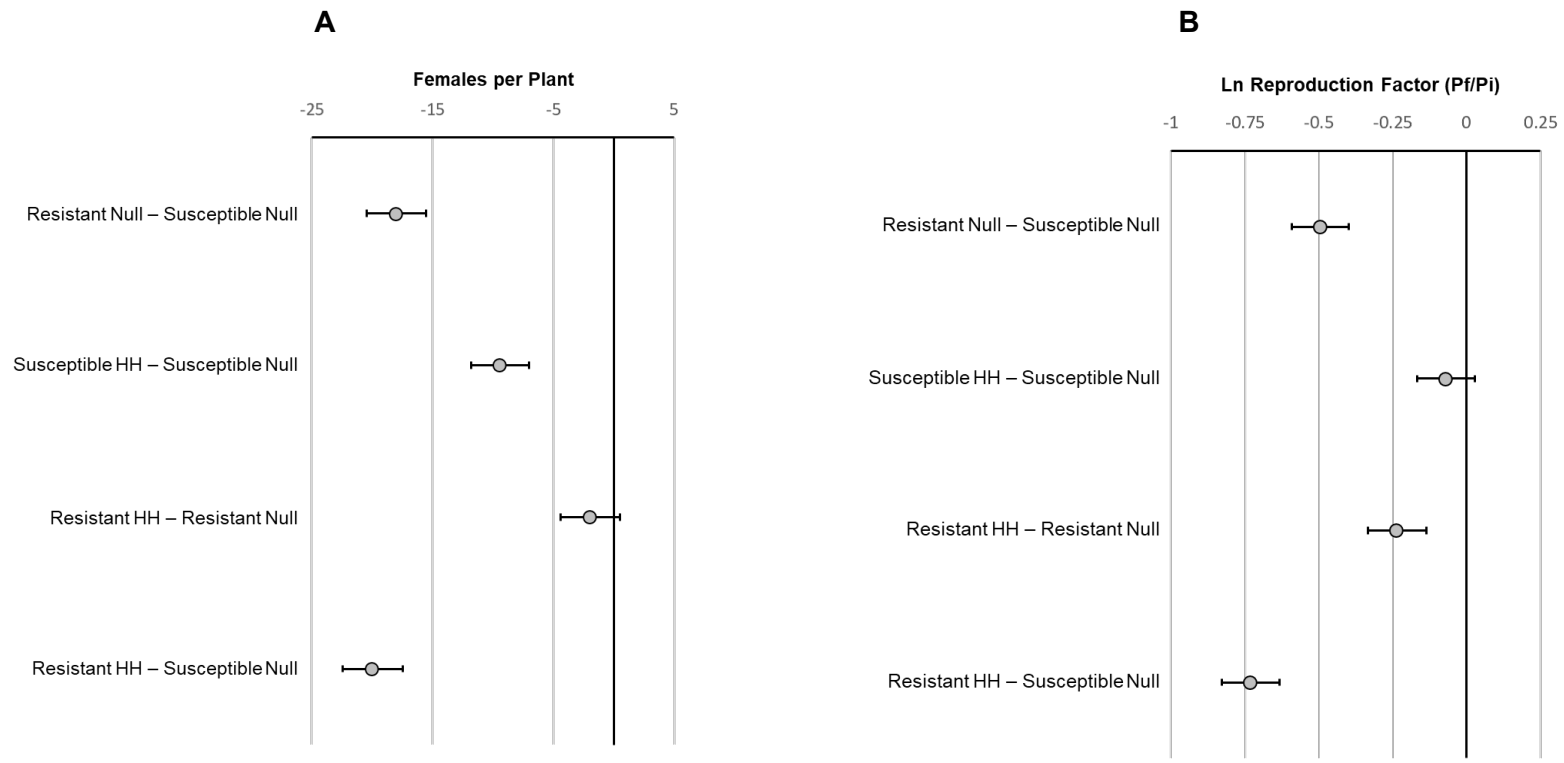


Figure A3.1. Soybean cyst nematode control provided by GMB151, native resistance, and their combination

SCN control was measured by (**A**) direct assessment of the number of females per root system after one generation, approximately 40–60 days after planting, and (**B**) assessing the change in population over the course of the season. A reproduction factor was calculated as the final SCN population density (Pf) divided by the initial SCN population density (Pi). The reproduction factor was natural log transformed for statistical analysis. Four experimental lines were tested within three different genetic backgrounds. The lines differed for the presence of the *rhg1b* allele conferring native resistance to SCN (Resistant or Susceptible) and the presence of the transgenic trait GMB151 conferring resistance to SCN (HH or Null).

A: The *rhg1b* allele reduced the number of SCN females per soybean plant by 63% early in the U.S. growing season. GMB151 soybean lowered the number of females per plant by 33% in the absence of the *rhg1b* allele. In the presence of the *rhg1b* allele, the presence of GMB151 soybean reduced the number of females by 11%, but this reduction was not statistically significant. The combination of *rhg1b* and GMB151 resulted in a total 70% reduction in SCN females per plant.

B: In the absence of *rhg1b* and GMB151, SCN populations on average increased more than six-fold during the course of the U.S. growing season, with an average end-of-season population density near 10,000 eggs/100 cc soil. Season-long SCN reproduction was dependent upon the presence of *rhg1b* by 39%. The effect of GMB151 on season-long reproduction was dependent upon the presence of *rhg1b*. In the presence of *rhg1b*, GMB151 reduced season-long reproduction by 21%. The combination of *rhg1b* and GMB151 reduced season-long SCN reproduction by 52%.

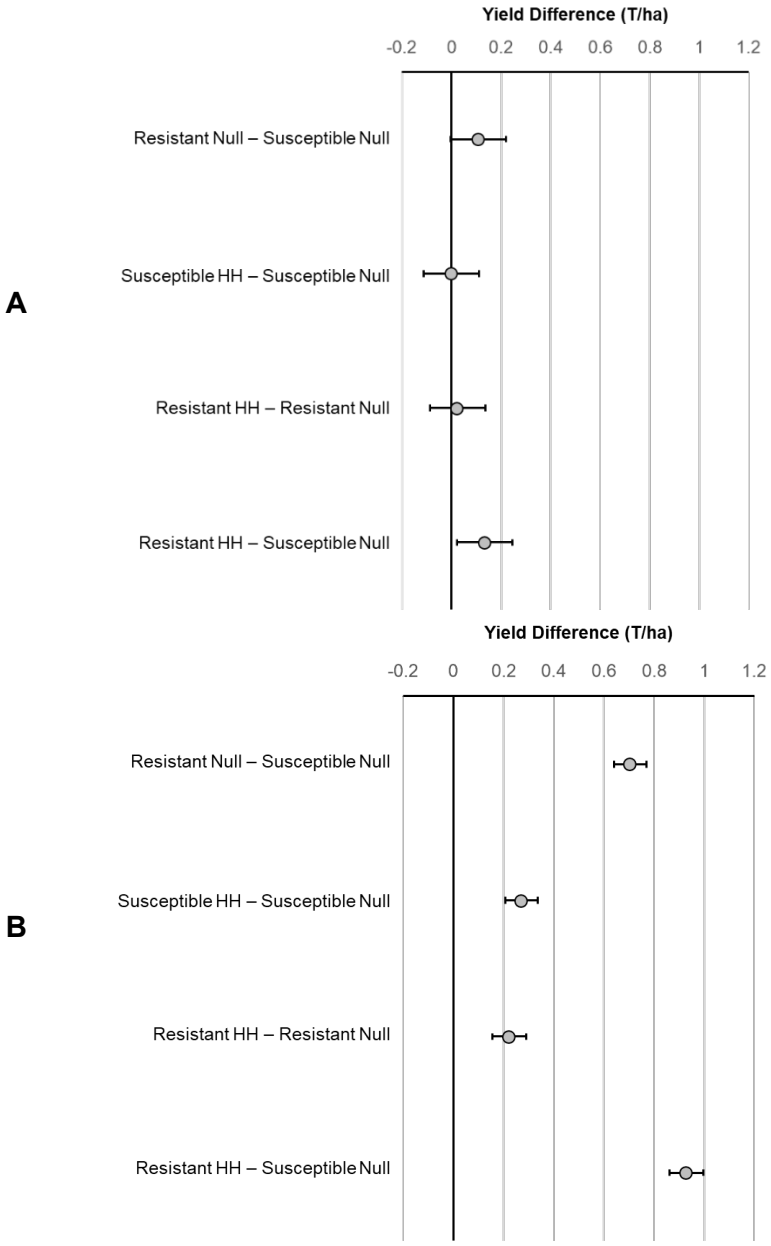


Figure A3.2. Yield protection from soybean cyst nematode injury provided by GMB151, native resistance, and their combination

Yield difference among the four experimental lines are depicted for the (**A**) SCN-free trial locations and the (**B**) SCN-infested trial locations. Four experimental lines were tested within three different genetic backgrounds. The lines differed for the presence of the *rhg1b* allele conferring native resistance to SCN (Resistant or Susceptible) and the presence of the transgenic trait GMB151 conferring resistance to SCN (HH or Null).

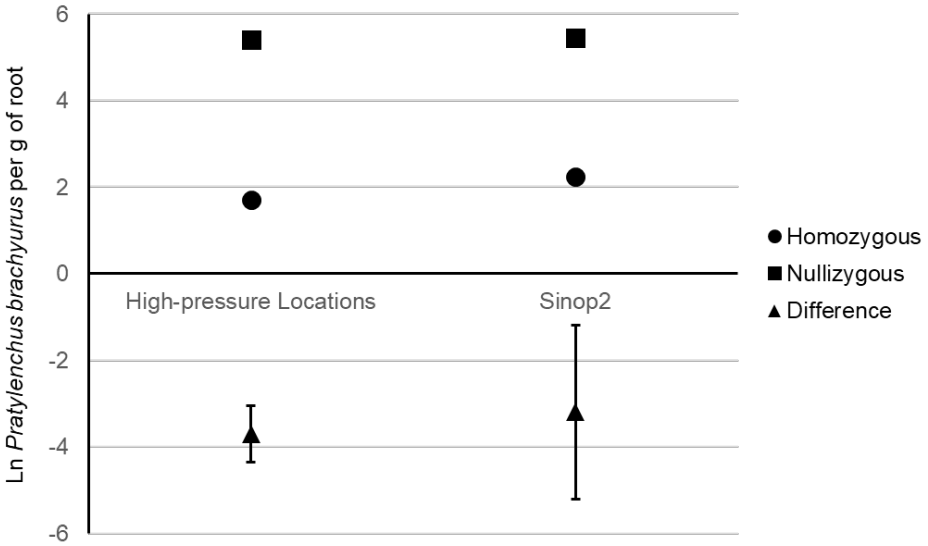


Figure A3.3. Effect of the GMB151 transgenic soybean trait on *Pratylenchus brachyurus* population densities

The population density of *P. brachyurus* was estimated for the GMB151 soybean homozygous and nullizygous lines at four locations in Brazil, the three anticipated high-pressure sites and the Sinop2 site that had higher *P. brachyurus* population densities than anticipated. The population density was natural log transformed for statistical analysis. The GBM151 transgenic soybean trait significantly reduced *P. brachyurus* populations. The 95% confidence error is represented by the error bars about the estimated difference between the homozygous and nullizygous lines. At the individual trial locations, average *P. brachyurus* reductions ranged from 90–97%.

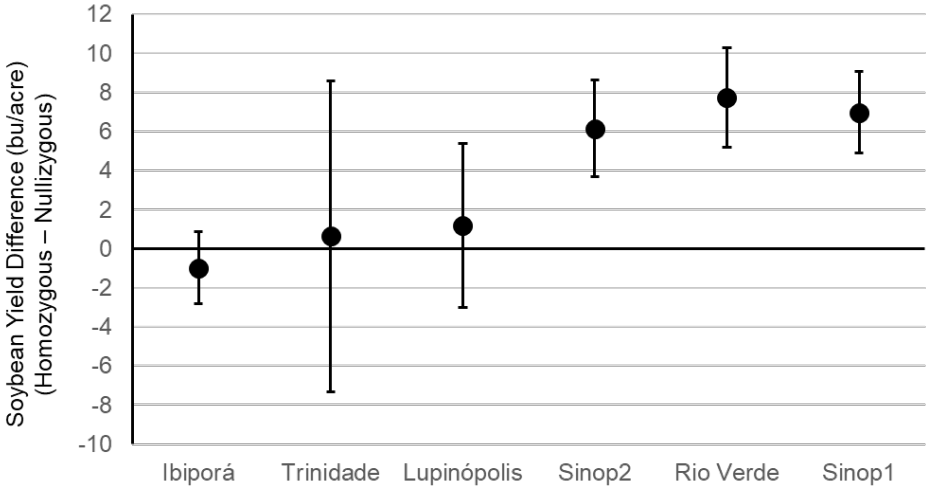


Figure A3.4. Effect of the GMB151 transgenic soybean trait on soybean yield

Soybean yield difference between the GMB151 homozygous and nullizygous lines at each of the Brazil research sites. Research sites are arranged from lowest to highest *P. brachyurus* population density. Error bars about the estimated yield differences represent the 95% confidence intervals. The GMB151 trait did not affect yield at the three locations with the lowest *P. brachyurus* population densities. At the three sites with the highest *P. brachyurus* population densities, GMB151 significantly improved soybean yield by an average of 7.0 bu/acre, or 21%.

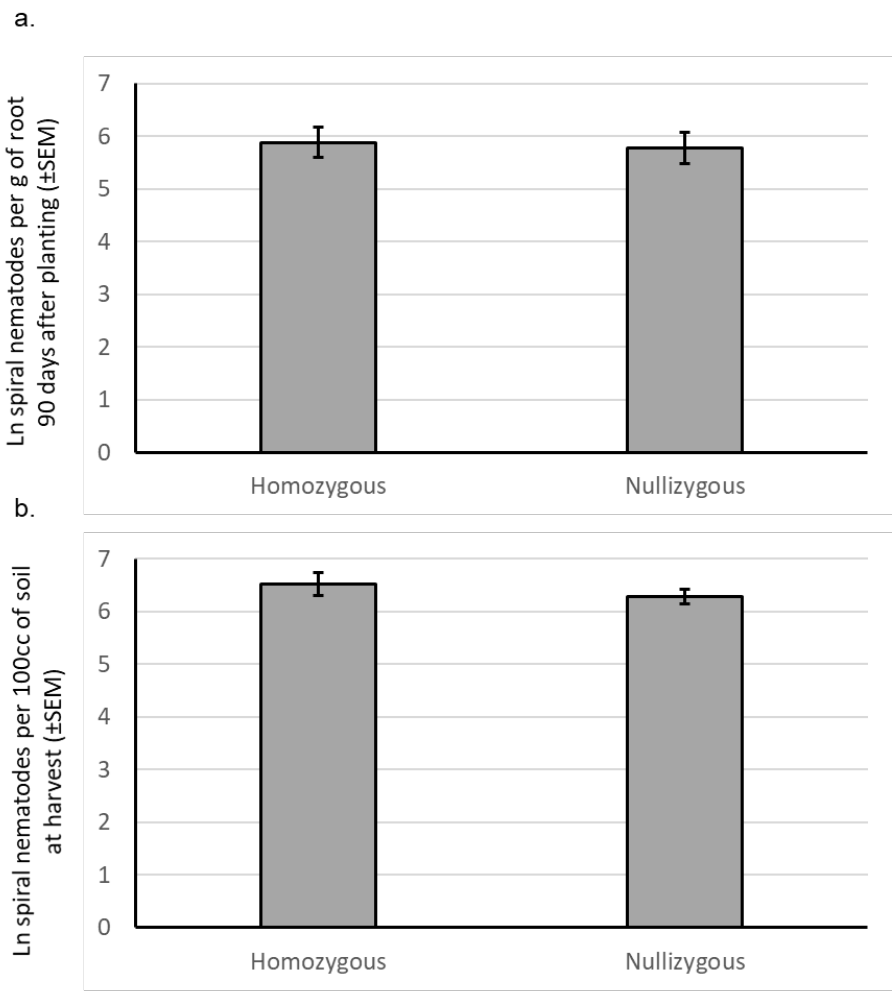


Figure A3.5. Effect of GMB151 on spiral nematode population densities

The effect of the GMB151 transgenic soybean trait on spiral nematodes was assessed in field trials. Two soybean lines, one homozygous for the GMB151 trait and the other a related nullizygous line, were grown in six field trials in Brazil conducted across the 2017/2018 and the 2018/2019 cropping seasons. The population density was natural log transformed for statistical analysis. Population densities of spiral nematodes were unaffected by the GMB151 transgenic soybean trait.

a: Soybean root systems were sampled 90 days after planting in both cropping seasons.

b: In the 2018/2019 cropping season, the population density of spiral nematodes in the soil was also assessed at soybean harvest.

Appendix 3 References

Bravo A, Gómez I, Porta H, García-Gómez BI, Rodriguez-Almazan C, Pardo L and Soberón M, 2013. Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity. Microbial Biotechnology 6, 17–26.

Crickmore N, Zeigler DR, Feitelson J, Schnepf E, Van Rie J, Lereclus D, Baum J and Dean DH, 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiology and Molecular Biology Reviews 62, 807–813.

Emmert EAB and Handelsman J, 1999. Biocontrol of plant disease: A (gram-) positive perspective. FEMS Microbiology Letters 171, 1–9.

Griffitts JS, Huffman DL, Whitacre JL, Barrows BD, Marroquin LD, Müller R, Brown JR, Hennet T, Esko JD and Aroian RV, 2003. Resistance to a bacterial toxin is mediated by removal of a conserved glycosylation pathway required for toxin-host interactions. Journal of Biological Chemistry 278, 45594–45602.

Griffitts JS and Aroian RV, 2005. Many roads to resistance: How invertebrates adapt to Bt toxins. Bioessays 27, 614–624.

Koenning SR, Overstreet C, Noling JW, Donald PA, Becker JO and Fortnum BA, 1999. Survey of crop losses in response to phytoparasitic nematodes in the United States for 1994. Journal of Nematology 31, 587–618.

Marroquin LD, Elyassnia D, Griffitts JS, Feitelson JS and Aroian RV, 2000. *Bacillus thuringiensis* (Bt) toxin susceptibility and isolation of resistance mutants in the nematode *Caenorhabditis elegans*. Genetics 155, 1693–1699.

Stirling GR, 1991. Chapter 2. Plant parasitic nematodes: The target of biological control. In: Biological Control of Plant Parasitic Nematodes: Progress, Problems and Prospects. CAB International, Wallingford, UK, 11–21.

Wei J-Z, Hale K, Carta L, Platzer E, Wong C, Fang S-C and Aroian RV, 2003. *Bacillus thuringiensis* crystal proteins that target nematodes. Proceedings of the National Academy of Sciences of the United States of America 100, 2760–2765.

Appendix 4. Detailed methods for protein expression study

List of Figures

1. Materials and methods

1.1. Test substance

The test substance was GMB151 soybean (OECD ID: BCS-GM151-6).

1.2. Control substances

There were no control substances. The non-GM counterpart Thorne was included for background correction during Cry14Ab-1 and HPPD-4 analyses.

1.3. Reference substances

The reference substances (see below) for the respective ELISA methods were microbially-produced proteins.

Reference substance (protein)	Batch No.	Purity
Cry14Ab-1	1514_Cry14Ab-1	91%
HPPD-4	1305_HPPD4Pa	88%

1.4. Test system

The test system was soybean. The test system was selected because the target crop, cultural practices, and management practices are all representative of the intended use pattern of GMB151 soybean.

1.5. Field production of protein expression samples

The GMB151 and non-GM counterpart Thorne soybean samples were obtained from field production trials in which tissue samples were harvested from plants grown in the U.S. in 2016 for protein expression analysis.

The soybean plants were grown at four field sites (Sites 01 through 04) that were located in areas representative of the commercial production of soybean in the U.S. Sample tissues were collected from each of the entries at different growth stages throughout the growing season. Three sites were randomly selected for protein expression analysis. The selected sites 02 (Germansville, PA), 03 (Fisk, MO), and 04 (Larned, KS) are representative of the locations of commercial soybean production with respect to cultural practices, soil type, and climatic conditions (Figure A4.1).

Each site had plots of GMB151 soybean that were treated or not treated with the trait-specific IFT herbicide as well as the non-GM counterpart Thorne. All entries were of the Thorne background.

The tissues harvested for protein expression analysis are summarized in Table 8 (Section 6.3 of the Petition).

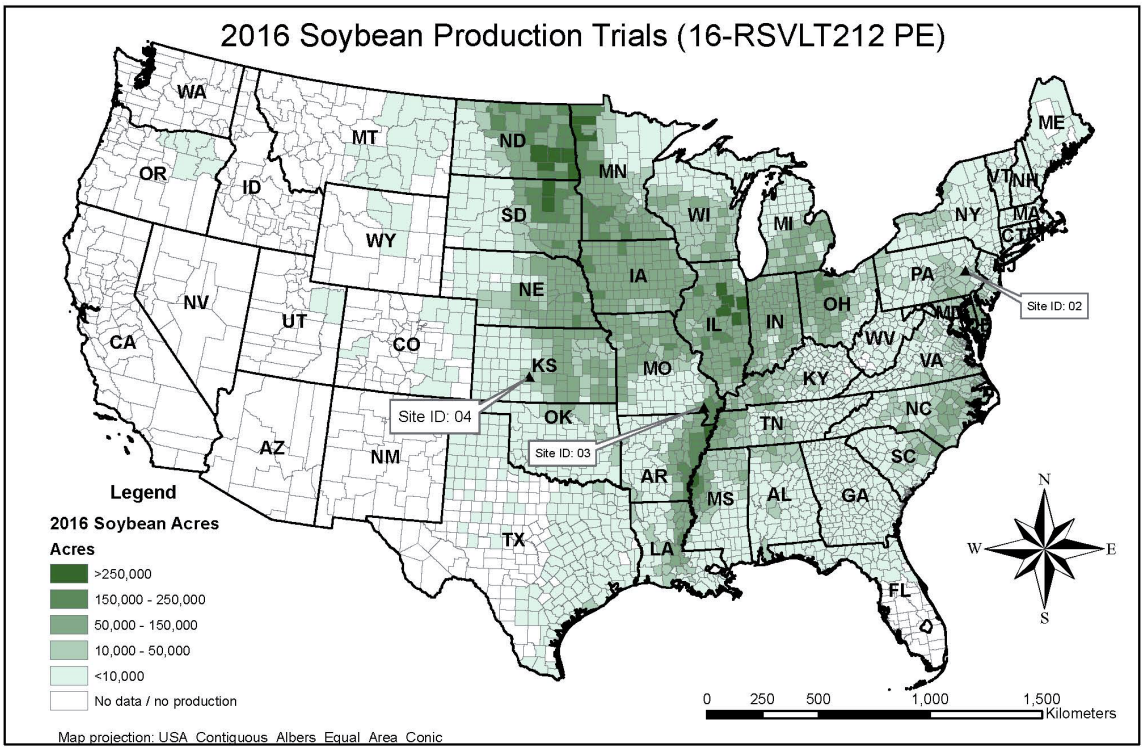


Figure A4.1. Field trial sites selected for protein expression analysis

1.6. Sample preparation

The samples were ground to a fine powder in the presence of liquid nitrogen and/or dry ice. After sample processing was completed, ground samples were stored at an average temperature of -60°C or lower until lyophilization.

Each frozen ground sample or a subsample of each frozen ground sample was lyophilized prior to protein extraction and ELISA analysis. The frozen ground material was maintained on dry ice before lyophilization. The % DW of each sample was determined from the FW of the sample before lyophilization and the DW of the sample after lyophilization by the following formula:

$$\% \text{ DW} = (\text{DW (g)} / \text{FW (g)}) \times 100$$

The lyophilized samples were stored at an average temperature of -60°C or lower until analysis.

1.7. Protein expression analysis

Four samples of a specific tissue in each entry per site were selected for expression analyses of Cry14Ab-1 and HPPD-4. In total, 12 replicate samples per entry were extracted, and expression levels of the event-specific proteins were quantified across the three selected sites.

1.7.1. Determination of Cry14Ab-1 expression in soybean samples

The quantitation of Cry14Ab-1 protein in the soybean samples was conducted with a validated Cry14Ab-1-specific ELISA method using the EnviroLogix QualiPlate™ Kit for Cry14Ab-1, which utilizes a mouse monoclonal antibody for capture and rabbit polyclonal antibody for detection. The Cry14Ab-1 was quantitated through interpolation of sample absorbance values on a standard curve generated in extraction buffer on each plate with Cry14Ab-1 reference protein. Protein was extracted from lyophilized powdered soybean samples by homogenization in extraction buffer at 20,000 rotations per minute (RPM) for 90 seconds with an Omni Prep homogenizer. The resulting protein extracts were clarified by centrifugation and diluted, if necessary, to ensure the absorbance values for the measured analyte were within the quantitative range of the ELISA method. Sample extracts were analyzed in triplicate, and the protein standard was measured in duplicate. Non-GM tissue blanks at the same dilutions as the samples and buffer blanks were included on each ELISA plate for background correction. The ELISA plates with standards and extracts were incubated for approximately 30 minutes at room temperature. The plates were then washed with 1X PBST (0.01 M phosphate buffered saline, 0.138 M NaCl, 0.0027 M KCl, 0.05% TWEEN® 20, pH 7.4). Antibody conjugate solution was dispensed into each well, and the plates were incubated for approximately 60 minutes at room temperature. The plates were then washed again with 1X PBST. Substrate solution was applied to each well, and the plates were incubated for approximately 30 minutes at room temperature. After incubation, stop solution was added, and absorbance at 450 nm was measured using a spectrophotometric plate reader.

1.7.2. Determination of HPPD-4 expression in soybean samples

The quantitation of HPPD-4 protein in the soybean samples was conducted with a validated HPPD-4-specific ELISA method using the EnviroLogix QuantiPlate™ Kit for HPPD-4, which utilizes mouse monoclonal antibodies for capture and for detection. The HPPD-4 was quantitated through interpolation of sample absorbance values on a standard curve generated in extraction buffer on each plate with HPPD-4 reference protein. Protein was extracted from lyophilized powdered soybean samples by homogenization in extraction buffer at 20,000 RPM for 90 seconds with an Omni Prep homogenizer. The resulting protein extracts were clarified by centrifugation and diluted, if necessary, to ensure the absorbance values for the measured analyte were within the quantitative range of the ELISA method. Sample extracts were analyzed in triplicate, and the protein standard was measured in duplicate. Non-GM tissue blanks at the same dilutions as the samples and buffer blanks were included on each ELISA plate for background correction. The ELISA plates with standards and extracts were incubated for approximately 30 minutes at room temperature. The plates were then washed with 1X PBST. Antibody conjugate solution was dispensed into each well, and the plates were incubated for approximately 60 minutes at room temperature. The plates were then washed again with 1X PBST. Substrate solution was applied to each well, and the plates were incubated for approximately 30 minutes at room temperature. After incubation, stop solution was added, and absorbance at 450 nm was measured using a spectrophotometer plate reader.

1.8. Data analyses

The standard curves for Cry14Ab-1 and HPPD-4 ELISA methods were generated using a 5-parameter logistic fit. The absorbance values obtained for all data points were interpolated on the standard curve to estimate the amount of protein detected in units of ng protein per mL of

extract (ng/mL). These values were converted to µg of protein per g of sample DW by the following formula:

$$\mu g \Big/ g \text{ DW in sample} = \left(\frac{\left(\frac{ng}{mL} \text{found}\right) \text{ x (mL extraction volume)}}{\text{g sample weight}} \right) \text{ x } \left(\frac{1\ \mu g}{1000\ ng}\right)$$

Protein expression levels were further converted from a DW to FW basis by the following formula:

$$\mu g/g \text{ FW in sample} = (\mu g/g \text{ DW in sample}) \text{ x } \left(\frac{\%DW}{100}\right)$$

1.9. Statistical analysis

SoftMax Pro™ GxP (Version 5.4.4) was used to calculate the mean concentration (ng/mL) of Cry14Ab-1 and HPPD-4 in each sample by averaging the replicate results for each sample extract.

The mean concentration and SD were determined for each entry within a site and across all three sites. The mean and SD were calculated using Microsoft Excel 2010. Statistics were calculated with full precision and rounded to two decimal places. The protein concentrations were not adjusted for extraction efficiency.

Unless otherwise noted, the DW and FW statistics calculated for each entry within a trial site were based on a population of n=4 (or four replicate samples per site). The DW and FW statistics calculated across all sites were based on a sample population of n=12 (three sites x four replicate samples/site).

Appendix 5. Detailed methods for aerobic soil degradation of Cry14Ab-1 protein

1. Materials

1.1. Test, control, and reference substances

The Cry14Ab-1 protein (Batch ID 1514_Cry14Ab-1; Purity 91%) was the test substance. There were no control or reference substances applicable to this study.

1.2. Test system

The test systems contained soils from California (CA), Iowa (IA), Kansas (KS), and Nebraska (NE). The test systems were selected because they were representative agricultural soil types for soybean cultivation in the U.S. and provided a diversity of physicochemical soil characteristics, as well as diverse geographical and climatic conditions.

1.3. Soil materials

Soils used in this study were collected from the top layer of the soil from Sanger, CA, Bagley, IA, Stilwell, KS, and Louisville, NE. Received soils were cleaned by removing the stones and plant parts and sieving through a 2-mm mesh sieve. The cleaned soils were stored at 2–8°C.

The moisture content of each soil was determined by heating three aliquots (approximately 10 g) in a microwave oven repeatedly until their sequential weights were constant.

Soils were characterized by Agvise Laboratories, Northwood, ND, for the following properties: percent sand, silt, and clay, maximum water holding capacity, soil moisture at 1/10 Bar and 1/3 Bar, organic carbon, pH, and cation exchange capacity.

The soil properties and the pre-study moisture content values are listed in Table 12 (Section 6.6 of this Petition).

1.4. ELISA kits

Cry14Ab-1 ELISA validation and analyses were performed using the EnviroLogix QuantiPlate™ Kit for Cry14Ab Poly-Poly ELISA.

2. Methods

2.1. Soil preparation and treatment

Soil preparation

An equivalent of 5 g of dry soil was weighed into each plastic centrifuge tube (test vessel) for each soil type, and Milli-Q water was added to allow for soil acclimation. The amount of water added was determined by considering the final moisture content required for the study and the current moisture content of the soil, taking into account the water volume of the treatment solutions.

The final soil moisture content was targeted at between pF 2.0 to 2.5 (1/10 to 1/3 Bar) and 55% of maximum water holding capacity, if possible.

Each test vessel was plugged with a foam cap and placed in a desiccator connected to a continuous supply of humidified air. The desiccator was placed in the environmental chamber at approximately 20°C (19–22°C) in the dark under aerobic conditions for five days to acclimate.

Soil treatment

To each acclimated soil, either 200 µL of 2 mg/mL Cry14Ab-1 protein solution (prepared in 50 mM Tris, 1 mM cysteine, pH 8.5 buffer) or buffer alone was added to prepare Cry14Ab-1 treated or blank soil samples. The dosing level for the Cry14Ab-1 treated soil sample was 80 µg Cry14Ab-1 protein per gram dry soil. Milli-Q water was added prior to and after soil treatment to ensure the final soil moisture content was between pF 2.0 to pF 2.5.

Twenty-seven Cry14Ab-1 treated soil samples (three for each of eight sampling intervals and three spares) and nine blank soil samples (one for each of eight sampling intervals and one spare) were prepared for each soil type.

2.2. Incubation and sampling

Following treatment and moisture adjustment, all samples, except Day 0 samples that were extracted right after treatment, were placed in a desiccator connected to a continuous supply of humidified air. The desiccator was placed in an environmental chamber at approximately 20°C (19–21°C) in the dark under aerobic conditions.

One blank soil sample and three Cry14Ab-1 treated soil samples were collected at Day 0, 0.1, 0.3, 1, 1.3, 2, 5, and 7 for each soil type.

2.3. Preparation of extracts from soil samples

At each sampling interval, soil samples were brought to the extraction lab at ambient temperature. The entire content of each test vessel was extracted according to the text box below. Rotated mixtures ware centrifuged at 4500 x *g* for 15 minutes at 2–8°C. Extracts were stored at 2–8°C until analysis.

Matrix	Extraction buffer	Buffer to soil ratio	Extraction method	Dilution buffer	Minimum dilution factor
CA soil	1% PVP-40 plus Buffer [a]	25 mL to 5 g dry weight soil	Rotated at speed 60–100% for 10 minutes in a cold room	1% PVP-40 [b]	4
IA soil					
KS soil					
NE soil					

[a] 22 mM potassium phosphate monobasic, 204 mM sodium phosphate dibasic, 137 mM sodium chloride, 27 mM potassium chloride, 1% PVP-40, 0.05% v/v Tween®20, and 0.4% octyl β-D-glucopyranoside

[b] 22 mM potassium phosphate monobasic, 204 mM sodium phosphate dibasic, 137 mM sodium chloride, 27 mM potassium chloride, 1% PVP-40, 0.05% v/v Tween®20

2.4. ELISA validation and analysis

ELISA validation

Parameters such as standard curve model and fit, intra-assay precision and accuracy, inter-assay precision and accuracy, limit of detection, quantitative range, specificity, dilutional linearity, extraction precision, and extract storage stability were evaluated to validate the method for each soil type.

Quantitative ELISA analysis of Cry14Ab-1 in soil extracts

At each sampling interval, one blank soil extract and three Cry14Ab-1 treated soil extracts of each soil type were analyzed by quantitative Cry14Ab-1 ELISA. Blank soil extracts were used for background correction of the ELISA results.

Data analysis and DT_{50} determination

The amounts of Cry14Ab-1 protein detected by ELISA at each sampling interval were used for DT_{50} determination. The software KinGUI 2 was used to evaluate the degradation kinetic model. Three different kinetic models, Single First Order Model (SFO), First Order Multi Compartment Model (FOMC), and Double First Order in Parallel Model (DFOP), were tested to determine the best fit kinetic model. The best-fit kinetic model was selected on the basis of the Chi-square scaled-error criterion and on the basis of a visual assessment of the goodness of the fits (diagrams of Measured & Predicted Residues vs. Time, diagrams of Residuals vs. Time). DT_{50} values were obtained from the analysis using the best-fit kinetic model.

2.5. Bioassay analysis

The 2 mg/mL Cry14Ab-1 protein treatment solution and the soil extracts from Day 0 and last sampling interval Day 7 were analyzed by the *C. elegans* bioassay. Each bioassay consisted of eight dose levels for a dose-response curve in six replicate wells and soil extract or treatment solution samples in at least twelve replicate wells. The dose-response curve was run concurrently to determine the LLOQ for that soil matrix. The detailed bioassay procedure is provided in Appendix 6.

2.6. Statistical analysis

Statistical analyses for this study included calculation of means, SDs, % coefficient of variation (%CV), % relative error (%RE), % relative difference (%RD), % relative change (%RC), and % inhibition using SoftMax Pro™ GxP and Microsoft Excel 2010. Statistics were calculated with full precision and rounded to simplify reporting.

Appendix 6. Description of the *Caenorhabditis elegans* bioassay

List of Tables

List of Figures

1. Methods

1.1. [

CBI-deleted

]

1.3. Validation parameters and acceptance criteria

The *C. elegans* bioassay validation results were evaluated using the following acceptance criteria.

Dose-Response Curve

The acceptability of the dose-response curve was evaluated based on the R^2 value of the curve and the number of non-zero dose levels used in the curve. It was considered to be acceptable if R^2 was ≥ 0.95 and there were at least five non-zero dose levels.

Intra-Assay Precision

Intra-assay precision was evaluated by the %CV at each dose level used to generate the dose-response curve. The %CV was calculated as a percent ratio of the replicate SD and the mean [ ] at Day 3 using the following formula:

$$\%\text{CV} = \left(\frac{\text{SD}}{\text{Mean}\ [\quad]}\right) \text{x}\ 100$$

CBI-deleted CBI-deleted

The precision acceptance criterion at each concentration level was %CV ≤ 35%.

Inter-Assay Precision

Inter-assay precision was evaluated by using the ratio of the highest and lowest ED_{50} values across all assays. It was considered to be acceptable if the highest ED_{50} value divided by the lowest ED_{50} value was ≤ 10.

Assay Pass Rate

At least 90% of the assays performed in this study had to meet the criteria set in this section.

1.4. Statistical analysis

Descriptive statistics (mean, SD, and %CV) were calculated using SoftMax Pro™ GxP. Statistics were calculated with full precision and rounded to simplify reporting.

2. Results and discussion

2.1. Dose-response curve

Figure A6.1 provides a representative dose-response curve that has been fitted with a semi-log curve fit. All curves had seven non-zero dose levels except three that had one dose level dropped to meet the $R^2 \geq 0.95$ criterion.

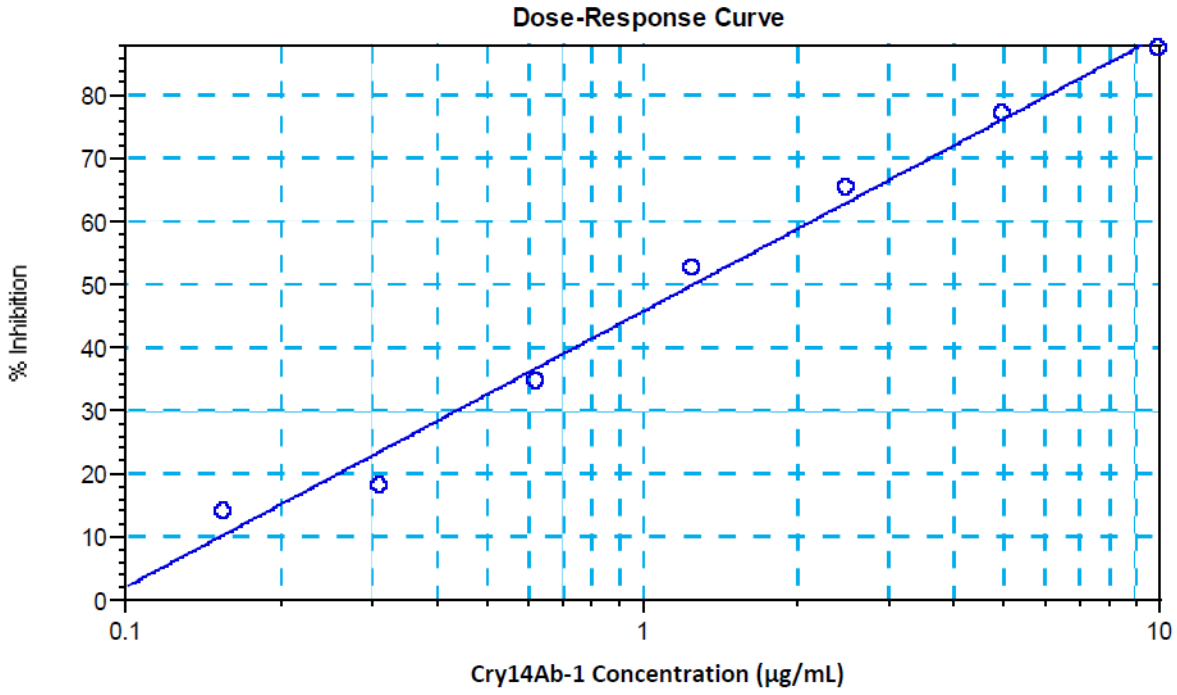


Figure A6.1. Representative dose-response curve with semi-log curve fit

2.2. Intra-assay precision

Table A6.1 provides a summary of the %CV values for the assays at Day 3. The results demonstrate acceptable intra-assay precision at all dose levels for the *C. elegans* bioassay using Cry14Ab-1.

Table A6.1. Mean %CV and the range of the %CV at each dose level on Day 3

Dose level (µg/mL Cry14Ab-1)	Mean %CV	%CV range
10	17.6%	10.0–29.4%
5	19.4%	10.2–28.3%
2.5	20.7%	13.2–34.8%
1.25	20.5%	8.5–30.7%
0.625	18.9%	10.1–31.5%
0.313	17.2%	7.6–34.8%
0.156	14.9%	7.2–27.6%
0	14.3%	5.8–34.6%

2.3. Inter-assay precision

Table A6.2 lists the mean, lowest, and highest ED_{50} values and the ratio of highest to lowest ED_{50} value across all assays. These results demonstrate acceptable inter-assay precision for the *C. elegans* bioassay using Cry14Ab-1.

Table A6.2. Mean ED_{50} and the ratio of the highest to lowest ED_{50} Value

	Mean	Lowest	Highest	Highest/Lowest
	µg Cry14Ab-1/mL solution			
ED_{50}	1.20	0.45	2.77	6.2

2.4. Assay pass rate

All 36 assays (100%) performed in this study met the criteria set in Section 1.3 of this Appendix.

3. Conclusions

The results of the validation of the *C. elegans* bioassay demonstrate that the method exhibits acceptable precision when performed with Cry14Ab-1.

Appendix 6 References

Stiernagle T, 2006. Maintenance of *C. elegans*. In: WormBook. Ed The *C. elegans* Research Community, doi/10.1895/wormbook.1.101.1, http://www.wormbook.org.

Appendix 7. Detailed information on non-target organism assessments

This appendix summarizes the list of studies conducted to assess potential effects of Cry14Ab-1 protein on NTOs, including larval and adult honey bees (*A. mellifera*), two species of ladybird beetles (*C. maculata* and *C. septempunctata*), green lacewings (*C. carnea*), water fleas (*D. magna*), Collembola (*F. candida*), and earthworms (*E. fetida*).

The concentrations of Cry14Ab-1 protein in each exposure diet/media were confirmed by ELISAs in either the diets or media to which the tested organisms were exposed or diets reconstituted using comparable formulations. If the concentrations measured by ELISA were within the acceptable accuracy range of quantitative ELISA (60–130%), the nominal concentrations were thus used for margin of exposure calculation. Otherwise, the measured concentration was used (e.g., *C. septempunctata*).

The bioactivity of different batches of Cry14Ab-1 protein used in the NTO studies was equivalent. The bioactivity was further validated by bioassays with *C. elegans* in select diets or exposure media.

1. Evaluation of potential adverse effects of Cry14Ab-1 protein to the honey bee (*Apis mellifera*) larvae in an *in vitro* study

The purpose of the study was to assess potential adverse effects of Cry14Ab-1 protein on the survival and development of honey bee larvae (*A. mellifera*) through dietary exposure over a 22-day duration. The study was designed according to the OECD Guidance Document for Testing of Chemicals, Honey Bee (*Apis mellifera*) Larval Toxicity Test, Repeated Exposure (Revision of April 2015) with modifications from U.S. EPA Office of Chemical Safety and Pollution Prevention (OCSPP) 850.supp.

The test was performed as a limit test with exposure to an artificial diet at a nominal test concentration of 1 mg Cry14Ab-1/g of diet and a total feeding dose of 154 µg Cry14Ab-1 per larva. There were four treatment groups, water control, buffer control, Cry14Ab-1, and dimethoate, each containing 48 larvae. The dimethoate was used as a positive control. The diets were composed of fresh royal jelly and an aqueous solution containing yeast extract, glucose, fructose, and deionized water. Diets were fed to honey bee larvae from Day 1 to Day 3. An aliquot (3.76% by volume) of the respective spiking solution containing either water, buffer, Cry14Ab-1, or dimethoate was added to the diet from Day 3 to Day 6.

During the test, larvae were maintained in the 48-well culture plates within commercially obtained grafting cells of a 9-mm internal diameter. The grafting cells were inserted on top of cellulosic dental rolls, and a single larva was placed onto the surface of the diet. Mortality was recorded on Day 4 through Day 8, Day 15, and Day 22.

The acceptance criteria of the study were met, with larval mortality in the control groups being ≤ 15% from Day 3 to Day 8, the emergence rate being ≥ 70% from Day 3 to Day 22, and the larval mortality in the positive control group being ≥ 50% from Day 3 to Day 8. No statistically significant differences in mortality and emergence rate were observed for the Cry14Ab-1 treatment relative to the control groups, and thus the NOED for Cry14Ab-1 protein to honey bee larvae is 1 mg/g diet or ≥ 154 µg/larva.

2. Evaluation of potential adverse effects of Cry14Ab-1 protein to the honey bee (*Apis mellifera*) adult in an oral toxicity study

The purpose of the study was to assess the potential chronic toxicity of Cry14Ab-1 protein on the survival and behavior of adult honey bee (*A. mellifera*) over a 10-day feeding duration. The study was designed according to OECD Guideline 2016 Proposal for a New Guideline for the Testing of Chemicals, Honey Bee (*Apis mellifera* L.), Chronic Oral Toxicity Test, 10-Day Feeding Test in the Laboratory, U.S. EPA OCSPP 885.4340 and OCSPP 850.supp, and Kling and Schmitzer (2015).

Two-day old female adult honey bees were exposed to a Cry14Ab-1 feeding solution at a concentration of 1000 mg/kg body weight, corresponding to 32.1 µg Cry14Ab-1/bee/day. The treatment solution was prepared by diluting the Cry14Ab-1 stock solution with sugar syrup to obtain a 50% syrup solution (water:syrup = 50:50, w:w). Water and buffer solutions were used for control groups. Dimethoate, used as a positive control reference, was fed to bees at a rate of 0.02 µg/bee/day.

The female worker bees in cages were fed with the treated food containing the treatment or buffer control solutions via syringes in *ad libitum* feeding. Syringes were weighed again after bees consumed the solutions. For each treatment, there were five replicates each with 10 bees in the stainless-steel cage. Mortality and behavioral abnormalities were assessed from Day 1 until the end of the study.

Cry14Ab-1 at a feeding rate of 32.1 µg/bee/day did not result in mortality or cause adverse behavioral effects to the adult honey bees. Therefore, the NOEC for lethal and behavioral effects to adult honey bee was ≥ 1000 mg Cry14Ab-1/kg feeding solution, and the NOED was ≥ 32.1 µg/bee/day.

3. Evaluation of the potential adverse effects of Cry14Ab-1 protein on ladybird beetle (*Coleomegilla maculata*)

The purpose of the study was to assess the potential adverse effects of Cry14Ab-1 protein on the survival and development of ladybird beetle (*C. maculata*) over a 21-day dietary exposure. The study was designed according to U.S. EPA OCSPP 885.4340.

Second instar larvae of *C. maculata* were exposed to Cry14Ab-1 protein at a single dose rate of 3.4 mg/g of diet in Petri dishes, which were approximately 50 mm diameter by 9 mm deep. The treatment diet was composed of Cry14Ab-1 protein together with bee pollen and *Ephestia kuehniella* eggs. The positive control diet contained 250 µg potassium arsenate/g of diet. The buffer control diet contained 0.46 mL sodium carbonate/bicarbonate/g of diet. A control group with untreated diet was also included.

There were 10 replicate vessels per testing group, each containing one insect. The vessels were maintained over 21 days in an incubator at 23–26°C, a relative humidity of 50–59%, and a light intensity of 670–770 lux with the light:dark cycle at 16 h: 8 h. Water was continuously supplied through moistened dental wicks. Approximately 50 µL of each diet treatment was provided to the test group insects daily from Day 0 until pupation occurred. The biological endpoints were evaluated on Day 15 due to abnormal mortality, which occurred in the control groups by Day 21.

The mean number of days to pupation was very similar among the water control, buffer control, and treatment groups. There were no detrimental effects on survival of *C. maculata* exposed to a diet containing 3.4 mg Cry14Ab-1 protein/g of diet compared to the control through Day 15 of the test. The NOEC for mortality was therefore 3.4 mg Cry14Ab-1 protein/g of diet.

4. Evaluation of the potential adverse effects of Cry14Ab-1 protein on ladybird beetle (*Coccinella septempunctata*)

The purpose of the study was to evaluate the potential adverse effects of Cry14Ab-1 protein on the survival and development of ladybird beetle (*C. septempunctata*) through dietary exposure over a period of 12 days. The study was designed according to U.S. EPA OCSPP 885.4340.

The 4-day old larvae of *C. septempunctata* were housed individually in closed Petri dishes with a diameter of 35 mm and exposed to the respective diet of each treatment group. There were four treatment groups, comprised of water and buffer controls as well as reference and test groups, each with 30 insects as replicates. The feeding diet was composed of bee pollen and *Ephestia* eggs. Cry14Ab-1 protein at a rate of 3.4 mg/g diet was added to the diet of the test group. Potassium arsenate at a rate of 65 µg/g diet was fed to the reference group, and Tris buffer at pH 8.5 was fed to the buffer control group. The study was conducted for 12 days in a controlled environment at 24.5–25.5°C and a relative humidity of 63–81%. Light was provided at an intensity of 2006–2884 lux with the light: dark cycle at 16 h: 8 h. The water was supplied through a moistened cotton wool pad within each test system. All exposure units were assessed daily, and the condition of each larva was recorded.

The preimaginal mortality of the insects in the untreated control group was 3.3%, and no mortality was observed either in the buffer control or the test group. There was 93.3% mortality in the positive control group on Day 12. The validity and sensitivity of the test system were indicated by the low observed control mortality (3.3%) and the mortality (93.3%) observed for the positive control group. The concentration of the Cry14Ab-1 protein in the diet was evaluated as 1.83 mg/g diet using ELISA and was lower than the normal accuracy range of the quantitative ELISA. Therefore, the NOED for preimaginal mortality of Cry14Ab-1 protein to *C. septempunctata* was reported as 1.83 mg/g.

5. Evaluation of the potential adverse effects of Cry14Ab-1 protein on green lacewing (*Chrysoperla carnea*)

The purpose of the study was to assess the potential adverse effects of Cry14Ab-1 protein on the survival of green lacewing (*C. carnea*) through dietary exposure over a period of 10 days. The study was designed according to U.S. EPA OCSPP 885.4340.

Two-day old larvae of *C. carnea* were housed individually in closed Petri dishes and exposed to diet of each of the four treatment groups, comprised of water and buffer controls as well as reference and test groups, each with 30 insects. The feeding diet was composed of mashed mealworms (*Tenebrio molitor*), honey, pollen, yeast, and dried egg powder. Cry14Ab-1 protein was added to the test diet at a rate of at 3.4 mg/g of diet, and potassium arsenate was added to the reference diet at a rate of 0.249 mg/g diet as a positive control. Diet was offered *ad libitum* on a small piece of plastic foil placed on the bottom of the Petri dish where larvae of *C. carnea* were exposed. The study was conducted for 10 days in a controlled environment room at 23.5–27.0°C and at a relative humidity of 67–73%. Light was provided at an intensity of 1297–2140 lux with

the light:dark cycle at 16 h: 8 h. All exposure units were assessed daily, and the condition of *C. carnea* was recorded.

The preimaginal mortality of the test organisms was 6.7% in the untreated control group and 3.3% in the buffer control group. No mortality was observed in the test group after correction based on the mortality in the control group. All larvae were dead in the reference control group on Day 10. The validity and sensitivity of the test system was indicated by the low observed control mortality (6.7%) and the mortality observed for the reference control group (100%). Therefore, the NOED for preimaginal mortality of Cry14Ab-1 protein to *C. carnea* was 3.4 mg/g.

6. Evaluation of the potential adverse effects of Cry14Ab-1 protein on water flea *Daphnia magna* in a static renewal laboratory study

The purpose of the study was to assess the potential adverse effects of Cry14Ab-1 protein on the survival, development, and reproduction of *D. magna* in a static renewal laboratory test system over a period of 21 days. The study was designed according to OECD Guideline for Testing of Chemicals, No. 211, U.S. EPA OCSPP 850.1300, and U.S. EPA OCSPP 885.4240.

The first instar (< 24 h old) of *D. magna* was exposed to the test solution made of 100 mL of Elendt M7 water mixed with Cry14Ab-1 at 500 µg/L in 250-mL glass beakers. The test group was compared with a buffer control group containing Tris buffer at pH 8.5. Ten replicates (one organism per replicate) were included for each treatment group. The exposure media were renewed on Days 2, 5, 7, 9, 12, 14, 16, and 19 by transferring the organisms to fresh test solution. The test solution was maintained at 18–22°C throughout the study, and light was provided at an intensity between 1000–1200 lux with a cycle of light: dark at 16 h: 8 h. The physical-chemical properties of the test solution were measured when renewing the test solution. Biological measurements were monitored, including the presence of aborted eggs and neonate mortality, parental survival, time of brood release, numbers of offspring, parental body length, and dry body weights at the end of the study. The validity and sensitivity of the study were verified by biological measurements.

At the end of the study, there was no mortality of the parental organisms (< 20%), and the average number of offspring per test organism was 119 neonates (> 60 neonates per female) in the control group. In addition, the well-being of the test organisms was further indicated by the average final body length of 4.19 mm for the parental organism and the average age for first offspring emergence at 9.9 days. No statistically significant differences were observed for developmental, reproductive, and survival endpoints of the test organisms between treatment and control groups. Therefore, the overall chronic NOEC for *D. magna* was 500 µg/L of exposure medium.

7. Evaluation of the potential adverse effects of Cry14Ab-1 protein on Collembola *Folsomia candida* tested on artificial substrate

The objective of the study was to assess the potential adverse effects of Cry14Ab-1 protein on survival and reproduction of the Collembolan species *F. candida* through a 28-day exposure on treated artificial substrate. The study was designed according to OECD Guidelines for Testing of Chemicals, Test No. 232, Collembolan Reproduction Test in Soil.

Collembolans at 10–12 days old were exposed to control and test diets in plastic vessels with a dimension of 8 x 6 x 4 cm (L × W × H). There were eight replicates per treatment group and

10 Collembolans per replicate. The artificial substrate within each vessel was composed of plaster of Paris, activated charcoal, and demineralized water. The diet for each treatment group was made of baker's dry yeast and respective solvent or protein. For the water and buffer control groups, deionized water or sodium carbonate buffer at pH 9.6 was added, respectively. For two test groups, Cry14Ab-1 protein solution was added at 1 mg/g and 10 mg/g, respectively. Diets were prepared at Day 0, 3, 7, 10, 14, 17, 21, and 24 and then inserted into the middle of each test vessel at a rate of 10 mg diet/vessel. The exposure was conducted for 28 days at 20 ± 2°C in the dark. Water was supplied to the Collembolans throughout the study. At the end of the study, Collembolans were extracted from each test vessel, and adult and juvenile Collembolans were counted individually.

A separate study with the positive control substance thiodicarb had a NOEC of 0.125 mg/g and indicated the test organisms were sufficiently sensitive. The validity of the study was indicated by 8.1% mortality of *F. candida* and a mean number of 834.6 juveniles per replicate in the control groups. No effects were observed for the test groups as compared to the control groups. Therefore, the NOEC for mortality and reproduction of the Cry14Ab-1 protein to *F. candida* was 10 mg/g.

8. Evaluation of the potential adverse effects of Cry14Ab-1 protein on earthworm *Eisenia fetida* tested in artificial soil

The objective of the study was to assess the potential adverse effects of Cry14Ab-1 protein on the survival and growth of the earthworm species *E. fetida* through a 14-day exposure in artificial soil. The study was designed according to OECD Guidelines for Testing of Chemicals, Test No. 207, Earthworm, Acute Toxicity Tests.

The adult earthworms, at approximately nine months old and weighing between 0.27–0.47 g, were exposed to 500 g of dry weight of artificial soil in a 1.5-L preserving jar. The artificial soil was made of 69.5% fine quartz sand, 10% Sphagnum peat, 20% Kaolin clay, and 0.5% calcium carbonate and was adjusted to 40–60% of the maximum water holding capacity. Three treatment groups were used including water, solvent control, and test groups, each with 10 earthworms per replicate and eight replicates per group. An aliquot (25 mL) of water or buffer solvent was added to the soils of control groups while the Cry14Ab-1 protein solution was added to the test group at a concentration of 50 mg/kg of soil at the beginning of the study. Earthworms were not fed throughout the exposure and were incubated under constant light with 400–800 lux at a temperature of 20 ± 2°C. On Day 7 and Day 14, the number of surviving earthworms was determined. In addition, weight, abnormal behavior, and observed symptoms were also determined on Day 14. A positive control substance was used for testing under the same conditions to demonstrate the sensitivity of the test systems and test organisms.

No mortality was observed among all treatment groups, which is in line with the quality criteria of the testing guideline. No statistically significant differences were observed on survival and growth of the earthworms between the test group and two control groups. Therefore, the NOEC of Cry14Ab-1 on the survival and growth of earthworms is 50 mg Cry14Ab-1 protein/kg of DW of the artificial soil.

Appendix 7 References

Kling A and Schmitzer S, 2015. Proposal for a new OECD guideline for the testing of chemical on adult honey bees (*Apis mellifera* L.) in a 10 day chronic feeding test in the laboratory and results of the recent ring test 2014. Hazards of pesticides to bees - 12th International Symposium of the ICP-BR Bee Protection Group, Ghent, Belgium, September 15–17, 2014. Julius-Kühn-Archiv, 450, pp. 69–74.

Appendix 8. Environmental safety assessment of the Cry14Ab-1 protein expressed in GMB151 soybean

List of Tables

1. Requirements for non-target organism studies with Cry proteins (test species, exposure routes)

1.1. General requirements for NTO studies and test species

For the environmental risk assessment of pesticidal Cry proteins expressed in GM plants, studies are performed with NTOs that are directly or indirectly exposed to these proteins but are not the target pests.

The safety assessment starts with the selection of NTOs exposed to the Cry protein, taking into account their importance and beneficial function in the environment and their potential sensitivity to the toxin based on their taxonomic relatedness to the target species (e.g., beneficial pollinators, endangered species, predators or parasites of the target pests, and species with a specific cultural value). Additional selection criteria consider that the test species can be reared under artificial conditions, an international established and validated GLP study protocol is available, and a well-characterized, specific toxic reference substance is available to demonstrate sensitivity of the test system (ILSI-CERA, 2013; Koch et al., 2015).

The test organisms selected for Cry14Ab-1 for these NTO studies are: adult and larval honey bees (*A. mellifera*) as representative pollinators, Collembola (*F. candida*) and earthworm (*E. fetida*) as representative soil-dwelling organisms, ladybird beetles (*C. maculata* or *C. septempunctata*) and green lacewings (*C. carnea*) as representative predator organisms, and water fleas (*D. magna*) as representative aquatic organisms.

1.2. Primary routes of exposure to Cry protein to NTOs

The primary route of exposure of the Cry protein to the NTOs (except Daphnids) is through dietary/oral ingestion. Dietary exposure to the Cry protein can happen in three different ways: exposure to pollen containing the Cry protein, exposure to the protein deposited in the soil, or exposure to prey that has fed on plant material from GM plants expressing the Cry protein. Daphnids may be exposed to the Cry protein through the media in which they live.

Assessing NTO impacts from GM crops follows a tiered testing regime in which the early tiers are done as laboratory tests under controlled conditions, with well-characterized test organisms and high concentrations of the Cry protein. The studies are designed to determine the actual hazard to a test species, usually using high doses to ensure a conservative margin of safety and certainty, and to give a maximum hazard result (Romeis, 2006; Rose, 2007).

1.3. Routes of exposure to Cry14Ab-1 protein to selected NTOs

To evaluate the environmental safety of the Cry14Ab-1 protein, standard single-species laboratory studies were conducted with the representative beneficial insect species and aquatic and soil invertebrate species as listed in Section 1.1 of this Appendix. The NTO exposure studies were conducted with the intention to maximize exposure of the Cry14Ab-1 protein to the testing species to ensure a conservative safety margin at realistic field conditions.

The concentration of Cry14Ab-1 in the diet/medium selected for each NTO study was based on the most relevant plant tissue that is responsible for the exposure of the tested NTO and the expression levels of Cry14Ab-1 in the respective plant tissue. A target of > 10X the safety margin was considered when conducting the NTO studies.

It was assumed that honeybee (*A. mellifera*) larvae and adults are exposed to Cry14Ab-1 protein via ingestion of GMB151 soybean pollen. For the soil-dwelling test organisms like Collembola (*F. candida*) and earthworms (*E. fetida*), the principal route of exposure was assumed to be the consumption of decomposing whole plant tissue (including roots) containing Cry14Ab-1 protein. For the ladybird beetles (*C. maculata* and *C. septempunctata*), the green lacewing (*C. carnea*), and the water flea (*D. magna)*, the plant tissue with the highest Cry14Ab-1 expression levels across different growth stages in the field expression study was used as a worst-case scenario approach. The assumption here would maximize the exposure estimate for each NTO under field conditions.

2. Source of the Cry14Ab-1 protein for NTO studies

The Cry14Ab-1 protein was produced in a recombinant *Bt* strain and provided as lyophilized powder.

3. Calculation of expected environmental concentrations or doses (EEC or EED) of the Cry14Ab-1 protein to each NTO

When calculating the EEC or EED of the Cry14Ab-1 protein to each NTO, the highest Cry14Ab-1 expression level in appropriate plant tissue was used, selecting from flowers, whole plants (including roots), and leaves of GMB151 soybean. The soybean flowers were sampled around full flowering stage, ranging from beginning of flowering to flower fading, which were indicated at phenological growth stage 60–66 on a BBCH scale (Munger et al*.,* 1997). The soybean whole plants (including roots) were sampled between post-full maturity and harvest, which was at stage BBCH 89–99. The soybean leaves were sampled throughout growth stages, and those at leaf development stage (BBCH 16–17) had the highest Cry14Ab-1 protein expression levels as shown in Table 9 of this Petition. In addition, the average consumption of the respective plant tissue by the respective test organism was used in the exposure calculations (Crailsheim et al., 1992; Babendreier et al*.,* 2004; U.S. EPA et al., 2014). A summary of the highest expression values of

Cry14Ab-1 in GMB151 soybean for flower, whole plant (including roots), and leaf is listed in Table A8.1.

Table A8.1. Highest expression levels of Cry14Ab-1 protein in GMB151 soybean flower, whole plant, and leaf

Plant tissue	Growth stage	Maximum value
Flower [a]	BBCH 60–66	72.10 µg/g DW
Whole plant [a]	BBCH 89–99	72.64 µg/g DW
Leaf [b]	BBCH 16–17	290.44 µg/g DW

DW = dry weight
[a] Flower and whole plant from GMB151 soybean plants treated with the trait-specific herbicide.
[b] Leaf from GMB151 soybean plants not treated with the trait-specific herbicide.

3.1. Calculation of EED of Cry14Ab-1 protein to honey bees

It was confirmed that Cry14Ab-1 expression levels in the plant reproductive structures (stamen/pistil) of GMB151 soybean were similar to those in whole flowers of GMB151 soybean during the flowering stage (Section 6.3.2 of this Petition). Thus, Cry14Ab-1 expression levels in the whole flower were used as a surrogate for its level in the pollen. Soybean is primarily self-pollinated, and cross-pollination by insects is less than 1% under field conditions (Ray et al., 2003). Therefore, the exposure estimate to honey bees would be maximized by using Cry14Ab-1 expression levels in the flower around full flowering stage at field testing conditions.

The EED of Cry14Ab-1 protein for honey bee larvae was calculated based on the highest Cry14Ab-1 protein expression level measured in flowers at growth stage BBCH 60–66 (Table A8.1) and the average consumption of pollen by honey bee larvae during development, which is 2 mg (Babendreier et al., 2004). Thus, the EED of Cry14Ab-1 protein for larval honey bee is 0.1442 µg as calculated below:

$$\text{EED (honeybee larvae)} = \frac{72.10\ \mu g\ \text{Cry14Ab} - 1}{g\ dw} \times \frac{g}{1000\ mg} \times 2\ mg\ \text{pollen} = 0.1442\ \mu g\ \text{Cry14Ab} - 1$$

The EED of Cry14Ab-1 protein for honey bee adults was calculated based on the highest Cry14Ab-1 protein expression levels measured in flowers at growth stage BBCH 60–66 (Table A8.1) and the highest daily consumption rate by adult worker honey bees, which is 12 mg pollen/day (Crailsheim et al., 1992; U.S. EPA et al., 2014). The EED of Cry14Ab-1 protein for adult honey bee is 0.8652 µg/day as calculated below:

$$\text{EED (honeybee adult)} = \frac{72.10\ \mu g\ \text{Cry14Ab} - 1}{g\ dw} \times \frac{g}{1000\ mg} \times \frac{12\ mg\ \text{pollen}}{day} = \frac{0.8652\ \mu g\ \text{Cry14Ab} - 1}{day}$$

3.2. EEC of Cry14Ab-1 protein to ladybird beetle and green lacewings

The highest Cry14Ab-1 expression level in the GMB151 soybean plant (290.44 µg/g DW in soybean leaves, Table A8.1) was used as the most conservative EEC of Cry14Ab-1 protein to

ladybird beetles (*C. maculata* or *C. septempunctata*) and green lacewings (*C. carnea*). Therefore, the EEC of Cry14Ab-1 protein to the tested ladybird beetles and green lacewings is 0.29044 mg/g based on DW of the diet.

3.3. EEC of Cry14Ab-1 protein to water flea

The EEC for *D. magna* was calculated using the expression level of Cry14Ab-1 in GMB151 soybean plant leaves, which contained the highest Cry14Ab-1 expression level. This will maximize the exposure estimate and ensure the safety assessment is conservative. According to Rosi-Marshal et al. (2007), the annual input rate of corn byproducts to the headwater stream is 8 g/m^2 of leaf and cob (0.1–7.9 g of ash-free dry mass), which corresponds to 0.08 g/L of water assuming a 0.1-m depth stream. Therefore, based on the input rate of the crop byproducts to the headwater stream, the EEC to *D. magna* may be estimated as follows:

$$\text{EEC } (Daphnia\ magna) = \frac{0.08\ g}{L} \times \frac{290.44\ \mu g\ \text{Cry14Ab}-1}{g\ leaf} = \frac{0.0232\ mg\ \text{Cry14Ab}-1}{L}$$

3.4. EEC of Cry14Ab-1 protein to Collembola and earthworms

The calculation of the EEC of Cry14Ab-1 for Collembola and earthworms adopted the method described in Prince et al. (2001) for deriving total biomass estimates from soybean yield and used the highest Cry14Ab-1 expression levels measured in whole plants to make bulk soil estimates of Cry14Ab-1 concentrations. The total biomass of a hypothetical soybean crop field per area was calculated using the mean soybean yields (44.2 bu/Ac, or 2,972 kg/ha) reported by the USDA National Agricultural Statistics Service survey for the years 2013–2017 (USDA-NASS, 2018). Subsequently, the Cry14Ab-1 protein expression level in whole plants (72.64 µg/g tissue; Section 6.3 of this Petition) was then applied to determine the amount of expressed protein per area of the soybean field. This value was then converted into the amount of Cry14Ab-1 per area of soil that contained soybean plants. During the process, it was assumed that the proteins from soybean plants only distributed to the top 6.7 inches (17 cm) of the soil. This is a reasonable and conservative assumption as the primary rooting zone is typically within the depth of a plow blade, which is around 6.7 inches. This results in a soil concentration of 0.32 µg Cry14Ab-1/cm^3, or an EEC of 0.24 µg/g soil assuming a bulk density of 1.33 g/cm^3 soil. Comparing the EEC of Cry14Ab-1 from GMB151 soybean to Cry proteins from other GM crops deregulated by the USDA, it is greater than some established values, but generally one to two orders of magnitude lower.

$$\frac{\text{kg Above Ground Biomass (AGB)}}{\text{Ha}} = \frac{2{,}972\ \text{kg grain}}{\text{Ha}} \text{ x } \frac{0.92\ \text{kg Dry Matter (DM) grain}}{1\ \text{kg grain}} \text{ x } \frac{1\ \text{kg AGB}}{0.42\ \text{kg DM grain}} = \frac{6{,}510\ \text{kg AGB}}{\text{Ha}}$$

$$\frac{\text{kg Total Biomass (TB)}}{\text{Ha}} = \left[\frac{6{,}510\ \text{kg AB}}{\text{Ha}} \text{ x } \frac{0.15\ \text{kg Root Biomass}}{1\ \text{kg AB}}\right] + \frac{6{,}510\ \text{kg AB}}{\text{Ha}} = \frac{7{,}487\ \text{kg TB}}{\text{Ha}}$$

$$\frac{\mu\text{g Cry14Ab-1}}{\text{cm}^3\ \text{soil}} = \frac{7{,}487\ \text{kg TB}}{\text{Ha}} \text{x } \frac{72{,}640\ \mu\text{g Cry14Ab-1}}{\text{kg TB}} \text{ x } \frac{1\ \text{Ha (at 17 cm depth)}}{1.7 \text{ x } 10^9\ \text{cm}^3} = \frac{0.32\ \mu\text{g}}{\text{cm}^3\text{soil}}$$

The details of each NTO study and the relevant Cry14Ab-1 expressing plant tissue used for calculating the EEC or EED are summarized in Table A8.2.

Table A8.2. Summary of NTO tests and Cry14Ab-1 expressing plant tissues used for EEC/EED calculations

Test organism	Species	Type of limit test	Plant tissues used for Cry14Ab-1 exposure estimate
Honey bee larvae	*Apis mellifera*	Repeated-dose 4 d; duration 22 d	Pollen: 72.10 µg/g DW
Honey bee adult	*Apis mellifera*	Chronic, 10 d	
Ladybird beetle	*Coleomegilla maculata*	Chronic, 12 d	Leaves: 290.44 µg/g DW
Ladybird beetle	*Coccinella septempunctata*	Chronic, 12 d	
Green lacewing	*Chrysoperla carnea*	Chronic, 12 d	
Water flea	*Daphnia magna*	Chronic, static test 21 d	
Collembola	*Folsomia candida*	Chronic, 28 d	Whole plants: 72.64 µg/g DW
Earthworm	*Eisenia fetida*	Chronic, 14 d	

DW = dry weight

4. Validation of the presence and biological activity of the Cry14Ab-1 protein in the tested diets or exposure media

The concentrations of Cry14Ab-1 protein were confirmed by ELISAs in either the diets or media to which the tested organisms were exposed (adult honeybee, *C. septempunctata*, and water flea) or diets reconstituted using comparable formulations (Collembola, green lacewing, earthworm, *C. maculata*, and honeybee larvae). If the concentrations measured by ELISA were within the acceptable accuracy range of quantitative ELISA (60–130%), the nominal concentrations were thus used for MOE calculation. Otherwise, the measured concentration was used (e.g., *C. septempunctata*, as described in Appendix 7).

The bioactivity of different batches of Cry14Ab-1 protein used in the NTO studies was equivalent (Section 6.4.1 of this Petition). The bioactivity was further validated by bioassays with *C. elegans* in select diets or exposure media (adult honeybee, *C. septempunctata*, and water flea).

5. No observed effect concentrations/doses and safety margins

No mortality or chronic adverse effects were observed after exposure to the Cry14Ab-1 protein in any of the NTO studies. Depending on the type of each NTO study, chronic endpoints were used related to the development, reproduction, and any potential adverse behavioral effects of the respective test organisms.

Since none of the test species showed any adverse effects after exposure to the Cry14Ab-1 protein, the overall chronic NOECs for mortality/survival were determined as the highest test concentrations, which ranged from 1–10 mg/g of diet or from 0.05–0.50 mg per gram of soil or per liter of exposure media (Table A8.3). The NOEDs calculated for the honeybee larvae and honeybee adult studies were 154 µg/larvae and 32.1 µg/bee/day (adult), respectively.

MOE were calculated for each NTO study as the ratio of the NOEC to EEC or NOED to EED. In all tested NTOs, a sufficiently large MOE could be established ranging from 6 to > 40000. The U.S. EPA published Level of Concern (LOC) is 50% mortality at 5X EEC (Rose, 2007; U.S. EPA, 2015). Therefore, although the MOE of *C. septempunctata* was 6, it is still conservative. Moreover, pollen is likely the main route for ladybird beetle to be exposed to Cry14Ab-1 protein. Thus, if the pollen was used to calculate the EEC of *C. septempunctata*, the more realistic MOE would be > 25. In summary, Cry14Ab-1 can be classified as non-toxic to all the tested NTOs at the tested levels and is not likely to produce any adverse effects on the tested NTOs at field exposure levels. Table A8.3 summarizes the results from the exposure estimates, NOEC or NOED, and calculated safety margins.

Table A8.3. Summary of expected environmental concentrations or doses (EECs or EEDs), no observed effect concentrations or doses (NOECs or NOEDs), and margins of exposure (MOE) of the Cry 14Ab-1 protein to each tested NTO

Test Organism	Species	Order	EEC/EED of Cry14Ab-1 (DW)	NOEC/NOED (mortality)	MOE [a]
Honey bee larvae	*Apis mellifera*	Hymenoptera	0.1442 µg/larvae	≥ 1 mg/g diet / 154 µg/larvae	1068
Honey bee adult	*Apis mellifera*		0.8652 µg/day	≥ 1 mg/g diet / 32.1 µg/day	37
Ladybird beetle	*Coleomegilla maculata*	Coleoptera	0.29044 mg/g	≥ 3.4 mg/g diet	12
Ladybird beetle	*Coccinella septempunctata*	Coleoptera	0.29044 mg/g	≥ 1.83 mg/g diet	6
Green Lacewing	*Chrysoperla carnea*	Neuroptera	0.29044 mg/g	≥ 3.4 mg/g diet	12
Water flea	*Daphnia magna*	Cladocera	0.0232 mg/L	≥ 0.500 mg/L	22
Collembola	*Folsomia candida*	Collembola	0.24 µg/g	≥ 10 mg/g diet	41667
Earthworm	*Eisenia fetida*	Haplotaxida	0.24 µg/g	≥ 0.05 mg/g soil	208

DW = dry weight; MOE = margin of exposure
[a] MOE were calculated as the result of NOED/EED for honey bees and NOEC/EEC for the other tested NTOs.

In addition, acute oral toxicity tests were conducted with one mammal (mouse) and one avian species (bobwhite quail) at the limit dose level of ≥ 2000 mg/kg body weight. Cry14Ab-1 exposure

did not result in any mortalities, treatment-related clinical signs, effects on body weight parameters, or any macroscopic changes at necropsy in the treated animals group.

6. Conclusion

Based on the NTO studies, no adverse effects were observed from the tested species including adult and larval honey bees, two soil-dwelling organisms (Collembola and earthworms), three predator organisms (two species of ladybird beetle and green lacewing), one aquatic organism (water flea), one mammal (mouse), and one avian species (bobwhite quail).

Based on the safety assessment of the tested NTOs, the calculated MOE are sufficiently large to conclude that the Cry14Ab-1 protein is not likely to pose any risk to the tested NTOs at realistic field exposure levels.

Appendix 8 References

Babendreier D, Kalberer N, Romeis J, Fluri P and Bigler F, 2004. Pollen consumption in honey bee larvae: A step forward in the risk assessment of transgenic plants. Apidologie 35, 293–300.

Crailsheim K, Schneider LHW, Hrassnigg N, Bühlmann G, Brosch U, Gmeinbauer R and Schöffmann B, 1992. Pollen consumption and utilization in worker honeybees (*Apis mellifera carnica*): Dependence on individual age and function. Journal of Insect Physiology 38, 409–419.

ILSI-CERA, 2013. Surrogate species selection for assessing potential adverse environmental impacts of genetically engineered plants on non-target organisms. Proceedings of the Center for Environmental Risk Assessment, ILSI Research Foundation, Washington, D.C., June 26–28, 2012.

Koch MS, Ward JM, Levine SL, Baum JA, Vicini JL and Hammond BG, 2015. The food and environmental safety of *Bt* crops. Frontiers in Plant Science 6, 283.

Munger P, Bleiholder H, Hack H, Hess M, Stauss R, van den Boom T and Weber E, 1997. Phenological growth stages of the soybean plant (*Glycine max* L. MERR.): Codification and description according to the BBCH scale. Journal of Agronomy and Crop Science 179, 209–217.

Prince SD, Haskett J, Steininger M, Strand H and Wright R, 2001. Net primary production of U.S. Midwest croplands from agricultural harvest yield data. Ecological Applications 11, 1194–1205.

Ray JD, Kilen TC, Abel CA and Paris RL, 2003. Soybean natural cross-pollination rates under field conditions. Environmental Biosafety Research 2, 133–138.

Romeis J, 2006. Non-target risk assessment of GM crops and regulation. IOBC/wprs Bulletin 29, 197–200.

Rose RI, 2007. White paper on tier-based testing for the effects of proteinaceous insecticidal plant-incorporated protectants on non-target arthropods for regulatory risk assessments. Biotechnology Regulatory Services, Biopesticides and Pollution Prevention Division, 40 pp.

Rosi-Marshall EJ, Tank JL, Royer TV, Whiles MR, Evans-White M, Chambers C, Griffiths NA, Pokelsek J and Stephen ML, 2007. Toxins in transgenic crop byproducts may affect headwater stream ecosystems. Proceedings of the National Academy of Sciences of the United States of America 104, 16204–16208.

USDA-NASS, 2018, online. Quick Stats. Accessed on July 6, 2018. Available at http://quickstats.nass.usda.gov/.

U.S. EPA, 2015. Draft Biopesticides Registration Action Document Plant-Incorporated Protectants *Bacillus thuringiensis* Cry1A.105 and Cry2Ab2 Proteins and the Genetic Material (Vector PV-GMIR13196) Necessary For Their Production in MON 87751 (OECD Unique ID. MON-87751-7) Soybean (PC Code 006614 and PC Code 006527).

U.S. EPA, Health Canada Pest Management Regulatory Agency and California Department of Pesticide Regulation, 2014. Guidance for assessing pesticide risks to bees. available at https://www.epa.gov/sites/production/files/2014-06/documents/pollinator_risk_assessment_guidance_06_19_14.pdf.

Appendix 9. Field assessment of potential impacts of GMB151 soybean on free-living nematodes

List of Tables

List of Figures

1. Introduction: Soil nematodes and their role as beneficial organisms in agriculture

The phylum Nematoda represents a significant portion of the quantity and diversity of microfauna in soils (Yeates, 1981; Bongers and Bongers, 1998). Soil nematode diversity is approached from a functional group perspective, based on the feeding behavior of the families present. Nematode taxonomy describes eight functional groups: 1) plant feeding, of which all PPNs are classified, 2) hyphal feeding or fungal feeders, 3) bacteria feeding, 4) substrate ingestion or incidental ingestion of bacteria, other microfauna, or unicellular eukaryotes, 5) animal predation or the consumption of other soil microfauna including nematodes, 6) unicellular eukaryote feeders including algae, diatoms, fungal spores, or whole yeast, 7) animal parasites in their dispersal or infective stages, and 8) omnivores, which is inclusive of functional groups 2, 3, 4, 5, and 6 (Yeates et al., 1993). Individual nematode species may occupy a functional group for their entire life cycle or one or more functional groups depending on life-stage. Of these functional groups, 2, 3, 4, 5, 6, and collectively 8 are considered FLNs because they do not depend on parasitism of plants or animals for survival. Annual cropping systems are typically dominated by two functional groups: 1) plant feeding and 3) bacteria feeding. Depending on the crop management system (e.g., organic/conventional, tillage intensity, fertilizer amendments, pesticide inputs), climate, or time of the cropping season, soil communities may have higher or lower proportions of other functional groups (Neher, 2010).

In soil ecosystems, FLNs affect the overall soil-flora-fauna continuum through several mechanisms. FLN consumption of soil bacteria, fungi, microfauna, and unicellular eukaryotes releases and cycles organically bound nitrogen into plant available forms (Ingham et al., 1985; Neher, 2010). There are also varying levels of evidence that robust FLN populations can suppress plant pathogens and diseases (Ishibashi, 1998; Neher, 2010; Lagerlöf et al., 2011; Löbmann et al., 2016) by consumption of pathogenic species (fungal, bacterial, or plant-parasitic microfauna),

competition, and eliciting plant resistance. Diversity, functional group population dynamics, and overall quantity can be indicators of soil health (Neher, 2001, 2010).

Cry14Ab-1 protein has been shown to be active against *C. elegans in vitro* (Appendix 3). Therefore, a field assessment of the potential impact of the cultivation of GMB151 soybean on FLNs was conducted.

2. Field assessment methods

The soil nematode community was sampled from a no-till soybean trial in Fruitland, IA. The field trial site in Fruitland, IA is a fine sand soil with a 7.5 pH and 1.2% organic matter. The soybean trial site had been planted as a randomized complete block design for two consecutive years. Treatments were not re-randomized in between years; therefore, the impact of multiple years of cultivation of GMB151 soybean on the soil nematode community could be assessed. The trial included four treatments, which were the fully crossed combination of two treatment factors, PI 88788-derived native SCN resistance (resistant or susceptible) and GMB151 zygosity (homozygous or nullizygous). Therefore, the trial contained four genetically related soybean lines that differed in the presence or absence of native SCN resistance and GMB151.

Soil samples were taken in 2019 during the podfill stage (R5/R6). This growth stage is just prior to peak biomass of the soybean plant (Pedersen, 2004). The nematode community was sampled from two distinct areas, bulk soil and rhizosphere soil. One bulk soil sample and one rhizosphere sample were collected from each plot. Bulk soil was sampled from the center of two soybean rows planted 76.2 cm apart. Ten soil cores 2.5 cm in diameter and 15–20 cm in depth were taken and thoroughly mixed for each bulk soil sample. This area was 38.1 cm from each soybean row and would have substantially less root growth than the soybean row. Secondary and tertiary roots from the current soybean crop would be present in low frequency, with > 90% of root biomass present in the area occurring in the top 15 cm of soil (Mitchell and Russell, 1971). SCN populations have been shown to be significantly lower in the fallow area between soybean rows, a good indication of lower root growth activity (Francl, 1986). Nematode communities in this region would still interact with soybean roots and be able to colonize the soybean rhizosphere around the taproot in the current or future seasons.

Rhizosphere soil samples were taken by digging up soybean plants with a shovel. The soybean plant was grasped by the stem and gently shaken to dislodge any bulk soil before placing inside a 20 L bucket. The root systems of the soybean plants were then vigorously shaken against the inside of the bucket to dislodge the soil adhered to the roots. The dislodged soil was considered rhizosphere soil. The rhizosphere soil was sampled from six plants per plot. The soil from each of the six plants was combined to form a single sample per plot.

Nematodes were extracted from soil samples using the sugar centrifugation method (Jenkins, 1964). All nematodes within a sample were identified to genus and enumerated under 40x magnification (Table A9.1). Each FLN genus was assigned to a trophic group according to Yeates et al. (1993) (Figure A9.1). Population densities for each trophic group are reported as nematodes per 100 cc of soil. In addition, potential changes in the nematode functional community were evaluated by assigning each genus of FLN to a functional guild according to the methods of Ferris et al. (2001). The maturity index (MI), enrichment index (EI), structure index (SI), and channel index (CI) were calculated to evaluate whether GMB151 functioned as a potential disturbance that would affect food web characteristics (Table A9.2). These indices delve deeper than simple

trophic group abundance and attempt to describe how soil food webs are functioning based on both trophic group abundance and the relative *r* versus *K* life strategy characteristics of the taxa present. *K*-selected taxa have few offspring to which many resources are devoted. These organisms perform better in stable environments. In contrast, *r*-selected strategists have many offspring with few resources devoted to each offspring. These taxa perform well in disturbed or unstable environments.

The maturity index measures the FLN community across the spectrum from *r* to *K* life strategies based on the abundance of each taxon within the community and its placement on the colonizer-persister index. Taxa are assigned a number one through five on the colonizer-persister index based on life strategy traits including duration of life, population growth rate, responsiveness to resources, and sensitivity to disturbances (Bongers and Bongers, 1998).

The enrichment index is a measure of how responsive the bacterivore and fungivore community is expected to be to nutrient enrichment (Ferris et al., 2001). Communities dominated by primary detrital consumers (Ba_1 and Fu_2) have higher enrichment indices.

The channel index compliments the enhancement index by indicating the predominant decomposition pathway at work in the FLN community. Higher channel index scores indicate greater participation of fungivores in primary decomposition pathways (Ferris et al., 2001).

The structural index attempts to measure the complexity of the soil food web based on the functional guilds and colonizer-persister classification of the members of the community. Communities that are considered more basal, indicating a stressed environment, have low structural index scores. In contrast, stable communities have high structural index scores.

Data on trophic group abundance were analyzed separately for each trophic group using a mixed effects ANOVA (Table A9.3). Each community index was also analyzed separately using the same model (Table A9.4). For each analysis, the effects of native SCN resistance, GMB151, and soil location (bulk or rhizosphere) were assessed. Experimental repetition, native SCN resistance, and GMB151 zygosity were considered whole plot effects while soil location was considered a split-plot effect.

3. Results and conclusions

For all four trophic groups, soil location significantly affected population densities (Table A9.3). FLN population densities of all four trophic groups were significantly higher in rhizosphere than bulk soil. Nematode population densities of all four trophic groups were unaffected by GMB151. None of the community indices were affected by the presence of GMB151 (Table A9.4). The interaction of soil location and native SCN resistance affected both the maturity index and structure index. In both cases, the presence of native SCN resistance led to a more basal community being present in the rhizosphere.

These results indicate that GMB151 does not significantly affect the FLN community in agricultural production fields.

Table A9.1. List of nematode genera identified in soil samples from Fruitland, IA field plots

Genera	FT [a]	c-p [b]	FG [c]	Genera	FT	c-p	FG
Achromadora	8	3	Om_3	*Ironus*	5a	4	Pr_4
Acrobeles	3	2	Ba_2	*Labronema*	8	4	Om_4
Acrobeloides	3	2	Ba_2	*Longidorus*	1d	PPI_5	PPI_5
Akrotonus	5	5	Pr_5	*Lordellonema*	8	4	Om_4
Alaimus	3	4	Ba_4	*Mesodorylaimus*	8	4	Om_4
Aphelenchoides	2	2	Fu_2	*Mesorhabditis*	3	1	Ba_1
Aphelenchus	2	2	Fu_2	*Metaporcelaimus*	8	5	Om_5
Aporcelaimellus	5	5	Pr_5	*Microdorylaimus*	8	4	Om_4
Aporcelaimus	5	5	Pr_5	*Mylonchulus*	5a	4	Pr_4
Aquatides	5	5	Pr_5	*Nygolaimus*	5	5	Pr_5
Axonchium	8	5	Om_5	*Panagrolaimus*	3	1	Ba_1
Belondira	8	5	Om_5	*Paraphelenchus*	2	2	Fu_2
Cephalobus	3	2	Ba_2	*Paraxonchium*	8	5	Om_5
Chiloplacus	3	2	Ba_2	*Pelodera*	3	1	Ba_1
Clarkus	5a	4	Pr_4	*Plectus*	3	2	Ba_2
Crassolabium	8	4	Om_4	*Pratylenchus*	1b	PPI_3	PPI_3
Deficephalobus	3	2	Ba_2	*Prismatolaimus*	8	3	Om_3
Discolaimus	5	5	Pr_5	*Pristionchus*	8	1	Om_1
Ditylenchus	2	2	Fu_2	*Protorhabditis*	3	1	Ba_1
Ecumenicus	8	4	Om_4	*Pseudacrobeles*	3	2	Ba_2
Eucephalobus	3	2	Ba_2	*Rhabditis*	3	1	Ba_1
Eudorylaimus	5	4	Pr_4	*Rotylenchulus*	1a	PPI_3	PPI_3
Eumonhystera	3	2	Ba_2	*Solididens*	5	5	Pr_5
Filenchus	1e	PPI_2 [d]	PPI_2	*Tripyla*	5a	3	Pr_3
Geomonhystera	3	2	Ba_2	*Tylencholaimus*	2	4	Fu_4
Helicotylenchus	1c	PPI_3	PPI_3	*Tylocephalus*	3	2	Ba_2
Heterodera	1a	PPI_3	PPI_3				

[a] FT = feeding type as described by Yeates et al. (1993). 1 = plant feeders (1a = sedentary parasites, 1b = migratory endoparasites, 1c = semi-endoparasites, 1d = ectoparasites, 1e = epidermal cell and root hair feeders). 2 = hyphal feeders, 3 = bacterial feeders, 5 = predators (5a = ingesters), 8 = omnivores.

[b] c-p = colonizer-persister index as defined by Bongers and Bongers (1998).

[c] FG = functional guild according to Ferris et al. (2001).

[d] Plant-parasitic nematodes were classified according to a plant-parasitic index (PPI) instead of a c-p index and feeding guilds.

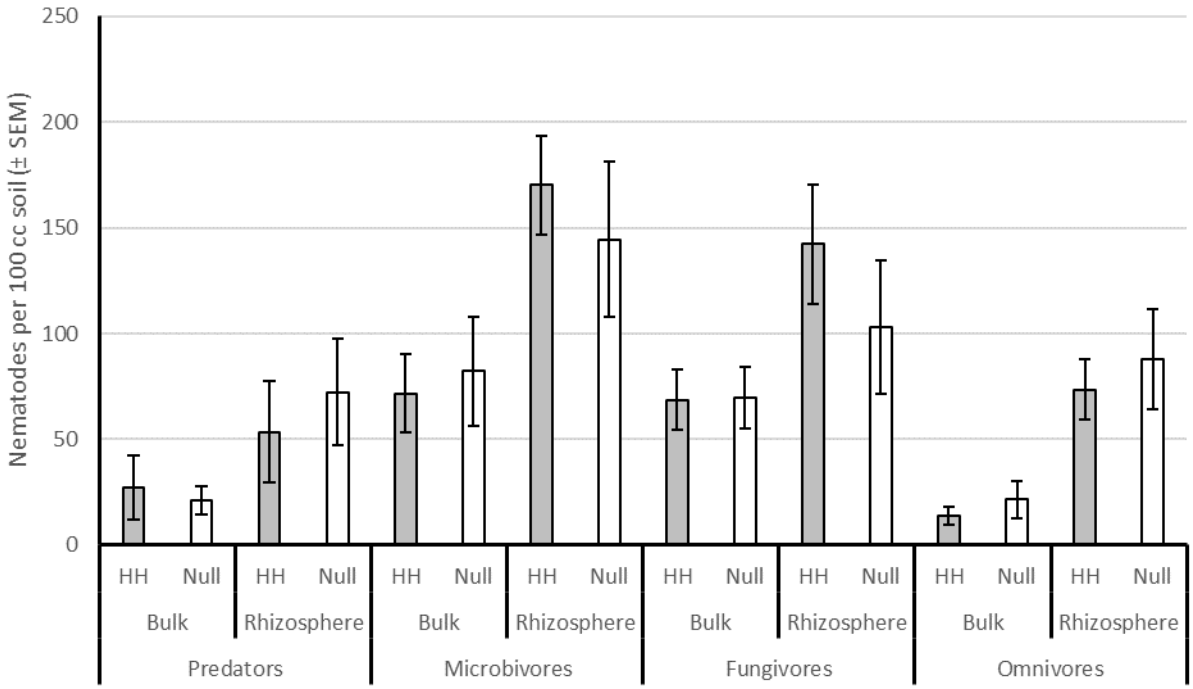


Figure A9.1. Population densities of free-living nematode trophic groups as affected by GMB151

Population densities of free-living nematodes were assessed by trophic group. Soil was sampled from a soybean field in Fruitland, IA planted with soybeans either homozygous (HH) or nullizygous (Null) for the transgenic trait GMB151. Soil was collected from both the soybean rhizosphere and bulk soil. Bulk soil was taken from the fallow ground equidistant between two soybean rows.

Table A9.2. Free-living nematode community indices as affected by soil location and the presence of GMB151

	Genera	MI [a]	EI [b]	SI [b]	CI [b]
Rhizosphere Soil [c]					
Homozygous	19.4 ± 5.2	3.07 ± 0.52	59.3 ± 25.9	87.1 ± 9.4	22.1 ± 33.1
Nullizygous	18.9 ± 2.3	3.11 ± 0.46	62.5 ± 21.7	89.8 ± 6.4	10.7 ± 15.1
Bulk Soil [c]					
Homozygous	16.0 ± 4.1	3.22 ± 0.29	58.3 ± 19.9	88.0 ± 6.1	21.3 ± 16.7
Nullizygous	16.6 ± 2.7	3.12 ± 0.52	63.1 ± 18.9	88.3 ± 5.7	12.5 ± 8.7

[a] Maturity index (MI) was calculated using the colonizer-persister scale according to Bongers and Bongers (1998).
[b] Enhancement index (EI), structure index (SI), and channel index (CI) were calculated by classifying nematode taxa into functional guilds according to Ferris et al. (2001), using the colonizer-persister scale of Bongers and Bongers (1998) and the feeding type classifications of Yeates et al. (1993).
[c] Rhizosphere soil was collected from the roots of soybean plants. Bulk soil was collected from the fallow ground between two planted soybean rows (38.1 cm distant from soybean rows).

Table A9.3. Analysis of variance tables of treatment effects on trophic group abundances of free-living nematodes

Effect	df [a]	F-statistic	p-value
Predator Abundance			
Repetition	3, 9	2.14	0.17
Native Resistance [b]	1, 9	0.30	0.60
Zygosity [c]	1, 9	0.09	0.78
Native Resistance*Zygosity	1, 9	2.19	0.17
Soil Location [d]	1, 12	12.00	< 0.01*
Soil Location*Native Resistance	1, 12	0.36	0.56
Soil Location*Zygosity	1, 12	1.23	0.29
Soil Location*Native Resistance*Zygosity	1, 12	1.32	0.27
Microbivore Abundance			
Repetition	3, 9	1.64	0.25
Native Resistance	1, 9	1.18	0.31
Zygosity	1, 9	0.06	0.82
Native Resistance*Zygosity	1, 9	0.31	0.59
Soil Location	1, 12	22.57	< 0.01*
Soil Location*Native Resistance	1, 12	0.74	0.41
Soil Location*Zygosity	1, 12	1.13	0.31
Soil Location*Native Resistance*Zygosity	1, 12	0.04	0.85
Fungivore Abundance			
Repetition	3, 9	4.25	0.04*
Native Resistance	1, 9	4.72	0.06
Zygosity	1, 9	0.82	0.39
Native Resistance*Zygosity	1, 9	0.09	0.77
Soil Location	1, 12	10.40	< 0.01*
Soil Location*Native Resistance	1, 12	0.69	0.42
Soil Location*Zygosity	1, 12	1.44	0.25
Soil Location*Native Resistance*Zygosity	1, 12	0.36	0.56
Omnivore Abundance			
Repetition	3, 9	1.48	0.29
Native Resistance	1, 9	0.60	0.46
Zygosity	1, 9	0.39	0.55
Native Resistance*Zygosity	1, 9	0.36	0.56
Soil Location	1, 12	33.60	< 0.0001*
Soil Location*Native Resistance	1, 12	0.82	0.38
Soil Location*Zygosity	1, 12	0.09	0.77
Soil Location*Native Resistance*Zygosity	1, 12	0.09	0.77

* Significant effect at $p < 0.05$.
[a] Numerator and denominator degrees of freedom.
[b] Presence or absence of the *rhg1b* SCN-resistance allele from PI 88788.
[c] Presence or absence of the transgenic trait GMB151.
[d] Location of soil sample taken (rhizosphere or bulk soil).

Table A9.4. Analysis of variance tables of treatment effects on community indices of free-living nematodes

Effect	df [a]	F-statistic	p-value
Maturity Index			
Repetition	3, 9	5.36	0.02*
Native Resistance [b]	1, 9	1.83	0.21
Zygosity [c]	1, 9	0.04	0.85
Native Resistance*Zygosity	1, 9	0.59	0.46
Soil Location [d]	1, 12	0.55	0.47
Soil Location*Native Resistance	1, 12	10.47	< 0.01*
Soil Location*Zygosity	1, 12	0.42	0.53
Soil Location*Native Resistance*Zygosity	1, 12	0.05	0.82
Enhancement Index			
Repetition	3, 9	9.76	< 0.01*
Native Resistance	1, 9	5.19	0.05
Zygosity	1, 9	0.86	0.38
Native Resistance*Zygosity	1, 9	0.54	0.48
Soil Location	1, 12	< 0.01	0.98
Soil Location*Native Resistance	1, 12	1.06	0.32
Soil Location*Zygosity	1, 12	0.01	0.91
Soil Location*Native Resistance*Zygosity	1, 12	0.75	0.40
Structure Index			
Repetition	3, 9	2.72	0.11
Native Resistance	1, 9	0.47	0.51
Zygosity	1, 9	0.37	0.56
Native Resistance*Zygosity	1, 9	0.22	0.65
Soil Location	1, 12	0.02	0.89
Soil Location*Native Resistance	1, 12	5.66	0.03*
Soil Location*Zygosity	1, 12	0.40	0.54
Soil Location*Native Resistance*Zygosity	1, 12	1.97	0.19
Channel Index			
Repetition	3, 9	1.26	0.34
Native Resistance	1, 9	0.58	0.47
Zygosity	1, 9	2.82	0.13
Native Resistance*Zygosity	1, 9	2.63	0.14
Soil Location	1, 12	0.78	0.40
Soil Location*Native Resistance	1, 12	0.34	0.57
Soil Location*Zygosity	1, 12	1.41	0.26
Soil Location*Native Resistance*Zygosity	1, 12	0.48	0.50

* Significant effect at $p < 0.05$.
[a] Numerator and denominator degrees of freedom.
[b] Presence or absence of the *rhg1b* SCN-resistance allele from PI 88788.
[c] Presence or absence of the transgenic trait GMB151.
[d] Location of soil sample taken (rhizosphere or bulk soil).

Appendix 9 References

Bongers T and Bongers M, 1998. Functional diversity of nematodes. Applied Soil Ecology 10, 239–251.

Ferris H, Bongers T and de Goede RGM, 2001. A framework for soil food web diagnostics: extension of the nematode faunal analysis concept. Applied Soil Ecology 18, 13–29.

Francl LJ, 1986. Spatial analysis of *Heterodera glycines* populations in field plots. Journal of Nematology 18, 183–189.

Ingham RE, Trofymow JA, Ingham ER and Coleman DC, 1985. Interactions of bacteria, fungi, and their nematode grazers: Effects on nutrient cycling and plant growth. Ecological Monographs 55, 119–140.

Ishibashi N, 1998. Possible integration of beneficial nematodes and microbial agents for biocontrol of soil pests. Japanese Journal of Nematology 28, 22–28.

Jenkins WR, 1964. A rapid centrifugal-flotation technique for separating nematodes from soil. Plant Disease Reporter 48, 692.

Lagerlöf J, Insunza V, Lundegårdh B and Rämert B, 2011. Interaction between a fungal plant disease, fungivorous nematodes and compost suppressiveness. Acta Agriculturae Scandinavica Section B – Soil and Plant Science 61, 372–377.

Löbmann MT, Vetukuri RR, de Zinger L, Alsanius BW, Grenville-Briggs LJ and Walter AJ, 2016. The occurrence of pathogen suppressive soils in Sweden in relation to soil biota, soil properties, and farming practices. Applied Soil Ecology 107, 57–65.

Mitchell RL and Russell WJ, 1971. Root development and rooting patterns of soybean (*Glycine max* (L.) Merrill) evaluated under field conditions. Agronomy Journal 63, 313–316.

Neher DA, 2001. Role of nematodes in soil health and their use as indicators. Journal of Nematology 33, 161–168.

Neher DA, 2010. Ecology of plant and free-living nematodes in natural and agricultural soil. Annual Review of Phytopathology 48, 371–394.

Pederson P, 2004, online. Soybean growth and development. Iowa State University Extension, Ames, IA. available at https://crops.extension.iastate.edu/files/article/SoybeanGrowthand Development_0.pdf

Yeates GW, 1981. Nematode populations in relation to soil environmental factors: A review. Pedobiologia 22, 312–338.

Yeates GW, Bongers T, de Goede RGM, Freckman DW and Georgieva SS, 1993. Feeding habits in soil nematode families and genera—An outline for soil ecologists. Journal of Nematology 25, 315–331.