Application to Amend the Australia New Zealand Food Standards Code
Schedule 26 - *Food Produced Using Gene Technology*

OECD Unique Identifier: COR-23134-4

COR23134 Soybean

Submitting company:
Corteva Agriscience Australia Pty Ltd

Submitted by:
[redacted]

[redacted]
Regulatory Manager ANZ
Corteva Agriscience Australia Pty Ltd (ABN 24 003 771 659)
Level 9, 67 Albert Avenue
CHATSWOOD NSW 2067 Australia
Telephone: [redacted]
Email: [redacted]

August 2024

SUMMARY

Corteva, Inc. is a publicly traded, global pure-play agriculture company that combines industry-leading innovation, high-touch customer engagement and operational execution to profitably deliver solutions for the world's most pressing agriculture challenges. Corteva generates advantaged market preference through its unique distribution strategy, together with its balanced and globally diverse mix of seed, crop protection, and digital products and services. With some of the most recognized brands in agriculture and a technology pipeline well positioned to drive growth, the company is committed to maximizing productivity for farmers, while working with stakeholders throughout the food system as it fulfills its promise to enrich the lives of those who produce and those who consume, ensuring progress for generations to come. More information can be found at www.corteva.com.

Corteva Agriscience Australia Pty Ltd, member of Corteva Agriscience group of companies, and its affiliated companies (herein referred to collectively as Corteva), is submitting this application to FSANZ to vary the Code to approve food uses of insect-resistant soybean (*Glycine max* [L.] Merrill) event COR-23134-4 (referred to as COR23134 soybean), a new food produced using gene technology.

COR23134 soybean was genetically modified to expresses the Cry1B.34.1, Cry1B.61.1, and IPD083Cb proteins for control of certain susceptible lepidopteran pests, and the GM-HRA protein that was used as a selectable marker. The GM-HRA protein present in COR23134 soybean is found in the approved soybean event DP-3Ø5423-1, which was subject to the A1018 application in 2009.

This application presents information supporting the safety and nutritional comparability of COR23134 soybean to conventional soybean. The molecular characterization analyses conducted on COR23134 soybean demonstrated that the introduced genes are integrated at a single locus, stably inherited across multiple generations, and segregate according to Mendel's law of genetics. The allergenic and toxic potential of the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins were evaluated, and these proteins were found unlikely to be allergenic or toxic to humans. A compositional comparability assessment demonstrated that the nutrient composition of COR23134 soybean forage and grain is comparable to that of conventional soybean, represented by non-genetically modified (non-GM) near-isoline soybean and non-GM commercial soybean.

Overall, data and information contained herein support the conclusion that COR23134 soybean, containing the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins, is as safe and nutritious as non-GM soybean.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHECKLISTS

		General requirements (3.1.1)
Check	**Page No.**	**Mandatory requirements**
☑		A Form of application ☑ *Application in English* ☑ *Executive Summary (separated from main application electronically)* ☑ *Relevant sections of Part 3 clearly identified* ☑ *Pages sequentially numbered* ☑ *Electronic copy (searchable)* ☑ *All references provided*
☑	14	B Applicant details
☑	14	C Purpose of the application
☑	15	D Justification for the application ☑ *Regulatory impact information* ☑ *Impact on international trade*
☑		E Information to support the application ☑ *Data requirements*
☑		F Assessment procedure ☑ *General* *Major* *Minor* *High level health claim variation*
☑		G Confidential commercial information ☑ *CCI material separated from other application material* ☑ *Formal request including reasons* ☑ *Non-confidential summary provided*
☑		H Other confidential information ☑ *Confidential material separated from other application material* ☑ *Formal request including reasons*
☐		I Exclusive Capturable Commercial Benefit ☐ *Justification provided*

		J International and other national standards
☑		☑ *International standards*
		Other national standards
☑	13	K Statutory Declaration
		L Checklist/s provided with application
☑	11	☑ *3.1.1 Checklist*
		☑ *All page number references from application included*
		☑ *Any other relevant checklists for Chapters 3.2–3.7*

		Foods produced using gene technology (3.5.1)
Check	**Page No.**	**Mandatory requirements**
☑	18	A.1 Nature and identity
☑	19	A.2 History of use of host and donor organisms
☑	24	A.3 Nature of genetic modification
☑	61	B.1 Characterisation and safety assessment
☑	61	B.2 New proteins
☑	154	B.3 Other (non-protein) new substances
☑	154	B.4 Novel herbicide metabolites in GM herbicide-tolerant plants
☑	154	B.5 Compositional analyses
☑	178	C Nutritional impact of GM food
☑	179	D Other information

STATUTORY DECLARATION

Statutory Declarations Act 1959 [1]

I, [redacted] Regulatory Manager of Corteva Agriscience, Level 9, 67 Albert Ave, Chatswood, NSW 2067 make the following declaration under the *Statutory Declarations Act 1959*:

1. the information provided in this application fully sets out the matters required
2. the information provided in this application is true to the best of my knowledge and belief
3. no information has been withheld that might prejudice this application, to the best of my knowledge and belief

I understand that a person who intentionally makes a false statement in a statutory declaration is guilty of an offence under section 11 of the *Statutory Declarations Act 1959*, and I believe that the statements in this declaration are true in every particular.

[redacted]

[*Signature of person making the declaration*]

Declared at Chatswood on 9th of February 2022

Before me,

[redacted]

Legal Practitioner
Corteva Agriscience, Level 9, 67 Albert Ave, Chatswood, NSW 2067

---

[1] http://www.comlaw.gov.au/Series/C1959A00052.

GENERAL INFORMATION ON THE APPLICATION

The chapter numbering follows section numbers from the FSANZ Application Handbook (Chapters 3.1 and 3.5.1).

B. Applicant

This application is submitted by:

Corteva Agriscience Australia Pty Ltd
Member of Corteva Agriscience group of companies
Level 9, 67 Albert Ave
Chatswood
NSW 2067

The primary contact is:

[redacted]
Regulatory Manager
Corteva Agriscience
Ph: [redacted]
Email: [redacted]

The Managing Director of Corteva Agriscience Australia Pty Ltd is:

[redacted]
Ph: [redacted]
Email: [redacted]

C. Purpose of the application

Corteva Agriscience Australia Pty Ltd, member of Corteva Agriscience group of companies, and its affiliated companies (herein referred to collectively as Corteva), has developed COR23134 soybean (OECD Unique Identifier COR-23134-4), a new event that has been transformed to express the Cry1B.34.1, Cry1B.61.1, and IPD083Cb proteins for control of certain susceptible lepidopteran pests, and the GM-HRA protein that was used as a selectable marker.

As a result of this application, Corteva seeks an amendment of Standard 1.5.2 *Food produced using gene technology* by inserting the following into table to Schedule 26-3(4) after the last entry: insect-protected soybean line COR23134.

D. Justification for the application

D(a) Need for the proposed change

Corteva is a member of Excellence Through Stewardship™ (ETS). Corteva has developed the new soybean event COR23134, which is being commercialized in accordance with the ETS Product Launch Stewardship Guidance and in compliance with Corteva polices regarding stewardship of GM products. In line with these guidelines, Corteva's process for launches of new products includes a longstanding process to evaluate export market information, value chain consultations, and regulatory functionality. Corteva's application to amend Standard 1.5.2 with respect to COR23134 soybean is in support of these policies.

D(b) Advantage of the genetically modified food

Soybean is a globally traded commodity produced in both temperate and tropical regions and serves as a key source of vegetable oils and protein. The introduction of insect-resistant COR23134 soybean is intended to help growers keep pace with increasing soybean demand globally. Soybean is grown as a commercial crop in over 35 countries. Brazil, United States, and Argentina together produce about 80% of the world's soybean. Soybean is grown primarily to produce grain for further processing, has a multitude of uses in the food, feed, and industrial sectors, and represents one of the major sources of edible vegetable oil and of proteins for livestock feed use. The total world production for soybeans in 2023 was approximately 397 million metric tonnes (t) (USDA-FAS, 2024).

Insect Resistance

Certain lepidopteran insects are serious pests of soybean in various geographies. Control of lepidopteran soybean pests has historically been managed with crop rotation, broad-spectrum insecticides, and in certain geographies transgenic crops expressing crystalline (Cry) proteins. As adoption of Bt soybean has increased, the selection pressure on target insects to develop resistance has become greater. Insect resistance to transgenic traits can reduce the efficacy of the traits over time, increasing costs of soybean production and/or reducing yield. The novel IPD083Cb protein in COR23134 soybean can serve as an alternative to traditional *Bt*-based insect control traits with potential to counter insect resistance to *Bt* proteins.

Selectable Marker

The GM-HRA protein was incorporated into COR23134 soybean to enable selection of plants containing the desired construct during the event development process.

D.1 Regulatory impact

Corteva have developed the new soybean line COR23134, which will be commercialized in accordance with the ETS Product Launch Stewardship Guidance and in compliance with Corteva polices regarding stewardship of GM products. In line with these guidelines, Corteva's approach to responsible launches of new products includes a longstanding process to evaluate export market information, value chain consultations, and regulatory functionality. Growers and end-users must take all steps within their control to follow appropriate stewardship requirements and confirm their buyer's acceptance of the seed or other material being purchased.

The area planted to soybeans worldwide has expanded rapidly due to development of varieties suited to regional planting conditions, rising yields, low production costs and global demand. Approximately 37 million hectares (ha) of soybeans were harvested in North America in 2022 (FAOSTAT 2024). South America has seen dramatic increases in soybean production:

Argentina soybean production increased from 8,637,500 ha in 2000 to 15,874,300 ha in 2022; Brazil soybean production increased from 13,656,800 ha to 40,895,000 ha in the same period (FAOSTAT 2024). China soybean production was 10,243,000 ha in 2022 (FAOSTAT 2024). Australia's soybean production is relatively small, 25,500 ha (57,200 t) in 2022 (FAOSTAT 2024). No soybean is grown in New Zealand (FAOSTAT 2024). Imports into Australia were 2,789 t of soybeans, 24,500 t of soybean oil and 22,150 t of soy sauce in 2022 (FAOSTAT 2024). Imports into New Zealand were 2,350 t of soybeans, 12,000 t of soybean oil and 4,010 t of soy sauce in 2022 (FAOSTAT 2024). Australia imports soybeans primarily from China and United States. New Zealand imports soybeans primarily from Canada and Australia (The Observatory of Economic Complexity (oec.world), 2024).

Soybeans are used for many different products, both edible and non-edible. About 85 percent of the world's soybeans are processed, or "crushed," annually into soybean meal and oil. Approximately 98 percent of the soybean meal that is crushed is further processed into animal feed with the balance used to make soy flour and proteins. Of the oil fraction, 95 percent is consumed as edible oil; the rest is used for industrial products such as fatty acids, soaps and biodiesel (Information About Soya, Soybeans (archive.org)). Soy-based human foods include tofu, miso, soy sauce, natto, tempeh, soymilk, soy flour, soy oil, concentrates and isolates, and soy sprouts (Soy Australia, 2024).

D.1.1 Costs and benefits for industry, consumers and government

Corteva launches new products in accordance with the Corteva Product Launch Policy and Excellence Through Stewardship Product Launch Guidance. Corteva's long-standing, multi-faceted approach includes evaluating export market information, performing value-chain consultations and consideration of regulatory functionality. Innovative technologies like COR23134 soybean are designed to deliver exceptional value and needed performance to the farmers that produce seed from these products, along with helping farmers provide enough safe, nutritious food to meet global demand. In line with these guidelines, Corteva's approach to responsible launches of new products includes a long-standing process to evaluate export market information, value chain consultations, regulatory functionality, preparedness to meet product ramp up and demand plans, and other factors. Corteva continues to advocate for a global synchronous, science-based and predictable regulatory system. Corteva also encourages farmers, industry, and consumer groups to continue to advocate for the acceptance of new, innovative technologies that help to improve farm productivity and profitability and contribute to the global economy and environmental sustainability.

Corteva does not develop nor import food or feed products into the Australian or New Zealand markets. The proposed amendment to the Standard, however, may result in increasing Australia and New Zealand's access to international soybean seed food markets while supporting Corteva's sale of seed in markets where COR23134 soybean is to be cultivated. In this sense, and in an effort to maintain transparency with FSANZ, Corteva acknowledges that there may be a capturable commercial benefit to Corteva as defined in Section 8 of the FSANZ Act. Any relevant local costs are made up of Corteva personnel time both locally and globally as well as of the direct fees associated with the submission.

Domestic production of soybean in Australia is supplemented by importation of soybean meal and oil. New Zealand fully relies on imports. The proposed variation to the Standard permits importation and use of food derived or developed from COR23134 soybean. This offers benefits to the industry and consumers in Australia and New Zealand, which result from the advantages of COR23134 soybean availability to growers in cultivation countries (see Section D(b) Advantage of the genetically modified food of the dossier above).

While Corteva does not possess quantitative data, which would allow it to estimate the benefits in monetary terms, COR23134 soybean is anticipated to contribute to the maintenance of stable global soybean supply, choice and affordability for consumers. No specific costs associated with the approval of COR23134 soybean for Australian and New Zealand consumers have been identified.

Similarly, an analysis in monetary terms for the food industry is hard to determine, however, Australian and New Zealand importers are expected to benefit from trade access, which the approval of COR23134 soybean will support (see also Section D.1.2 *Impact on international trade* below). Compliance with import requirements is also anticipated to be simplified when sourcing from markets in which COR23134 soybean is commercialized. The only identified cost associated with the approval of COR23134 soybean for Australian and New Zealand industry is meeting their GM labelling requirements for those foods derived from COR23134 soybean which trigger them, similarly to other existing GM soybean varieties.

No dollar value of the costs and benefits for the governments can be assigned with the available information. However, from the government perspective, approval of COR23134 soybean will support global regulatory harmonization and limit potential instances of non-compliance related to the regulation of GM foods. No costs associated with the approval of COR23134 soybean for the Australian and New Zealand governments have been identified.

D.1.2 Impact on international trade

The addition of COR23134 soybean to Schedule 26 is anticipated to facilitate import access to soybean and its products from the applicable cultivation countries. Without such an approval, grain handlers may undertake a scientifically unnecessary and costly activities to segregate COR23134 soybean and food products derived from it for Australian and New Zealand markets. Therefore, amending the Food Code to include COR23134 soybean is anticipated to have a positive impact on Australian and New Zealand access to international commodity trade markets.

A. TECHNICAL INFORMATION ON THE FOOD PRODUCED USING GENE TECHNOLOGY

A.1 Nature and identity of the genetically of the genetically modified food

A.1 (a) Description of the GM organism, nature and purpose of the genetic modification

Soybean (*Glycine max* [L.] Merr.) event COR-23134-4 (referred to as COR23134 soybean) was genetically modified to expresses the Cry1B.34.1, Cry1B.61.1, and IPD083Cb proteins for control of certain susceptible lepidopteran pests and the GM-HRA protein that was used as a selectable marker.

The Cry1B.34.1 protein is encoded by the *cry1B.34.1* gene, a gene composed of sequences from a *cry1B*-class gene and the *cry1Ca1* gene, both derived from *Bacillus thuringiensis* (*Bt*). The expressed Cry1B.34.1 protein binds to receptors in the brush border membrane of certain susceptible lepidopteran pests and causes cell death through the formation of non-specific, ion conducting pores in the apical membrane of the midgut epithelial cells.

The Cry1B.61.1 protein is encoded by the *cry1B.61.1* gene, a modified *cry1B*-class gene derived from *Bacillus thuringiensis*. The expressed Cry1B.61.1 protein is effective against certain susceptible lepidopteran pests by causing disruption of the midgut epithelial cells.

The IPD083Cb protein is encoded by the insecticidal protein gene, *ipd083Cb*, from giant maidenhair fern (*Adiantum trapeziforme* var. *braziliense*). The expressed IPD083Cb protein is effective against certain susceptible lepidopteran pests by causing disruption of the midgut epithelial cells, providing an alternative to traditional *Bt*-based insect control traits.

The GM-HRA protein is encoded by the *gm-hra_1* gene, a modified acetolactate synthase (*als*) gene from soybean. The expressed GM-HRA protein serves as a selectable marker during transformation which allows for the growth of tissue in the presence of ALS-inhibiting herbicides, e.g., sulfonylureas and triazolopyrimidine.

A.1 (b) GM Organism Identification

In accordance with OECD's "Guidance for the Designation of a Unique Identifier for Transgenic Plants", this event has an OECD identifier of COR-23134-4, also referred to as COR23134 soybean.

A.1 (c) Trade name

Soybean event COR23134 is at a pre-commercialization stage and has not yet been assigned a commercial product name.

A.2 History of use of the host and donor organisms

A.2 (a) Donor organisms

***Bacillus thuringiensis* (*Bt*): donor of the *cry1B.34.1* and *cry1B.61.1* genes**

- Class: Bacillus/Clostridium group (low G+C Gram-positive bacteria)
- Order: Bacillales
- Family: Bacillaceae
- Genus: *Bacillus*
- Species: *B. thuringiensis*

Bt is a diverse group of Gram-positive, spore-forming bacteria that has a history of safe use as a pesticide over several decades (US-EPA, 1998; US-EPA, 2001). It occurs ubiquitously in the soil and on plants including vegetables, cotton, tobacco, tree crops, and forest crops (Schnepf et al., 1998; Shelton, 2012). Several Cry proteins have been deployed as safe and effective pest control agents in microbial *Bt* formulations for almost 40 years. Several Cry proteins have also been effectively deployed as safe and effective pest control agents and have a history of safe use in genetically modified crops (ISAAA, 2023).

***Adiantum trapeziforme*: donor of the *ipd083Cb* gene**

- Class: Polypodiopsida
- Order: Polypodiales
- Family: Pteridaceae
- Genus: *Adiantum L.*
- Species: *A. trapeziforme L.*
- Sub-species: *braziliense*

Adiantum trapeziforme is known as the giant maidenhair fern or diamond maidenhair fern. Ferns are among the oldest living organisms on the planet and, with the exception of Antarctica, are globally distributed (Fernández, 2011). Ferns of the genus *Adiantum L.* are found in temperate and tropical regions worldwide. *A. trapeziforme L.* is native to the tropical rainforests of Central and South America (Kew Science, 2020) and has been introduced in the state of Florida in the United States (USDA-NRCS, 2023).

Humans have used ferns for many applications including occasional sources of food (Simmons and Herman, 2023). Members of the maidenhair fern family and other non-seed plants have been utilized for ethnomedicinal purposes from treating respiratory infections such as cough, colds, to pneumonia with research continuing into the potential benefits of compounds from members of this genus (Rastogi *et al.*, 2018). Many species of genus *Adiantum L.* are used in traditional medicine as infusions, decoctions, or pastes (Rastogi *et al.*, 2018). There are no reports of *A. trapeziforme* being poisonous to humans or livestock.

***Glycine max*: donor of the *gm-hra_1* gene**

- Class: Magnoliopsida
- Order: Fabales

- Family: Fabaceae (also referred to as Leguminosae)
- Genus: *Glycine* (also referred to as *Glycine* Willd.)
- Species: *G. max* (L.) Merr.

Soybean is the world's leading oilseed crop with a long history of use (OECD, 2000; OECD, 2012). Historical and geographical evidence suggests that soybeans were first domesticated in the eastern half of China between the 17th and 11th century B.C. Soybeans were first introduced into the United States, now a major producer, in 1765. Today, soybeans are grown as a commercial crop in over 35 countries worldwide (OECD, 2000). Soybeans have a multitude of uses in the human food, animal feed, and industrial sectors, and represent one of the major sources of edible vegetable oil and of proteins for livestock feed use (CFIA, 2021; OECD, 2000).

A.2 (b) Host organism

This section describes soybean and its characteristics that are relevant to the safety of COR23134 soybean. Discussed within this section include brief overviews of the morphology, natural habitat, mode of reproduction and dispersal, outcrossing rate, and weediness potential of soybean.

Soybean is extensively cultivated in many different agricultural areas worldwide and is the world's leading oilseed crop with a long history of use (OECD, 2000; OECD, 2012). Soybeans or processed fractions are consumed in many different human foods and animal feeds and used for numerous industrial applications (CFIA, 2021; OECD, 2006).

Taxonomy

- Class: Magnoliopsida
- Order: Fabales
- Family: Fabaceae (also referred to as Leguminosae)
- Genus: *Glycine* (also referred to as *Glycine* Willd.)
- Species: *Glycine max* (L.) Merr.
- Common name: soybean; soya bean

Cultivated soybean, *Glycine max* (L.) Merr., is a diploidized tetraploid species with a chromosome number of 2n = 2x = 40, and it is a domesticated species of the wild soybean (*G. soja*) or a *G. soja*/*G. max* complex (Hymowitz, 1970; Kim *et al.*, 2012; Wang *et al.*, 2016). The genus *Glycine* is divided into the subgenera *Glycine* and *Soja*. *G. max* belongs to the subgenus *Soja*, which also contains *G. soja* (Siebold & Zucc.), a wild species of soybean that grows in many Asian countries.

Morphology

Cultivated soybean (*Glycine max* [L.] Merr.) is an annual erect and bushy-type herbaceous plant which can grow up to approximately 60 inches (1.5 meters) in height and has cultivars of all determinate types. Soybean plants have 4 types of leaves: simple cotyledons (seed leaves), primary leaves, foliage leaves, and prophylls (flower bracteole). The pair of cotyledons occur first and are oppositely arranged. The two primary leaves are unifoliate, ovate, and positioned opposite each other at the same node above the cotyledons. Subsequent foliage leaves are trifoliolate and are on alternate sides along the stem but compound leaves with four or more leaflets are also occasionally present. The prophyll is the first leaf of the lateral shoot and occurs as small pairs of simple leaves, found at the base of lateral branches and the lower part of the pedicel of each flower. Soybean has a simple rooting system, consisting of a taproot root that may reach a depth of 80 inches (2 meters) and many lateral roots that may reach 100 inches (2.5 meters) in length. Under typical field conditions, the root system is less extensive and mostly in the top 6 inches (15 centimeters) of the soil. The roots have nodules which are formed by means of nitrogen-fixing bacteria and are characteristic of the family Leguminosae. Soybean has a typical papilionaceous flower with a tubular calyx of five unequal sepal lobes and a five-parted corolla consisting of one posterior standard petal, two lateral wing petals, and two anterior keel petals in contact with each other but not fused. Enclosed in the corolla are one pistil, nine fused stamens, and a single posterior stamen. Soybean pods are usually straight or slightly curved with oval-shaped seeds whose shape depending upon the cultivar can range from round to elongated and flattened (CFIA, 2021; OECD, 2000; OECD, 2006).

Center of Origin

Historical and genetic evidence suggests that soybean was first domesticated in China (Hymowitz, 1970; Sedivy *et al.*, 2017). Wild soybean, *Glycine soja*, is endemic in China, Korea, Japan, Taiwan, and Russia (OECD, 2000; OECD, 2006). *G. soja* is considered to be the closest wild relative of *G. max* (Hymowitz, 1970; Kim *et al.*, 2012; Wang *et al.*, 2016).

Natural Habitat and Generation Time

Soybean is a photoperiod sensitive, short-day plant, and flowering is quicker under short days. Temperature response and photoperiodism determine the cultivar adaptation. Water requirements vary significantly but are highest from flowering to seed filling. Soybeans grow best in neutral or slightly acidic soil and can tolerate a pH range of approximately 5.5 to 7.8. Soybeans are typically grown where growing-season temperatures are 50-100 ºF (10-40 °C). Soybean seeds germinate in 5-7 days under soil temperatures of at least 50 ºF (10 ºC). A root zone temperature of at least 59 ºF (15 °C) is needed for adequate soybean nodulation and nitrogen fixation, while 77 ºF (25 °C) is optimal. Soybean is an annual crop with a cultural cycle of 2-5 months depending on the variety and area of production. The cultivars are classified into 13 maturity groups (MGs) according to the length of time from planting to maturity, with lower-numbered MGs representing earlier maturing varieties. These groupings were identified based on cultivar adaptation within certain latitudes and day length. (CFIA, 2021; OECD, 2000; OECD, 2006). There are distinct soybean cultivars that are well adapted to tropical, sub-tropical, and temperate regions.

Mode of Reproduction and Dispersal

Soybean is a self-pollinated crop, i.e., the anthers mature in the bud and directly pollinate the stigma of the same flower and is propagated by seed. Flowering usually begins 25 to 50 days after planting and lasts 20 to 40 days. The soybean stigma is receptive to pollen approximately

24 hours before anthesis and remains receptive after 48 hours of anthesis (CFIA, 2021; OECD, 2000). As a result, soybean exhibits a high percentage of self-fertilization, and cross-pollination is usually less than 1% (CFIA, 2021). Cross-pollination of soybean plants that are more than 33 feet (10 meters) from the source pollen is absent or very rare. Insects are believed to be responsible for some cross-pollination. As they feed, insects may transfer pollen between soybean flowers. Soybean plants release very little airborne pollen, which does not travel long distances. Thus, wind dispersal is expected to be negligible. A soybean plant can produce as many as 400 pods, with 2-20 pods at a single node and 1-5 seeds each pod. Mature seeds develop from 30 to 50 days after fertilization and can vary in shape and color based on plant genetics and environment. Neither the seed pods nor the seeds have morphological characteristics that would encourage animal transportation (CFIA, 2021; OECD, 2000).

Inter-Specific and Intra-Specific Crosses/ Gene Flow

Cultivated soybean is an annual herb and cannot cross with other perennial species of *Glycine* (subgenus *Glycine*). Soybean can only cross with other members of *Glycine* subgenus *Soja*; however, the potential for such gene flow is limited by geographic isolation (OECD, 2000). Additionally, soybean is a highly self-pollinating plant species, limiting the chances of cross-pollination. Natural cross-pollination to neighboring plants is typically below 1%, making gene introgression very difficult (CFIA, 2021).

Survival, Dormancy, and Weediness/ Invasiveness Potential

Generally, soybeans are incapable of sustained reproduction outside of domestic cultivation and is non-invasive of natural habitats (CFIA, 2021). It is generally recognized that the domestication of crop plants over thousands of generations has resulted in modern crop cultivars that have lost common distinctive attributes of weeds and rarely grow without human intervention. Dissemination of soybean seed can occur either by mechanical harvesting or transportation, but again these seeds generally do not survive without human intervention. Cultivated soybean rarely exhibits seed dormancy and will only grow as a volunteer under certain environmental conditions (OECD, 2000). A lack of dormancy is selected for in commercial soybean seeds, therefore commercial soybean seeds germinate quickly. Frost and/or cold weather often kills volunteer soybean. Even then volunteer soybean are considered weak competitors with the succeeding crop and are readily controlled by mechanical means or other management practices (CFIA, 2021; OECD, 2000).

Information on the host plant's genotype and phenotype relevant to its safety

Soybean is extensively cultivated in many different agricultural areas worldwide and is the world's leading oilseed crop with a long history of use (OECD, 2000; OECD, 2012). Soybeans or processed fractions are consumed in many different human foods and animal feeds and used for numerous industrial applications (CFIA, 2021; OECD, 2012).

Unprocessed soybean is not typically used as human food sources because in part it contains anti-nutrients such as trypsin inhibitors and lectins. However, the processing methods applied to soybean are well known and have a long history of safe use. In fact, soybean is one of the oldest cultivated crops (OECD, 2000).

Conclusion

Soybean is a commonly cultivated crop around the world, and its biology and history of safe use demonstrate that the unmodified organism is safe for human and animal consumption.

A.3 Nature of the genetic modification

A.3 (a) Transformation Method

COR23134 soybean was created by *Agrobacterium*-mediated transformation of a Corteva Agriscience elite soybean variety 93Y21 with plasmid [redacted] (Figure 1; Table 1). Immature soybean cotyledons were inoculated with *Agrobacterium tumefaciens* strain [redacted] containing plasmid [redacted]. *Agrobacterium tumefaciens* strain [redacted] is a disarmed strain that contains the *vir* genes and enables efficient transfer of the transfer DNA (T-DNA) region of the transformation plasmid to the inoculated host plant tissue. After 4 days of co-cultivation with *Agrobacterium* on filter paper, the cotyledons were transferred to a liquid culture medium containing the antibiotic [redacted] for 7 days to kill *Agrobacterium*. Following the recovery period, the cotyledons were transferred to a liquid medium with [redacted] herbicide selection and containing [redacted] to kill residual *Agrobacterium*. After 6-8 weeks in the selection medium, healthy green callus was transferred to a solid maturation medium containing [redacted] and incubated for 3-4 weeks, followed by desiccation of the resulting embryos for 4-7 days. Embryos were then transferred to a solid germination medium containing [redacted] to initiate shoot and root development for 4-6 weeks. Once shoots and roots were established, healthy plants were selected, and PCR was used to confirm the presence of the [redacted] T-DNA insert. Plants that were regenerated from transformation and tissue culture (designated T0 plants) were selected for further characterization. Refer to Figure 9 for a schematic overview of the transformation and event development process for COR23134 soybean. The subsequent breeding of COR23134 soybean proceeded as indicated in Figure 10 to produce specific generations for the characterization and assessments conducted, as well as for the development of commercial soybean lines.

The T-DNA region of plasmid [redacted] contains four gene cassettes (Figure 2; Table 2). The first gene cassette (*cry1B.34.1*) contains the *cry1B.34.1* gene, a gene composed of sequences from a *cry1B*-class gene and the *cry1Ca1* gene, both derived from *Bacillus thuringiensis* (WO Patent 2016061197 (Izumi Wilcoxon and Yamamoto, 2016); GenBank accession CAA30396.1, respectively). The expressed Cry1B.34.1 protein confers control of certain susceptible lepidopteran pests. The Cry1B.34.1 protein is 665 amino acids (aa) in length and has a molecular weight of approximately 75 kDa. Expression of the *cry1B.34.1* gene is controlled by the [redacted] The terminator for the *cry1B.34.1* gene is the [redacted]. Two additional terminator regions from the [redacted]) are present between the first and second gene cassettes. These additional terminators are intended to prevent any potential transcriptional interference. Transcriptional interference is defined as the transcriptional suppression of one gene on another when both are in proximity (Shearwin *et al.*, 2005). The placement of one or multiple transcriptional terminators between gene cassettes has been shown to reduce the occurrence of transcriptional interference (Greger *et al.*, 1998).

The second gene cassette (*cry1B.61.1*) contains the *cry1B.61.1* gene, a modified *cry1B*-class gene, derived from *Bacillus thuringiensis* (WO Patent 2017180715 (Horn *et al.*, 2017)). The expressed Cry1B.61.1 protein confers control of certain susceptible lepidopteran pests. The Cry1B.61.1 protein is 656 aa in length and has a molecular weight of approximately 74 kDa. Expression of the *cry1B.61.1* gene is controlled by the [redacted] The [redacted] terminator [redacted] for the *cry1B.61.1* gene is the [redacted] Two additional terminator regions from the [redacted] are present between the second and third gene cassettes to prevent possible transcriptional interference.

The third gene cassette (*ipd083Cb*) contains the insecticidal protein gene, *ipd083Cb*, from giant maidenhair fern (*Adiantum trapeziforme* var. *braziliense*) (US Patent 10227608 (Barry *et al.*, 2019)). The expressed IPD083Cb protein confers control of certain susceptible lepidopteran pests. The IPD083Cb protein is 853 aa in length and has a molecular weight of approximately 95 kDa. Expression of the *ipd083Cb* gene is controlled by the [redacted] The [redacted] terminator [redacted] for the *ipd083Cb* gene is the [redacted]

The fourth gene cassette (*gm-hra_1*) contains the *gm-hra_1* gene, a modified acetolactate synthase (*als*) gene, from *Glycine max* (US Patent 7834242 (Falco and Li, 2010)). The expressed GM-HRA protein in plant tissue serves as a selectable marker during transformation which allows for the growth of tissue in the presence of ALS-inhibiting herbicides, e.g., sulfonylurea and triazolopyrimidine. The GM-HRA protein is 651 aa in length and has a molecular weight of approximately 70 kDa. Expression of the *gm-hra_1* gene is controlled by the [redacted] The terminator for the *gm-hra_1* gene is the [redacted]

Additional terminator sequences are present adjacent to the Right Border and Left Border within the T-DNA region: the [redacted] respectively.

The [redacted] T-DNA contains one flippase recombinase target sequence, FRT1 (Proteau *et al.*, 1986), and four *att*B recombination sites (Cheo *et al.*, 2004; Hartley *et al.*, 2000; Katzen, 2007). The presence of these sites alone does not cause any recombination, since to function, these sites need a specific recombinase enzyme that is not naturally present in plants (Cox, 1988; Dale and Ow, 1990; Thorpe and Smith, 1998).

[redacted]

Figure 1. Map of Plasmid [redacted]
Schematic diagram of plasmid [redacted] indicating the *cry1B.34.1*, *cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* gene cassettes. The size of plasmid [redacted] is [redacted] bp.

Table 1. Description of the Genetic Elements in Plasmid [redacted]

Region	Location on Plasmid (bp to bp)	Genetic Element	Size (bp)	Description
T-DNA	[redacted]		[redacted]	**See Table 2 for information on the elements in this region**
Plasmid Construct	[redacted]	**Includes Elements Below**	[redacted]	**DNA from various sources for plasmid construction and plasmid replication**
	28,137 – 28,640	CEN6 ARS	504	Sequence composed of a centromere and an autonomously replicating sequence from *Saccharomyces cerevisiae* (baker's yeast) (Sikorski and Hieter, 1989)
	28,910 – 29,125	URA3 Promoter	216	Promoter region from the *Saccharomyces cerevisiae* (baker's yeast) orotidine-5'-phosphate decarboxylase gene (Flynn and Reece, 1999)
	29,126 – 29,929	URA3	804	Orotidine-5'-phosphate decarboxylase gene from *Saccharomyces cerevisiae* (baker's yeast) (Flynn and Reece, 1999)
	29,966 – 32,524	pVS1 *ori*	2,559	Origin of replication from *Pseudomonas aeruginosa* pVS1 plasmid (Itoh *et al.*, 1984)
	32,912 – 33,922	*spc*	1,011	Spectinomycin resistance gene from bacteria (Fling *et al.*, 1985)
	34,092 – 34,886	*nptIII*	795	Neomycin phosphotransferase III (neomycin/kanamycin resistance) gene from *Streptococcus faecalis* (Trieu-Cuot and Courvalin, 1983)
	35,118 – 35,706 (complementary)	pUC *ori*	589	Origin of replication from *Escherichia coli* pUC19 plasmid (GenBank accession KP700956.1 (Yanisch-Perron *et al.*, 1985))
Ti Plasmid Backbone	**35,838 – 35,909**	**Includes Elements Below**	**72**	**Overdrive and intergenic regions from the *Agrobacterium tumefaciens* Ti plasmid**
	35,838 – 35,871	Ti Plasmid Region	34	Sequence from the *Agrobacterium tumefaciens* Ti plasmid (GenBank accession KX986282.1 (Komari *et al.*, 1996))
	35,872 – 35,895	Overdrive	24	T-DNA transmission enhancer from the *Agrobacterium tumefaciens* Ti plasmid (Peralta *et al.*, 1986)
	35,896 – 35,909	Ti Plasmid Region	14	Sequence from the *Agrobacterium tumefaciens* Ti plasmid (GenBank accession KX986282.1 (Komari *et al.*, 1996))

[redacted]

Figure 2. Map of the T-DNA Region from Plasmid [redacted]
Schematic diagram of the [redacted] T-DNA region indicating the *cry1B.34.1*, *cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* gene cassettes. The size of the T-DNA is [redacted] bp.

Table 2. Description of Genetic Elements in the T-DNA Region from Plasmid [redacted]

Gene Cassette	Location on T-DNA (bp to bp)	Genetic Element	Size (bp)	Description
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]

[1] Nomenclature defined according to Bacterial Pesticidal Protein Resource Center (www.bpprc.org Crickmore *et al.* (2021))

Table 2. Description of Genetic Elements in the T-DNA Region from Plasmid [redacted] (continued)

Gene Cassette	Location on T-DNA (bp to bp)	Genetic Element	Size (bp)	Description
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]

Table 2. Description of Genetic Elements in the T-DNA Region from Plasmid [redacted] (continued)

Gene Cassette	Location on T-DNA (bp to bp)	Genetic Element	Size (bp)	Description
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]

<u>A.3 (b) Description of the construct and the transformation vectors used</u>

Please refer to Section A.3 (a) *Transformation Method* for the vector used in transformation, Tables 1-2 for the description of the genetic elements in plasmid [redacted] and Figure 1 for the map of plasmid [redacted].

A.3 (c) Molecular characterization

Molecular characterization of COR23134 soybean plants was conducted using Southern-by-Sequencing (SbS™ technology, referred to as SbS) to determine insertion copy number and organization within the plant genome and to confirm the absence of plasmid backbone sequences. Southern blot analysis was performed to confirm stable genetic inheritance of the inserted *cry1B.34.1, cry1B.61.1*, *ipd083Cb,* and *gm-hra_1* gene cassettes across multiple generations during the breeding process (see also A.3 (e) *Stability of the genetic changes*). Segregation analysis was conducted for six generations of COR23134 soybean to confirm stable Mendelian inheritance (presented in A.3 (e) *Stability of the genetic changes*). Sanger sequencing of the insert and its flanking genomic regions, bioinformatics assessments of the flanking genomic sequences for chromosomal location of the insert and for potential endogenous gene disruption, and bioinformatics assessments of translated stop codon-bracketed frames for allergenicity and toxicity were conducted to characterize the inserted DNA in COR23134 soybean. Additionally, an event-specific quantitative real-time PCR detection method was developed and validated for COR23134 soybean.

Southern-by-Sequencing (SbS) Analysis to Determine Insertion Copy Number and Organization and Confirm the Absence of Plasmid Backbone Sequences

SbS analysis utilizes probe-based sequence capture, Next Generation Sequencing (NGS) techniques, and bioinformatics procedures to capture, sequence, and identify inserted DNA within the soybean genome. By compiling a large number of unique sequencing reads and mapping them against the sequences of the transformation plasmid and control soybean genome, unique junctions due to inserted DNA are identified in the bioinformatics analysis and used to determine the insertion copy number and organization within the plant genome and confirm the absence of plasmid backbone sequences.

The SbS technique utilizes capture probes homologous to the transformation plasmid to isolate genomic DNA that hybridizes to the probe sequences (Zastrow-Hayes et al., 2015). Captured DNA is then sequenced using a Next Generation Sequencing (NGS) procedure and the results are analyzed using bioinformatics tools. During the analysis, junction reads are identified as those sequencing reads where part of the read shows exact homology to the plasmid DNA sequence while the rest of the read does not match the contiguous plasmid. Junctions may occur between inserted DNA and genomic DNA (plasmid-genome junction), or between two plasmid-derived DNA sequences (plasmid-plasmid junction) that are not contiguous in the transformation plasmid. Multiple sequencing reads are generated for each junction and are compiled into a consensus sequence for the junction. By compiling a large number of unique sequencing reads and comparing them to the transformation plasmid and control soybean genome, unique junctions due to inserted DNA are identified. A unique junction is defined as one in which the 20-bp plasmid-derived sequence and the 30-bp adjacent sequence are the same across multiple reads, although the overall length of the multiple reads for that junction may vary due to the sequencing process. The number of unique junctions is related to the number of plasmid insertions present in the soybean genome (for example, a single T-DNA insertion is expected to have two unique plasmid-genome junctions). Detection of additional unique junctions beyond the two expected for a single insertion would indicate the presence of rearrangements or deletions within the insertion, or additional insertions derived from plasmid DNA. The absence of any junctions indicates there are no detectable insertions within the soybean genome.

The segregating T1 generation of COR23134 soybean was analyzed by SbS, using capture probes targeting all sequences of plasmid [redacted], to determine the insertion copy number and organization and to confirm the absence of plasmid backbone sequences. SbS was also performed on one 93Y21 control soybean plant, and on a positive control sample to confirm that the assay could reliably detect [redacted] plasmid DNA diluted in control soybean DNA. Based on the results obtained for transgenic COR23134 soybean, a schematic diagram of the COR23134 insertion was developed and is provided in Figure 3.

Several genetic elements in plasmid [redacted] are derived from soybean, and thus the homologous elements in the 93Y21 control soybean genome will be captured by the full-coverage probes used in the SbS analysis. These endogenous elements [redacted], *gm-hra_1*, and [redacted]; Table 3; Figure 1 and Figure 2) will have sequencing reads in the SbS results due to the homologous elements in the 93Y21 control soybean genome. However, if no junctions are detected, these sequencing reads only indicate the presence of the endogenous elements in their normal context of the soybean genome and are not from inserted DNA.

SbS analysis results for the control soybean are shown in Figure 4 and the positive control sample is presented in Figure 5. Results from the segregating T1 generation of COR23134 soybean are presented in Figure 6 for a representative transgenic COR23134 soybean plant, Figure 7 for a representative null segregant plant, and additional figures in Appendix A for the other transgenic COR23134 soybean plants.

SbS Analysis of the 93Y21 Control Soybean

Sequencing reads of the 93Y21 control soybean were aligned to the [redacted] T-DNA and plasmid sequences (Figure 4); however, coverage was obtained only for the endogenous genetic elements derived from the soybean genome. These sequencing reads were due to the capture and sequencing of these genetic elements in their normal context within the 93Y21 control soybean genome (Table 3). Variation in coverage of the soybean endogenous elements is due to sequence variations between the 93Y21 control soybean and the soybean varieties from which the genetic elements in the plasmid were derived. No junctions were detected between plasmid sequence and the soybean genome sequence (Table 4), indicating that there are no plasmid insertions in the control soybean, and the sequencing reads were solely due to the endogenous genetic elements present in the 93Y21 control soybean genome.

SbS Analysis of the Positive Control Sample Containing [redacted] Plasmid DNA

SbS analysis of the positive control sample ([redacted] plasmid DNA diluted in control soybean DNA) resulted in sequence coverage across the entire length of plasmid [redacted] (Figure 5), indicating that the SbS assay utilizing the full-coverage probe library is sensitive enough to detect [redacted] sequence. No junctions were detected between plasmid sequence and the soybean genome sequence (Table 4). The plasmid-plasmid junctions identified in the positive control sample (Figure 5) are artifacts of mapping a circular plasmid to a linear map and show the start and end points of the plasmid sequence but do not indicate insertions in the soybean genome.

SbS Analysis of the T1 Generation of COR23134 Soybean

SbS analysis of the segregating T1 generation of COR23134 soybean showed four positive plants (plant IDs 437164754, 437164755, 437164756, and 437164757) that contained the inserted DNA (Table 4; a representative plant ID 437164754 is presented in Figure 6; additional figures in Appendix A. Each of these plants contained two unique plasmid-genome junctions, one at each end of the insertion, that were identical across the four plants. The insertion, derived from the [redacted] T-DNA, starts with the 5' junction at bp [redacted] and ends with the 3' junction at bp [redacted] (Figure 3). The number of sequencing reads at the 5' and 3' junctions is provided in Table 4. In each of the four positive plants, a singular and identical plasmid-plasmid junction was found, suggesting that these plants harbor a deletion within the insertion. As a result, a 21-bp deletion was identified at position [redacted] in all four plants in the [redacted] promoter of the [redacted] that differs from the [redacted] T-DNA sequence (Junction panel of panel A of Figure 6; additional figures in Appendix A). The number of sequencing reads at the 21-bp deletion junction is provided in Table 5. There were no additional junctions between plasmid [redacted] and the soybean genome detected in the plants, indicating that there are no additional plasmid-derived insertions present in COR23134 soybean. There were no additional unexpected junctions between non-contiguous regions of the [redacted] T-DNA identified, indicating that there are no additional rearrangements, deletions, or duplications in the inserted DNA. Furthermore, there were no junctions between the backbone sequence of [redacted] and the soybean genome sequences, demonstrating that no plasmid backbone sequences were incorporated into COR23134 soybean.

Each of the six null segregant COR23134 soybean plants from the T1 generation that was determined to be negative for the COR23134 insertion (plant IDs 437164750, 437164751, 437164752, 437164753, 437164758, and 437164759) yielded sequencing reads for the endogenous genetic elements derived from the soybean genome (a representative plant ID 437164750 is presented in Figure 7). There were no junctions between the plasmid [redacted] sequence and the soybean genome sequence detected in these plants (Table 4), indicating that these plants did not contain any insertions derived from [redacted].

SbS analysis of the segregating T1 generation of COR23134 soybean demonstrated that COR23134 soybean contains a single copy of the inserted DNA derived from the [redacted] T-DNA, with the expected organization except for a 21-bp deletion in the [redacted] promoter, and that no additional insertions or plasmid backbone sequences are present in its genome.

Additional details regarding analytical methods for SbS analysis are provided in Appendix A.

Table 3. Soybean Endogenous Elements in Plasmid [redacted] and [redacted] T-DNA

Number[a]	Name of Endogenous Element[b]	Present in Plasmid or T-DNA
1	[redacted]	[redacted] T-DNA
2	[redacted]	[redacted] T-DNA
3	*gm-hra*_1, [redacted]	[redacted] T-DNA

Note: Transfer DNA (T-DNA); untranslated region (UTR).

[a] The numbers indicating soybean endogenous genetic elements are shown as circled numbers found below the linear maps in Figure 4 - Figure 7.

[b] As shown in the plasmid and T-DNA maps in Figure 1 and Figure 2, respectively.

Table 4. SbS Junction Reads of the COR23134 Insertion

Sample Description	Total Reads at 5′ Junction[a]	Unique Reads at 5′ Junction[b]	Total Reads at 3′ Junction[c]	Unique Reads at 3′ Junction[d]	COR23134 Insertion
T1 Generation Plant ID 437164750	0	0	0	0	-
T1 Generation Plant ID 437164751	0	0	0	0	-
T1 Generation Plant ID 437164752	0	0	0	0	-
T1 Generation Plant ID 437164753	0	0	0	0	-
T1 Generation Plant ID 437164754	2768	225	1720	149	+
T1 Generation Plant ID 437164755	2227	222	1349	137	+
T1 Generation Plant ID 437164756	2956	208	1886	153	+
T1 Generation Plant ID 437164757	2815	229	1525	144	+
T1 Generation Plant ID 437164758	0	0	0	0	-
T1 Generation Plant ID 437164759	0	0	0	0	-
93Y21 Control Soybean	0	0	0	0	-
[redacted] Positive Control	0	0	0	0	-

Total number of sequencing reads across the 5′ genome-plasmid junction of the COR23134 insertion.

[b] Unique sequencing reads establishing the location of the 5′ genome-plasmid junction of the COR23134 insert (Figure 3). Multiple identical NGS-supporting reads are condensed into each unique read.

[c] Total number of sequencing reads across the 3′ plasmid-genome junction of the COR23134 insertion.

[d] Unique sequencing reads establishing the location of the 3′ plasmid-genome junction of the COR23134 insert (Figure 3). Multiple identical NGS-supporting reads are condensed into each unique read.

Table 5. SbS Junction Reads of the 21-bp Deletion in the COR23134 Insertion

Sample Description	Total Reads at Junction[a]	Unique Reads at Junction[b]	COR23134 Insertion
T1 Generation Plant ID 437164750	0	0	-
T1 Generation Plant ID 437164751	0	0	-
T1 Generation Plant ID 437164752	0	0	-
T1 Generation Plant ID 437164753	0	0	-
T1 Generation Plant ID 437164754	381	79	+
T1 Generation Plant ID 437164755	333	79	+
T1 Generation Plant ID 437164756	428	78	+
T1 Generation Plant ID 437164757	393	82	+
T1 Generation Plant ID 437164758	0	0	-
T1 Generation Plant ID 437164759	0	0	-
93Y21 Control Soybean	0	0	-
[redacted] Positive Control	0	0	-

[a] Total number of sequencing reads across the plasmid-plasmid junction of the 21-bp deletion in COR23134 insertion.
[b] Unique sequencing reads establishing the location of the plasmid-plasmid junction of the 21-bp deletion in COR23134 insertion. Multiple identical NGS-supporting reads are condensed into each unique read.

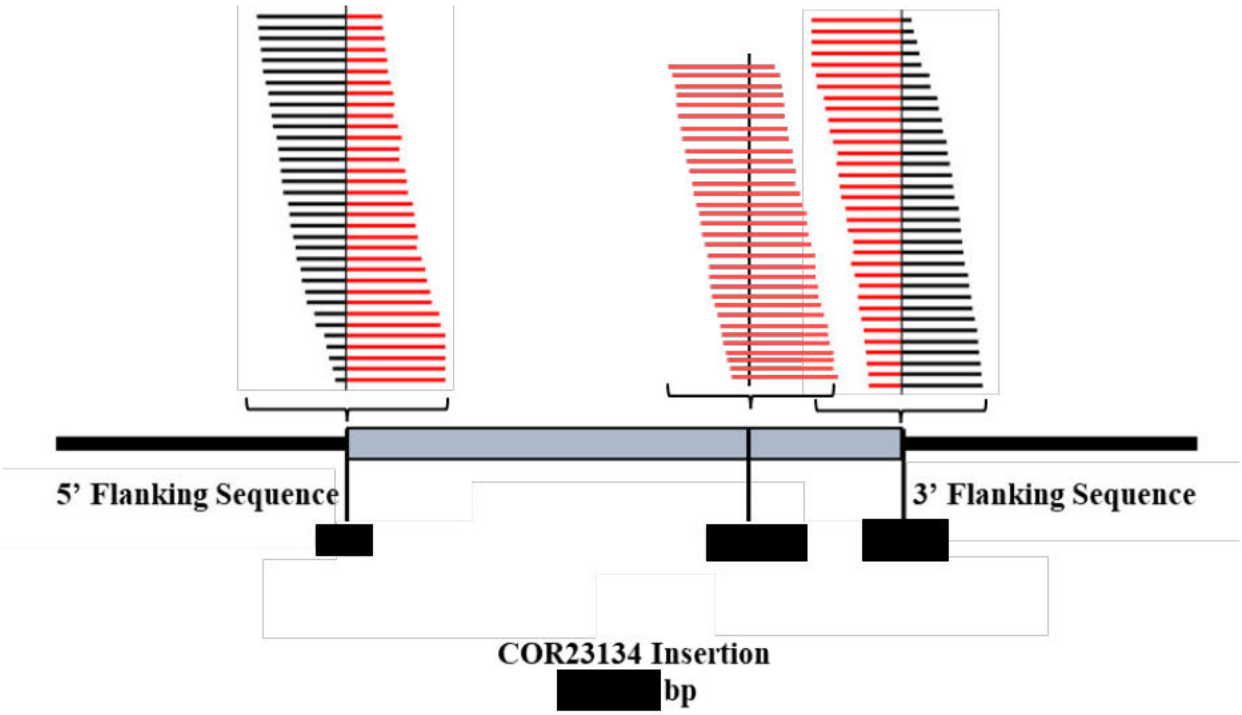


Figure 3. Map of the Insertion in COR23134 Soybean
Schematic diagram of the DNA insertion in COR23134 soybean based on the SbS analysis described. The flanking soybean genomic regions are indicated in the map by black bars. A single copy of the insertion, derived from the [redacted] T-DNA (Figure 2) and shown by the gray box, is integrated into the COR23134 soybean genome. Vertical lines show the locations of the two unique plasmid-genome junctions and one unique plasmid-plasmid junction. The numbers below the map indicate the bp location of the junction nucleotide in reference to the sequence of the [redacted] T-DNA. Representative individual sequencing reads across the junctions are shown as horizontally stacked lines above each junction (not to scale); black indicates flanking genomic sequence and red indicates inserted plasmid DNA sequence within each sequencing read.

A. Alignment of NGS Reads to [redacted] T-DNA Region

[redacted]

B. Alignment of NGS Reads to [redacted]

[redacted]

Figure 4. SbS Analysis for Control Soybean

The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] T-DNA or [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in each plasmid derived from the soybean genome (identified by numbers; Table 3), while tan bars indicate genetic elements derived from other sources. The absence of any junctions between plasmid and genomic sequences indicates that there are no insertions or plasmid backbone sequence present in the control soybean. **A)** SbS results for the control soybean aligned against the [redacted] T-DNA region intended for insertion [redacted] bp; Figure 2). Coverage was obtained only for regions derived from soybean endogenous elements. Variation in coverage of the endogenous elements is due to sequence variations between the control soybean and the source of the corresponding genetic elements. As no junctions were detected between the [redacted] T-DNA sequence and the soybean genome, there are no DNA insertions identified in the 93Y21 control soybean, and the sequencing reads are solely due to the endogenous elements present in the 93Y21 soybean genome. **B)** SbS results for 93Y21control soybean aligned against the plasmid [redacted] sequence [redacted] bp; Figure 1). Coverage was obtained only for the endogenous elements.

[redacted]

Figure 5. SbS Analysis for [redacted] Positive Control Sample
The positive control sample consisted of [redacted] plasmid DNA diluted in the control soybean genomic DNA. Shown are the SbS results of the positive control sample aligned against [redacted] [redacted] bp; Figure 1). The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in the plasmid derived from the soybean genome (identified by numbers; Table 3), while tan bars indicate genetic elements derived from other sources. The two plasmid-plasmid junctions (red arrows) shown at the bottom of the graph are artifacts of aligning reads from a circular plasmid to a linear map. They show the start and end points (Junctions 1 and [redacted]) of the plasmid sequence but do not indicate insertions in genomic DNA of control soybean. Coverage was obtained across the entire length of the plasmid, indicating successful capture of [redacted] fragments by the SbS probe library.

A. Alignment of NGS Reads to [redacted] T-DNA Region

[redacted]

B. Alignment of NGS Reads to [redacted]

Figure 6. SbS Analysis for Representative Transgenic COR23134 Soybean Plant (Plant ID 437164754)

The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] T-DNA or [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in each plasmid derived from the soybean genome (identified by numbers; Table 3), while tan bars indicate genetic elements derived from other sources. **A)** SbS results for transgenic COR23134 soybean aligned against the [redacted] T-DNA region intended for insertion [redacted] bp; Figure 2) indicating that this plant contains the insertion. Arrows in the Junctions panel indicate the two plasmid-genome junctions (black arrows) and one plasmid-plasmid junction (red arrow) identified by SbS; the numbers below the arrows refer to the bp location of the junction relative to the [redacted] T-DNA. The insertion comprises bp [redacted] of the [redacted] T-DNA shown in Figure 2. The presence of two plasmid-genome junctions (Junctions [redacted]) demonstrates the presence of a single insertion in the COR23134 soybean genome. One plasmid-plasmid junction (Junction [redacted]) indicates the location of a 21-bp deletion [redacted] bp) identified in all plants containing the COR23134 insertion. **B)** SbS results for transgenic COR23134 soybean aligned against the plasmid [redacted] sequence ([redacted] bp; Figure 1). Coverage was obtained for the elements between the Right and Left Borders transferred into COR23134 soybean; however, for clarity the junctions identified in panel A are not shown in this view. The absence of any other junctions to the [redacted] sequence shows that there are no additional insertions or backbone sequence present in COR23134 soybean.

A. Alignment of NGS Reads to [redacted] T-DNA Region

[redacted]

B. Alignment of NGS Reads to [redacted]

[redacted]

Figure 7. SbS Analysis for Representative Null Segregant Plant (Plant ID 437164750)
The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] T-DNA or [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in each plasmid derived from the soybean genome (identified by numbers; Figure 3), while tan bars indicate genetic elements derived from other sources. A) SbS results aligned against the [redacted] T-DNA region intended for insertion [redacted] bp; Figure 2), indicating that this plant does not contain the insertion. Coverage was obtained only for regions derived from soybean endogenous elements. Variation in coverage of the endogenous elements is due to sequence variations between the control soybean and the source of the corresponding genetic elements. As no junctions were detected between the [redacted] T-DNA sequence and the soybean genome, there are no DNA insertions identified in this plant, and the sequencing reads are solely due to the endogenous elements present in the 93Y21 soybean genome. B) SbS results aligned against the plasmid [redacted] sequence [redacted] bp; Figure 1). Coverage was obtained only for the endogenous elements. The absence of any junctions between plasmid and genomic sequences indicates that there are no insertions or backbone sequence present in this plant from the T1 generation of COR23134 soybean.

Sequence of Insert and Genomic Border Regions Using Sanger Sequencing

Sequence characterization analysis was performed to determine the DNA sequence of the COR23134 insert and flanking genomic regions. It should be noted that while DNA sequencing provides certain molecular information, the exact nucleotide sequence should not be viewed as static. Spontaneous mutations are a very common phenomenon in plants, presenting a biological mechanism of adaptation to constantly changing environment (Weber *et al.*, 2012). Spontaneous mutations can occur in any part of the plant genome and in both non-GM and GM plants (Waigmann *et al.*, 2013). In GM plants, there is no scientific basis to expect that the frequency of spontaneous mutations in transgenic insert or flanking genomic regions would be greater than in the rest of the plant genome, or that they would have a differential impact on safety (La Paz *et al.*, 2010; Waigmann *et al.*, 2013).

The sequence of the insert and its flanking genomic regions was determined to confirm the integrity of the inserted DNA in COR23134d soybean. PCR primers were designed to amplify seven overlapping PCR fragments spanning the insert and the 5' and 3' flanking genomic regions (Figure 8). At least six plasmids (three from each of two independent PCR reactions) for each PCR fragment were sequenced in both forward and reverse directions to cover every nucleotide by Sanger sequencing, and the resulting sequencing reads were used to determine the consensus sequence for each PCR fragment. The consensus sequences from all seven overlapping PCR fragments were combined to determine the sequence for COR23134 soybean. The total length of sequence determined for COR23134 soybean is [redacted] base pairs (bp), composed of [redacted] bp of 5' flanking genomic sequence, [redacted] bp of 3' flanking genomic sequence, and [redacted] bp of the inserted DNA.

In comparison with the sequence of the [redacted] T-DNA, the COR23134 insert is composed of bps [redacted] of the [redacted] T-DNA, except for a 21-bp deletion in the [redacted] promoter at bps [redacted] relative to the [redacted] T-DNA sequence. All remaining sequence is intact and identical to the [redacted] T-DNA sequence.

Additional details regarding analytical methods for Sanger sequencing analysis are provided in Appendix D.

[redacted]

COR23134 Soybean

Figure 8. Map of the Insert and Flanking Genomic Regions in COR23134 Soybean
Seven overlapping PCR fragments (A, B, C, D, E, F, and G) spanning the insert and its flanking genomic regions were amplified from genomic DNA of COR23134 soybean. Each black horizontal bar represents the relative position of the PCR fragment, and the vertical dash lines represent the flanking genomic regions and insert junctions. The total length of sequence determined for COR23134 soybean is [redacted] base pairs (bp), composed of [redacted] bp of 5' flanking genomic sequence, [redacted] bp of 3' flanking genomic sequence, and [redacted] bp of inserted DNA derived from the [redacted] T-DNA.

Genomic Border Region Analysis

The [redacted]-bp 5' flanking genomic sequence and the [redacted]-bp 3' flanking genomic sequence of the COR23134 insert were individually subjected to pairwise sequence alignment analysis using BLASTN v2.11.0+ to search against a *Glycine max* reference database [redacted] to identify the genomic location of the insert. The 5' flanking genomic sequence aligns perfectly (E-value = 0, identity = 100%, length = [redacted] nt) to the direct strand of chromosome [redacted]. The 3' flanking genomic sequence aligns near-perfectly (E-value = 0, identity = 98%, gaps = 20, length = [redacted] nt) to the direct strand of chromosome [redacted] Consequently, there is a very high likelihood that the COR23134 soybean insert is located on chromosome [redacted]. The assessment for potential gene disruption by the COR23134 insert concludes that no potential gene disruption was indicated as a result of this bioinformatics analysis.

Reading Frame Analysis of the Insert/Border Junctions and Bioinformatic Assessments for Allergens and Toxins

A bioinformatics assessment of potentially expressed peptides (i.e., translated stop codon-bracketed frames) was conducted following established international criteria (Codex Alimentarius Commission, 2009; EFSA, 2010; EFSA Panel on Genetically Modified Organisms (GMO), 2011; FAO/WHO, 2001). All translated stop codon-bracketed frames of length ≥ eight amino acids (aa) in COR23134 soybean that are within the insertion or that cross the boundary between the insertion and its flanking genomic regions were identified and evaluated for similarity of allergens and toxins. A total of 1,695 translated stop codon-bracketed frames ≥ eight aa, either contained entirely within the insertion or crossing the boundary between the insertion and its flanking genomic regions, were identified for the COR23134 soybean sequence.

Allergenicity Analysis

In COR23134 soybean, searches of the translated stop codon-bracketed frames against the COMPARE allergen database revealed one [redacted]-aa frame (COR23134_784) in the [redacted] coding sequence, displaying > 35% identity with five known allergens over a "sliding window" of 80 aa. None of the translated stop codon-bracketed frames in COR23134 soybean produced an eight-contiguous amino acid match to an allergen. While transcription of COR23134_784 would be expected given an upstream promoter element, COR23134_784 [redacted] and therefore one would expect preferential translation of the intended [redacted] protein in this region. The lack of possible expression, coupled with the fact that the percent identities of the various alignments to allergens are low and that the E-values are high – indicating that these are low-significance alignments – shows no allergenicity concern arose from the bioinformatics analysis of COR23134 soybean.

Toxicity Analysis

In COR23134 soybean, searches of the translated stop codon-bracketed frames against the internal database of protein toxin sequences found no alignments. Searches of the translated stop codon-bracketed frames against the NCBI nr protein sequences found 17 frames with alignments, none to protein toxins or other proteins harmful to humans or animals. These data indicate that no toxicity concern arose from the bioinformatics analysis of COR23134 soybean.

Bioinformatics evaluation of the COR23134 soybean insert did not generate biologically relevant amino acid sequence similarities to allergens and protein toxins that are harmful to humans or animals.

Event-Specific Detection Methodology

The event-specific quantitative real-time PCR (qPCR) method was developed and validated for the detection of soybean event COR-23134-4 through quantification of the relative content of soybean event COR-23134-4 to total soybean DNA in alignment with internationally accepted criteria for method validation (EU-RL-GMFF, 2015). Standard curves were used for both the taxon-specific *Lectin* and COR-23134-4 assays. Relative quantification of COR23134 soybean was calculated using the ratio between the mean copy number of COR23134 soybean in comparison to the haploid soybean genome.

The event-specific assay for COR23134 soybean is designed to amplify the target sequence at the 5' junction between the flanking soybean genomic sequence and the COR23134 insertion. The binding site of the forward primer is within the flanking soybean genomic sequence, the binding site of the reverse primer is within the COR23134 insertion sequence, and the binding site of the probe spans the junction of the flanking soybean genomic sequence and the COR23134 insertion sequence. The event-specific qPCR assay for COR23134 soybean amplifies an 80-bp product. The *Lectin* (*Le1*) taxon-specific PCR assay is a validated soybean-specific PCR assay for the lectin gene (GenBank Accession ID XM_028337035) (QT-TAX-GM-002) with the replacement of the TAMRA™ (6-carboxytetramethylrhodamine, succinimidyl ester) quencher with a Black Hole Quencher 1 (BHQ1™) quencher (EU-RL-GMFF, 2013). The assay amplifies a 74-bp product.

Absence of Genes that Code Resistance to Antibiotics

COR23134 soybean was analyzed by SbS to confirm the absence of antibiotic resistance gene sequences from plasmid [redacted] used in transformation steps during event development, including the *spc* and *ntpIII* antibiotic resistance genes.

SbS analysis of the positive control samples [redacted] plasmid DNA diluted in non-GM control 93Y21 soybean DNA) resulted in sequence coverage across the entire length of the [redacted] plasmid (Figure 5), indicating that the SbS assay utilizing the full-coverage probe library, inclusive of antibiotic resistance genes, is sensitive enough to detect sequences from plasmid [redacted].

SbS analysis of COR23134 soybean plants showed that there were no junctions identified between soybean genome sequences and the plasmid backbone sequence of [redacted] in any of the plants analyzed (Figure 6 and additional figures in Appendix A). These results confirm that no plasmid backbone sequences containing antibiotic resistance genes from plasmid [redacted] were incorporated into COR23134 soybean during transformation.

Conclusion on the Molecular Characterization of COR23134 Soybean

SbS, Southern blot, multi-generation segregation, Sanger sequencing of the insert and its flanking genomic regions, bioinformatics assessments of the flanking genomic sequences for chromosomal location of the insert and for potential endogenous gene disruption, and bioinformatics assessments of translated stop codon-bracketed frames for allergenicity and toxicity, were conducted to characterize the inserted DNA in COR23134 soybean. (See also section A.3 (e) *Stability of the genetic changes*.)

SbS analysis confirmed that COR23134 soybean contains a single copy of the inserted DNA with the expected organization except for a 21-bp deletion in the [redacted] promoter, and that no additional insertions or plasmid backbone sequences are present in COR23134 soybean genome. Southern blot analysis of five generations of COR23134 soybean confirmed that the inserted DNA in COR23134 soybean is consistent and stable across multiple generations during the breeding process. Segregation analysis of six generations of COR23134 soybean confirmed that the inserted DNA segregated as a single locus in accordance with Mendelian rules of inheritance and stably integrated into the soybean genome across generations during the breeding process. Sanger sequencing of the insert and its flanking genomic regions confirmed the integrity of the inserted DNA from the [redacted] T-DNA in COR23134 soybean except for a 21-bp deletion in the [redacted] promoter. Bioinformatic assessments of the genomic sequences flanking the COR23134 soybean insert confirmed the chromosomal location of the insert and no disruption of endogenous genes. Bioinformatics assessments of translated stop codon-bracketed frames that are within the insertion or that cross the boundary between the insertion and its flanking genomic regions confirmed that the COR23134 soybean insert did not generate biologically relevant amino acid sequence similarities to allergens and protein toxins that are harmful to humans or animals.

Together, these analyses confirmed that a single copy of the inserted DNA, with no plasmid backbone sequences, is present in the COR23134 soybean genome. The introduced genes are stably inherited across multiple generations and segregated according to Mendel's law of inheritance during the breeding process. Sanger sequencing determined the sequences of the inserted DNA and flanking genomic regions in COR23134 soybean, and bioinformatic assessments of the genomic sequences flanking the COR23134 soybean insert confirmed the chromosomal location of the insert and no disruption of endogenous genes. Bioinformatic assessments for allergenicity or toxicity support the conclusion that COR23134 soybean was found unlikely to be toxic or allergenic to humans or animals. Additionally, an event-specific quantitative real-time PCR detection method was developed and validated for COR23134 soybean.

A.3 (d) Breeding process

Plants that were regenerated from transformation and tissue culture (designated T0 plants) were selected for further characterization. A schematic overview of the transformation and event development process for COR23134 soybean is provided in Figure 9.

[redacted]

Figure 9. Event Development Process of COR23134 Soybean

The subsequent breeding of COR23134 soybean proceeded as indicated in Figure 10 to produce specific generations for the characterization and assessments conducted, as well as for the development of commercial soybean lines. Table 6 provides the generations used for each characterization study.

[redacted]

Figure 10. Breeding Diagram for COR23134 Soybean and Generations Used for Analysis
The breeding steps to produce the generations used for characterization and safety assessments are shown schematically. A Corteva Agriscience elite soybean variety 93Y21 was used for transformation to produce COR23134 soybean.

Table 6. Generations and Comparators Used for Analysis of COR23134 Soybean

Analysis	Seed Generation(s) Used	Comparators
Insertion copy number, insertion organization, and the absence of plasmid backbone sequences by SbS	T1	93Y21
Insertion organization and stability by Southern blot	T1, T2, T3, T4, T5	93Y21
Mendelian inheritance by multi-generation segregation analysis	T1, T2, T3, T4, T5, T6	93Y21
Sequence determination of Insert and its flanking genomic regions by Sanger sequencing	T4	93Y21
Composition and expression analysis	T5	93Y21

Selection of Comparators

For the characterization of COR23134 soybean, the Corteva Agriscience elite soybean variety 93Y21 was used as an experimental control (Table 6). The control line selected is non-genetically modified (non-GM) and represent the same genetic background of the soybean line used to produce the COR23134 soybean generations used in analysis (Figure 10).

In addition, conventional non-GM soybean lines (i.e., reference lines), were used to obtain tolerance intervals for compositional analyses. These soybean lines were chosen to represent a wide range of conventional non-GM varieties. These tolerance intervals help represent the biological variation of the soybean crop for compositional analytes and further helped to determine the comparability of COR23134 soybean to conventional non-GM soybean.

A.3 (e) Stability of the genetic changes

Southern Analysis to Determine Stable Genetic Inheritance across Generations

Southern blot analysis was performed on five generations of COR23134 soybean to evaluate the stability of the inserted *cry1B.34.1, cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* gene cassettes across multiple generations.

Restriction enzyme *Bst*1107 I was selected to verify the stability of the COR23134 soybean insertion across the five generations (T1, T2, T3, T4, and T5) of COR23134 soybean plants. *Bst*1107 I has two adjacent restriction sites within the COR23134 soybean insertion (Figure 11 and Figure 12), which provides a means to uniquely identify the event, as additional sites would be in the adjacent flanking genomic sequences (Figure 13).

Genomic DNA samples extracted from leaf tissue of the five generations of COR23134 soybean and control 93Y21 soybean plants were digested with *Bst*1107 I and hybridized with the *cry1B.34.1*, *cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* probes for Southern analysis. Hybridization patterns of these probes would exhibit event-specific bands unique to the COR23134 soybean insertion, and thus provide a means of verification that the genomic border regions of the COR23134 soybean insertion were not changed across the five generations during breeding. Plasmid [redacted] was added to control soybean DNA, digested with *Bst*1107 I, and included on the blot to verify successful probe hybridization. The probes used for Southern hybridization are described in Table 7 and shown in Figure 12. The predicted and observed hybridization bands on Southern blots are provided in Table 8.

Hybridization of the *cry1B.34.1* probe to *Bst*1107 I-digested genomic DNA resulted in two consistent bands of approximately [redacted] bp and [redacted] bp in all five generations of COR23134 soybean (Table 8; Figure 14). A single 5' border band of greater than [redacted] bp was predicted based on the [redacted] T-DNA map (Figure 12); however, due to the high homology of the *cry1B.34.1* and *cry1B.61.1* genes it is not unexpected to also detect the 3' border band of greater than [redacted] bp with the *cry1B.34.1* probe. Detection of the approximately [redacted]-bp band confirmed that the 5' border fragment, containing the *cry1B.34.1* gene, is intact and stable across the five generations of COR23134 soybean. Furthermore, detection of the approximately [redacted]-bp band infers that the 3' border fragment, containing the *cry1B.61.1* gene in the COR23134 soybean insertion, is also intact and stable. Positive control lanes containing plasmid DNA showed two bands of [redacted] bp and [redacted] bp, confirming successful hybridization of the *cry1B.34.1* probe. No bands were observed in the control soybean plant.

Hybridization of the *cry1B.61.1* probe to *Bst*1107 I-digested genomic DNA resulted in two consistent bands of approximately [redacted] bp and [redacted] bp in all five generations of COR23134 soybean (Table 8; Figure 15). A single 3' border band of greater than [redacted] bp was predicted based on the [redacted] T-DNA map (Figure 12); however, due to the high homology of the *cry1B.61.1* and *cry1B.34.1* genes it is not unexpected to also detect the 5' border band of greater than [redacted] bp with the *cry1B.61.1* probe. Detection of the approximately [redacted]-bp band confirmed that the 3' border fragment, containing the *cry1B.61.1* gene, is intact and stable across the five generations of COR23134 soybean. Detection of the approximately [redacted]-bp band provides additional evidence that the 5' border fragment, containing the *cry1B.34.1* gene in the COR23134 soybean insertion, is also intact and stable. Positive control lanes containing plasmid DNA showed two bands of [redacted] bp and [redacted] bp, confirming successful hybridization of the *cry1B.61.1* probe. No bands were observed in the control soybean plant.

Hybridization of the *ipd083Cb* and *gm-hra_1* probes to *Bst*1107 I-digested genomic DNA each resulted in a consistent band of approximately [redacted] bp in all five generations of COR23134 soybean analyzed (Table 8; Figure 16 and Figure 17, respectively), verifying that this band is due to the 3' border fragment. These results confirmed that the 3' border fragment, containing the *ipd083Cb* and *gm-hra_1* genes in the COR23134 soybean insertion, are intact and stable across the five generations of COR23134 soybean. In addition to the insertion-derived bands, hybridization of the *gm-hra_1* probe resulted in four endogenous bands of approximately [redacted] bp, [redacted] bp, [redacted] bp, and [redacted] bp across the COR23134 soybean and control soybean (Figure 17). These bands can be attributed to hybridization of the probe to endogenous sequences in the soybean genome that are homologous to the *gm-hra_1* probe. The positive control lanes containing plasmid DNA showed the expected band of [redacted] bp, confirming successful hybridization of the *ipd083Cb* and *gm-hra_1* probes.

The presence of equivalent bands from hybridization with each of the *cry1B.34.1*, *cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* probes within the plants from all five generations analyzed confirms that the inserted DNA in COR23134 soybean is consistent and stable across multiple generations during the breeding process.

Additional details regarding analytical methods for Southern analysis are provided in Appendix B.

Table 7. Description of DNA Probes Used for Southern Hybridization

Probe Lot Number	Genetic Element/ Probe Name	Probe Length (bp)	Position on [redacted] T-DNA (bp to bp)[a]
[redacted]	*cry1B.34.1*[b]	[redacted]	[redacted]
[redacted]	*cry1B.61.1*[c]	[redacted]	[redacted]
[redacted]	*ipd083Cb*[d]	[redacted]	[redacted]
[redacted]	*gm-hra_1*[e]	[redacted]	[redacted]

a The probe position is based on the [redacted] T-DNA map (Figure 12).
b The *cry1B.34.1* probe comprises three fragments that are combined in a single hybridization solution.
c The *cry1B.61.1* probe comprises two fragments that are combined in a single hybridization solution.
d The *ipd083Cb* probe comprises three fragments that are combined in a single hybridization solution.
e The *gm-hra_1* probe comprises two fragments that are combined in a single hybridization solution.

Table 8. Predicted and Observed Hybridization Bands on Southern Blots; *Bst*1107 I Digest

Probe Name	Predicted and Observed Fragment Size from Plasmid [redacted] (bp)[a]	Predicted Fragment Size from [redacted] T-DNA (bp)[b]	Observed Fragment Size in COR23134 Soybean (bp)[c]	Figure
cry1B.34.1	[redacted]	[redacted]	[redacted]	Figure 14
cry1B.61.1	[redacted]	[redacted]	[redacted]	Figure 15
ipd083Cb	[redacted]	[redacted]	[redacted]	Figure 16
gm-hra_1	[redacted]	[redacted]	[redacted]	Figure 17

Note: An Asterisk (*) and gray shadings indicates that the designated bands are due to hybridization to endogenous sequences. These bands were identified in both COR23134 soybean and 93Y21 control soybean.
[a] Predicted and observed fragment sizes based on the [redacted] plasmid map (Figure 11).
[b] Predicted sizes based on the [redacted] T-DNA (Figure 12).
[c] Observed fragment sizes are approximated from the DIG-labeled DNA Molecular Weight Marker III and VII fragments on the Southern blots. Due to inability to determine the exact sizes on the blot, all approximated values are rounded to the nearest 100 bp.
[d] The band of [redacted] bp is the predicted band based on the [redacted] plasmid map (Figure 11), while the band of [redacted] bp is due to high levels of homology between the *cry1B.34.1* probe and the *cry1B.61.1* gene on the plasmid.
[e] The band of [redacted] bp is the 5' genomic border band predicted by the [redacted] T-DNA map Figure 12), while the band of [redacted] bp is due to high levels of homology between the *cry1B.34.1* probe and the *cry1B.61.1* gene in the COR23134 insertion.
[f] The band of [redacted] bp is the predicted band based on the [redacted] plasmid map (Figure 11), while the band of [redacted] bp is due to high levels of homology between the *cry1B.61.1* probe and the *cry1B.34.1* gene on the plasmid.
[g] The band of [redacted] bp is the 3' genomic border band predicted by the [redacted] T-DNA map (Figure 12), while the band of [redacted] bp is due to high levels of homology between the *cry1B.61.1* probe and the *cry1B.34.1* gene in the COR23134 insertion.

[redacted]

Figure 11. Map of Plasmid [redacted] for Southern Analysis
Schematic diagram of plasmid [redacted] indicating the *Bst*1107 I restriction enzyme sites with base pair positions. The size of plasmid [redacted] is [redacted] bp. The Right Border and Left Border flank the T-DNA (Figure 12) that was transferred during *Agrobacterium*-mediated transformation to create COR23134 soybean.

[redacted]

Figure 12. Map of the [redacted] T-DNA for Southern Analysis
Schematic diagram of the [redacted] T-DNA indicating the *Bst*1107 I restriction enzyme sites with base pair positions. The size of the [redacted] T-DNA is [redacted] bp. The locations of the Southern blot probes are shown by the boxes below the map.

[redacted]

Figure 13. Map of the COR23134 Insertion for Southern Analysis

Schematic diagram of the COR23134 soybean insertion indicating the *Bst*1107 I restriction enzyme sites. The locations of the Southern blot probes are shown by the boxes below the map. The flanking soybean genomic sequences are represented by the horizontal black rectangular bars. The *Bst*1107 I restriction sites are indicated with the sizes of observed fragments on Southern blots shown below the map in base pairs (bp). The locations of restriction enzyme sites in the flanking soybean genomic regions are not to scale.

[redacted]

Lane	Sample	Lane	Sample
1	DIG-labeled DNA marker III	8	COR23134 soybean T4 generation
2	1 copy of [redacted] + 93Y21 control soybean	9	COR23134 soybean T5 generation
3	93Y21 control soybean	10	Blank
4	Blank	11	93Y21 control soybean
5	COR23134 soybean T1 generation	12	1 copy of [redacted] + 93Y21 control soybean
6	COR23134 soybean T2 generation	13	DIG-labeled DNA marker VII
7	COR23134 soybean T3 generation		

Figure 14. Southern Blot Analysis of COR23134 Soybean; *Bst*1107 I Digest with *cry1B.34.1* Probe

Genomic DNA isolated from leaf tissues of COR23134 soybean from T1, T2, T3, T4, and T5 generations and 93Y21 control soybean plants was digested with *Bst*1107 I and hybridized to the *cry1B.34.1* probe. Approximately 10 µg of genomic DNA was digested and loaded per lane. Positive control lanes include [redacted] plasmid DNA at approximately one gene copy number and 10 µg of control soybean DNA. The arrows indicate the COR23134-specific bands. Sizes of the DIG-labeled DNA Molecular Weight Marker III and VII are indicated adjacent to the blot image in kilobases (kb).

Lanes 5 to 9: The band of [redacted] bp is the 5' genomic border band predicted by the [redacted] T-DNA map, while the band of [redacted] bp is due to high levels of homology between the *cry1B.34.1* probe and the *cry1B.61.1* gene in the COR23134 insertion.

Lanes 2 and 12: The [redacted] bp band is the predicted band based on the [redacted] plasmid map, while the [redacted] bp band is due to high levels of homology between the *cry1B.34.1* probe and the *cry1B.61.1* gene on the plasmid.

[redacted]

Lane	Sample
1	DIG-labeled DNA marker III
2	1 copy of [redacted] + 93Y21 control soybean
3	93Y21 control soybean
4	Blank
5	COR23134 soybean T1 generation
6	COR23134 soybean T2 generation
7	COR23134 soybean T3 generation
8	COR23134 soybean T4 generation
9	COR23134 soybean T5 generation
10	Blank
11	93Y21 control soybean
12	1 copy of [redacted] + 93Y21 control soybean
13	DIG-labeled DNA marker VII

Figure 15. Southern Blot Analysis of COR23134 Soybean; *Bst*1107 I Digest with *cry1B.61.1* Probe

Genomic DNA isolated from leaf tissues of COR23134 soybean from T1, T2, T3, T4, and T5 generations and 93Y21 control soybean plants was digested with *Bst*1107 I and hybridized to the *cry1B.61.1* probe. Approximately 10 µg of genomic DNA was digested and loaded per lane. Positive control lanes include [redacted] plasmid DNA at approximately one gene copy number and 10 µg of control soybean DNA. The arrows indicate the COR23134-specific bands. Sizes of the DIG-labeled DNA Molecular Weight Marker III and VII are indicated adjacent to the blot image in kilobases (kb).

Lanes 5 to 9: The band of [redacted] bp is the 3' genomic border band predicted by the [redacted] T-DNA map, while the band of [redacted] bp is due to high levels of homology between the *cry1B.61.1* probe and the *cry1B.34.1* gene in the COR23134 insertion.

Lanes 2 and 12: The [redacted] bp band is the predicted band based on the [redacted] plasmid map, while the [redacted] bp band is due to high levels of homology between the *cry1B.61.1* probe and the *cry1B.34.1* gene on the plasmid.

[redacted]

Lane	Sample
1	DIG-labeled DNA marker III
2	1 copy of [redacted] + 93Y21 control soybean
3	93Y21 control soybean
4	Blank
5	COR23134 soybean T1 generation
6	COR23134 soybean T2 generation
7	COR23134 soybean T3 generation
8	COR23134 soybean T4 generation
9	COR23134 soybean T5 generation
10	Blank
11	93Y21 control soybean
12	1 copy of [redacted] + 93Y21 control soybean
13	DIG-labeled DNA marker VII

Figure 16. Southern Blot Analysis of COR23134 Soybean; *Bst*1107 I Digest with *ipd083Cb* Probe

Genomic DNA isolated from leaf tissues of COR23134 soybean from T1, T2, T3, T4, and T5 generations and 93Y21 control soybean plants was digested with *Bst*1107 I and hybridized to the *ipd083Cb* probe. Approximately 10 µg of genomic DNA was digested and loaded per lane. Positive control lanes include [redacted] plasmid DNA at approximately one gene copy number and 10 µg of control soybean DNA. The arrow indicates the COR23134-specific band. Sizes of the DIG-labeled DNA Molecular Weight Marker III and VII are indicated adjacent to the blot image in kilobases (kb).

[redacted]

Lane	Sample	Lane	Sample
1	DIG-labeled DNA marker III	8	COR23134 soybean T4 generation
2	1 copy of [redacted] + 93Y21 control soybean	9	COR23134 soybean T5 generation
3	93Y21 control soybean	10	Blank
4	Blank	11	93Y21 control soybean
5	COR23134 soybean T1 generation	12	1 copy of [redacted] + 93Y21 control soybean
6	COR23134 soybean T2 generation	13	DIG-labeled DNA marker VII
7	COR23134 soybean T3 generation		

Figure 17. Southern Blot Analysis of COR23134 Soybean; *Bst*1107 I Digest with *gm-hra_1* Probe

Genomic DNA isolated from leaf tissues of COR23134 soybean from T1, T2, T3, T4, and T5 generations and 93Y21 control soybean plants was digested with *Bst*1107 I and hybridized to the *gm-hra_1* probe. Approximately 10 µg of genomic DNA was digested and loaded per lane. Positive control lanes include [redacted] plasmid DNA at approximately one gene copy number and 10 µg of control soybean DNA. The arrow indicates the COR23134-specific band. The other bands in Lanes 5 to 9 were due to hybridization to endogenous sequences, as these bands were identified in both COR23134 soybean and 93Y21 control soybean (Lanes 2, 3, 11, and 12). Sizes of the DIG-labeled DNA Molecular Weight Marker III and VII are indicated adjacent to the blot image in kilobases (kb).

Multi-Generation Segregation Analysis

Segregation analysis was performed on six generations of COR23134 soybean (T1, T2, T3, T4, T5, and T6) to confirm the Mendelian inheritance pattern of the inserted DNA during the breeding process. The observed inheritance pattern predicts the segregation of these genes within COR23134 soybean as a single genetic locus throughout the commercial breeding process.

The genotypic analysis was performed by endpoint qualitative real-time polymerase chain reaction (referred to as endpoint PCR) on individual seed chip samples to determine the presence or absence of event COR-23134-4. Except for the T1 generation which used 25 seed chips, 100 seed chips from each of T2, T3, T4, T5, and T6 generations of COR23134 soybean were tested prior to planting. Remnant seed corresponding to each seed chip tested by endpoint PCR were planted for later use including herbicide tolerance evaluation. The phenotypic analysis was performed by a visual herbicide injury evaluation to confirm the presence or absence of tolerance to herbicide treatment for each individual plant. Diclosulam (an active ingredient for triazolopyrimidine herbicides) was applied to 100 plants from each of T4, T5, and T6 generations of COR23134 soybean at the V3 growth stage. The individual results for each plant were compared to the genotypic endpoint PCR results to verify co-segregation of both genotype and phenotype. A chi-square test was performed at the 0.05 significance level to compare the observed segregation ratios of T1, T2, and T3 generations of COR23134 soybean. A chi-square test was not performed for T4, T5 and T6 generations of COR23134 soybean as all plants were identified as positive as expected for a homozygous generation.

A summary of segregation results for COR23134 soybean (T1, T2, T3, T4, T5, and T6 generations) is provided in Table 9. No statistically significant deviation from the expected segregation ratio was found in the T1, T2, and T3 generations of COR23134 soybean, and the T4, T5, and T6 generations were confirmed to be non-segregating. The genotypic analyses based on endpoint PCR results demonstrated that the observed segregation ratios matched the expected segregation ratios for all six generations. The phenotypic analysis for tolerance to diclosulam herbicide aligned with the results for genotypic analyses.

The results of the multi-generation segregation analysis demonstrated that the inserted DNA in COR23134 soybean segregated as a single locus in accordance with Mendelian rules of inheritance for a single genetic locus, indicating stable integration of the insert into the soybean genome and a stable genetic inheritance pattern across generations during the breeding process.

Additional details regarding analytical methods for the multi-generation segregation analysis are provided in Appendix C.

Table 9. Summary of Genotypic and Phenotypic Segregation Analyses for Six Generations of COR23134 Soybean

Generation	Expected Segregation Ratio	Observed Segregation[a]			Statistical Analysis	
	(Positive:Negative)	Positive	Negative	Total	Chi-Square[b]	P-Value
T1[c]	3:1	20	5	25	0.33	0.5637
T2	3:1	71	29	100	0.85	0.3556
T3	3:1	78	22	100	0.48	0.4884
T4	Homozygous	100	0	100	--	--
T5	Homozygous	100	0	100	--	--
T6	Homozygous	100	0	100	--	--

a The observed segregation ratio was determined using qualitative endpoint PCR results.
b Degrees of freedom = 1. A chi-square value greater than 3.84 (P-value less than 0.05) would indicate a significant difference.
c T1 generation consisted of 25 seed and was included in the Mendelian segregation statistical analysis; however, the plants were only used for Southern molecular analysis.

B. CHARACTERIZATION AND SAFETY ASSESSMENT OF NEW SUBSTANCES

B.1 Characterization and safety assessment of new substances

There are no new substances associated with COR23134 soybean other than the proteins encoded by the new genes (see section B.2 *New proteins* below).

B.2 New proteins

Cry1B.34.1 Protein

Amino Acid Sequence of the Cry1B.34.1 Protein

The deduced amino acid sequence from the translation of the *cry1B.34.1* gene encodes the Cry1B.34.1 protein that is 665 amino acids in length and has a molecular weight of approximately 75 kDa (Figure 18).

```
  1  MAPSNRKNEN EIINAVSNHS AQMDLSLDAR IEDSLCVAEV NNIDPFVSAS
 51  TVQTGISIAG RILGVLGVPF AGQLASFYSF LVGELWPSGR DPWEIFLEHV
101  EQLIRQQVTE NTRNTAIARL EGLGRGYRSY QQALETWLDN RNDARSRSII
151  LERYVALELD ITTAIPLFSI RNQEVPLLMV YAQAANLHLL LLRDASLFGS
201  EWGMSSADVN QYYQEQIRYT EEYSNHCVQW YNTGLNNLRG TNAESWLRYN
251  QFRRDLTLGV LDLVALFPSY DTRTYPINTS AQLTREIYTD PIGRTNAPSG
301  FASTNWFNNN APSFSAIEAA IFRPPHLLDF PEQLTIYSAS SRWSSTQHMN
351  YWVGHRLNFR PIGGTLNTST QGLTNNTSIN PVTLQFTSRD VYRTESNAGT
401  NILFTTPVNG VPWARFNFIN PQNIYERGAT TYSQPYQGVG IQLFDSETEL
451  PPETTERPNY ESYSHRLSHI GLIIGNTLRA PVYSWTHRSA TLTNTIDPER
501  INQIPLVKGF RVWGGTSVIT GPGFTGGDIL RRNTFGDFVS LQVNINSPIT
551  QRYRLRFRYA SSRDARVIVL TGAASTGVGG QVSVNMPLQK TMEIGENLTS
601  RTFRYTDFSN PFSFRANPDI IGISEQPLFG AGSISSGELY IDKIEIILAD
651  ATFEAESDLE RAQKA*
```

Figure 18. Deduced Amino Acid Sequence of the Cry1B.34.1 Protein

The deduced amino acid sequence from the translation of the *cry1B.34.1* gene from plasmid [redacted] The asterisk (*) indicates the translational stop codon. The Cry1B.34.1 protein is 665 amino acids in length and has a molecular weight of approximately 75 kDa.

Function and Activity of the Cry1B.34.1 Protein

Cry1B.34.1 Expressed in COR23134 Soybean

The Cry1B.34.1 protein expressed in COR23134 soybean is encoded by the *cry1B.34.1* gene, a gene composed of sequences from a *cry1B*-class gene and the *cry1Ca1* gene, all derived from *Bacillus thuringiensis* (*Bt*). The expressed Cry1B.34.1 protein confers control of certain susceptible lepidopteran pests by causing disruption of the midgut epithelium. The Cry1B.34.1 protein binds to specific receptors in the brush border membrane of certain susceptible lepidopteran pests and causes cell death through the formation of non-specific, ion-conducting pores in the apical membrane of the midgut epithelial cells.

Full-length Cry1B.34 and Truncated Cry1B.34.1

There are two versions of the Cry1B.34 protein for deployment in different crops. The Cry1B.34 protein is expressed in maize event DP-91Ø521-2 (referred to as DP910521 maize; recently evaluated by FSANZ in the A1281 application), and the Cry1B.34.1 protein is expressed in COR23134 soybean. Both the Cry1B.34 and Cry1B.34.1 proteins, when expressed in planta, confer control of certain susceptible lepidopteran pests of maize and soybean, respectively. DP910521 maize contains the *cry1B.34* gene expressing the full-length Cry1B.34 protein that is 1149 amino acids in length with a molecular weight of approximately 129 kDa (Figure 19). COR23134 soybean contains the *cry1B.34.1* gene expressing the truncated Cry1B.34.1 protein that is 665 amino acids in length with a molecular weight of approximately 75 kDa (Figure 19). The *cry1B.34.1* gene was created by the removal of the region from the *cry1B.34* gene that encodes the C-terminal 484-amino acids crystal forming domain of the full-length Cry1B.34 protein. The amino acid sequences of the full-length Cry1B.34 protein in DP910521 maize and the truncated Cry1B.34.1 protein in COR23134 soybean are identical except the Cry1B.34.1 protein in COR23134 soybean lacks the last 484 amino acids at the C-terminus, Figure 20).

Cry Protein Structure and Mode of Action

Cry proteins have a well characterized and multi-step mode of action (Jurat-Fuentes and Crickmore, 2017; Soberón *et al.*, 2016; Tetreau *et al.*, 2021). The bacterium *Bt* produces a large quantity of these Cry proteins that are packaged into a distinct crystalline form during a specific stage of the lifecycle of the bacterium. The purpose of the crystal form in nature is to help stabilize the proteins in the environment after the bacterial cell has died or entered a spore state (Tetreau *et al.*, 2021). The Cry proteins produced and crystallized by *Bt* (referred to as protoxins) contain protoxin portions at the N-terminus and C-terminus in addition to the three-domain toxin core.

In order for Cry proteins to elicit an effect on a sensitive organism, several steps need to occur. First, the Cry protoxin must be ingested by an insect. Once ingested, the alkaline environment of the insect digestive system solubilizes the ingested protein. The Cry protein becomes activated by specific proteases within the insect midgut that cleave the protoxin portions of the N-terminal peptide and the C-terminal crystal forming domain. Once activated, specific regions of the activated protein are then able to bind to unique and corresponding receptors located within the insect midgut. Once bound to those specific receptors, the protein then undergoes a conformational change and is able to form a pore through the insect epithelial cells. These pores cause midgut epithelial cell disruption, loss of ion regulation, and ultimately result in the death of the insect.

Functional Equivalency between Full-length Cry1B.34 and Truncated Cry1B.34.1

Genetically modified crops expressing Cry proteins generally produce either the protoxin form of the protein or a truncated form containing the toxin core portion without the crystal forming domain. For Cry1 proteins, the "crystal forming domain" at the C-terminus generally comprises about half of the protoxin size (Tetreau *et al.*, 2021). This holds true for the "full-length" Cry1B.34 protein expressed in DP910521 maize. The truncated version of the Cry1B.34.1 protein expressed in COR23134 soybean lacks this crystal forming domain; however, its toxin core sequence is identical with the full-length Cry1B.34 protein (Figure 20). As the crystal forming domain region of Cry proteins is cleaved within the insect midgut before the protein is rendered active, and before the protein can bind to specific receptors within the insect midgut, it has no role in the toxic effects of Cry proteins (Bravo *et al.*, 2007).

Based on the significant weight of evidence provided above, it is concluded that it is appropriate to reference and rely upon the safety studies performed using the full-length Cry1B.34 protein to assess the safety of COR23134 soybean. Therefore, the microbially derived Cry1B.34 protein, containing the identical toxin core amino acid sequence as the COR23134 soybean-derived Cry1B.34.1 protein, was utilized for the safety assessments of thermolability analysis, digestibility analyses using *in vitro* gastric and intestinal digestion models, and acute oral toxicity.

[redacted]

Figure 19. Domain Organization of the Full-length Cry1B.34 and Truncated Cry1B.34.1 Proteins

The Cry1B.34 protein encoded by the *cry1B.34* gene in DP910521 maize is a full-length (also referred to as the protoxin) Cry protein consisting of a short peptide at the N-terminus (aa 1-30), a three-domain toxin core (aa 31-647), and a crystal forming domain at the C-terminus (aa 665-1149). The Cry1B.34.1 protein encoded by the *cry1B.34.1* gene in COR23134 soybean is identical to the Cry1B.34 protein except it lacks the last 484 amino acids (aa 666-1149) that comprise the C-terminal crystal forming domain of the full-length Cry1B.34 protein.

```
Cry1B.34     MAPSNRKNEN EIINAVSNHS AQMDLSLDAR IEDSLCVAEV NNIDPFVSAS    50
Cry1B.34.1   MAPSNRKNEN EIINAVSNHS AQMDLSLDAR IEDSLCVAEV NNIDPFVSAS    50

Cry1B.34     TVQTGISIAG RILGVLGVPF AGQLASFYSF LVGELWPSGR DPWEIFLEHV   100
Cry1B.34.1   TVQTGISIAG RILGVLGVPF AGQLASFYSF LVGELWPSGR DPWEIFLEHV   100

Cry1B.34     EQLIRQQVTE NTRNTAIARL EGLGRGYRSY QQALETWLDN RNDARSRSII   150
Cry1B.34.1   EQLIRQQVTE NTRNTAIARL EGLGRGYRSY QQALETWLDN RNDARSRSII   150

Cry1B.34     LERYVALELD ITTAIPLFSI RNQEVPLLMV YAQAANLHLL LLRDASLFGS   200
Cry1B.34.1   LERYVALELD ITTAIPLFSI RNQEVPLLMV YAQAANLHLL LLRDASLFGS   200

Cry1B.34     EWGMSSADVN QYYQEQIRYT EEYSNHCVQW YNTGLNNLRG TNAESWLRYN   250
Cry1B.34.1   EWGMSSADVN QYYQEQIRYT EEYSNHCVQW YNTGLNNLRG TNAESWLRYN   250

Cry1B.34     QFRRDLTLGV LDLVALFPSY DTRTYPINTS AQLTREIYTD PIGRTNAPSG   300
Cry1B.34.1   QFRRDLTLGV LDLVALFPSY DTRTYPINTS AQLTREIYTD PIGRTNAPSG   300

Cry1B.34     FASTNWFNNN APSFSAIEAA IFRPPHLLDF PEQLTIYSAS SRWSSTQHMN   350
Cry1B.34.1   FASTNWFNNN APSFSAIEAA IFRPPHLLDF PEQLTIYSAS SRWSSTQHMN   350

Cry1B.34     YWVGHRLNFR PIGGTLNTST QGLTNNTSIN PVTLQFTSRD VYRTESNAGT   400
Cry1B.34.1   YWVGHRLNFR PIGGTLNTST QGLTNNTSIN PVTLQFTSRD VYRTESNAGT   400

Cry1B.34     NILFTTPVNG VPWARFNFIN PQNIYERGAT TYSQPYQGVG IQLFDSETEL   450
Cry1B.34.1   NILFTTPVNG VPWARFNFIN PQNIYERGAT TYSQPYQGVG IQLFDSETEL   450

Cry1B.34     PPETTERPNY ESYSHRLSHI GLIIGNTLRA PVYSWTHRSA TLTNTIDPER   500
Cry1B.34.1   PPETTERPNY ESYSHRLSHI GLIIGNTLRA PVYSWTHRSA TLTNTIDPER   500

Cry1B.34     INQIPLVKGF RVWGGTSVIT GPGFTGGDIL RRNTFGDFVS LQVNINSPIT   550
Cry1B.34.1   INQIPLVKGF RVWGGTSVIT GPGFTGGDIL RRNTFGDFVS LQVNINSPIT   550

Cry1B.34     QRYRLRFRYA SSRDARVIVL TGAASTGVGG QVSVNMPLQK TMEIGENLTS   600
Cry1B.34.1   QRYRLRFRYA SSRDARVIVL TGAASTGVGG QVSVNMPLQK TMEIGENLTS   600

Cry1B.34     RTFRYTDFSN PFSFRANPDI IGISEQPLFG AGSISSGELY IDKIEIILAD   650
Cry1B.34.1   RTFRYTDFSN PFSFRANPDI IGISEQPLFG AGSISSGELY IDKIEIILAD   650

Cry1B.34     ATFEAESDLE RAQKAGAGLF TRTRDGLQVN VTDYQVDRAA NLVSCLSDEQ   700
Cry1B.34.1   ATFEAESDLE RAQKA*---- ---------- ---------- ----------   665

Cry1B.34     YSHDKKMLME AVRAAKRLSR ERNLLQDPDF NEINSTEENG WKASNGIIIS   750
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     EGGPFFKGRV LQLASARENY PTYIYQKVDA SVLKPYTRYR LDGFVKSSED   800
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     LEIDLVHQHK VHLVKNVPDN LVSDTYPDGS CRGVNRCDEQ HQVDVQIDTE   850
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     HHPMDCCEAA QTHEFSSYIN TGDLNSSVDQ GIWVVLKVRT ADGYATLGNL   900
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     ELVEVGPLSG ESLEREQRDN AKWNAELGRE RAETDRVYLA AKQAINHLFV   950
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     DYQDQQLNPE IGLAEINEAS NLVESITGVY SDTVLQIPGI SYEIYTELSD  1000
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     RLQQASYLYT SRNAVQNGDF DSGLDSWNAT TDASVQQDGN MHFLVLSHWD  1050
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     AQVTQQLRVN PNCKYVLRVT ARKVGGGDGY VTIRDGAHHR ETLTFNACDY  1100
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------

Cry1B.34     DVNGTYVNDN TYITKEVVFY PHTEHTWVEV SESEGAFYID SIELIETQE*  1149
Cry1B.34.1   ---------- ---------- ---------- ---------- ----------
```

Figure 20. Alignments of the Deduced Amino Acid Sequence of the Cry1B.34 and Cry1B.34.1 Proteins Encoded by the *cry1B.34* and *cry1B.34.1* Genes

Deduced amino acid sequence alignments show the Cry1B.34.1 protein sequence encoded by the *cry1B.34.1* gene in COR23134 soybean are identical to the Cry1B.34 sequence encoded by the *cry1B.34* gene in DP910521 maize except for the Cry1B.34.1 protein in COR23134 soybean lacking the last 484 amino acids (aa 666-1149), the crystal forming domain portion at the C-terminus. Asterisks (*) indicate the translational stop codons.

Characterization of the Cry1B.34.1 Protein Derived from COR23134 Soybean

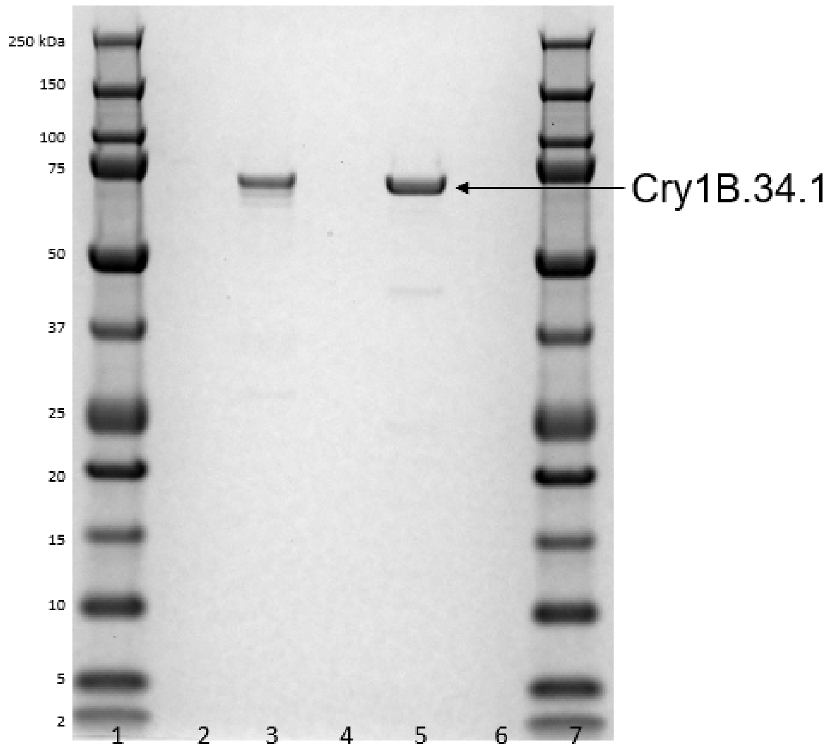
The Cry1B.34.1 protein expressed in COR23134 soybean was purified from the whole plant tissue using immunoaffinity chromatography.

The biochemical characteristics of the COR23134 soybean-derived Cry1B.34.1 protein were characterized using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot, protein glycosylation analysis, peptide mapping by liquid chromatography mass spectrometry (LC-MS), and N-terminal amino acid sequencing. The results showed that the COR23134 soybean-derived Cry1B.34.1 protein has the expected molecular weight, immunoreactivity, and amino acid sequence, and is not glycosylated.

SDS-PAGE Analysis

The sample of the COR23134 soybean-derived Cry1B.34.1 protein, along with a microbially derived Cry1B.34.1 protein, was analyzed by SDS-PAGE. As expected, the Cry1B.34.1 proteins derived from COR23134 soybean and the microbial system migrated as a predominant band consistent with the expected molecular weight (Figure 21).

Additional details regarding SDS-PAGE analytical methods are provided in Appendix E.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	COR23134 Soybean-Derived Cry1B.34.1 Protein
4	1X LDS/DTT Sample Buffer Blank
5	Microbially Derived Cry1B.34.1 Protein (1 μg)
6	1X LDS/DTT Sample Buffer Blank
7	Pre-stained Protein Molecular Weight Marker[a]

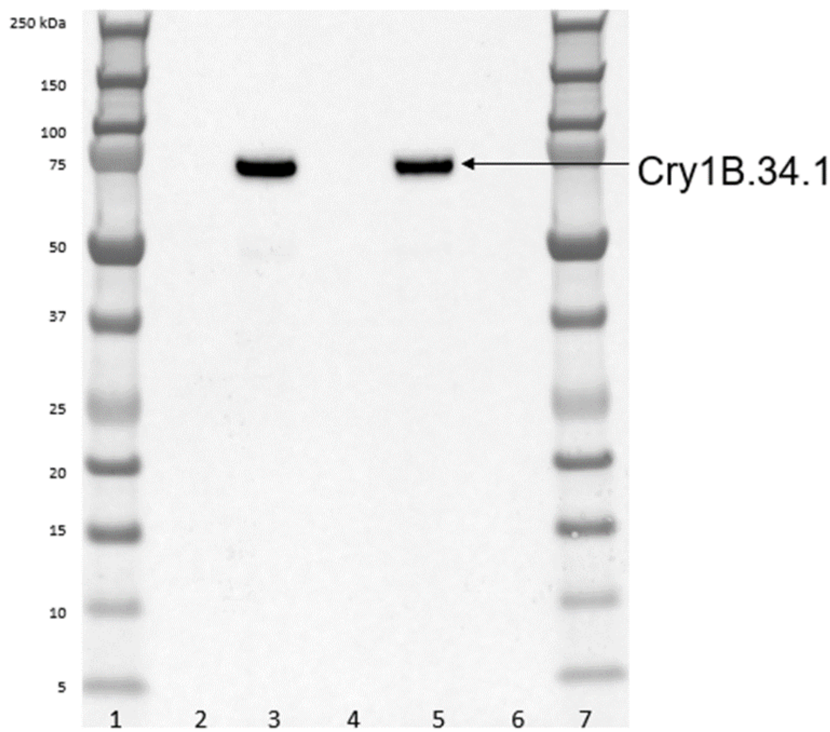
Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 21. SDS-PAGE Analysis of the Cry1B.34.1 Protein
Coomassie blue staining of the SDS-PAGE gel demonstrated the protein migrated as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived Cry1B.34.1 protein (Lane 3).

Western Blot Analysis

The sample of the COR23134 soybean-derived Cry1B.34.1 protein, along with a microbially derived Cry1B.34.1 protein, was analyzed by Western blot. As expected, the Cry1B.34.1 proteins derived from COR23134 soybean and the microbial system are immunoreactive and have the expected molecular weight (Figure 22).

Additional details regarding Western blot analytical methods are provided in Appendix E.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Microbially Derived Cry1B.34.1 Protein (10 ng)
4	1X LDS/DTT Sample Buffer Blank
5	COR23134 Soybean-Derived Cry1B.34.1 Protein
6	1X LDS/DTT Sample Buffer Blank
7	Pre-stained Protein Molecular Weight Marker[a]

Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 22. Western Blot Analysis of the Cry1B.34.1 Protein
Western blot analysis demonstrated that the Cry1B.34.1 protein was immunoreactive to monoclonal antibody and visible as a predominant band consistent with the expected molecular weight (Lane 5).

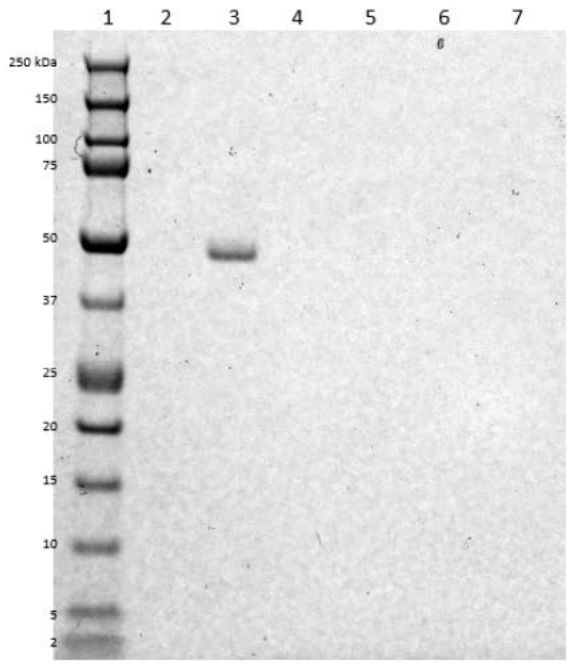
Protein Glycosylation Analysis

The sample of the COR23134 soybean-derived Cry1B.34.1 protein was analyzed by SDS-PAGE followed by the glycoprotein staining for glycosylation analysis. The gel also included a positive control (horseradish peroxidase) and a negative control (soybean trypsin inhibitor). The gel was first stained using a Pierce Glycoprotein Staining Kit to visualize any glycoproteins, imaged, and then stained with the Coomassie blue reagent to visualize all protein bands.

Glycosylation was determined to be negative for the COR23134 soybean-derived protein (Figure 23). The horseradish peroxidase positive control was clearly visible as a stained band. The soybean trypsin inhibitor negative control was not stained by the glycoprotein stain.

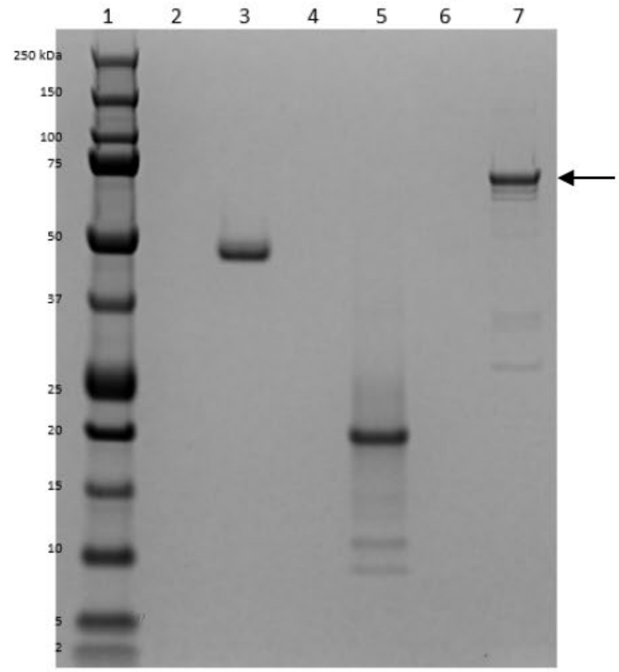
Additional details regarding glycoprotein analytical methods are provided in Appendix E.

A) Glycoprotein Staining



Glycoprotein Stain

B) Total Protein Staining



Total Protein Stain

Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Horseradish Peroxidase Positive Control (1.0 µg)
4	1X LDS/DTT Sample Buffer Blank
5	Soybean Trypsin Inhibitor Negative Control (1.0 µg)
6	1X LDS/DTT Sample Buffer Blank
7	COR23134 Soybean-Derived Cry1B.34.1 Protein

Note: The glycoprotein gel was stained with glycoprotein staining reagent. The total protein stain gel was stained with glycoprotein staining reagent followed by staining with Coomassie blue reagent for total proteins. Kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 23. Glycosylation Analysis of the COR23134 Soybean-Derived Cry1B.34.1 Protein
A) Glycoprotein staining: Glycosylation was not detected for the COR23134 soybean-derived Cry1B.34.1 protein (Lane 7). The horseradish peroxidase positive control was stained (Lane 3), and the soybean trypsin inhibitor negative control was not stained (Lane 5). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the COR23134 soybean-derived Cry1B.34.1 protein (Lane 7) and both the positive (Lane 3) and negative (Lane 5) control proteins.

Mass Spectrometry Peptide Mapping Analysis

The sample of the COR23134 soybean-derived Cry1B.34.1 protein was analyzed by SDS-PAGE. The gel was stained with Coomassie blue reagent, and the bands containing the Cry1B.34.1 protein were excised. The excised Cry1B.34.1 protein bands were digested with trypsin or chymotrypsin. Digested samples were analyzed using liquid chromatography-mass spectrometry (LC-MS). The resulting MS data were used to search and match the peptides from the Cry1B.34.1 protein sequence, and the combined sequence coverage was calculated for the COR23134 soybean-derived Cry1B.34.1 protein.

The combined sequence coverage of the identified tryptic and chymotryptic peptides for COR23134 soybean-derived Cry1B.34.1 protein accounts for 93.7% (622/664) of the amino acid sequence (Figure 24).

Additional details regarding peptide mapping analytical methods are provided in Appendix E.

1 APSNR**KNENE** **IINAVSNHSA** **QMDLSLDARI** **EDSLCVAEVN** **NIDPFVSAST**
51 **VQTGISIAGR** **IL**GVLGVPFA GQLASFYSF**L** **VGELWPSGRD** **PWEIFLEHVE**
101 **QLIRQQVTEN** **TRNTAIARLE** **GLGRGYRSYQ** **QALETWLDNR** **NDARSRSIIL**
151 **ERYVALELDI** **TTAIPLFSIR** **NQEVPLLMVY** **AQAANLHLLL** **LRDASLFGSE**
201 **WGMSSADVNQ** **YYQEQIRYTE** **EYSNHCVQWY** **NTGLNNLRGT** **NAESWLRY**NQ
251 F**RRDLTLGVL** **DLVALFPSYD** **TRTYPINTSA** **QLTREIYTDP** **IGRTNAPSGF**
301 ASTNW**FNNNA** **PSFSAIEAAI** **FRPPHLLDFP** **EQL**TIYSASS R**WSSTQHMNY**
351 **WVGHRLNFRP** **IGGTLNTSTQ** **GLTNNTSINP** **VTLQFTSRDV** **YRTESNAGTN**
401 **ILFTTPVNGV** **PWARFNFINP** **QNIYERGATT** **YSQPYQGVGI** **QLFDSETELP**
451 **PETTERPNYE** **SYSHRLSHIG** **LIIGNTLRAP** **VYSWTHRSAT** **LTNTIDPERI**
501 **NQIPLVKGFR** **VWGGTSVITG** **PGFTGGDILR** **RNTFGDFVSL** **QVNINSPITQ**
551 **RY**RLRF**RYAS** **SRDARVIVLT** **GAASTGVGGQ** **VSVNMPLQKT** **MEIGENLTSR**
601 **TFRYTDFSNP** **FSFRANPDII** **GISEQPLFGA** **GSISSGELYI** **DKIEIILADA**
651 **TFEAESDLER** **AQKA**

Red type	Bold red type indicates soybean-derived Cry1B.34.1 peptides identified using LC-MS analysis against the Cry1B.34.1 protein sequence.
Amino acid residue abbreviations	alanine (A), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

Note: The Cry1B.34.1 protein sequence does not include the N-terminal methionine as it is anticipated to be absent.

Figure 24. Identified Tryptic and Chymotryptic Peptide Amino Acid Sequence of the COR23134 Soybean-Derived Cry1B.34.1 Protein Using LC-MS Analysis

N-Terminal Amino Acid Sequence Analysis

The Edman sequencing analysis of the COR23134 soybean-derived Cry1B.34.1 protein sample identified an N-terminal sequence (APS), matching the amino acid residues 1-3 of the Cry1B.34.1 protein sequence without the N-terminal methionine as expected (Dummitt *et al.*, 2003; Sherman *et al.*, 1985).

Additional details regarding N-terminal amino acid sequencing analytical methods are provided in Appendix E.

Bioactivity Assay

The biological activity of the microbially derived Cry1B.34.1 protein was evaluated by conducting a 7-day bioassay using *Spodoptera frugiperda* (fall armyworm; Lepidoptera: Noctuidae), a species sensitive to the Cry1B.34.1 protein.

Bioactivity analysis demonstrated that the microbially derived Cry1B.34.1 protein had insecticidal activity toward a target insect, *S. frugiperda* (Table 10). The biological activity of the test diet containing 15 ng of the Cry1B.34.1 protein per mg wet diet was demonstrated by 100% mortality in *S. frugiperda* compared to 0% mortality in *S. frugiperda* fed the buffer control diet.

Additional details regarding bioactivity assay methods are provided in Appendix E.

Table 10. Summary of the Microbially Derived Cry1B.34.1 Protein Bioactivity Assay Using *Spodoptera frugiperda*

Treatment	Treatment Description	Concentration (ng Cry1B.34.1 Protein/mg Wet Diet)	Total Number of Observations	Total Number of Dead Organisms	Mortality (%)
1	Buffer Control Diet	0	20	0	0
2	Test Diet	15	20	20	100

Note: The summary of *Spodoptera frugiperda* mortality data consisted of the calculation of dead larvae divided by the total number of observed larvae at the end of the study and multiplied by 100. The concentration of the Cry1B.34.1 protein in Treatment 2 was based on the wet weight of the sensitive insect artificial diet.

Allergenicity and Toxicity Analyses of the Cry1B.34.1 Protein

A weight-of-evidence approach was applied to determine the allergenic and toxic potential of the Cry1B.34.1 protein expressed in COR23134 soybean, including an assessment of the following: the history of safe use of the source organism, a bioinformatic comparison of the amino acid sequence of the Cry1B.34.1 protein to known or putative allergen and protein toxin sequences, an evaluation of the heat lability of the Cry1B.34 protein using a sensitive insect bioassay, evaluations of the stability of the Cry1B.34 protein using *in vitro* gastric and intestinal digestion models, determination of the Cry1B.34.1 protein glycosylation status, and an evaluation of acute toxicity in mice following oral exposure to the Cry1B.34 protein.

Source Organism of the Cry1B.34.1 Protein

The Cry1B.34.1 protein expressed in COR23134 soybean is encoded by the *cry1B.34.1* gene, a gene composed of sequences from a *cry1B*-class gene and the *cry1Ca1* gene, all derived from *Bacillus thuringiensis* (*Bt*).

Bt is a diverse group of Gram-positive, spore-forming bacteria that has a history of safe use as a pesticide over several decades (US-EPA, 1998; US-EPA, 2001). It occurs ubiquitously in the soil and on plants including vegetables, cotton, tobacco, tree crops, and forest crops (Schnepf et al., 1998; Shelton, 2012). Several Cry proteins have been deployed as safe and effective pest control agents in microbial *Bt* formulations for almost 40 years. Several Cry proteins have also been effectively deployed as safe and effective pest control agents and have a history of safe use in genetically modified crops (ISAAA, 2023).

Bioinformatic Analysis of Cry1B.34.1 Homology to Known and Putative Allergens

Assessing newly expressed proteins for potential cross-reactivity with known and putative allergens is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). A bioinformatic assessment of the Cry1B.34.1 protein sequence (665 amino acids [aa]) for potential cross-reactivity with allergens was conducted by following established international criteria (Codex Alimentarius Commission, 2009; FAO/WHO, 2001).

Two separate searches for the Cry1B.34.1 protein sequence were performed using the Comprehensive Protein Allergen Resource (COMPARE) 2023 database (January 26, 2023 (van Ree *et al.*, 2021)). This peer-reviewed database is a collaborative effort of the Health and Environmental Sciences Institute (HESI) Protein Allergens, Toxins, and Bioinformatics (PATB) Committee and contains 2,631 sequences.

The first search was the sliding 80-mer window search, accomplished with an internally developed Perl script running FASTA v35.04 (Pearson and Lipman, 1988) with an *E*-score cutoff set to 100. In a sliding window search, each sequentially overlapping 80 aa sub-sequence of the overall Cry1B.34.1 protein sequence is used as a query against the COMPARE allergen database sequences. The script examined all alignments generated from the query and reported any possessing > 35% identity over an alignment length of ≥ 80 aa. Additionally, the script rescaled the percent identity to an 80-mer window for any alignments possessing an alignment length shorter than 80 aa; the number of identities in these alignments would be divided by 80, then multiplied by 100, and would report any alignment possessing an adjusted percent identity > 35%.

The second search used EMBOSS fuzzpro v6.6.0 (Rice *et al.*, 2000) to identify any eight or greater contiguous identical amino acid matches between the Cry1B.34.1 protein sequence and the COMPARE allergen sequences.

Results of the search of the Cry1B.34.1 protein sequence against the COMPARE allergen database sequences found no alignments that were a length of 80 aa or greater with a sequence identity of > 35% and no alignments shorter than 80 aa with a sequence identity > 35% when normalized to an 80-mer window. No contiguous 8-residue exact matches between the Cry1B.34.1 protein sequence and the allergen sequences were identified in the second search. Collectively, these data indicate that no allergenicity concern arose from the bioinformatics assessment of the Cry1B.34.1 protein.

Bioinformatics evaluation of the Cry1B.34.1 protein sequence did not generate biologically relevant amino acid sequence similarities to allergens that are harmful to humans or animals.

Bioinformatic Analysis of Cry1B.34.1 Homology to Known and Putative Protein Toxins

Assessing newly expressed proteins for potential sequence similarity with protein toxins is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). The potential toxicity of the Cry1B.34.1 protein was assessed by comparison of its sequence 1) an internal toxin database, and 2) the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein database.

The internal toxin database is a subset of sequences found in UniProtKB/Swiss-Prot (The UniProt Consortium, 2023). UniProtKB/Swiss-Prot is a curated database of non-redundant proteins containing functional information for over 550,000 sequences. To produce the internal toxin database, the proteins in UniProtKB/Swiss Prot are filtered for molecular function by keywords that could imply toxicity or adverse health effects (e.g., toxin, hemagglutinin, vasoactive). The internal toxin database is updated annually and contains 8,858 sequences.

The search between the Cry1B.34.1 protein sequence and protein sequences in the internal toxin database was conducted using BLASTP v2.10.0+ with an E-value set to 10^{-4}. No alignments were returned between the Cry1B.34.1 protein sequence and any protein sequence in the internal toxin database. Therefore, no toxicity concern arose from the bioinformatics assessment of the Cry1B.34.1 protein.

The BLASTP search of the Cry1B.34.1 protein against the NCBI nr protein database returned the maximum number of alignment descriptions/sequence alignments, 500/250, with an E-value from 0 to 9×10^{-117} to various insecticidal crystal proteins from different *Bacillus* species. This is expected since Cry1B.34.1 is an insecticidal protein. None of the accessions returned by the BLASTP search are proteins known to be toxic to humans or animals.

Bioinformatics evaluation of the Cry1B.34.1 protein sequence did not generate biologically relevant amino acid sequence similarities to protein toxins that are harmful to humans or animals.

Thermolability Analysis

Thermal stability of the Cry1B.34 protein was characterized by determining the biological activity of the heat-treated Cry1B.34 protein when incorporated in an artificial diet and fed to *Spodoptera frugiperda* (*S. frugiperda*), an insect sensitive to the Cry1B.34 protein. The Cry1B.34 protein was incubated at various temperatures (25 °C, 50 °C, 75 °C, and 95 °C) for approximately 30 minutes before incorporation into the artificial diet. Each test diet contained a targeted concentration of 25 ng Cry1B.34 protein per mg diet wet weight. Larvae were exposed via oral ingestion to the diets in a 7-day bioassay. A positive control diet containing the unheated Cry1B.34 protein and a bioassay control diet containing buffer were included in the bioassay to verify assay performance. After seven days, statistical analyses were conducted to evaluate *S. frugiperda* mortality of the heat-treated test groups relative to the unheated test group.

The results demonstrated that the Cry1B.34 protein heated for approximately 30 minutes at temperatures of 75 °C and 95 °C (Treatments 5 and 6, respectively) was inactive against *S. frugiperda* (P values < 0.0001) compared to the unheated Cry1B.34 control (Treatment 2) when incorporated in an artificial insect diet. No statistically significant decreases in protein activity were observed for the Cry1B.34 protein heated for approximately 30 minutes at 25 °C or 50 °C (Treatments 3 and 4, respectively) when compared to the unheated Cry1B.34 control (Table 11).

Additional details regarding thermolability analytical methods are provided in Appendix E.

Table 11. Biological Activity of the Heat-Treated Cry1B.34 Protein in Artificial Diet Fed to *Spodoptera frugiperda*

Treatment	Treatment Description	Test Dosing Solution Incubation Condition	Total Number of Observations	Mortality (%)	Fisher's Test P-Value	Number of Surviving Organisms	Weight of Surviving Organisms (mg)	
							Mean ± Standard Deviation	Range
1	Buffer Control Diet	NA	20	0	--	20	36.7 ± 13.9	2.3 - 50.1
2	Unheated Control Diet	Unheated	20	100	--	0	NA	NA
3	Test Diet	25 °C	20	100	1.0000	0	NA	NA
4	Test Diet	50 °C	20	95.0	0.5000	1	0.100[a]	NA
5	Test Diet	75 °C	20	0	<0.0001[b]	20	13.8 ± 7.88	0.2 - 34.7
6	Test Diet	95 °C	20	0	<0.0001[b]	20	41.1 ± 8.55	24.0 - 59.4

Note: The unheated control diet and the test diets contained a targeted concentration of 25 ng Cry1B.34 protein per mg diet wet weight. Not applicable (NA).

[a] The reported mean is the weight value of the one surviving larva after the 7-day feeding period.

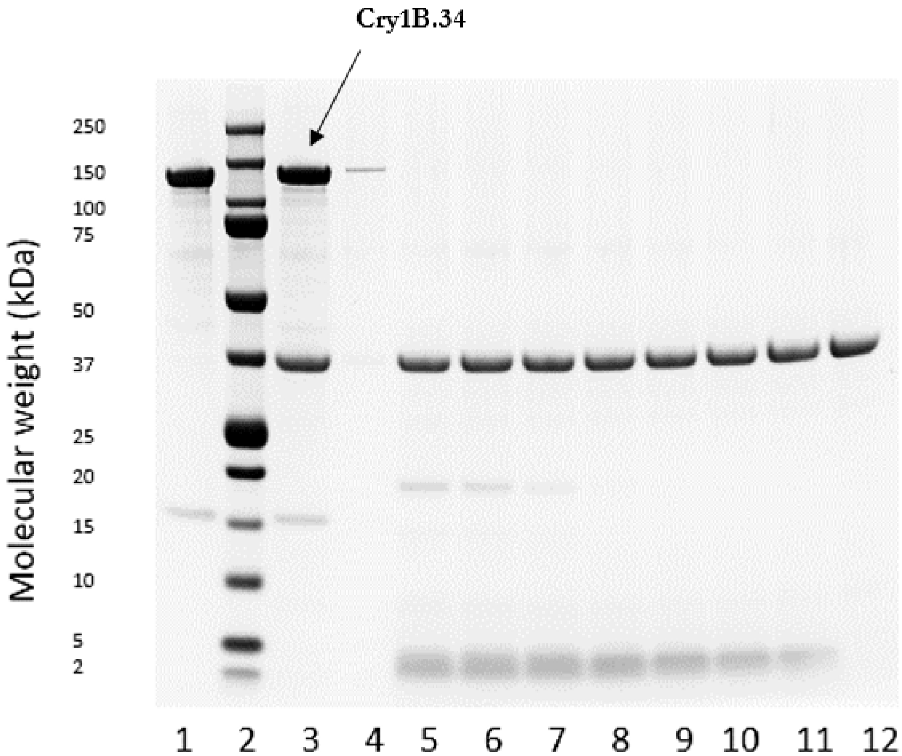
[b] A statistically significant difference (P-value < 0.05) was observed in comparison to Treatment 2.

Digestibility Analysis with Simulated Gastric Fluid (SGF)

Simulated gastric fluid (SGF) containing pepsin at pH ~1.2 was used to assess the susceptibility of the Cry1B.34 protein to proteolytic digestion by pepsin *in vitro*. The Cry1B.34 protein was incubated in SGF for 0, 0.5, 1, 2, 5, 10, 20, and 60 minutes. A positive control (bovine serum albumin) and a negative control (β-lactoglobulin) were included in the assay and were incubated in SGF for 0, 1, and 60 minutes. After incubation in SGF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SGF digestibility results showed that the Cry1B.34 protein migrating at approximately 129 kDa was digested within 0.5 minutes in SGF as demonstrated by both SDS-PAGE and western blot analysis (Figure 25 and Figure 26, respectively). A band migrating at approximately 20 kDa was digested within 5 minutes in SGF. On the SDS-PAGE gel, low molecular weight bands (~2-5 kDa) remained detectable in the Cry1B.34 protein samples for up to 60 minutes in SGF.

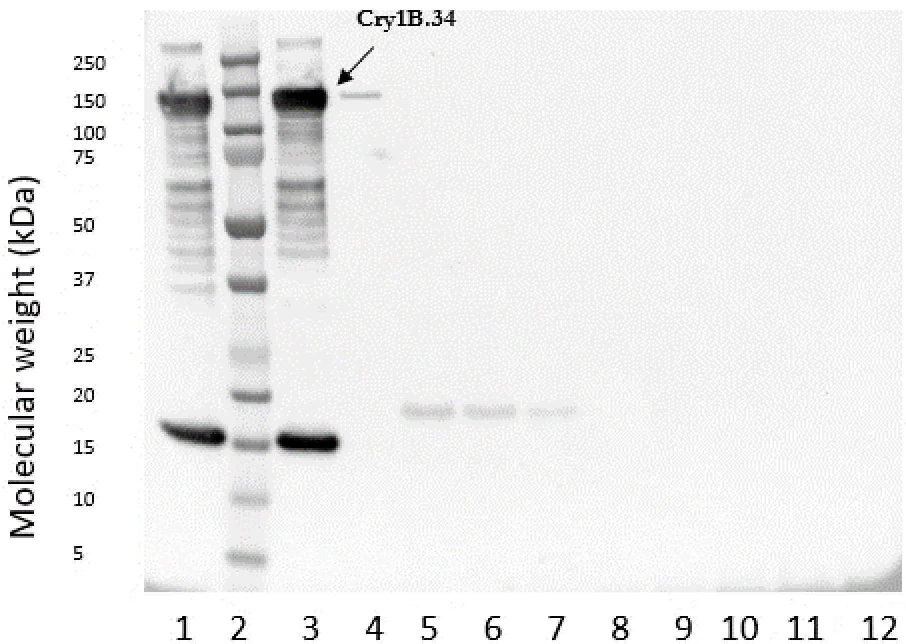
Additional details regarding SGF analytical methods are provided in Appendix E.



Lane	Sample Descriptions
1	Cry1B.34 protein in 10 mM CAPS buffer (no SGF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.34 protein in SGF, Time 0
4	Cry1B.34 protein in SGF, Time 0; 1:20 dilution
5	Cry1B.34 protein in SGF, 0.5 minutes
6	Cry1B.34 protein in SGF, 1 minute
7	Cry1B.34 protein in SGF, 2 minutes
8	Cry1B.34 protein in SGF, 5 minutes
9	Cry1B.34 protein in SGF, 10 minutes
10	Cry1B.34 protein in SGF, 20 minutes
11	Cry1B.34 protein in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: kilodalton (kDa), simulated gastric fluid (SGF), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).
[a] Molecular weight markers were included to provide a visual estimate that migration was within the expected range of the predicted molecular weight.

Figure 25. SDS-PAGE Analysis of the Cry1B.34 Protein in Simulated Gastric Fluid Digestion Time Course



Lane	Sample Descriptions
1	Cry1B.34 protein in 10 mM CAPS buffer (no SGF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.34 protein in SGF, Time 0
4	Cry1B.34 protein in SGF, Time 0; 1:200 dilution
5	Cry1B.34 protein in SGF, 0.5 minutes
6	Cry1B.34 protein in SGF, 1 minute
7	Cry1B.34 protein in SGF, 2 minutes
8	Cry1B.34 protein in SGF, 5 minutes
9	Cry1B.34 protein in SGF, 10 minutes
10	Cry1B.34 protein in SGF, 20 minutes
11	Cry1B.34 protein in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: kilodalton (kDa) and simulated gastric fluid (SGF).
[a] Molecular weight markers were included to provide a visual estimate that migration was within the expected range of the predicted molecular weight.

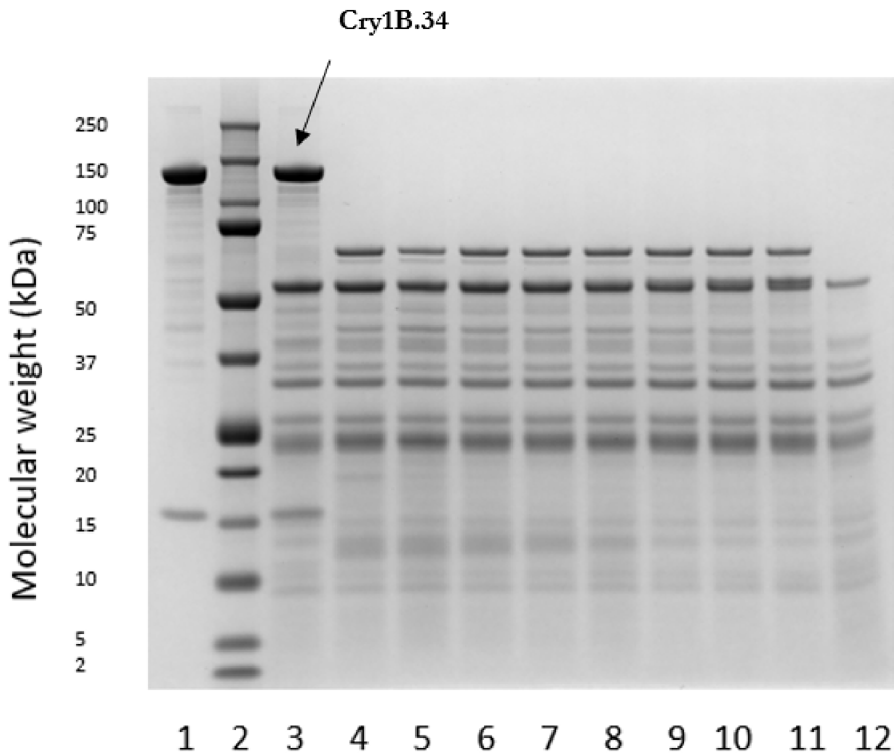
Figure 26. Western Blot Analysis of the Cry1B.34 Protein in Simulated Gastric Fluid Digestion Time Course

Digestibility Analysis with Simulated Intestinal Fluid (SIF)

Simulated intestinal fluid (SIF) containing pancreatin at ~pH 7.5 was used to assess the susceptibility of the Cry1B.34 protein to proteolytic digestion by pancreatin in vitro. The Cry1B.34 protein was incubated in SIF for 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. A positive control (β-lactoglobulin) and a negative control (bovine serum albumin) were included in the assay and were incubated in SIF for 0, 1, and 60 minutes. After incubation in SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SIF digestibility results showed that the Cry1B.34 protein migrating at approximately 129 kDa was digested into smaller fragments migrating at less than 75 kDa within 0.5 minutes in SIF as demonstrated by SDS-PAGE and within 1 minute as demonstrated by the western blot (Figure 27 and Figure 28, respectively). These smaller fragments remained detectable for up to 60 minutes.

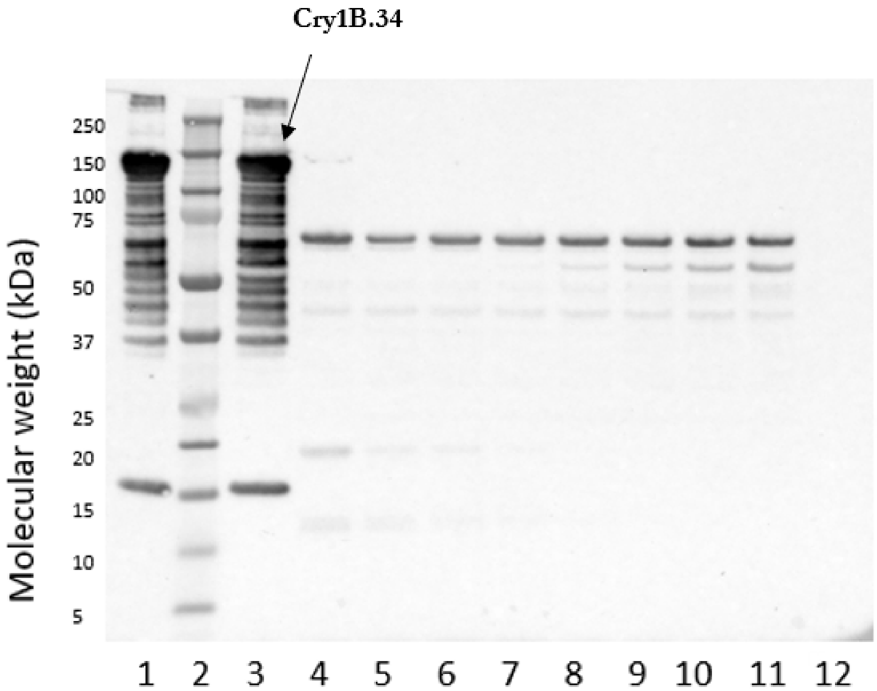
Additional details regarding SIF analytical methods are provided in Appendix E.



Lane	Sample Descriptions
1	Cry1B.34 protein in 10 mM CAPS buffer (no SIF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.34 protein in SIF, Time 0
4	Cry1B.34 protein in SIF, 0.5 minutes
5	Cry1B.34 protein in SIF, 1 minute
6	Cry1B.34 protein in SIF, 2 minutes
7	Cry1B.34 protein in SIF, 5 minutes
8	Cry1B.34 protein in SIF, 10 minutes
9	Cry1B.34 protein in SIF, 20 minutes
10	Cry1B.34 protein in SIF, 30 minutes
11	Cry1B.34 protein in SIF, 60 minutes

Note: kilodalton (kDa), simulated intestinal fluid (SIF), and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).
[a] Molecular weight markers were included to provide a visual estimate that migration was within the expected range of the predicted molecular weight

Figure 27. SDS-PAGE Analysis of the Cry1B.34 Protein in Simulated Intestinal Fluid Digestion Time Course



Lane	Sample Descriptions
1	Cry1B.34 protein in 10 mM CAPS buffer (no SIF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.34 protein in SIF, Time 0
4	Cry1B.34 protein in SIF, 0.5 minutes
5	Cry1B.34 protein in SIF, 1 minute
6	Cry1B.34 protein in SIF, 2 minutes
7	Cry1B.34 protein in SIF, 5 minutes
8	Cry1B.34 protein in SIF, 10 minutes
9	Cry1B.34 protein in SIF, 20 minutes
10	Cry1B.34 protein in SIF, 30 minutes
11	Cry1B.34 protein in SIF, 60 minutes
12	SIF Control, 60 minutes

Note: kilodalton (kDa) and simulated intestinal fluid (SIF).
[a] Molecular weight markers were included to provide a visual estimate that migration was within the expected range of the predicted molecular weight.

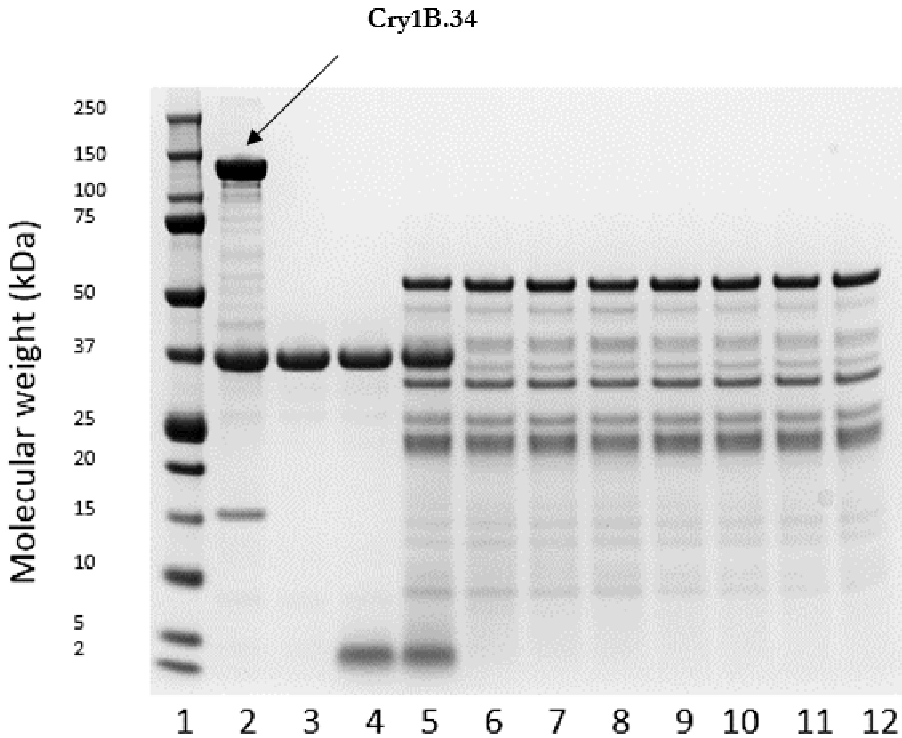
Figure 28. Western Blot Analysis of the Cry1B.34 Protein in Simulated Intestinal Fluid Digestion Time Course

Sequential Digestibility Analysis with Simulated Gastric Fluid (SGF) and Simulated Intestinal Fluid (SIF)

Sequential digestion in simulated intestinal fluid (SIF) following a digestion in SGF was used to assess the susceptibility of the low molecular weight SGF fragments (~2-5 kDa, Figure 25) of the Cry1B.34 protein. The Cry1B.34 protein was incubated for 10 minutes in SGF containing pepsin at PH ~1.2 and then incubated for 0, 0.5, 1, 2, 5, 10, 20, and 30 minutes in SIF containing pancreatin at PH~7.5. After incubation in SGF/ SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain was used to detect protein bands.

The sequential pepsin (SGF) and pancreatin (SIF) digestibility results showed that the low molecular weight bands (~2-5 kDa) observed in SGF digestion (Figure 25) were digested within 0.5 minutes during sequential SIF digestion (Figure 29).

Additional details regarding sequential digestibility analytical methods are provided in Appendix E.



Lane	Sample Descriptions
1	Pre-stained protein molecular weight marker[a]
2	Cry1B.34 Protein in SGF, Time 0
3	SGF Control, 10 minutes
4	Cry1B.34 Protein in SGF, 10 minutes
5	Cry1B.34 Protein in SGF 10 minutes, SIF Time 0
6	Cry1B.34 Protein in SGF 10 minutes, SIF 0.5 minutes
7	Cry1B.34 Protein in SGF 10 minutes, SIF 1 minute
8	Cry1B.34 Protein in SGF 10 minutes, SIF 2 minutes
9	Cry1B.34 Protein in SGF 10 minutes, SIF 5 minutes
10	Cry1B.34 Protein in SGF 10 minutes, SIF 10 minutes
11	Cry1B.34 Protein in SGF 10 minutes, SIF 20 minutes
12	Cry1B.34 Protein in SGF 10 minutes, SIF 30 minutes

Note: kilodalton (kDa), simulated gastric fluid (SGF), simulated intestinal fluid (SIF), and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

[a] Molecular weight markers were included to provide a visual estimate that migration was within the expected range of the predicted molecular weight.

Figure 29. SDS-PAGE Analysis of Cry1B.34 Protein in a Sequential Digestion with Simulated Gastric Fluid and Simulated Intestinal Fluid

Protein Glycosylation Analysis

As stated previously in the characterization section, the results from glycoprotein staining analysis confirmed the absence of glycosylation for the COR23134 soybean-derived Cry1B.34.1 protein.

Evaluation of the Acute Toxicity of the Cry1B.34 Protein

A study was conducted to evaluate the potential acute toxicity of the test substance, Cry1B.34, in [redacted] mice following oral exposure at the limit dose (5000 mg/kg body weight, adjusted for Cry1B.34 content). The Cry1B.34 protein and bovine serum albumin (BSA) protein were each reconstituted in deionized water. Vehicle control, BSA comparative control, and the Cry1B.34 test substance formulations were administered orally by gavage in three split doses, separated by approximately four hours; the BSA comparative control was administered at an equivalent target dose to that of the test substance. The mice were fasted prior to and throughout the dosing procedure.

Body weights were evaluated on test day 1 (prior to fasting and shortly prior to administration of the first dose), 2, 3, 5, 8, and 15. Clinical signs were evaluated ten times on test day 1 (distributed before and after each dose) and daily thereafter. On test day 15, all surviving mice were euthanized and given a gross pathological examination.

All animals survived to scheduled euthanasia. There were no test substance-related clinical observations and all animals gained weight during the 2-week observation period prior to euthanasia. No gross lesions were observed.

Under the conditions of this study, intragastric exposure of the Cry1B.34 protein to male and female mice at 5000 mg/kg body weight did not result in mortality or other evidence of acute oral toxicity, based on evaluation of body weight, clinical signs, and gross pathology. Therefore, the acute oral toxicity tolerant dose and the LD_{50} of the Cry1B.34 protein were determined to be greater than 5000 mg/kg body weight. After adjustment for molecular weight, the LD_{50} was determined to be greater than 2920 mg Cry1B.34.1/kg body weight.

Conclusion on Safety of the Cry1B.34.1 Protein in COR23134 Soybean

There are two versions of the Cry1B.34 protein for deployment in different crops. The Cry1B.34.1 protein encoded by the *cry1B.34.1* gene in COR23134 soybean as well as the Cry1B.34 protein encoded by the *cry1B.34* gene in DP910521 maize, when expressed in planta, confers control of certain susceptible lepidopteran pests. The *cry1B.34.1* gene was created by the removal of the region from the *cry1B.34* gene that encodes the C-terminal 484-amino acids crystal forming domain of the full-length Cry1B.34 protein. The truncated version of the Cry1B.34.1 protein expressed in COR23134 soybean lacks this crystal forming domain; however, its toxin core sequence is identical with the full-length Cry1B.34 protein. As the crystal forming domain region of Cry proteins is cleaved within the insect midgut before the protein is rendered active, and before the protein can bind to specific receptors within the insect midgut, it has no role in the toxic effects of Cry proteins. Based on the significant weight of evidence provided above, it is concluded that it is appropriate to reference and rely upon the safety studies performed using the full-length Cry1B.34 protein to assess the safety of COR23134 soybean. Therefore, the microbially derived

Cry1B.34 protein, containing the identical toxin core amino acid sequence as the COR23134 soybean-derived Cry1B.34.1 protein, was utilized for the safety assessments of thermolability analysis, digestibility analyses using in vitro gastric and intestinal digestion models, and acute oral toxicity.

In conclusion, protein characterization results via SDS-PAGE, western blot, glycosylation analysis, mass spectrometry peptide mapping analysis, and N-terminal amino acid sequence analysis have demonstrated that the Cry1B.34.1 protein derived from COR23134 soybean has the expected molecular weight, immunoreactivity, and amino acid sequence, and is not glycosylated.

The allergenic potential of the Cry1B.34.1 protein was evaluated by assessing the Cry1B.34.1 protein source organism and history of safe use, a bioinformatic comparison of the amino acid sequence of the Cry1B.34.1 protein with known and putative allergen sequences, evaluation of the heat lability of the Cry1B.34 protein using a sensitive insect bioassay, evaluation of the stability of the Cry1B.34 protein using *in vitro* gastric and intestinal digestion models, and determination of the Cry1B.34 protein glycosylation status. The toxicity potential of the Cry1B.34.1 protein was evaluated by a bioinformatic comparison of the Cry1B.34.1 amino acid sequence to known and putative protein toxins and by an acute toxicity in mice following oral exposure to the Cry1B.34 protein.

The bioinformatic comparison of the Cry1B.34.1 protein sequence to known and putative allergen and protein toxin sequences showed that the Cry1B.34.1 protein is unlikely to be allergenic or toxic for humans or animals. The Cry1B.34 protein was digested within 0.5 minutes in SGF. The band migrating at ~20 kDa was digested within 5 minutes in SGF, and some low molecular weight bands (~2-5 kDa) remained detectable for up to 60 minutes in SGF. The Cry1B.34 protein was digested within 1 minute in SIF, and some lower molecular weight bands remained visible after 60 minutes. The low molecular weight bands remaining from SGF digestion were digested within 0.5 minutes in sequential SIF. The Cry1B.34 protein was not glycosylated. The Cry1B.34 protein heated for approximately 30 minutes at 75 °C and 95 °C was inactive against *Spodoptera frugiperda* when incorporated in an artificial diet. The acute oral toxicity assessment in mice determined the LD_{50} of the Cry1B.34 protein to be greater than 5000 mg/kg. After adjustment for molecular weight, the LD_{50} was determined to be greater than 2920 mg Cry1B.34.1/kg body weight. These data support the conclusion that COR23134 soybean, expressing the Cry1B.34.1 protein, is as safe as conventional soybean for the food and feed supply.

Based on this weight of evidence, consumption of the Cry1B.34.1 protein from COR23134 soybean is unlikely to cause an adverse effect on humans or animals.

Cry1B.61.1 Protein

Amino Acid Sequence of the Cry1B.61.1 Protein

The deduced amino acid sequence from the translation of the *cry1B.61.1* gene encodes the Cry1B.61.1 protein that is 656 amino acids in length and has a molecular weight of approximately 74 kDa (Figure 30).

```
  1 MPSNRKNENE IINALSIPAV SNHSAQMDLS LDARIEDSLC IAEGNNINPL
 51 VSASTVQTGI NIAGRILGVL GVPFAGQLAS FYSFIVGELW PSGRDPWEIF
101 MEHVEQLVRQ HITMNARNTA LARLQGLGAS FRAYQQSLED WLENRDNART
151 RSVLYTQYIA LELDFLNAMP LFAINNQQVP LLMVYAQAAN LHLLLLRDAS
201 LFGSEFGLTS QEIQRYYERQ AEKTREYSDY CARWYNTGLN NLRGTNAESW
251 LRYNQFRRDL TLGVLDLVAL FPSYDTRIYP INTSAQLTRE IYTDPIGRTN
301 APSGFASTNW FNNNAPSFSA IEAAIFRPPH LLDFPEQLTI YSASSRWSST
351 QHMNYWVGHR LNFRPIGGTL NTSTHGATNT SINPVTLQFT SRDVYRTESY
401 AGINILLTTP VNGVPWARFN WRNPLNSLRG SLLYTIGYTG VGTQLFDSET
451 ELPPETTERP NYESYSHRLS NIRLIIGGTL RAPVYSWTHR SADRTNTIAT
501 NIITQIPAVK GNFLFNGLVI SGPGFTGGDL VRLNNSGNNI QNRGYIEVPI
551 QFRSTSTRYR VRVRYASVTP IRLSVNWGNS NIFSSIVPAT ATSLDNLQSR
601 NFGYFESRNA FTSATGNVVG VRNFSENAGV IIDRFEFIPV TATFEAEYDL
651 ERAQEA*
```

Figure 30. Deduced Amino Acid Sequence of the Cry1B.61.1 Protein
The deduced amino acid sequence from the translation of the *cry1B.61.1* gene from plasmid [redacted] The asterisk (*) indicates the translational stop codon. The Cry1B.61.1 protein is 656 amino acids in length and has an approximate molecular weight of 74 kDa.

Function and Activity of the Cry1B.61.1 Protein

The Cry1B.61.1 protein expressed in COR23134 soybean is encoded by the *cry1B.61.1* gene, a modified *cry1B*-class gene, derived from *Bacillus thuringiensis*. The expressed Cry1B.61.1 protein confers control of certain susceptible lepidopteran pests by causing disruption of the midgut epithelium. The Cry1B.61.1 protein binds to specific receptors in the brush border membrane of certain susceptible lepidopteran pests and causes cell death through the formation of non-specific, ion-conducting pores in the apical membrane of the midgut epithelial cells.

Characterization of the Cry1B.61.1 Protein Derived from COR23134 Soybean and the Microbial System

The Cry1B.61.1 protein expressed in COR23134 soybean was purified from the whole plant tissue using immunoaffinity chromatography.

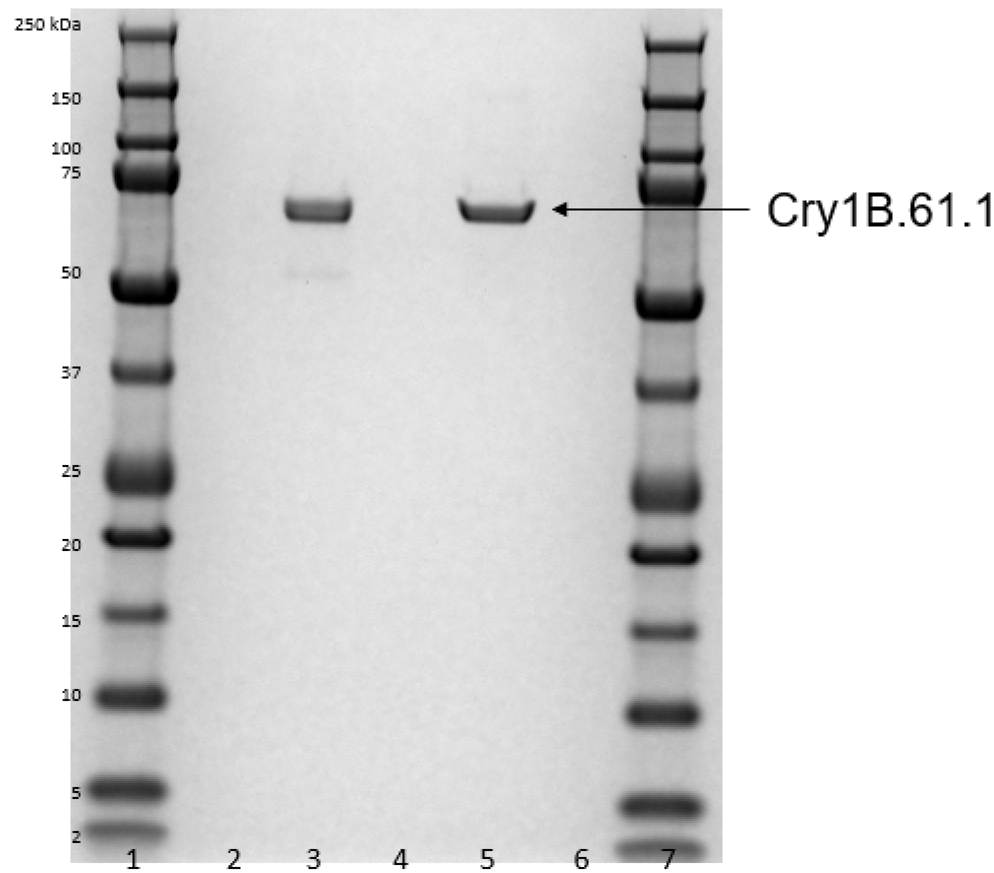
In order to have sufficient amounts of the purified Cry1B.61.1 protein for the multiple studies required to assess its safety, the Cry1B.61.1 protein was expressed in an *Escherichia coli* protein expression system. The microbially derived protein was purified using immobilized metal affinity chromatography.

The biochemical characteristics of the COR23134 soybean-derived and microbially derived Cry1B.61.1 proteins were characterized using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot, glycosylation analysis, peptide mapping by liquid chromatography mass spectrometry (LC-MS), and N-terminal amino acid sequencing. For the microbially derived Cry1B.61.1 protein, the bioactivity was verified by a sensitive insect bioassay. The results demonstrated that the COR23134 soybean-derived and microbially derived Cry1B.61.1 proteins have the expected molecular weight, immunoreactivity, and amino acid sequence, and are not glycosylated. The microbially derived Cry1B.61.1 protein was demonstrated to be an appropriate test substance for use in safety studies.

SDS-PAGE Analysis

Samples of the COR23134 soybean-derived Cry1B.61.1 protein and the microbially derived Cry1B.61.1 protein were analyzed by SDS-PAGE. As expected, the Cry1B.61.1 proteins, derived from both COR23134 soybean and the microbial system, migrated as a predominant band consistent with the expected molecular weight (Figure 31).

Additional details regarding SDS-PAGE analytical methods are provided in Appendix F.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	COR23134 Soybean-Derived Cry1B.61.1 Protein
4	1X LDS/DTT Sample Buffer Blank
5	Microbially Derived Cry1B.61.1 Protein (1 µg)
6	1X LDS/DTT Sample Buffer Blank
7	Pre-stained Protein Molecular Weight Marker[a]

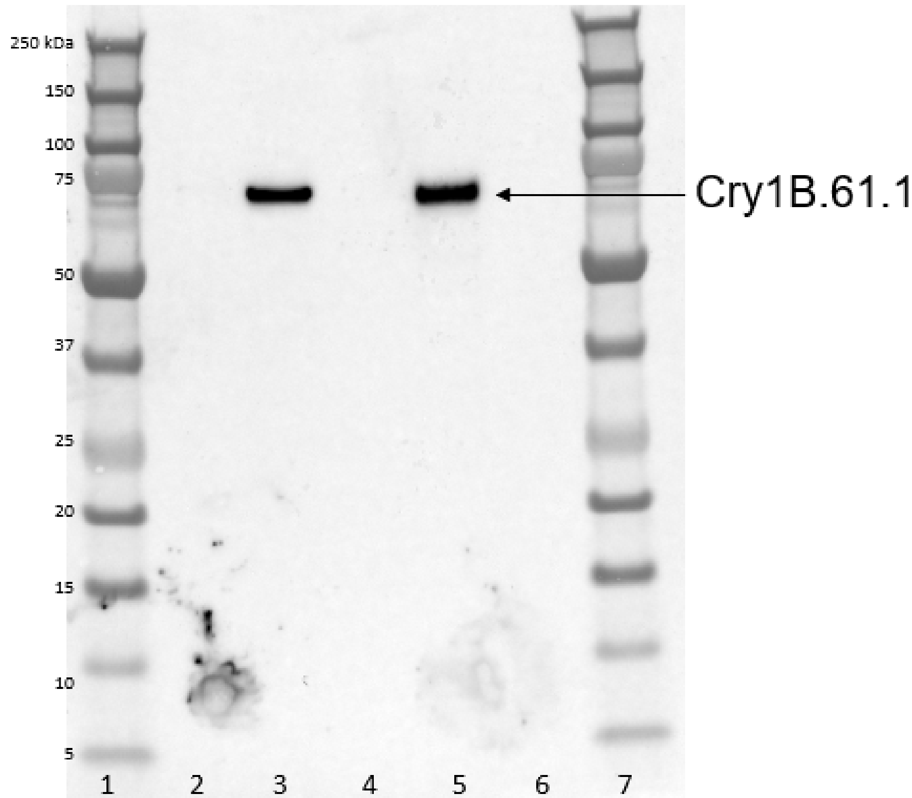
Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 31. SDS-PAGE Analysis of the Cry1B.61.1 Protein
Coomassie blue staining of the SDS-PAGE gel demonstrated the protein migrated as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived Cry1B.61.1 protein (Lane 3) and the microbially-derived Cry1B.61.1 protein (Lane 5).

Western Blot Analysis

Samples of the COR23134 soybean-derived Cry1B.61.1 protein and the microbially derived Cry1B.61.1 protein were analyzed by Western blot. As expected, the Cry1B.61.1 proteins derived from both COR23134 soybean and the microbial system are immunoreactive and have the expected molecular weight (Figure 32).

Additional details regarding Western blot analytical methods are provided in Appendix F.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Microbially Derived Cry1B.61.1 Protein (10 ng)
4	1X LDS/DTT Sample Buffer Blank
5	COR23134 Soybean-Derived Cry1B.61.1 Protein
6	1X LDS/DTT Sample Buffer Blank
7	Pre-stained Protein Molecular Weight Marker[a]

Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 32. Western Blot Analysis of the Cry1B.61.1 Protein

Western blot analysis demonstrated that the Cry1B.61.1 protein was immunoreactive to polyclonal antibody and visible as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived Cry1B.61.1 protein (Lane 5) and the microbially-derived Cry1B.61.1 protein (Lane 3).

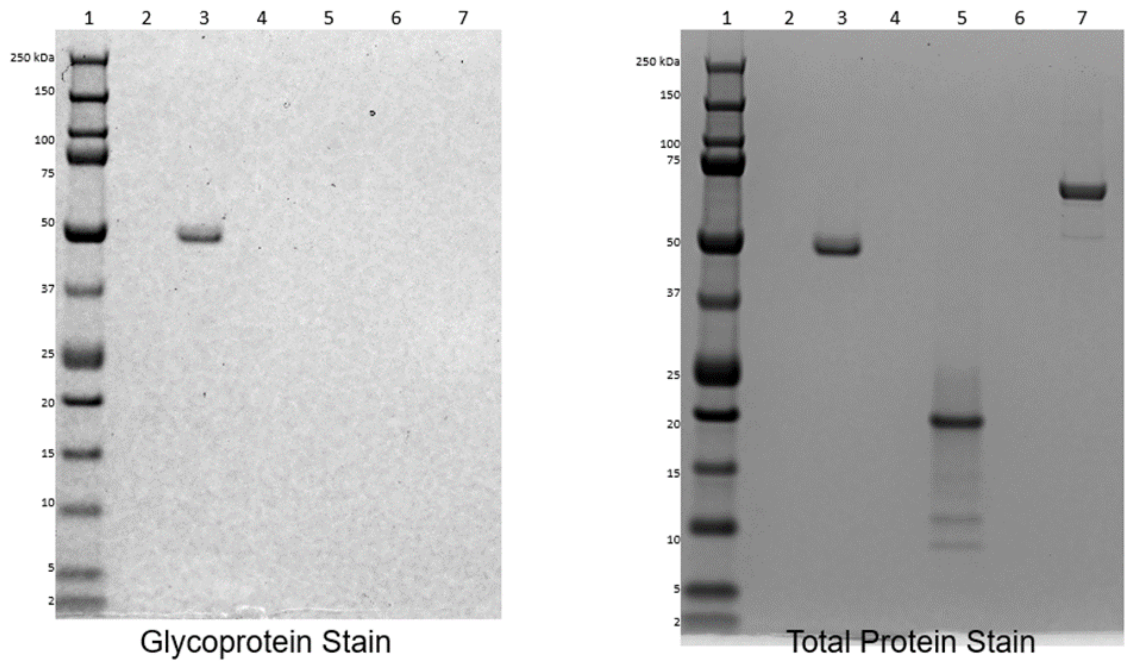
Protein Glycosylation Analysis

Samples of the COR23134 soybean-derived Cry1B.61.1 protein and the microbially derived Cry1B.61.1 protein were analyzed by SDS-PAGE followed by the glycoprotein staining for glycosylation analysis. Each gel also included a positive control (horseradish peroxidase) and a negative control (soybean trypsin inhibitor). The gel was first stained using a Pierce Glycoprotein Staining Kit to visualize any glycoproteins, imaged, and then stained with the Coomassie blue reagent to visualize all protein bands.

Glycosylation was determined to be negative for both the COR23134 soybean-derived and microbially derived Cry1B.61.1 proteins (Figure 33 and Figure 34, respectively). The horseradish peroxidase positive control was clearly visible as a stained band. The soybean trypsin inhibitor negative control was not stained by the glycoprotein stain.

Additional details regarding glycosylation analytical methods are provided in Appendix F.

A) Glycoprotein Staining **B) Total Protein Staining**

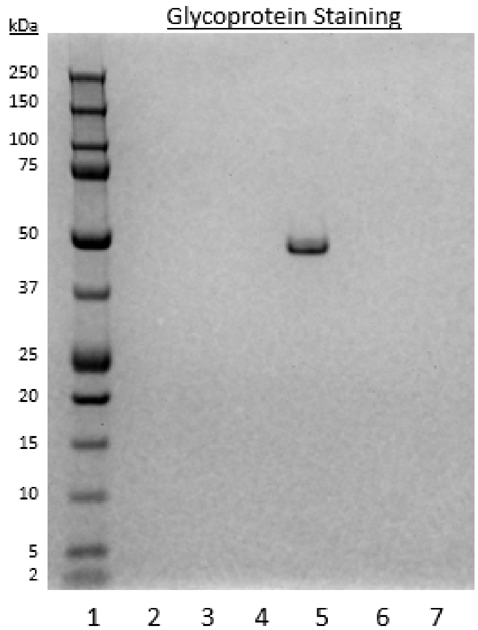


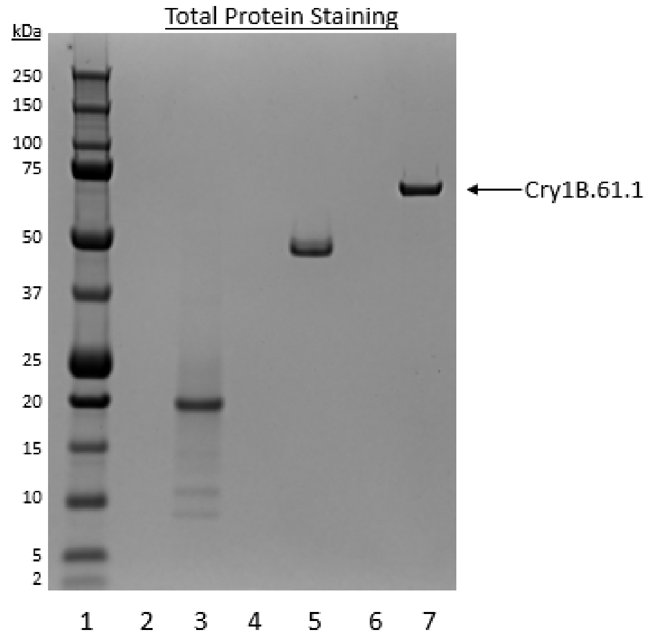
Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Horseradish Peroxidase Positive Control (1.0 µg)
4	1X LDS/DTT Sample Buffer Blank
5	Soybean Trypsin Inhibitor Negative Control (1.0 µg)
6	1X LDS/DTT Sample Buffer Blank
7	COR23134 Soybean-Derived Cry1B.61.1 Protein

Note: The glycoprotein gel was stained with glycoprotein staining reagent. The total protein stain gel was stained with glycoprotein staining reagent followed by staining with Coomassie blue reagent for total proteins. Kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 33. Glycosylation Analysis of the COR23134 Soybean-Derived Cry1B.61.1 Protein
A) Glycoprotein staining: Glycosylation was not detected for the COR23134 soybean-derived Cry1B.61.1 protein (Lane 7). The horseradish peroxidase positive control was stained (Lane 3), and the soybean trypsin inhibitor negative control was not stained (Lane 5). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the COR23134 soybean-derived Cry1B.61.1 protein (Lane 7) and both the positive (Lane 3) and negative (Lane 5) control proteins.

A) Glycoprotein Staining **B) Total Protein Staining**





Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Soybean Trypsin Inhibitor Negative Control (1.0 µg)
4	1X LDS/DTT Sample Buffer Blank
5	Horseradish Peroxidase Positive Control (1.0 µg)
6	1X LDS/DTT Sample Buffer Blank
7	Microbially Derived Cry1B.61.1 Protein

Note: The glycoprotein gel was stained with glycoprotein staining reagent. The total protein stain gel was stained with glycoprotein staining reagent followed by staining with Coomassie blue reagent for total proteins. Kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 34. Glycosylation Analysis of the Microbially Derived Cry1B.61.1 Protein
A) Glycoprotein staining: Glycosylation was not detected for the microbially derived Cry1B.61.1 protein (Lane 7). The horseradish peroxidase positive control was stained (Lane 5), and the soybean trypsin inhibitor negative control was not stained (Lane 3). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the microbially derived Cry1B.61.1 protein (Lane 7) and both the positive (Lane 5) and negative (Lane 3) control proteins.

Mass Spectrometry Peptide Mapping Analysis

Samples of the COR23134 soybean-derived Cry1B.61.1 protein and the microbially derived Cry1B.61.1 protein were analyzed by SDS-PAGE. The gel was stained with Coomassie blue reagent, and the bands containing the Cry1B.61.1 protein were excised for each sample. The excised Cry1B.61.1 protein bands were digested with trypsin or chymotrypsin. Digested samples were analyzed using liquid chromatography-mass spectrometry (LC-MS). The resulting MS data were used to search and match the peptides from the Cry1B.61.1 protein sequence, and the combined sequence coverage was calculated.

The combined sequence coverage of the identified tryptic and chymotryptic peptides for the COR23134 soybean-derived Cry1B.61.1 protein accounts for 90.2% (591/655) of the amino acid sequence (Figure 35). The combined sequence coverage of the identified tryptic and chymotryptic peptides for the microbially derived Cry1B.61.1 protein accounts for 90.2% (597/662) of the amino acid sequence (Figure 36).

Additional details regarding peptide mapping analytical methods are provided in Appendix F.

1 **PSNRKNENEI INALSIPAVS NHSAQMDLSL DARIEDSLCI AEGNNINPLV**
51 **SASTVQTGIN IAGRIL**GVLG VPFAGQLASF Y**SFIVGELWP SGRDPWEIFM**
101 **EHVEQLVRQH ITMNARNTAL ARLQGLGASF RAYQQSLEDW LENRDNAR**TR
151 SVLYTQYIAL ELDFLNAMPL **FAINNQQVPL** LMVY**AQAANL HLLLLRDASL**
201 **FGSEFGLTSQ EIQRYYERQA EKTREYSDYC ARWYNTGLNN LRGTNAESWL**
251 **RY**NQFR**RDLT LGVLDLVALF PSYDTRIYPI NTSAQLTREI YTDPIGRTNA**
301 **PSGFASTNWF NNNAPSFSAI EAAIFRPPHL LDFPEQL**TIY SASSR**WSSTQ**
351 **HMNYWVGHRL NFRPIGGTLN TSTHGATNTS INPVTLQFTS RDVYRTESYA**
401 **GINILLTTPV NGVPWARFNW RNPLNSLRGS LLYTIGYTGV GTQLFDSETE**
451 **LPPETTERPN YESYSHRLSN IRLIIGGTLR APVYSWTHRS ADRTNTIATN**
501 **IITQIPAVKG NFLFNGLVIS GPGFTGGDLV RLNNSGNNIQ NRGYIEVPIQ**
551 **FR**STSTRYRV RVR**YASVTPI RLSVNWGNSN IFSSIVPATA TSLDNLQSRN**
601 **FGYFESRNAF TSATGNVVGV RNFSENAGVI IDRFEFIPVT ATFEAEYDLE**
651 **RAQEA**

Red type	Bold red type indicates soybean-derived Cry1B.61.1 peptides identified using LC-MS analysis against the expected Cry1B.61.1 protein sequence.
Amino acid residue abbreviations	alanine (A), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

Note: The Cry1B.61.1 protein sequence does not include the N-terminal methionine as it is anticipated to be absent.

Figure 35. Identified Tryptic and Chymotryptic Peptide Amino Acid Sequence of the COR23134 Soybean-Derived Cry1B.61.1 Protein Using LC-MS Analysis

1 **PSNRKNENEI INALSIPAVS NHSAQMDLSL DARIEDSLCI AEGNNINPLV**
51 **SASTVQTGIN IAGRIL**GVLG VPFAGQLASF Y**SFIVGELWP SGRDPWEIFM**
101 **EHVEQLVRQH ITMNAR**NTAL **ARLQGLGASF RAYQQSLEDW LENRDNAR**TR
151 SVLYTQYIAL ELDFLNAMPL **FAINNQQVPL LMVYAQAANL HLL**L**LRDASL**
201 **FGSEFGLTSQ EIQRYYERQA EKTREYSDYC ARWYNTGLNN LRGTNAESWL**
251 **RY**NQFR**RDLT LGVLDLVALF PSYDTRIYPI NTSAQLTREI YTDPIGRTNA**
301 **PSGFASTNWF NNNAPSFSAI EAAIFRPPHL LDFPEQL**TIY SASSR**WSSTQ**
351 **HMNYWVGHRL NFRPIGGTLN TSTHGATNTS INPVTLQFTS RDVYRTESYA**
401 **GINILLTTPV NGVPWAR**F**NW RNPLNSLRGS LLYTIGYTGV GTQLFDSETE**
451 **LPPETTERPN YESYSHR**L**SN IRLIIGGTLR APVYSWTHRS ADRTNTIATN**
501 **IITQIPAVKG NFLFNGLVIS GPGFTGGDLV RLNNSGNNIQ NRGYIEVPIQ**
551 **FR**STSTRYRV R**VRYASVTPI RLSVNWGNSN IFSSIVPATA TSLDNLQSRN**
601 **FGYFESRNAF TSATGNVVGV RNFSENAGVI IDRFEFIPVT ATFEAEYDLE**
651 **RAQEAGHHHH HH**

Red bold type	Red bold type indicates microbially derived Cry1B.61.1 peptides identified using LC-MS analysis against the expected microbially derived Cry1B.61.1 protein sequence.
Amino acid residue abbreviations	Alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

Note: The Cry1B.61.1 sequence does not include the N-terminal methionine as it is anticipated to be absent.

Figure 36. Identified Tryptic and Chymotryptic Peptide Amino Acid Sequence of Microbially Derived Cry1B.61.1 Protein Using LC-MS Analysis

N-Terminal Amino Acid Sequence Analysis

The N-terminal peptide for the COR23134 soybean-derived Cry1B.61.1 protein was identified by LC-MS as PSNRKNENEIINAL from the chymotryptic digestion, matching the amino acid residues 1-14 of the Cry1B.61.1 protein sequence without the N-terminal methionine. The results confirmed the N-terminal methionine was absent as expected (Dummitt *et al.*, 2003; Sherman *et al.*, 1985).

The Edman sequencing analysis of the microbially derived Cry1B.61.1 protein sample identified an N-terminal sequence (PSNRKNENEI), matching the amino acid residues 1-10 of the Cry1B.61.1 protein sequence without the N-terminal methionine as expected (Dummitt *et al.*, 2003; Sherman *et al.*, 1985).

Additional details regarding N-terminal amino acid sequencing analytical methods are provided in Appendix F.

Bioactivity Assay

The biological activity of the microbially derived Cry1B.61.1 protein was evaluated by conducting a 7-day bioassay using *Chrysodeixis includens* (soybean looper; Lepidoptera: Noctuidae), a species sensitive to the Cry1B.61.1 protein.

Bioactivity analysis demonstrated that the microbially derived Cry1B.61.1 protein had insecticidal activity toward a target insect, *C. includens* (Table 12). The biological activity of the test diets containing 50 and 500 ng of the Cry1B.61.1 protein per mg wet diet was demonstrated by 100% mortality in *C. includens* compared to *C. includens* fed the buffer control diet. The biological activity of the test diet containing 5 ng of the Cry1B.61.1 protein per mg wet diet was demonstrated by increased mortality and decreased weight in *C. includens* compared to *C. includens* fed the buffer control diet.

Additional details regarding bioactivity assay methods are provided in Appendix F.

Table 12. Summary of the Microbially Derived Cry1B.61.1 Protein Bioactivity Assay Using *Chrysodeixis includens*

Treatment	Treatment Description	Concentration (ng Cry1B.61.1 Protein/mg Wet Diet)	Total Number of Observations	Mortality (%)	Number of Surviving Organisms	Weight of Surviving Organisms (mg)	
						Mean ± Standard Deviation	Range
1	Buffer Control Diet	0	17[a]	17.6	14	2.9 ± 2.4	0.1-9.3
2	Test Diet	5	20	40.0	12	1.7 ± 0.8	0.5-3.0
3	Test Diet	50	20	100	0	NA	NA
4	Test Diet	500	20	100	0	NA	NA

Note: The summary of *Chrysodeixis includens* mortality data consisted of the calculation of dead larvae divided by the total number of observed larvae at the end of the study and multiplied by 100. Not applicable (NA); there were no surviving organisms in this treatment.

[a] Organisms counted as missing during the bioassay were not included in the total number of observations for a given treatment.

Allergenicity and Toxicity Analyses of the Cry1B.61.1 Protein

A weight-of-evidence approach was applied to determine the allergenic and toxic potential of the Cry1B.61.1 protein expressed in COR23134 soybean, including an assessment of the following: the history of safe use of the source organism, a bioinformatic comparison of the amino acid sequence of the Cry1B.61.1 protein to known or putative allergen and protein toxin sequences, an evaluation of the heat lability of the Cry1B.61.1 protein using a sensitive insect bioassay, evaluations of the stability of the Cry1B.61.1 protein using *in vitro* gastric and intestinal digestion models, determination of the Cry1B.61.1 protein glycosylation status, and an evaluation of acute toxicity in mice following oral exposure to the Cry1B.61.1 protein.

Source Organism of the Cry1B.61.1 Protein

The Cry1B.61.1 protein expressed in COR23134 soybean is encoded by the *cry1B.61.1* gene, a modified *cry1B* class gene, derived from *Bacillus thuringiensis* (*Bt*).

Bt is a diverse group of Gram-positive, spore-forming bacteria that has a history of safe use as a pesticide over several decades (US-EPA, 1998; US-EPA, 2001). It occurs ubiquitously in the soil and on plants including vegetables, cotton, tobacco, tree crops, and forest crops (Schnepf et al., 1998; Shelton, 2012). Several Cry proteins have been deployed as safe and effective pest control agents in microbial *Bt* formulations for almost 40 years. Several Cry proteins have also been effectively deployed as safe and effective pest control agents and have a history of safe use in genetically modified crops (ISAAA, 2023).

Bioinformatic Analysis of Cry1B.61.1 Homology to Known and Putative Allergens

Assessing newly expressed proteins for potential cross-reactivity with known and putative allergens is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). A bioinformatic assessment of the Cry1B.61.1 protein sequence (656 amino acids [aa]) for potential cross-reactivity with allergens was conducted by following established international criteria (Codex Alimentarius Commission, 2009; FAO/WHO, 2001).

Two separate searches for the Cry1B.61.1 protein sequence were performed using the Comprehensive Protein Allergen Resource (COMPARE) 2023 database (January 26, 2023 (van Ree *et al.*, 2021)). This peer-reviewed database is a collaborative effort of the Health and Environmental Sciences Institute (HESI) Protein Allergens, Toxins, and Bioinformatics (PATB) Committee and contains 2,631 sequences.

The first search was the sliding 80-mer window search, accomplished with an internally developed Perl script running FASTA v35.04 (Pearson and Lipman, 1988) with an *E*-score cutoff set to 100. In a sliding window search, each sequentially overlapping 80 aa subsequence of the overall Cry1B.61.1 protein sequence is used as a query against the COMPARE allergen database sequences. The script examined all alignments generated from the query and reported any possessing > 35% identity over an alignment length of ≥ 80 aa. Additionally, the script rescaled the percent identity to an 80-mer window for any alignments possessing an alignment length shorter than 80 aa; the number of identities in these alignments would be divided by 80, then multiplied by 100, and would report any alignment possessing an adjusted percent identity > 35%.

The second search used EMBOSS fuzzpro v6.6.0 (Rice *et al.*, 2000) to identify any eight or greater contiguous identical amino acid matches between the Cry1B.61.1 protein sequence and the COMPARE allergen sequences.

Results of the search of the Cry1B.61.1 protein sequence against the COMPARE allergen database sequences found no alignments that were a length of 80 aa or greater with a sequence identity of > 35% and no alignments shorter than 80 aa with a sequence identity > 35% when normalized to an 80-mer window. No contiguous 8-residue exact matches between the Cry1B.61.1 protein sequence and the allergen sequences were identified in the second search. Collectively, these data indicate that no allergenicity concern arose from the bioinformatics assessment of the Cry1B.61.1 protein.

Bioinformatics evaluation of the Cry1B.61.1 protein sequence did not generate biologically relevant amino acid sequence similarities to allergens that are harmful to humans or animals.

Bioinformatic Analysis of Cry1B.61.1 Homology to Known and Putative Protein Toxins

Assessing newly expressed proteins for potential sequence similarity with protein toxin is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). The potential toxicity of the Cry1B.61.1 protein was assessed by comparison of its sequence 1) an internal toxin database, and 2) the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein database.

The internal toxin database is a subset of sequences found in UniProtKB/Swiss-Prot (The UniProt Consortium, 2023). UniProtKB/Swiss-Prot is a curated database of non-redundant proteins containing functional information for over 550,000 sequences. To produce the internal toxin database, the proteins in UniProtKB/Swiss Prot are filtered for molecular function by keywords that could imply toxicity or adverse health effects (e.g., toxin, hemagglutinin, vasoactive). The internal toxin database is updated annually and contains 8,858 sequences.

The search between the Cry1B.61.1 protein sequence and protein sequences in the internal toxin database was conducted using BLASTP v2.10.0+ with an E-value set to 10^{-4}. No alignments were returned between the Cry1B.61.1 protein sequence and any protein sequence in the internal toxin database. Therefore, no toxicity concern arose from the bioinformatics assessment of the Cry1B.61.1 protein.

The BLASTP search of the Cry1B.61.1 protein against the NCBI nr protein database returned the maximum number of alignment descriptions/sequence alignments, 500/250, with an E-value from 0 to 2×10^{-101} to various insecticidal crystal proteins from different Bacillus species. This is expected since Cry1B.61.1 is an insecticidal protein. None of the accessions returned by the BLASTP search are proteins known to be toxic to humans or animals.

Bioinformatics evaluation of the Cry1B.61.1 protein sequence did not generate biologically relevant amino acid sequence similarities to protein toxins that are harmful to humans or animals.

Thermolability Analysis

Thermal stability of the Cry1B.61.1 protein was characterized by determining the biological activity of the heat-treated Cry1B.61.1 protein when incorporated in an artificial diet and fed to

Chrysodeixis includens (*C. includens*), an insect sensitive to the Cry1B.61.1 protein. The Cry1B.61.1 protein was incubated at various temperatures (25 °C, 50 °C, 75 °C, and 95 °C) for approximately 30 minutes before incorporation into the artificial diet. Each test diet contained a targeted concentration of 100 ng Cry1B.61.1 protein per mg diet wet weight. Larvae were exposed via oral ingestion to the diets in a 7-day bioassay. A positive control diet containing the unheated Cry1B.61.1 protein and a bioassay control diet containing buffer were included in the bioassay to verify assay performance. After seven days, statistical analyses were conducted to evaluate *C. includens* mortality of the heat-treated test groups relative to the unheated test group.

The results demonstrated that the Cry1B.61.1 protein heated for approximately 30 minutes at 75 °C and 95 °C (Treatments 5 and 6, respectively) had significantly reduced activity against *C. includens* (P values < 0.0001) compared to the unheated Cry1B.61.1 control (Treatment 2) when incorporated in an artificial diet. No statistically significant decreases in protein activity were observed for the Cry1B.61.1 protein heated for approximately 30 minutes at 25 °C and 50 °C (Treatments 3 and 4, respectively) when compared to the unheated Cry1B.61.1 control (Table 13).

Additional details regarding thermolability analytical methods are provided in Appendix F.

Table 13. Biological Activity of Heat-Treated Cry1B.61.1 Protein in Artificial Diet Fed to *Chrysodeixis includens*

Treatment	Treatment Description	Test Dosing Solution Incubation Condition	Total Number of Observations	Mortality (%)	Fisher's Test P-Value	Number of Surviving Organisms	Weight of Surviving Organisms (mg)	
							Mean ± Standard Deviation	Range
1	Buffer Control Diet	NA	20	0	--	20	15.6 ± 5.98	6.9 - 30.7
2	Control Diet	Unheated	20	100	--	0	--	--
3	Test Diet	25 °C	20	100	1.0000	0	--	--
4	Test Diet	50 °C	20	100	1.0000	0	--	--
5	Test Diet	75 °C	20	20.0	<0.0001[a]	16	14.0 ± 7.33	4.6 - 28.1
6	Test Diet	95 °C	20	15.0	<0.0001[a]	17	14.5 ± 5.01	8.1 - 28.1

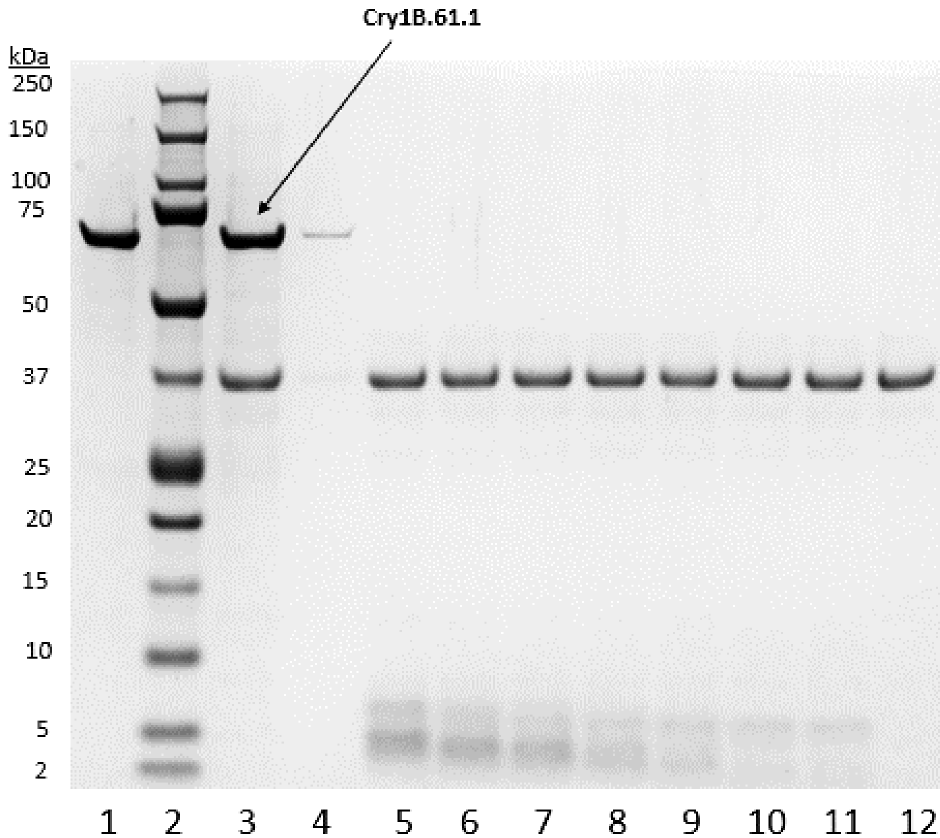
Note: Not applicable (NA); the buffer control diet did not contain the Cry1B.61.1 protein and the dosing solution was not incubated. The unheated control diet (Treatment 2) and the test diets (Treatments 3-6) contained a targeted concentration of 100 ng Cry1B.61.1 protein per mg diet wet weight.
[a] A statistically significant difference (P-value < 0.05) was observed in comparison to Treatment 2.

Digestibility Analysis with Simulated Gastric Fluid (SGF)

Simulated gastric fluid (SGF) containing pepsin at pH ~1.2 was used to assess the susceptibility of the Cry1B.61.1 protein to proteolytic digestion by pepsin *in vitro*. The Cry1B.61.1 protein was incubated in SGF for 0, 0.25, 1, 2, 5, 10, 30, and 60 minutes. A positive control (bovine serum albumin) and a negative control (β-lactoglobulin) were included in the assay and were incubated in SGF for 0, 1, and 60 minutes. After incubation in SGF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SGF digestibility results showed that the Cry1B.61.1 protein migrating at approximately 74 kDa was digested within 0.25 minutes in SGF as demonstrated by both SDS-PAGE and western blot analysis (Figure 37 and Figure 38, respectively). On the SDS-PAGE gel, low molecular weight bands (~2-10 kDa) remained detectable in the Cry1B.61.1 protein samples for 60 minutes in SGF.

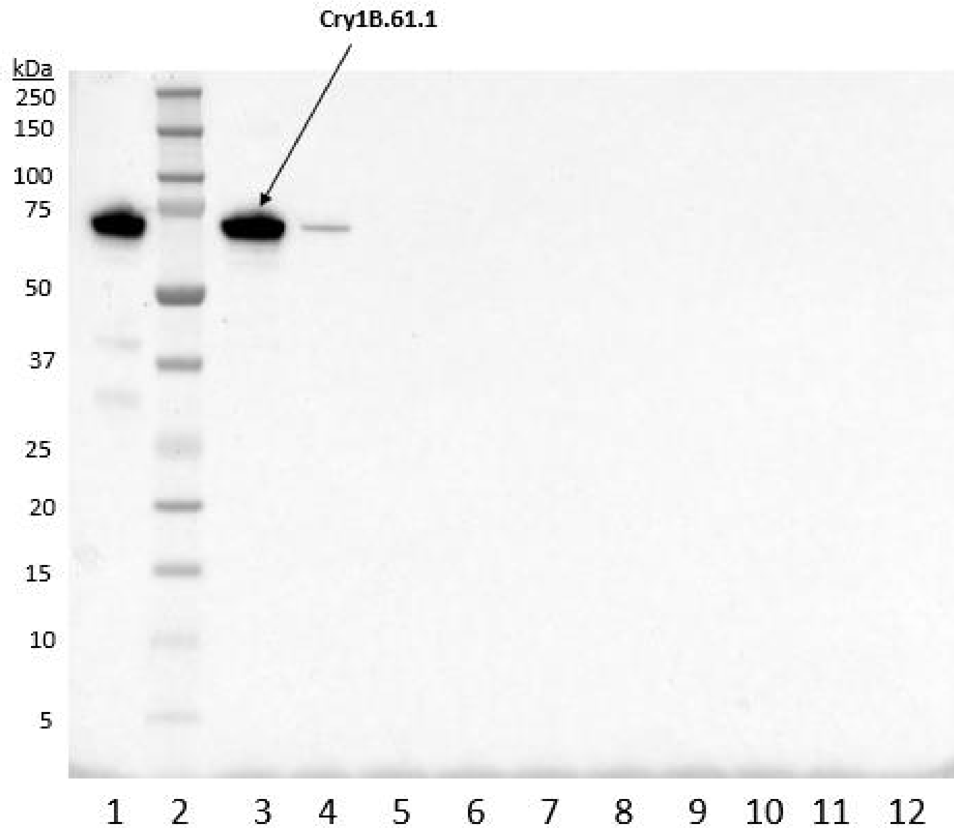
Additional details regarding SGF analytical methods are provided in Appendix F.



Lane	Sample Descriptions
1	Cry1B.61.1 protein in buffer (no SGF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.61.1 protein in SGF, Time 0
4	Cry1B.61.1 protein in SGF, Time 0; 1:20 dilution
5	Cry1B.61.1 protein in SGF, 0.25 minutes
6	Cry1B.61.1 protein in SGF, 1 minute
7	Cry1B.61.1 protein in SGF, 2 minutes
8	Cry1B.61.1 protein in SGF, 5 minutes
9	Cry1B.61.1 protein in SGF, 10 minutes
10	Cry1B.61.1 protein in SGF, 30 minutes
11	Cry1B.61.1 protein in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: kilodalton (kDa), simulated gastric fluid (SGF), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

Figure 37. SDS-PAGE Analysis of the Cry1B.61.1 Protein in Simulated Gastric Fluid Digestion Time Course



Lane	Sample Descriptions
1	Cry1B.61.1 protein in buffer (no SGF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.61.1 protein in SGF, Time 0
4	Cry1B.61.1 protein in SGF, Time 0; 1:50 dilution
5	Cry1B.61.1 protein in SGF, 0.25 minutes
6	Cry1B.61.1 protein in SGF, 1 minute
7	Cry1B.61.1 protein in SGF, 2 minutes
8	Cry1B.61.1 protein in SGF, 5 minutes
9	Cry1B.61.1 protein in SGF, 10 minutes
10	Cry1B.61.1 protein in SGF, 30 minutes
11	Cry1B.61.1 protein in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: kilodalton (kDa) and simulated gastric fluid (SGF).
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

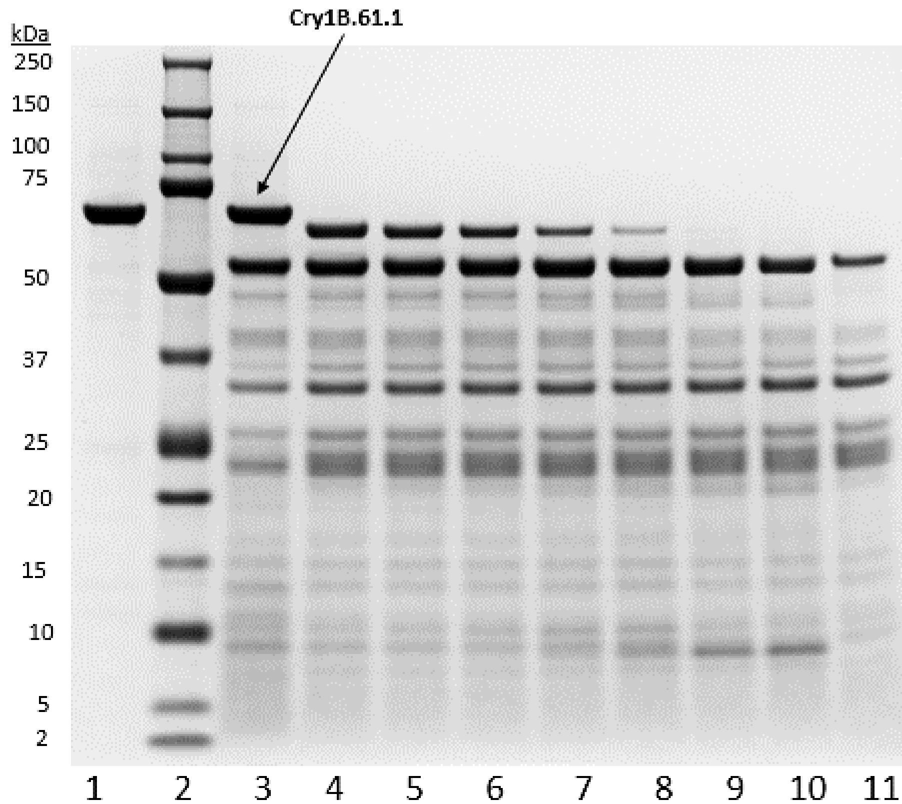
Figure 38. Western Blot Analysis of the Cry1B.61.1 Protein in Simulated Gastric Fluid Digestion Time Course

Digestibility Analysis with Simulated Intestinal Fluid (SIF)

Simulated intestinal fluid (SIF) containing pancreatin at ~pH 7.5 was used to assess the susceptibility of the Cry1B.61.1 protein to proteolytic digestion by pancreatin in vitro. The Cry1B.61.1 protein was incubated in SIF for 0, 0.25, 1, 2, 5, 10, 30, and 60 minutes. A positive control (β-lactoglobulin) and a negative control (bovine serum albumin) were included in the assay and were incubated in SIF for 0, 1, and 60 minutes. After incubation in SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SIF digestibility results showed that the Cry1B.61.1 protein migrating at approximately 74 kDa was digested into smaller fragments within 0.25 minutes in SIF as demonstrated by both SDS-PAGE and western blot analysis (Figure 39 and Figure 40, respectively). These smaller fragments remained detectable via western blot for 60 minutes.

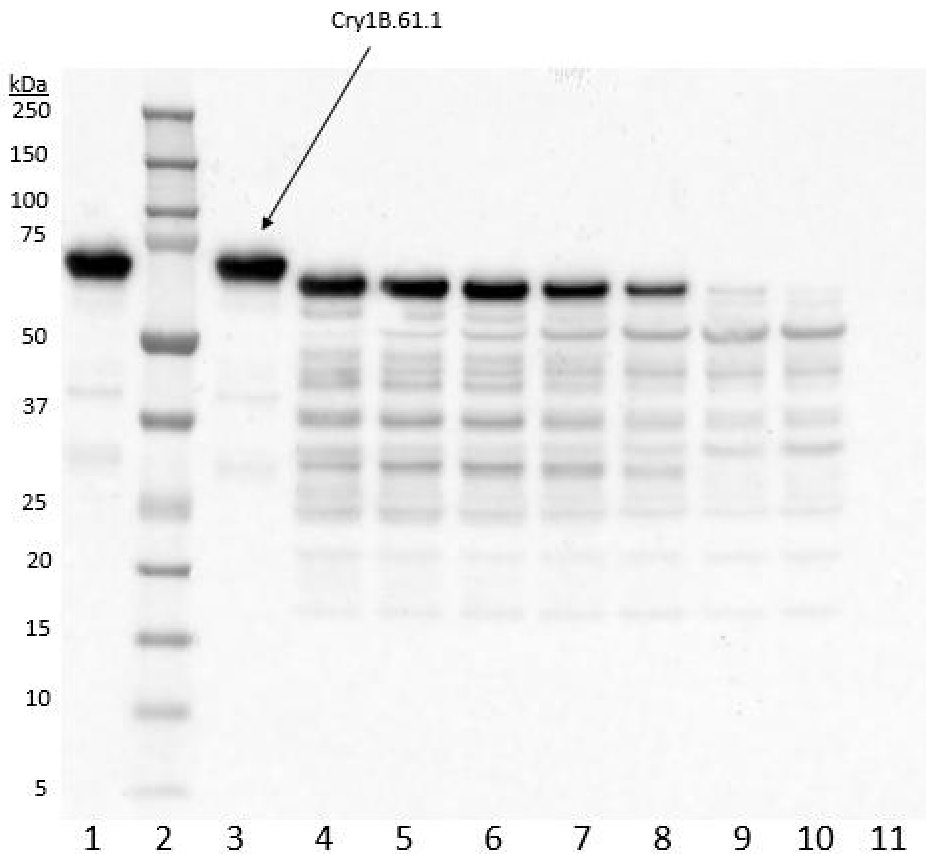
Additional details regarding SIF analytical methods are provided in Appendix F.



Lane	Sample Descriptions
1	Cry1B.61.1 protein in water (no SIF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.61.1 protein in SIF, Time 0
4	Cry1B.61.1 protein in SIF, 0.25 minutes
5	Cry1B.61.1 protein in SIF, 1 minute
6	Cry1B.61.1 protein in SIF, 2 minutes
7	Cry1B.61.1 protein in SIF, 5 minutes
8	Cry1B.61.1 protein in SIF, 10 minutes
9	Cry1B.61.1 protein in SIF, 30 minutes
10	Cry1B.61.1 protein in SIF, 60 minutes
11	SIF Control, 60 minutes

Note: kilodalton (kDa) and simulated intestinal fluid (SIF)
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

Figure 39. SDS-PAGE Analysis of the Cry1B.61.1 Protein in Simulated Intestinal Fluid Digestion Time Course



Lane	Sample Descriptions
1	Cry1B.61.1 protein in water (no SIF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	Cry1B.61.1 protein in SIF, Time 0
4	Cry1B.61.1 protein in SIF, 0.25 minutes
5	Cry1B.61.1 protein in SIF, 1 minute
6	Cry1B.61.1 protein in SIF, 2 minutes
7	Cry1B.61.1 protein in SIF, 5 minutes
8	Cry1B.61.1 protein in SIF, 10 minutes
9	Cry1B.61.1 protein in SIF, 30 minutes
10	Cry1B.61.1 protein in SIF, 60 minutes
11	SIF Control, 60 minutes

Note: kilodalton (kDa) and simulated intestinal fluid (SIF)
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

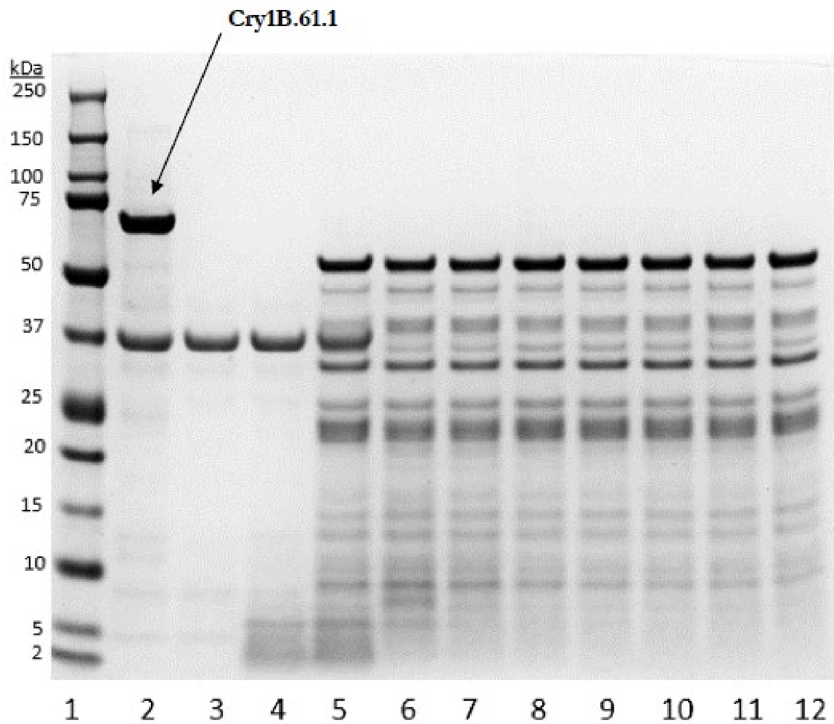
Figure 40. Western Blot Analysis of the Cry1B.61.1 Protein in Simulated Intestinal Fluid Digestion Time Course

Sequential Digestibility Analysis with Simulated Gastric Fluid (SGF) and Simulated Intestinal Fluid (SIF)

Sequential digestion in simulated intestinal fluid (SIF) following a digestion in SGF was used to assess the susceptibility of the low molecular weight SGF fragments (~2-10 kDa, Figure 37) of the Cry1B.61.1 protein. The Cry1B.61.1 protein was incubated for 5 minutes in SGF containing pepsin at pH ~1.2 and then incubated for 0, 0.25, 1, 2, 5, 10, 20, and 30 minutes in SIF containing pancreatin at pH ~7.5. After incubation in SGF/ SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain was used to detect protein bands.

The sequential pepsin (SGF) and pancreatin (SIF) digestibility results showed that the low molecular weight bands (~2-10 kDa) observed in SGF digestion (Figure 37) were digested within 1 minute during sequential SIF digestion (Figure 41).

Additional details regarding sequential digestibility analytical methods are provided in Appendix F.



Lane	Sample Descriptions
1	Pre-stained protein molecular weight marker[a]
2	Cry1B.61.1 Protein in SGF, Time 0
3	SGF Control, 5 minutes
4	Cry1B.61.1 Protein in SGF, 5 minutes
5	Cry1B.61.1 Protein in SGF 5 minutes, SIF Time 0
6	Cry1B.61.1 Protein in SGF 5 minutes, SIF 0.25 minutes
7	Cry1B.61.1 Protein in SGF 5 minutes, SIF 1 minute
8	Cry1B.61.1 Protein in SGF 5 minutes, SIF 2 minutes
9	Cry1B.61.1 Protein in SGF 5 minutes, SIF 5 minutes
10	Cry1B.61.1 Protein in SGF 5 minutes, SIF 10 minutes
11	Cry1B.61.1 Protein in SGF 5 minutes, SIF 20 minutes
12	Cry1B.61.1 Protein in SGF 5 minutes, SIF 30 minutes

Note: kilodalton (kDa), simulated gastric fluid (SGF), simulated intestinal fluid (SIF), and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

Figure 41. SDS-PAGE Analysis of the Cry1B.61.1 Protein in a Sequential Digestion with Simulated Gastric Fluid and Simulated Intestinal Fluid

Protein Glycosylation Analysis

As stated previously in characterization section, the results from glycoprotein staining analysis confirmed the absence of glycosylation for the COR23134 soybean-derived and the microbially derived Cry1B.61.1 proteins.

Evaluation of the Acute Toxicity of the Cry1B.61.1 Protein

A study was conducted to evaluate the potential acute toxicity of the test substance, Cry1B.61.1, in [redacted] mice following oral exposure at the limit dose (5000 mg/kg body weight, adjusted for Cry1B.61.1 content). The Cry1B.61.1 protein and bovine serum albumin (BSA) protein were each reconstituted in deionized water. Vehicle control, BSA comparative control, and the Cry1B.61.1 test substance formulations were administered orally by gavage in three split doses, separated by approximately four hours; the BSA comparative control was administered at an equivalent target dose to that of the Cry1B.61.1 protein. The mice were fasted prior to and throughout the dosing procedure.

Body weights were evaluated on test day 1 (prior to fasting and shortly prior to administration of the first dose), 2, 3, 5, 8, and 15. Clinical signs were evaluated ten times on test day 1 (distributed before and after each dose) and daily thereafter. On test day 15, all surviving mice were euthanized and given a gross pathological examination.

All animals survived to scheduled euthanasia. There were no toxicologically significant clinical abnormalities and all animals gained body weight over the course of the study (test days 1-15). No gross lesions were observed.

Under the conditions of this study, oral exposure via intragastric administration of the Cry1B.61.1 protein to male and female mice at 5000 mg/kg body weight did not result in mortality or other evidence of acute oral toxicity, based on evaluation of body weight, clinical signs, and gross pathology. Therefore, the acute oral toxicity tolerant dose and the LD_{50} of the Cry1B.61.1 protein were determined to be greater than 5000 mg/kg body weight.

Conclusion on Safety of the Cry1B.61.1 Protein in COR23134 Soybean

In conclusion, protein characterization results via SDS-PAGE, western blot, glycosylation analysis, mass spectrometry peptide mapping analysis, and N-terminal amino acid sequence analysis have demonstrated that the Cry1B.61.1 protein derived from COR23134 soybean and the microbial system has the expected molecular weight, immunoreactivity, and amino acid sequence, and is not glycosylated. Characterization of the microbially derived Cry1B.61.1 protein demonstrated that it is an appropriate test substance for use in safety studies.

The allergenic potential of the Cry1B.61.1 protein was evaluated by assessing the Cry1B.61.1 protein source organism and history of safe use, a bioinformatic comparison of the amino acid sequence of the Cry1B.61.1 protein with known and putative allergen sequences, an evaluation of the heat lability of the Cry1B.61.1 protein using a sensitive insect bioassay, evaluations of the stability of the Cry1B.61.1 protein using *in vitro* gastric and intestinal digestion models, and determination of the Cry1B.61.1 protein glycosylation status. The toxicity potential of the Cry1B.61.1 protein was evaluated by a bioinformatic comparison of the Cry1B.61.1 amino acid sequence to known and putative protein toxins and by an acute toxicity in mice following oral exposure to the Cry1B.61.1 protein.

The bioinformatic comparison of the Cry1B.61.1 protein sequence to known and putative allergen and protein toxin sequences showed that the Cry1B.61.1 protein is unlikely to be allergenic or toxic for humans or animals. The Cry1B.61.1 protein was digested within 0.25 minutes in SGF, and some low molecular weight bands (~2-10 kDa) remained detectable via SDS-PAGE for 60 minutes in SGF. The Cry1B.61.1 protein was digested within 0.25 minutes in SIF, and some smaller fragments remained detectable via western blot for 60 minutes. The low molecular weight bands remaining from SGF digestion were digested within 1 minute during sequential SIF digestion. The Cry1B.61.1 protein was not glycosylated. The Cry1B.61.1 protein heated for approximately 30 minutes at 75 °C and 95 °C had significantly reduced activity against *Chrysodeixis includens* when incorporated in an artificial diet. The acute oral toxicity assessment in mice determined the LD_{50} of the Cry1B.61.1 protein to be greater than 5000 mg/kg. These data support the conclusion that COR23134 soybean, expressing the Cry1B.61.1 protein, is as safe as conventional soybean for the food and feed supply.

Based on this weight of evidence, consumption of the Cry1B.61.1 protein from COR23134 soybean is unlikely to cause an adverse effect on humans or animals.

IPD083Cb Protein

Amino Acid Sequence of the IPD083Cb Protein

The deduced amino acid sequence from the translation of the *ipd083Cb* gene encodes the IPD083Cb protein that is 853 amino acids in length and has a molecular weight of approximately 95 kDa (Figure 30).

```
  1   MADYSTLYRD LNQISMPLDR VEFSEVMVIH RMYLRLSDLN VGELPGAERV
 51   KRLYVLADVV ELATFAHPQL LNTRMPGSVT VIILCRLLQF PTDGSFAAWL
101   ELPFMELHTL IEQYRSEIKA ADDAKWGTYV HAEEVQLSPL FNGWPYLVVE
151   AQRCIITAAM HNTFNRPGWV RSITQFTTDQ SGRVDTTLLA RTEFGHIDLP
201   LETDSPTAFS VSHRQSTNLP VEYTGIPVEV VTDPNILMGM QTSVHIAELV
251   KACYPSPELV SAVGVHVNWL NEVLLRVVQK ESQLQGTEAY NECLALLGRI
301   QCVMKMGPFV SVVPQLQYRM YGSLIRQMAQ VAQNYDQDFR QLKLFIAQNQ
351   ILGSYLLQQN KAFADREVQM ESFHSAVISQ RRQELDDAIA KMDRLSLQME
401   EEDRAMEQAR KEMEEGLKQF QNEQVARAVF AVLKSVAMIA LAFVTAGATA
451   PGAAASAAQA VNIAGQAAQA LRRVVEILEG LEAVMEVVAA IKHLVDALDQ
501   VSQIVDAPPM PDMPSEADWS IFVNEIEAVA EGMPTEVSEV PAWKAKCKNV
551   AALGREMCIT AEQISQLQYD IWVQGLLRDI AQSHADRLAA IQPANLTNYL
601   EMAIQMDMRT TRILIGLLNI MRIQNAALMY EYLLTPTQLT AWPLRMDTVA
651   NLLITHESAA LSGLAQLGPP SDFTSRHVVK GIPVSLLLDG GDWEFEIPVQ
701   GGMSSFPSSW TRVRIRHLEM HFVQEASGGG EIIHQPATQT GTIYILLQGS
751   TVFHDRRREE VMTFQAAVPL NYHYAYRLDT GEATLTNEPS EQFANTFMQM
801   TPFTHWRLRL SASAAENKGL AFPTATAPDS TTEIAITFHV TAIRQIDWRQ
851   EEE*
```

Figure 42. Deduced Amino Acid Sequence of the IPD083Cb Protein
The deduced amino acid sequence from the translation of the *ipd083Cb* gene from plasmid [redacted] The asterisk (*) indicates the translational stop codon. The IPD083Cb protein is 853 amino acids in length and has an molecular weight of approximately 95 kDa.

Function and Activity of the IPD083Cb Protein

The IPD083Cb protein expressed in COR23134 soybean is encoded by the insecticidal protein gene, *ipd083Cb*, from giant maidenhair fern (*Adiantum trapeziforme* var. *braziliense*). The expressed IPD083Cb protein confers control of certain susceptible lepidopteran pests (Liu *et al.*, 2019), by causing disruption of the midgut epithelium. The site of action of the IPD083Cb protein appears to be similar to that of Cry proteins derived from *Bacillus thuringiensis* (*Bt*). The IPD083Cb protein binds to target sites located on the midgut of susceptible insects, leading to insect death. The competitive binding assays using brush border membrane vesicles from insect midguts demonstrated that IPD083Cb does not bind to the same target sites as *Bt*-derived insecticidal proteins including Cry2A.127, Vip3Aa, or a variant of Cry1Ab, indicating that IPD083Cb is unlikely to share cross-resistance with insects that are resistant to these proteins.

Characterization of the IPD083Cb Protein Derived from COR23134 Soybean and the Heterologous Plant System

The IPD083Cb protein expressed in COR23134 soybean was purified from the whole plant tissue using immunoaffinity chromatography.

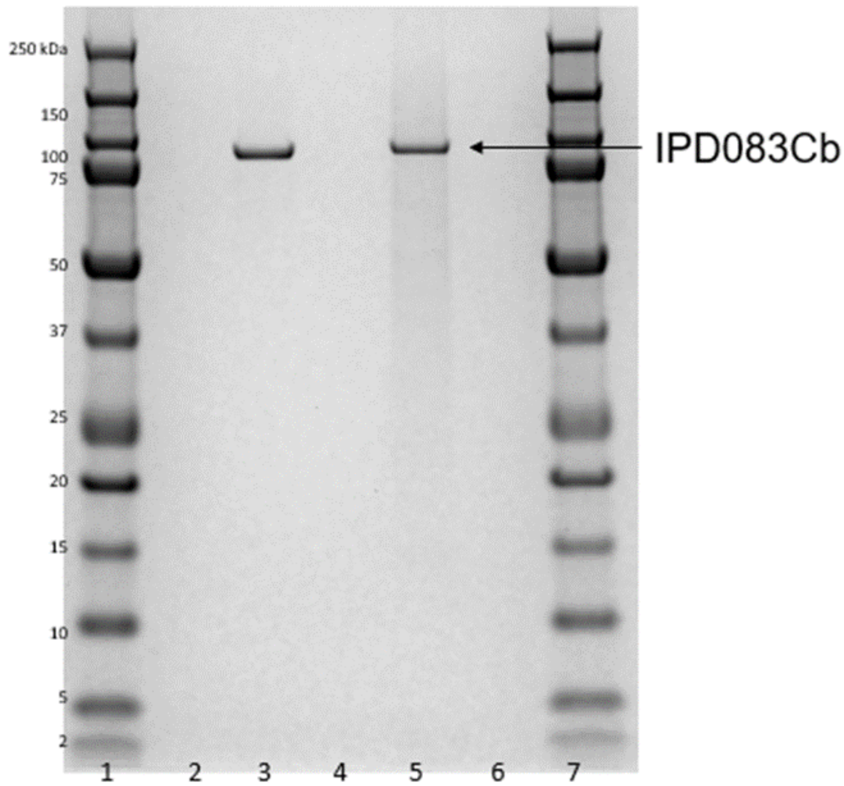
In order to have sufficient amounts of the purified IPD083Cb protein for the multiple studies required to assess its safety, the IPD083Cb protein was expressed in a tobacco-based protein expression system. The tobacco-expressed IPD083Cb protein was purified using immobilized metal affinity chromatography.

The biochemical characteristics of the COR23134 soybean-derived and tobacco-expressed IPD083Cb proteins were characterized using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot, glycosylation analysis, peptide mapping by liquid chromatography mass spectrometry (LC-MS), and N-terminal amino acid sequencing. For the tobacco-expressed IPD083Cb protein, the bioactivity was verified by a sensitive insect bioassay. The results demonstrated that the COR23134 soybean-derived and tobacco-expressed IPD083Cb proteins have the expected molecular weight, immunoreactivity, and amino acid sequence, and are not glycosylated. The tobacco-expressed IPD083Cb protein was demonstrated to be an appropriate test substance for use in safety studies.

SDS-PAGE Analysis

Samples of the COR23134 soybean-derived IPD083Cb protein and the tobacco-expressed IPD083Cb protein were analyzed by SDS-PAGE. As expected, the IPD083Cb proteins derived from both COR23134 soybean and the tobacco expression system migrated as a predominant band consistent with the expected molecular weight (Figure 43).

Additional details regarding SDS-PAGE analytical methods are provided in Appendix G.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	COR23134 Soybean-Derived IPD083Cb Protein
4	1X LDS/DTT Sample Buffer Blank
5	Tobacco-Expressed IPD083Cb Protein (1 µg)
6	1X LDS/DTT Sample Buffer Blank
7	Pre-stained Protein Molecular Weight Marker[a]

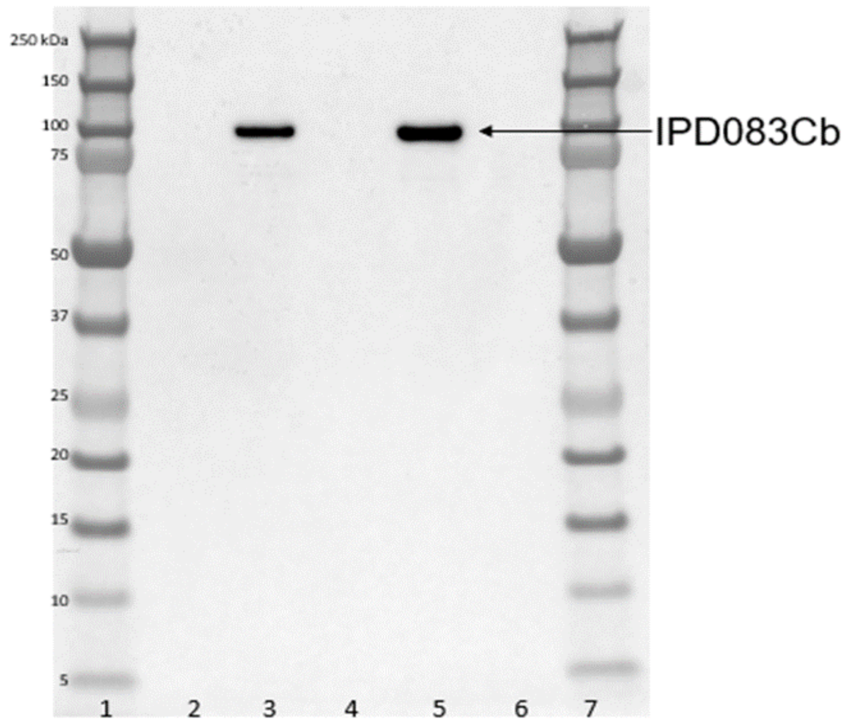
Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 43. SDS-PAGE Analysis of the IPD083Cb Protein
Coomassie blue staining of the SDS-PAGE gel demonstrated the protein migrated as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived IPD083Cb protein (Lane 3) and the tobacco-expressed IPD083Cb protein (Lane 5).

Western Blot Analysis

Samples of the COR23134 soybean-derived IPD083Cb protein and the tobacco-expressed IPD083Cb protein were analyzed by Western blot. As expected, the IPD083Cb proteins derived from both COR23134 soybean and the tobacco expression system are immunoreactive and have the expected molecular weight (Figure 44).

Additional details regarding Western blot analytical methods are provided in Appendix G.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Tobacco-Expressed IPD083Cb Protein (10 ng)
4	1X LDS/DTT Sample Buffer Blank
5	COR23134 Soybean-Derived IPD083Cb Protein
6	1X LDS/DTT Sample Buffer Blank
7	Pre-stained Protein Molecular Weight Marker[a]

Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 44. Western Blot Analysis of the IPD083C Protein
Western blot analysis demonstrated that the IPD083Cb protein was immunoreactive to polyclonal antibody and visible as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived IPD083Cb protein (Lane 5) and the tobacco-expressed IPD083Cb protein (Lane 3).

Protein Glycosylation Analysis

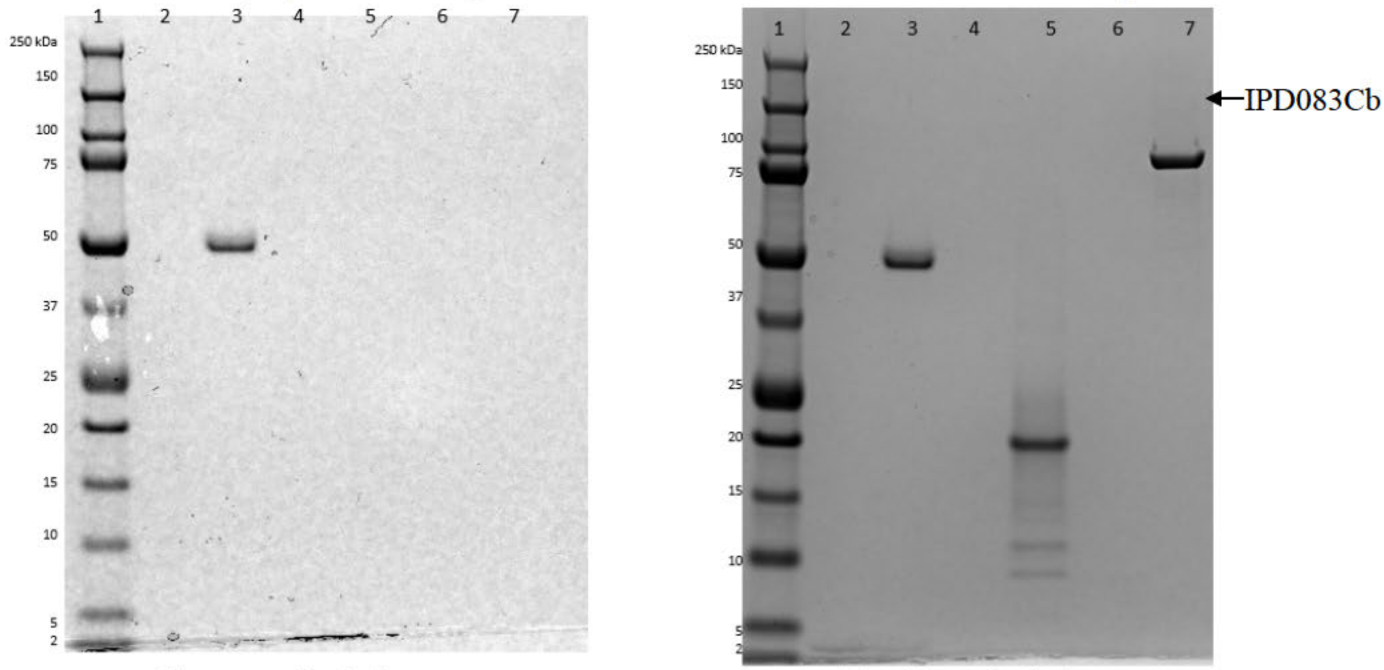
Samples of the COR23134 soybean-derived IPD083Cb protein and the tobacco-expressed IPD083Cb protein were analyzed by SDS-PAGE followed by the glycoprotein staining for glycosylation analysis. Each gel also included a positive control (horseradish peroxidase) and a negative control (soybean trypsin inhibitor). The gel was first stained using a Pierce Glycoprotein Staining Kit to visualize any glycoproteins, imaged, and then stained with the Coomassie blue reagent to visualize all protein bands.

Glycosylation was determined to be negative for both the COR23134 soybean-derived and tobacco-expressed IPD083Cb proteins (Figure 45 and Figure 46, respectively). The horseradish peroxidase positive control was clearly visible as a stained band. The soybean trypsin inhibitor negative control was not stained by the glycoprotein stain.

Additional details regarding glycosylation analytical methods are provided in Appendix G.

A) Glycoprotein Staining

B) Total Protein Staining



Glycoprotein Stain

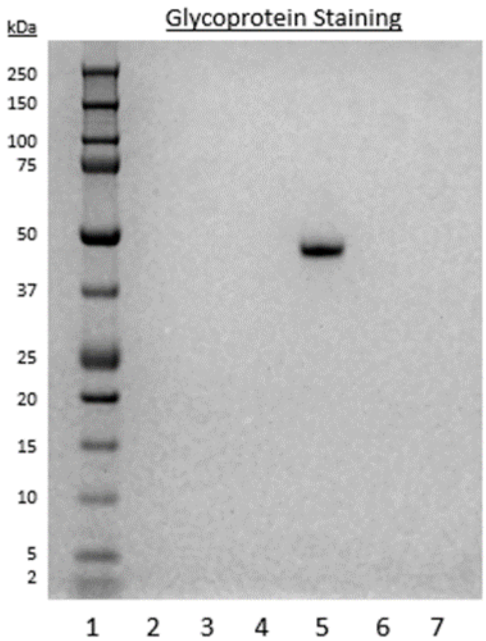
Total Protein Stain

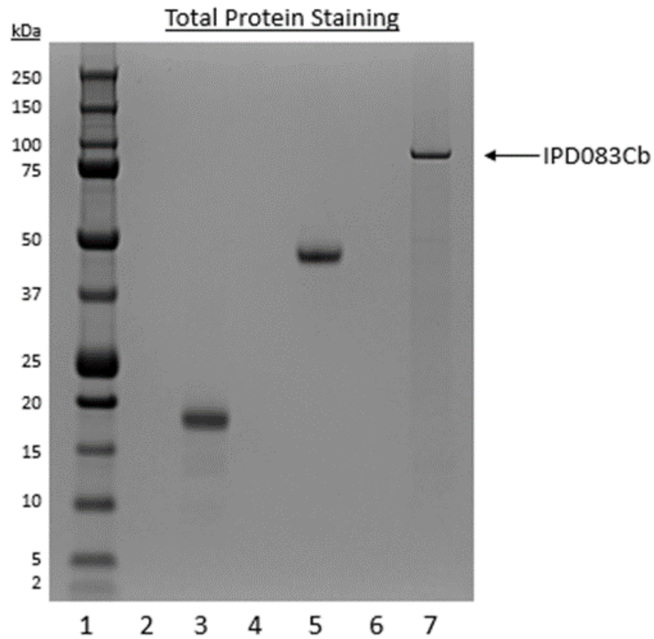
Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Horseradish Peroxidase Positive Control (1.0 µg)
4	1X LDS/DTT Sample Buffer Blank
5	Soybean Trypsin Inhibitor Negative Control (1.0 µg)
6	1X LDS/DTT Sample Buffer Blank
7	COR23134 Soybean-Derived IPD083Cb Protein

Note: The glycoprotein gel was stained with glycoprotein staining reagent. The total protein stain gel was stained with glycoprotein staining reagent followed by staining with Coomassie blue reagent for total proteins. Kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 45. Glycosylation Analysis of the COR23134 Soybean-Derived IPD083Cb Protein
A) Glycoprotein staining: Glycosylation was not detected for the COR23134 soybean-derived IPD083Cb protein (Lane 7). The horseradish peroxidase positive control was stained (Lane 3), and the soybean trypsin inhibitor negative control was not stained (Lane 5). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the COR23134 soybean-derived IPD083Cb protein (Lane 7) and both the positive (Lane 3) and negative (Lane 5) control proteins.

A) Glycoprotein Staining **B) Total Protein Staining**





Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Soybean Trypsin Inhibitor Negative Control (1.0 µg)
4	1X LDS/DTT Sample Buffer Blank
5	Horseradish Peroxidase Positive Control (1.0 µg)
6	1X LDS/DTT Sample Buffer Blank
7	Tobacco-Expressed IPD083Cb Protein

Note: The glycoprotein gel was stained with glycoprotein staining reagent. The total protein stain gel was stained with glycoprotein staining reagent followed by staining with Coomassie blue reagent for total proteins. Kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 46. Glycosylation Analysis of the Tobacco-Expressed IPD083Cb Protein
A) Glycoprotein staining: Glycosylation was not detected for the tobacco-expressed IPD083Cb protein (Lane 7). The horseradish peroxidase positive control was stained (Lane 5), and the soybean trypsin inhibitor negative control was not stained (Lane 3). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the tobacco-expressed IPD083Cb protein (Lane 7) and both the positive (Lane 5) and negative (Lane 3) control proteins.

Mass Spectrometry Peptide Mapping Analysis

Samples of the COR23134 soybean-derived IPD083Cb protein and the tobacco-expressed IPD083Cb protein were analyzed by SDS-PAGE. The gel was stained with Coomassie blue reagent, and the bands containing the IPD083Cb protein were excised for each sample. The excised IPD083Cb protein bands were digested with trypsin or chymotrypsin. The digested samples were analyzed using liquid chromatography-mass spectrometry (LC-MS). The resulting MS data were used to search and match the peptides from the IPD083Cb protein sequence, and the combined sequence coverage was calculated.

The combined sequence coverage of the identified tryptic and chymotryptic peptides for the COR23134 soybean-derived IPD083Cb protein accounts for 95.7% (815/852) of the amino acid sequence (Figure 47). The combined sequence coverage of the identified tryptic and chymotryptic peptides for the tobacco-expressed IPD083Cb protein accounts for 87.5% (751/858) of the amino acid sequence (Figure 48).

Additional details regarding peptide mapping analytical methods are provided in Appendix G.

1 **ADYSTLYRDL NQISMPLDRV EFSEVMVIHR MYLRLSDLNV GELPGAERVK**
51 **RLYVLADVVE LATFAHPQLL NTRMPGSVTV IILCRLLQFP TDGSFAAWLE**
101 **LPFMELHTLI EQYRSEIKAA DDAKWGTYVH AEEVQLSPLF NGWPYLVVEA**
151 **QRCIITAAMH NTFNRPGWVR SITQFTTDQS GRVDTTLLAR TEFGHIDLPL**
201 **ETDSPTAFSV SHRQSTNLPV EYTGIPVEVV TDPNILMGMQ TSVHIAELVK**
251 **ACYPSPELVS AVGVHVNWLN EVLLRVVQKE SQLQGTEAYN ECLALLGRIQ**
301 **CVMKMGPFVS VVPQLQYRMY GSLIRQMAQV AQNYDQDFRQ L**K**LFIAQNQI**
351 **LGSYLLQQNK AFADREVQME SFHSAVISQR RQELDDAIAK MDRLSLQMEE**
401 **EDRAMEQARK EMEEGLKQFQ NEQVARAVFA VLKSVAMIAL AFVTAGATAP**
451 **GAAASAAQAV NIAGQAAQAL RRVVEILEGL EAVMEVVAAI K**HLVDALDQV
501 SQIVDAPPMP DMPSEADW**SI FVNEIEAVAE GMPTEVSEVP AW**KAKCK**NVA**
551 **ALGREMCITA EQISQLQYDI WVQGLLRDIA QSHADRLAAI QPANLTNYLE**
601 **MAIQMDMR**TT R**ILIGLLNIM RIQNAALMYE YLLTPTQLTA WPLRMDTVAN**
651 **LLITHESAAL SGLAQLGPPS DFTSRHVVKG IPVSLLLDGG DWEFEIPVQG**
701 **GMSSFPSSWT RVRIRHLEMH FVQEASGGGE IIHQPATQTG TIYILLQGST**
751 **VFHDR**R**REEV MTFQAAVPLN YHYAYRLDTG EATLTNEPSE QFANTFMQMT**
801 **PFTHWRLRLS ASAAENKGLA FPTATAPDST TEIAITFHVT AIRQIDWRQE**
851 **EE**

Red type	Bold red type indicates soybean-derived IPD083Cb peptides identified using LC-MS analysis against the expected IPD083Cb protein sequence.
Amino acid residue abbreviations	alanine (A), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

Note: The IPD083Cb protein sequence does not include the N-terminal methionine as it is anticipated to be absent, and the N-terminal alanine residue of the protein was acetylated.

Figure 47. Identified Tryptic and Chymotryptic Peptide Amino Acid Sequence of the COR23134 Soybean-Derived IPD083Cb Protein Using LC-MS Analysis

1 **ADYSTLYRDL NQISMPLDRV EFSEVMVIHR MY**L**RLSDLNV GELPGAERVK**
51 **RLYVLADVVE LATFAHPQLL NTRMPGSVTV IILCRLLQFP TDGSF**AAWLE
101 LPF**MELHTLI EQYRSEIKAA DDAKWGTYVH AEEVQLSPLF NGWPYLVVEA**
151 **QRCIITAAMH NTFNRPGWVR SITQFTTDQS GRVDTTLLAR TEFGHIDLPL**
201 **ETDSPTAFSV SHRQSTNLPV EYTGIPVEVV TDPNILMGMQ TSVHIAELVK**
251 ACYPSPELVS AVGVHVNWLN EVL**LRVVQKE SQLQGTEAYN ECLALLGR**IQ
301 CVMK**MGPFVS VVPQLQYRMY GSLIRQMAQV AQNYDQDFRQ LKLFIAQNQI**
351 **LGSYLLQQNK AFADREVQME SFHSAVISQR RQELDDAIAK** MDR**LSLQMEE**
401 **EDRAMEQARK EMEEGLKQFQ NEQVARAVFA VLKSVAMIAL AFVTAGATAP**
451 **GAAASAAQAV NIAGQAAQAL** R**RVVEILEGL EAVMEVVAAI K**HLVDALDQV
501 SQIVDAPPMP DMPSEADW**SI FVNEIEAVAE GMPTEVSEVP AW**KAKCKNVA
551 AL**GREMCITA EQISQLQYDI WVQGLLRDIA QSHADRLAAI QPANLTNYLE**
601 **MAIQMDMR**TT R**ILIGLLNIM RIQNAALMYE YLLTPTQLTA WPLRMDTVAN**
651 **LLITHESAAL SGLAQLGPPS DFTSRHVVKG IPVSLLLDGG DWEFEIPVQG**
701 **GMSSFPSSWT R**VRIRHLEMH F**VQEASGGGE IIHQPATQTG TIY**ILLQGST
751 VFHDRR**REEV MTFQAAVPLN YHYAYRLDTG EATLTNEPSE QFANTFMQMT**
801 **PFTHWR**L**RLS ASAAENKGLA FPTATAPDST TEIAITFHVT AIRQIDWRQE**
851 **EEHHHHHH**

Red bold type	Red bold type indicates tobacco-expressed IPD083Cb peptides identified using LC-MS analysis against the expected IPD083Cb protein sequence.
Amino acid residue abbreviations	Alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

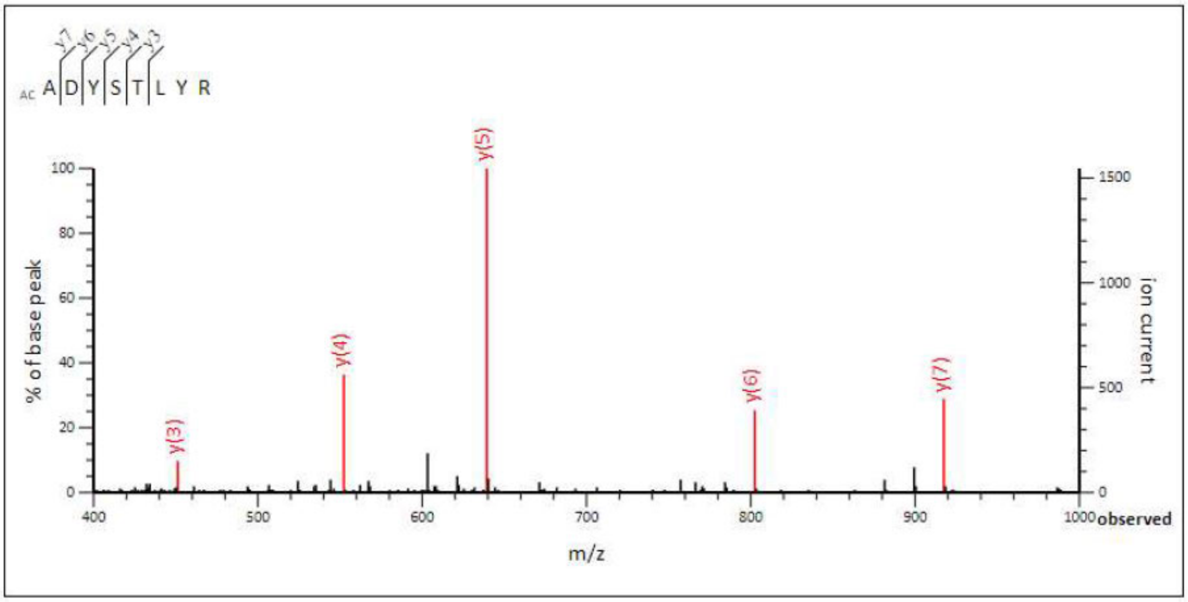
Note: The IPD083Cb protein sequence does not include the N-terminal methionine as it is anticipated to be absent, and the N-terminal alanine residue of the protein was acetylated.

Figure 48. Identified Tryptic and Chymotryptic Peptide Amino Acid Sequence of the Tobacco-Expressed IPD083Cb Protein Using LC-MS Analysis

N-Terminal Amino Acid Sequence Analysis

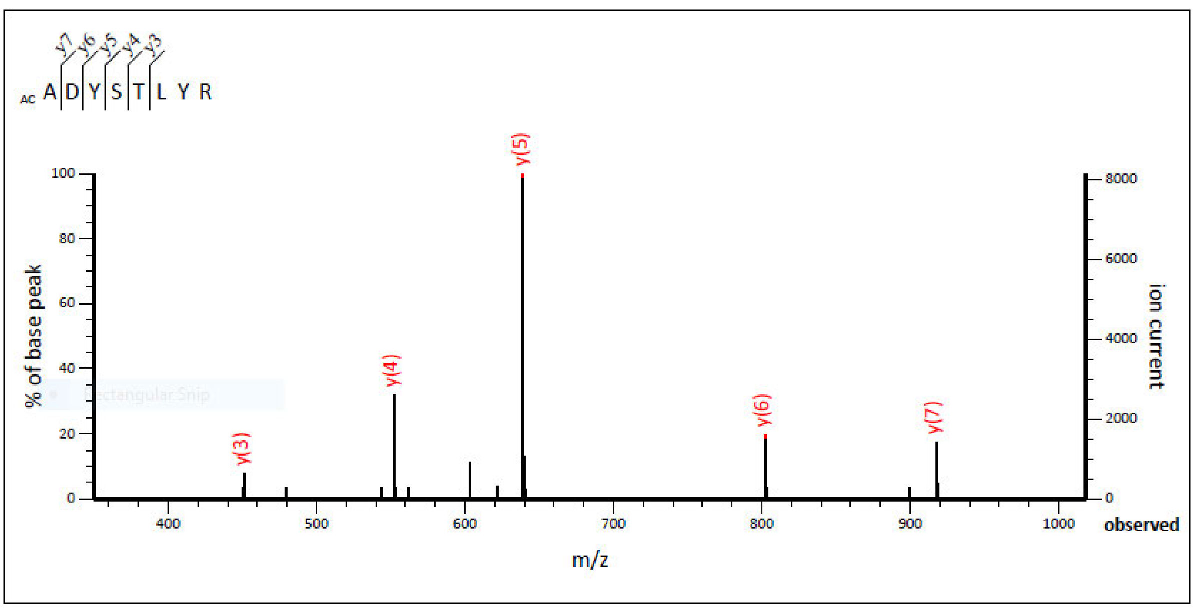
The Edman N-terminal amino acid sequence analysis of the COR23134 soybean-derived and tobacco-expressed IPD083Cb proteins did not obtain sequence data, suggesting the N-terminus of the protein was blocked. The N-terminal peptide was identified as ADYSTLYR from the tryptic digestion of the LC-MS analysis. The results indicated the N-terminal methionine was absent as expected (Dummitt *et al.*, 2003; Sherman *et al.*, 1985) and the N-terminal alanine residue of the protein was acetylated for the COR23134 soybean-derived and tobacco-expressed IPD083Cb proteins (Figure 49 and Figure 50, respectively).

Additional details regarding N-terminal amino acid sequencing analytical methods are provided in Appendix G.



Note: The N-terminal peptide was identified as ADYSTLYR from the tryptic digestion of the COR23134 soybean-derived IPD083Cb protein by LC-MS analysis. The N-terminal methionine was missing as expected and the alanine residue of the protein was acetylated (Ac). The mass spectrometry data show the mass to charge ratio (m/z) versus the intensity of the observed peptide fragment ions. The peptides were fragmented at the amide bond yielding b- and/or y-ions. Peaks labeled "b" and/or "y" represent ions where the charge is retained on the N-terminus or C-terminus, respectively. The number after the b- or y-ion corresponds to the peptide fragmentation site.

Figure 49. N-Terminal Peptide Identification of the COR23134 Soybean-Derived IPD083Cb Protein Using LC-MS Analysis



Note: The N-terminal peptide was identified as ADYSTLYR from the tryptic digestion of the IPD083Cb protein by LC-MS analysis. The N-terminal methionine residue of the protein was absent as expected and the N-terminal peptide was acetylated (Ac). The mass spectrometry data show the mass to charge ratio (m/z) versus the intensity of the observed peptide fragment ions relative to the intensity of the base peak. The peptides were fragmented at the amide bond yielding y-ions. Peaks labeled "y" represent ions where the charge is retained on the C-terminus. The number after the y-ion corresponds to the peptide fragmentation site.

Figure 50. N-Terminal Peptide Identification of the Tobacco-Expressed IPD083Cb Protein Using LC-MS Analysis

Bioactivity Assay

The biological activity of the tobacco-expressed IPD083Cb protein was evaluated by conducting a 7-day bioassay using *Chrysodeixis includens* (soybean looper; Lepidoptera: Noctuidae), a species sensitive to the IPD083Cb protein.

Bioactivity analysis demonstrated that the tobacco-expressed IPD083Cb protein had insecticidal activity toward a target insect, *C. includens* (Table 14). The biological activity of the test diet containing 50 µg of the IPD083Cb protein per cm^2 agar-based diet was demonstrated increased mortality and decreased weight in *C. includens* compared to *C. includens* fed the buffer control diet.

Additional details regarding bioactivity assay methods are provided in Appendix G.

Table 14. Summary of the Tobacco-Expressed IPD083Cb Protein Bioactivity Assay Using *Chrysodeixis includens*

Treatment	Treatment Description	Concentration (µg IPD083Cb Protein/cm^2 agar-based diet	Total Number of Observations	Mortality (%)	Number of Surviving Organisms	Weight of Surviving Organisms (mg)	
						Mean ± Standard Deviation	Range
1	Buffer Control Diet	0	24	16.7	20	15.2 ± 4.6	0.9-22.6
2	Test Diet	50.0	24	45.8	13	0.7 ± 0.5	0.1-1.6

Note: The summary of *Chrysodeixis includens* mortality data consisted of the calculation of dead larvae divided by the total number of observed larvae at the end of the study and multiplied by 100.

Allergenicity and Toxicity Analyses of the IPD083Cb Protein

A weight-of-evidence approach was applied to determine the allergenic and toxic potential of the IPD083Cb protein expressed in COR23134 soybean, including an assessment of the following: the history of safe use of the source organism, a bioinformatic comparison of the amino acid sequence of the IPD083Cb protein to known or putative allergen and protein toxin sequences, an evaluation of the heat lability of the IPD083Cb protein using a sensitive insect bioassay, evaluations of the stability of the IPD083Cb protein using *in vitro* gastric and intestinal digestion models, determination of the IPD083Cb protein glycosylation status, and an evaluation of acute toxicity in mice following oral exposure to the IPD083Cb protein.

Source Organism of the IPD083Cb Protein

The IPD083Cb protein expressed in COR23134 soybean is encoded by the insecticidal protein gene, *ipd083Cb*, from giant maidenhair fern (*Adiantum trapeziforme* var. *braziliense*).

Adiantum trapeziforme is known as the giant maidenhair fern or diamond maidenhair fern. Ferns are among the oldest living organisms on the planet and, with the exception of Antarctica, are globally distributed (Fernández, 2011). Ferns of the genus *Adiantum L.* are found in temperate and tropical regions worldwide. *A. trapeziforme L.* is native to the tropical rainforests of Central and South America (Kew Science, 2020) and has been introduced in the state of Florida in the United States (USDA-NRCS, 2023).

Humans have used ferns for many applications including occasional sources of food (Simmons and Herman, 2023). Members of the maidenhair fern family and other non-seed plants have been utilized for ethnomedicinal purposes from treating respiratory infections such as cough, colds, to pneumonia with research continuing into the potential benefits of compounds from members of this genus (Rastogi *et al.*, 2018). Many species of genus *Adiantum L.* are used in traditional medicine as infusions, decoctions, or pastes (Rastogi *et al.*, 2018). There are no reports of *A. trapeziforme* being poisonous to humans or livestock.

Bioinformatic Analysis of IPD083Cb Homology to Known and Putative Allergens

Assessing newly expressed proteins for potential cross-reactivity with known and putative allergens is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). A bioinformatic assessment of the IPD083Cb protein sequence (853 amino acids [aa]) for potential cross-reactivity with allergens was conducted by following established international criteria (Codex Alimentarius Commission, 2009; FAO/WHO, 2001).

Two separate searches for the IPD083Cb protein sequence were performed using the Comprehensive Protein Allergen Resource (COMPARE) 2023 database (January 26, 2023 (van Ree et al., 2021)). This peer-reviewed database is a collaborative effort of the Health and Environmental Sciences Institute (HESI) Protein Allergens, Toxins, and Bioinformatics (PATB) Committee and contains 2,631 sequences.

The first search was the sliding 80-mer window search, accomplished with an internally developed Perl script running FASTA v35.04 (Pearson and Lipman, 1988) with an E-score cutoff set to 100. In a sliding window search, each sequentially overlapping 80 aa subsequence of the overall

IPD083Cb protein sequence is used as a query against the COMPARE allergen database sequences. The script examined all alignments generated from the query and reported any possessing > 35% identity over an alignment length of ≥ 80 aa. Additionally, the script rescaled the percent identity to an 80-mer window for any alignments possessing an alignment length shorter than 80 aa; the number of identities in these alignments would be divided by 80, then multiplied by 100, and would report any alignment possessing an adjusted percent identity > 35%.

The second search used EMBOSS fuzzpro v6.6.0 (Rice et al., 2000) to identify any eight or greater contiguous identical amino acid matches between the IPD083Cb protein sequence and the COMPARE allergen sequences.

Results of the search of the IPD083Cb protein sequence against the COMPARE allergen database sequences found no alignments that were a length of 80 aa or greater with a sequence identity of > 35% and no alignments shorter than 80 aa with a sequence identity > 35% when normalized to an 80-mer window. No contiguous 8-residue exact matches between the IPD083Cb protein sequence and the allergen sequences were identified in the second search. Collectively, these data indicate that no allergenicity concern arose from the bioinformatics assessment of the IPD083Cb protein.

Bioinformatics evaluation of the IPD083Cb protein sequence did not generate biologically relevant amino acid sequence similarities to allergens that are harmful to humans or animals.

Bioinformatic Analysis of IPD083Cb Homology to Known and Putative Protein Toxins

Assessing newly expressed proteins for potential sequence similarity with protein toxin is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). The potential toxicity of the IPD083Cb protein was assessed by comparison of its sequence 1) an internal toxin database, and 2) the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein database.

The internal toxin database is a subset of sequences found in UniProtKB/Swiss-Prot (The UniProt Consortium, 2023). UniProtKB/Swiss-Prot is a curated database of non-redundant proteins containing functional information for over 550,000 sequences. To produce the internal toxin database, the proteins in UniProtKB/Swiss Prot are filtered for molecular function by keywords that could imply toxicity or adverse health effects (e.g., toxin, hemagglutinin, vasoactive, etc.). The internal toxin database is updated annually and contains 8,858 sequences.

The search between the IPD083Cb protein sequence and protein sequences in the internal toxin database was conducted using BLASTP v2.10.0+ with an *E*-value set to 10^{-4}. No alignments were returned between the IPD083Cb protein sequence and any protein sequence in the internal toxin database. Therefore, no toxicity concern arose from the bioinformatics assessment of the IPD083Cb protein.

The BLASTP search of the IPD083Cb protein against the NCBI nr protein database returned the 105 alignments with an *E*-value from 0 to 1×10^{-4} to various closely related IPD083Cb proteins or hypothetical proteins. None of the accessions returned by the BLASTP search are proteins known to be toxic to humans or animals.

Bioinformatics evaluation of the IPD083Cb protein sequence did not generate biologically relevant amino acid sequence similarities to protein toxins that are harmful to humans or animals.

Thermolability Analysis

Thermal stability of the IPD083Cb protein was characterized by determining the biological activity of the heat-treated IPD083Cb protein when applied to an artificial diet and fed to *Anticarsia gemmatalis* (*A. gemmatalis*), an insect sensitive to the IPD083Cb protein. The IPD083Cb protein was incubated at various temperatures (25 °C, 50 °C, 75 °C, and 95 °C) for approximately 30 minutes before incorporation into the artificial diet. Each test diet contained a targeted concentration of 50 µg IPD083Cb protein per cm^2 agar-based diet. Larvae were exposed via oral ingestion to the treatments prepared by surface application of dosing solutions to an agar-based artificial diet in a 7-day bioassay. A positive control diet containing the unheated IPD083Cb protein and a bioassay control diet containing ultrapure water were included in the bioassay to verify assay performance. After seven days, statistical analyses were conducted to evaluate *A. gemmatalis* mortality of the heat-treated test groups relative to the unheated test group.

The results demonstrated that the IPD083Cb protein heated for approximately 30 minutes at 75 °C and 95 °C (Treatments 5 and 6, respectively) had significantly reduced activity against *A. gemmatalis* (P values < 0.0001) compared to the unheated IPD083Cb control when applied to an artificial diet. No statistically significant decreases in protein activity were observed for the IPD083Cb protein heated for approximately 30 minutes at 25 °C and 50 °C (Treatments 3 and 4, respectively) when compared to the unheated IPD083Cb control (Table 15).

Additional details regarding thermolability analytical methods are provided in Appendix G.

Table 15. Biological Activity of the Heat-Treated IPD083Cb Protein in Artificial Diet Fed to *Anticarsia gemmatalis*

Treatment	Treatment Description	Test Dosing Solution Incubation Condition	Total Number of Observations	Mortality (%)	Fisher's Test P-Value	Number of Surviving Organisms	Weight of Surviving Organisms (mg)	
							Mean ± Standard Deviation	Range
1	Bioassay Control Diet	NA	24	0	--	24	33.6 ± 7.71	18.8 - 50.7
2	Control Diet	Unheated	24	100	--	0	--	--
3	Test Diet	25 °C	24	95.8	0.5000	1	0.00[c,d]	--
4	Test Diet	50 °C	23[a]	95.7	0.4894	1	0.200[c]	--
5	Test Diet	75 °C	24	16.7	<0.0001[b]	20	20.3 ± 9.56	5.0 - 40.5
6	Test Diet	95 °C	24	12.5	<0.0001[b]	21	21.9 ± 9.90	3.5 - 40.7

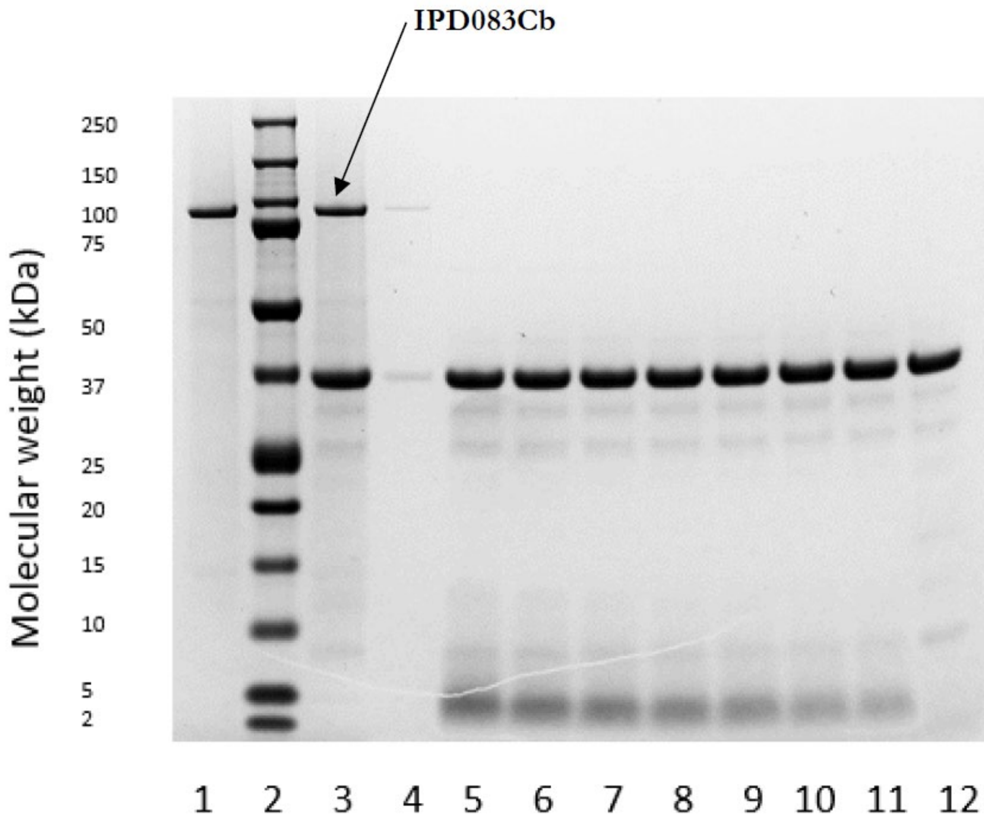
Note: Not applicable (NA); the bioassay control diet did not contain the IPD083Cb protein and the dosing solution was not incubated. Dosing solutions used to prepare the bioassay control and unheated control diets were maintained on wet ice until applied to diets.
[a] Organisms counted as missing during the bioassay were not included in the total number of observations for a given treatment.
[b] A statistically significant difference (P-value < 0.05) was observed in comparison to Treatment 2.
[c] The reported mean is the weight value of the one surviving organism.
[d] A weight of 0 mg is possible for some live organisms due to limitations of the balance.

Digestibility Analysis with Simulated Gastric Fluid (SGF)

Simulated gastric fluid (SGF) containing pepsin at pH ~1.2 was used to assess the susceptibility of the IPD083Cb protein to proteolytic digestion by pepsin *in vitro*. The IPD083Cb protein was incubated in SGF for 0, 0.25, 1, 2, 5, 10, 30, and 60 minutes. A positive control (bovine serum albumin) and a negative control (β-lactoglobulin) were included in the assay and were incubated in SGF for 0, 1, and 60 minutes. After incubation in SGF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SGF digestibility results showed that the IPD083Cb protein migrating at approximately 96 kDa was digested within 0.25 minutes in SGF as demonstrated by both SDS-PAGE and western blot analysis (Figure 51 and Figure 52, respectively). A band migrating at ~60 kDa was detected by western blot and > 99% was digested within 5 minutes. The band was undetectable within 30 minutes in SGF. On the SDS-PAGE gel, low molecular weight bands (~2-5 kDa) remained detectable in the IPD083Cb protein samples for 60 minutes in SGF.

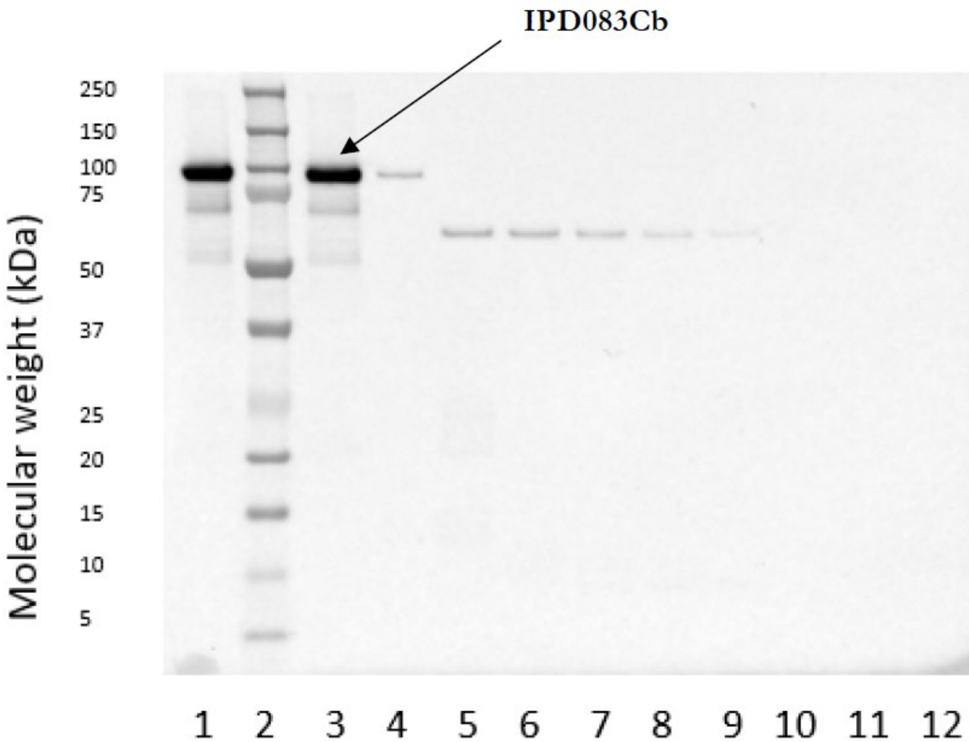
Additional details regarding SGF analytical methods are provided in Appendix G.



Lane	Sample Descriptions
1	IPD083Cb protein in water (no SGF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	IPD083Cb protein in SGF, Time 0
4	IPD083Cb protein in SGF, Time 0 (1:20 dilution)
5	IPD083Cb protein in SGF, 15 seconds
6	IPD083Cb protein in SGF, 1 minute
7	IPD083Cb protein in SGF, 2 minutes
8	IPD083Cb protein in SGF, 5 minutes
9	IPD083Cb protein in SGF, 10 minutes
10	IPD083Cb protein in SGF, 30 minutes
11	IPD083Cb protein in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: kilodalton (kDa), simulated gastric fluid (SGF), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

Figure 51. SDS-PAGE Analysis of the IPD083Cb Protein in Simulated Gastric Fluid Digestion Time Course



Lane	Sample Descriptions
1	IPD083Cb protein in water (no SGF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	IPD083Cb protein in SGF, Time 0
4	IPD083Cb protein in SGF, Time 0 (1:100 dilution)
5	IPD083Cb protein in SGF, 15 seconds
6	IPD083Cb protein in SGF, 1 minute
7	IPD083Cb protein in SGF, 2 minutes
8	IPD083Cb protein in SGF, 5 minutes
9	IPD083Cb protein in SGF, 10 minutes
10	IPD083Cb protein in SGF, 30 minutes
11	IPD083Cb protein in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: kilodalton (kDa) and simulated gastric fluid (SGF).
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

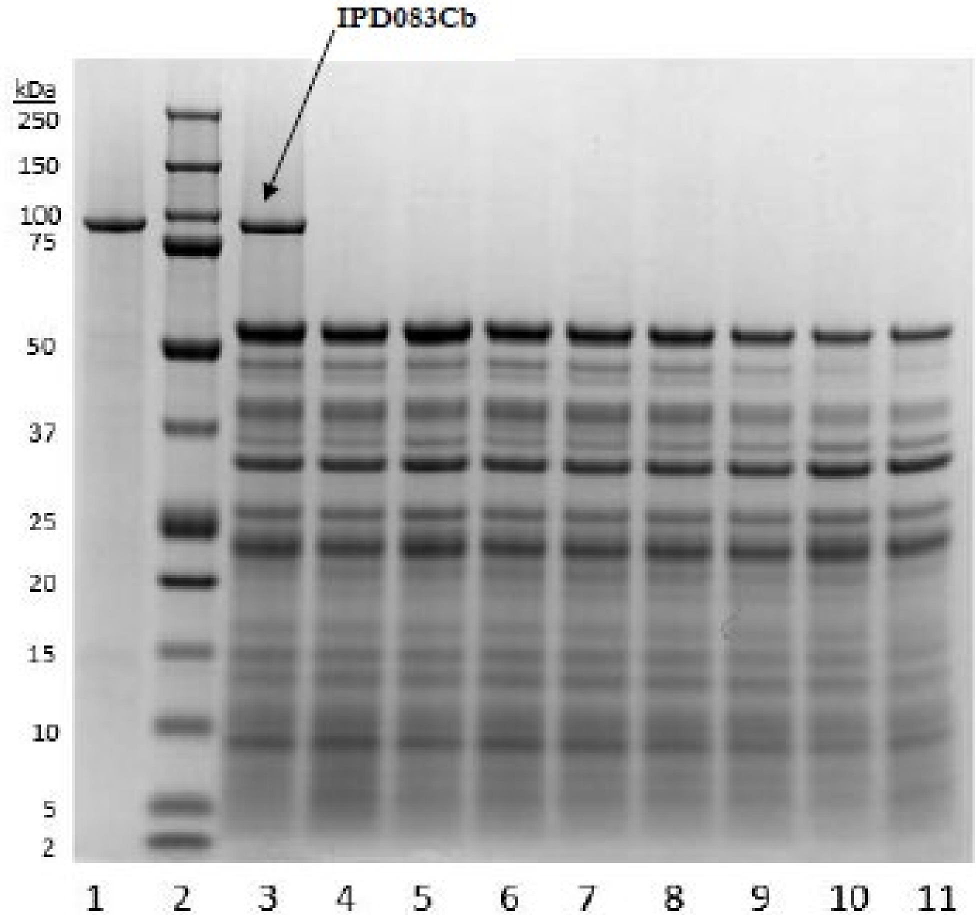
Figure 52. Western Blot Analysis of the IPD083Cb Protein in Simulated Gastric Fluid Digestion Time Course

Digestibility Analysis with Simulated Intestinal Fluid (SIF)

Simulated intestinal fluid (SIF) containing pancreatin at ~pH 7.5 was used to assess the susceptibility of the IPD083Cb protein to proteolytic digestion by pancreatin in vitro. The IPD083Cb protein was incubated in SIF for 0, 0.25, 1, 2, 5, 10, 30, and 60 minutes. A positive control (β-lactoglobulin) and a negative control (bovine serum albumin) were included in the assay and were incubated in SIF for 0, 1, and 60 minutes. After incubation in SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SIF digestibility results showed that the IPD083Cb protein migrating at approximately 96 kDa was digested into smaller fragments within 0.25 minutes in SIF as demonstrated by both SDS-PAGE and western blot analysis (Figure 53 and Figure 54, respectively). These smaller fragments migrating at approximately 20-60 kDa remained detectable via western blot for 60 minutes.

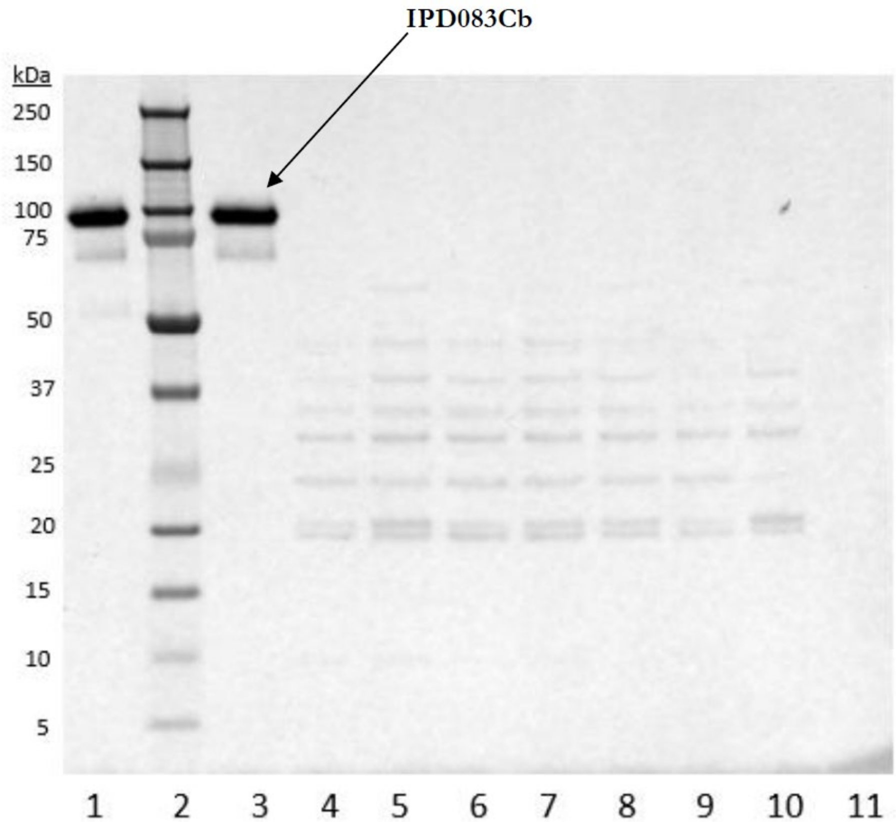
Additional details regarding SIF analytical methods are provided in Appendix G.



Lane	Sample Descriptions
1	IPD083Cb protein in water (no SIF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	IPD083Cb protein in SIF, Time 0
4	IPD083Cb protein in SIF, 15 seconds
5	IPD083Cb protein in SIF, 1 minute
6	IPD083Cb protein in SIF, 2 minutes
7	IPD083Cb protein in SIF, 5 minutes
8	IPD083Cb protein in SIF, 10 minutes
9	IPD083Cb protein in SIF, 30 minutes
10	IPD083Cb protein in SIF, 60 minutes
11	SIF Control, 60 minutes

Note: kilodalton (kDa) and simulated intestinal fluid (SIF)
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

Figure 53. SDS-PAGE Analysis of the IPD083Cb Protein in Simulated Intestinal Fluid Digestion Time Course



Lane	Sample Descriptions
1	IPD083Cb protein in water (no SIF), Time 0
2	Pre-stained protein molecular weight marker[a]
3	IPD083Cb protein in SIF, Time 0
4	IPD083Cb protein in SIF, 15 seconds
5	IPD083Cb protein in SIF, 1 minute
6	IPD083Cb protein in SIF, 2 minutes
7	IPD083Cb protein in SIF, 5 minutes
8	IPD083Cb protein in SIF, 10 minutes
9	IPD083Cb protein in SIF, 30 minutes
10	IPD083Cb protein in SIF, 60 minutes
11	SIF Control, 60 minutes

Note: kilodalton (kDa) and simulated intestinal fluid (SIF)
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

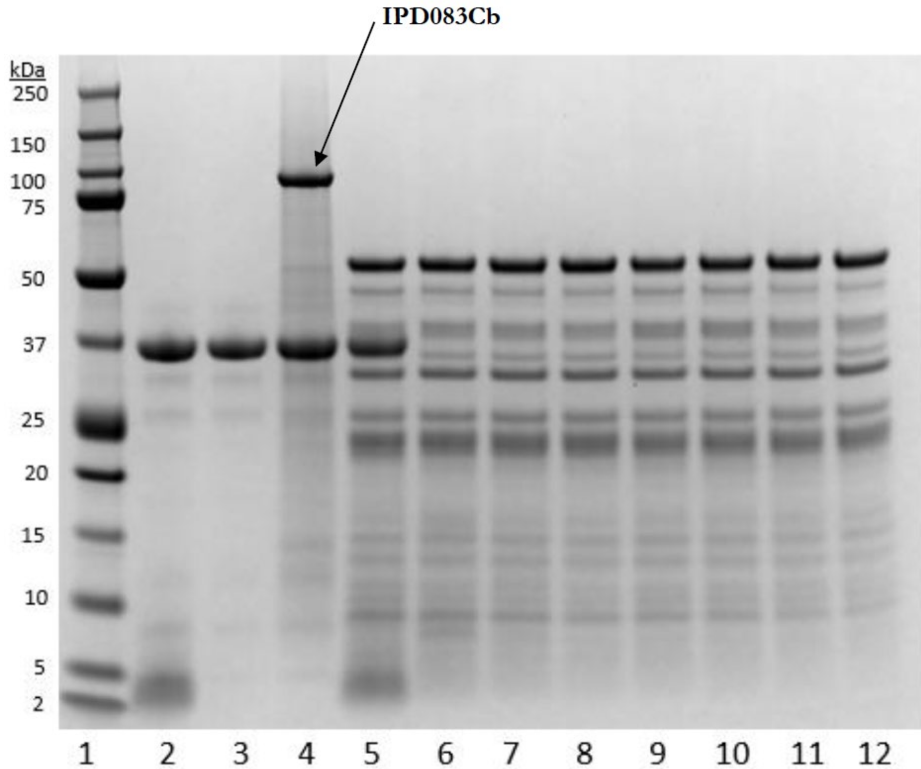
Figure 54. Western Blot Analysis of the IPD083Cb Protein in Simulated Intestinal Fluid Digestion Time Course

Sequential Digestibility Analysis with Simulated Gastric Fluid (SGF) and Simulated Intestinal Fluid (SIF)

Sequential digestion in simulated intestinal fluid (SIF) following a digestion in SGF was used to assess the susceptibility of the low molecular weight SGF fragments (~2-5 kDa, Figure 51) of the IPD083Cb protein. The IPD083Cb protein was incubated for 10 minutes in SGF containing pepsin at pH ~1.2 and then incubated for 0, 0.25, 1, 2, 5, 10, 20, and 30 minutes in SIF containing pancreatin at pH ~7.5. After incubation in SGF/ SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain was used to detect protein bands.

The sequential pepsin (SGF) and pancreatin (SIF) digestibility results showed that the low molecular weight bands (~2-5 kDa) observed in SGF digestion (Figure 51) were digested within 0.25 minute during sequential SIF digestion (Figure 55).

Additional details regarding sequential digestibility analytical methods are provided in Appendix G.



Lane	Sample Descriptions
1	Pre-stained protein molecular weight marker[a]
2	IPD083Cb Protein in SGF, 10 minutes
3	SGF Control, 10 minutes
4	IPD083Cb Protein in SGF, Time 0
5	IPD083Cb Protein in SGF 10 minutes, SIF Time 0
6	IPD083Cb Protein in SGF 10 minutes, SIF 0.25 minutes
7	IPD083Cb Protein in SGF 10 minutes, SIF 1 minute
8	IPD083Cb Protein in SGF 10 minutes, SIF 2 minutes
9	IPD083Cb Protein in SGF 10 minutes, SIF 5 minutes
10	IPD083Cb Protein in SGF 10 minutes, SIF 10 minutes
11	IPD083Cb Protein in SGF 10 minutes, SIF 20 minutes
12	IPD083Cb Protein in SGF 10 minutes, SIF 30 minutes

Note: kilodalton (kDa), simulated gastric fluid (SGF), simulated intestinal fluid (SIF), and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).
[a] Molecular weight markers were included to provide a visual estimate of protein migration.

Figure 55. SDS-PAGE Analysis of the IPD083Cb Protein in a Sequential Digestion with Simulated Gastric Fluid and Simulated Intestinal Fluid

Protein Glycosylation Analysis

As stated previously in the characterization section, the results from glycoprotein staining analysis confirmed the absence of glycosylation for the COR23134 soybean-derived and the tabacco-expressed IPD083Cb proteins.

Evaluation of the Acute Toxicity of the IPD083Cb Protein

A study was conducted to evaluate the potential acute toxicity of the test substance, IPD083Cb, in [redacted] mice following oral exposure at the limit dose (5000 mg/kg body weight, adjusted for IPD083Cb content). The IPD083Cb protein and bovine serum albumin (BSA) protein were each reconstituted in deionized water. Vehicle control, BSA comparative control, and the IPD083Cb test substance formulations were administered orally by gavage in three split doses, separated by approximately four hours; the BSA comparative control was administered at an equivalent target dose to that of the IPD083Cb protein. The mice were fasted prior to and throughout the dosing procedure.

Body weights were evaluated on test day 1 (prior to fasting and shortly prior to administration of the first dose), 2, 3, 5, 8, and 15. Clinical signs were evaluated ten times on test day 1 (distributed before and after each dose) and daily thereafter. On test day 15, all surviving mice were euthanized and given a gross pathological examination.

There were no test substance-related incidents of mortality or clinical signs. All animals in the test substance-treated group survived to scheduled euthanasia. A total of 4 unscheduled deaths occurred in the vehicle control or BSA control groups and were attributed to technical error during dose administration, based on perforation of the esophagus noted in the gross pathology evaluation for all 4 of these animals. All surviving animals gained weight during the 2-week observation period prior to euthanasia. No test substance-related gross lesions were observed.

Under the conditions of this study, intragastric exposure of IPD083Cb protein to male and female mice at 5000 mg/kg body weight did not result in mortality or other evidence of acute oral toxicity, based on evaluation of body weight, clinical signs, and gross pathology. Therefore, the acute oral toxicity tolerant dose and the LD_{50} of the IPD083Cb protein were determined to be greater than 5000 mg/kg body weight.

Conclusion on Safety of the IPD083Cb Protein in COR23134 Soybean

In conclusion, protein characterization results via SDS-PAGE, western blot, glycosylation analysis, mass spectrometry peptide mapping analysis, and N-terminal amino acid sequence analysis have demonstrated that the IPD083Cb protein derived from COR23134 soybean and the tobacco expression system has the expected molecular weight, immunoreactivity, and amino acid sequence, and is not glycosylated. Characterization of the tobacco-expressed IPD083Cb protein demonstrated that it is an appropriate test substance for use in safety studies.

The allergenic potential of the IPD083Cb protein was evaluated by assessing the IPD083Cb protein source organism and history of safe use, a bioinformatic comparison of the amino acid sequence of the IPD083Cb protein with known and putative allergen sequences, an evaluation of the heat lability of the IPD083Cb protein using a sensitive insect bioassay, evaluations of the stability of the IPD083Cb protein using *in vitro* gastric and intestinal digestion models, and determination of the IPD083Cb protein glycosylation status. The toxicity potential of the IPD083Cb protein was evaluated by a bioinformatic comparison of the IPD083Cb amino acid sequence to known and putative protein toxins and by an acute toxicity in mice following oral exposure to the IPD083Cb protein.

The bioinformatic comparison of the IPD083Cb protein sequence to known and putative allergen and protein toxin sequences showed that the IPD083Cb protein is unlikely to be allergenic or toxic for humans or animals. The IPD083Cb protein was digested within 0.25 minutes in SGF, and some low molecular weight bands (~2-5 kDa) remained detectable via SDS-PAGE for 60 minutes in SGF. The IPD083Cb protein was digested within 0.25 minutes in SIF, and some smaller fragments (~20-60 kDa) remained detectable via western blot for 60 minutes. The low molecular weight bands remaining from SGF digestion were digested within 0.25 minute during sequential SIF digestion. The IPD083Cb protein was not glycosylated. The IPD083Cb protein heated for approximately 30 minutes at 75 °C and 95 °C had significantly reduced activity against *Anticarsia gemmatalis* when applied to an artificial diet. The acute oral toxicity assessment in mice determined the LD_{50} of the IPD083Cb protein to be greater than 5000 mg/kg. These data support the conclusion that COR23134 soybean, expressing the IPD083Cb protein, is as safe as conventional soybean for the food and feed supply.

Based on this weight of evidence, consumption of the IPD083Cb protein from COR23134 soybean is unlikely to cause an adverse effect on humans or animals.

GM-HRA Protein

Amino Acid Sequence of the GM-HRA Protein

The deduced amino acid sequence from the translation of the *gm-hra_1* gene encodes the GM-HRA- protein that is 651 amino acids in length and has a molecular weight of approximately 70 kDa for the full-length precursor protein (Figure 56). The first 47 amino acids at the N-terminus (i.e., chloroplast transit peptide) is cleaved from the precursor protein during processing in planta, resulting in the mature GM-HRA protein of 604 amino acids with a molecular weight of approximately 65 kDa. The first amino acid of the mature GM-HRA protein is the serine (S) located at the 48th aa position of the precursor protein (Figure 56).

```
  1  MAATASRTTR FSSSSSHPTF PKRITRSTLP LSHQTLTKPN HALKIKCSIS
 51  KPPTAAPFTK EAPTTEPFVS RFASGEPRKG ADILVEALER QGVTTVFAYP
101  GGASMEIHQA LTRSAAIRNV LPRHEQGGVF AAEGYARSSG LPGVCIATSG
151  PGATNLVSGL ADALMDSVPV VAITGQVARR MIGTDAFQET PIVEVSRSIT
201  KHNYLILDVD DIPRVVAEAF FVATSGRPGP VLIDIPKDVQ QQLAVPNWDE
251  PVNLPGYLAR LPRPPAEAQL EHIVRLIMEA QKPVLYVGGG SLNSSAELRR
301  FVELTGIPVA STLMGLGTFP IGDEYSLQML GMHGTVYANY AVDNSDLLLA
351  FGVRFDDRVT GKLEAFASRA KIVHIDIDSA EIGKNKQAHV SVCADLKLAL
401  KGINMILEEK GVEGKFDLGG WREEINVQKH KFPLGYKTFQ DAISPQHAIE
451  VLDELTNGDA IVSTGVGQHQ MWAAQFYKYK RPRQWLTSGG LGAMGFGLPA
501  AIGAAVANPG AVVVDIDGDG SFIMNVQELA TIRVENLPVK ILLLNNQHLG
551  MVVQLEDRFY KSNRAHTYLG DPSSESEIFP NMLKFADACG IPAARVTKKE
601  ELRAAIQRML DTPGPYLLDV IVPHQEHVLP MIPSNGSFKD VITEGDGRTR
651  Y*
```

Figure 56. Deduced Amino Acid Sequence of the GM-HRA Protein

The deduced amino acid sequence from the translation of the *gm-hra_1* gene from plasmid [redacted] The asterisk (*) indicates the translational stop codon. The full-length precursor GM-HRA protein is 651 amino acids in length and has a molecular weight of approximately 70 kDa. The first 47 amino acids at the N-terminus (i.e., chloroplast transit peptide) is cleaved from the precursor protein during processing in planta, resulting in the mature GM-HRA protein of 604 amino acids with a molecular weight of approximately 65 kDa. The bold, shaded amino acid residue, serine (S) at the 48th aa position, is the first amino acid of the mature GM-HRA protein sequence. The two bold, underlined amino acid residues, alanine (A) and leucine (L) at the 178th and 555th aa positions, respectively, are the differences from the soybean endogenous acetolactate synthase (ALS) protein sequence (proline and tryptophan, respectively), which render GM-HRA tolerant to ALS-inhibiting herbicides.

Function and Activity of the GM-HRA Protein

GM-HRA Expressed in COR23134 Soybean

The GM-HRA protein expressed in COR23134 soybean is encoded by the *gm-hra_1* gene which was derived from the soybean acetolactate synthase (*als*) gene by introducing two mutations in the amino acid sequence: the alanine (A) at the 178th aa position and the leucine (L) at the 555th aa position (Figure 56) in relation to proline (P) and tryptophan (W), respectively in the native ALS protein. These two amino acid changes are responsible for GM-HRA tolerance to ALS-inhibiting herbicides (e.g., sulfonylureas and triazolopyrimidine) (Falco and Li, 2010). The expressed GM-HRA protein in COR23134 soybean was used as a selectable marker during the event development.

GM-HRA Mode of Action

The mode of action of GM-HRA has been characterized and described (Lee *et al.*, 1988). The GM-HRA protein confers tolerance to the ALS-inhibiting herbicides, which act by inhibition of the acetolactate synthase (ALS) enzyme required for biosynthesis of the branched chain amino acids (isoleucine, leucine, and valine). ALS-inhibiting herbicides act by blocking access to the ALS enzyme active site, thus preventing the biosynthetic reaction from occurring. The GM-HRA protein contains mutations first identified in sulfonylurea-tolerant tobacco mutants that were subsequently engineered into the soybean *als* gene to result in expression of GM-HRA.

GM-HRA Encoded by the gm-hra and gm-hra_1 Genes

There are two versions of the modified *als* gene sequence for deployment in different soybean products: 1) the *gm-hra* gene in soybean event DP-3Ø5423-1 (referred to as DP305423 soybean; evaluated by FSANZ in the A1018 application in 2009) and 2) the *gm-hra_1* gene in COR23134 soybean. Both the *gm-hra* and *gm-hra_1* genes, when expressed in planta, produce the same mature GM-HRA protein after the cleavage of chloroplast transit peptides (Figure 57).

Sequences of the 1,956-bp *gm-hra_1* gene in COR23134 soybean are identical to those of the 1,971-bp *gm-hra* gene in DP305423 soybean except the *gm-hra* gene contains 15 additional nucleotides from the soybean *als* 5′ UTR at the 5′ end. Therefore, the deduced amino acid sequence from translation of the *gm-hra_1* and *gm-hra* genes is identical, with the exception that the *gm-hra* gene in DP305423 soybean encodes the five additional amino acids (MPHNT) at the N-terminus of the protein (Figure 57). However, as the N-terminus of the precursor protein sequence contains a chloroplast transit peptide, the mature GM-HRA proteins encoded from both the *gm-hra_1* and *gm-hra* genes are identical following cleavage of the chloroplast transit peptide upon import into the plastid.

The mature GM-HRA protein encoded by the *gm-hra_1* gene in COR23134 soybean is identical, after import into the chloroplast and removal of the transit peptide, to the mature GM-HRA protein encoded by the *gm-hra* gene in DP305423 soybean. The GM-HRA protein has been previously risk-assessed for potential allergenicity and toxicity by numerous regulatory agencies and is unlikely to present significant risks to the environment, human, or animal health (Mathesius *et al.*, 2009). DP305423 soybean, expressing the GM-HRA protein, has been authorized for food and/or feed use by regulatory authorities in 19 different countries and/or regions (ISAAA, 2023). These previous assessments of the GM-HRA protein are also relevant for the assessment of the GM-HRA protein in COR23134 soybean.

Based on the significant weight of evidence provided above, it is concluded that it is appropriate to reference and rely upon the related GM-HRA safety studies used for DP305423 soybean to assess the safety of COR23134 soybean.

```
gm-hra      MPHNTMAATA SRTTRFSSSS SHPTFPKRIT RSTLPLSHQT LTKPNHALKI   50
gm-hra_1    -----MAATA SRTTRFSSSS SHPTFPKRIT RSTLPLSHQT LTKPNHALKI   45

gm-hra      KCSISKPPTA APFTKEAPTT EPFVSRFASG EPRKGADILV EALERQGVTT  100
gm-hra_1    KCSISKPPTA APFTKEAPTT EPFVSRFASG EPRKGADILV EALERQGVTT   95

gm-hra      VFAYPGGASM EIHQALTRSA AIRNVLPRHE QGGVFAAEGY ARSSGLPGVC  150
gm-hra_1    VFAYPGGASM EIHQALTRSA AIRNVLPRHE QGGVFAAEGY ARSSGLPGVC  145

gm-hra      IATSGPGATN LVSGLADALM DSVPVVAITG QVARRMIGTD AFQETPIVEV  200
gm-hra_1    IATSGPGATN LVSGLADALM DSVPVVAITG QVARRMIGTD AFQETPIVEV  195

gm-hra      SRSITKHNYL ILDVDDIPRV VAEAFFVATS GRPGPVLIDI PKDVQQQLAV  250
gm-hra_1    SRSITKHNYL ILDVDDIPRV VAEAFFVATS GRPGPVLIDI PKDVQQQLAV  245

gm-hra      PNWDEPVNLP GYLARLPRPP AEAQLEHIVR LIMEAQKPVL YVGGGSLNSS  300
gm-hra_1    PNWDEPVNLP GYLARLPRPP AEAQLEHIVR LIMEAQKPVL YVGGGSLNSS  295

gm-hra      AELRRFVELT GIPVASTLMG LGTFPIGDEY SLQMLGMHGT VYANYAVDNS  350
gm-hra_1    AELRRFVELT GIPVASTLMG LGTFPIGDEY SLQMLGMHGT VYANYAVDNS  345

gm-hra      DLLLAFGVRF DDRVTGKLEA FASRAKIVHI DIDSAEIGKN KQAHVSVCAD  400
gm-hra_1    DLLLAFGVRF DDRVTGKLEA FASRAKIVHI DIDSAEIGKN KQAHVSVCAD  395

gm-hra      LKLALKGINM ILEEKGVEGK FDLGGWREEI NVQKHKFPLG YKTFQDAISP  450
gm-hra_1    LKLALKGINM ILEEKGVEGK FDLGGWREEI NVQKHKFPLG YKTFQDAISP  445

gm-hra      QHAIEVLDEL TNGDAIVSTG VGQHQMWAAQ FYKYKRPRQW LTSGGLGAMG  500
gm-hra_1    QHAIEVLDEL TNGDAIVSTG VGQHQMWAAQ FYKYKRPRQW LTSGGLGAMG  495

gm-hra      FGLPAAIGAA VANPGAVVVD IDGDGSFIMN VQELATIRVE NLPVKILLLN  550
gm-hra_1    FGLPAAIGAA VANPGAVVVD IDGDGSFIMN VQELATIRVE NLPVKILLLN  545

gm-hra      NQHLGMVVQL EDRFYKSNRA HTYLGDPSSE SEIFPNMLKF ADACGIPAAR  600
gm-hra_1    NQHLGMVVQL EDRFYKSNRA HTYLGDPSSE SEIFPNMLKF ADACGIPAAR  595

gm-hra      VTKKEELRAA IQRMLDTPGP YLLDVIVPHQ EHVLPMIPSN GSFKDVITEG  650
gm-hra_1    VTKKEELRAA IQRMLDTPGP YLLDVIVPHQ EHVLPMIPSN GSFKDVITEG  645

gm-hra      DGRTRY*                                                 656
gm-hra_1    DGRTRY*                                                 651
```

Figure 57. Alignments of the Deduced Amino Acid Sequence of the GM-HRA Protein Encoded by the *gm-hra* and *gm-hra_1* Genes

Sequence alignments of the GM-HRA protein show that the deduced amino acid sequences from the translation of the *gm-hra* gene in DP305423 soybean and the *gm-hra_1* gene in COR23134 soybean are identical with the exception that the *gm-hra* gene in DP305423 soybean encodes the five additional amino acids (MPHNT) at the N-terminus. The chloroplast transit peptide sequence that is removed from the precursor protein upon import into the plastid to form the mature GM-HRA protein is underlined. Asterisks (*) indicate the translational stop codons.

Characterization of the GM-HRA Protein Derived from COR23134 Soybean and the Microbial System

The GM-HRA protein expressed in COR23134 soybean was purified from the whole plant tissue using immunoaffinity chromatography.

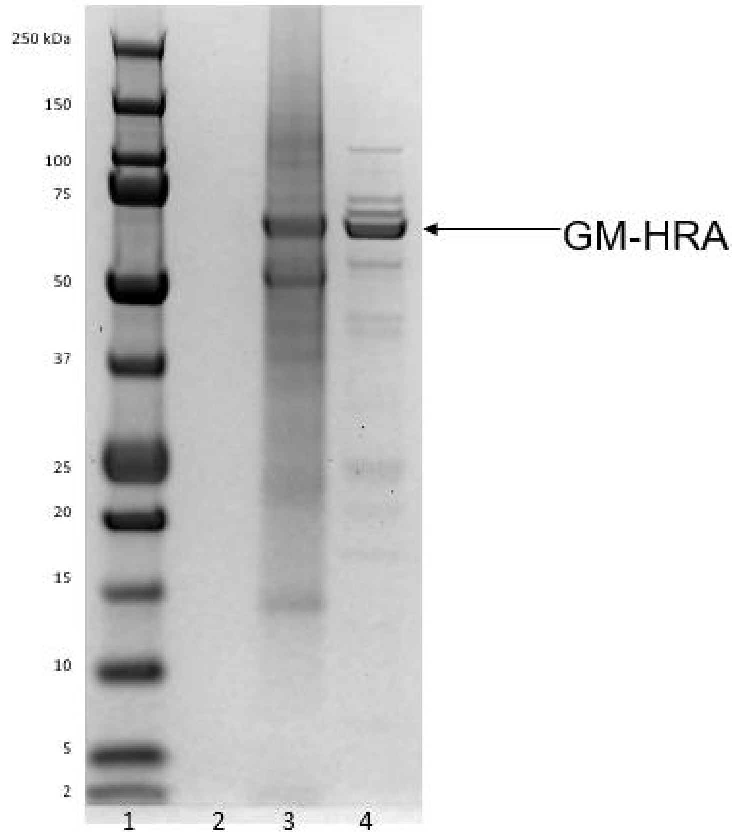
In order to have sufficient amounts of the purified GM-HRA protein for the multiple studies required to assess its safety, the GM-HRA protein was expressed with an N-terminal His-T7 fusion tag in an *Escherichia coli* protein expression system. The microbially derived GM-HRA protein was purified using immobilized metal affinity chromatography and the fusion tag was removed through thrombin cleavage. Thrombin cleavage of the His-T7 fusion tag resulted in one additional N-terminal amino acid residue, glycine (G), for the microbially derived GM-HRA protein.

The biochemical characteristics of the COR23134 soybean-derived and microbially derived GM-HRA proteins were characterized using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot, glycosylation analysis, peptide mapping by liquid chromatography mass spectrometry (LC-MS), and N-terminal amino acid sequencing. For the microbially derived GM-HRA protein, the functional activity was verified by a spectrophotometric assay for enzymatic measurement of acetolactate synthase (ALS). The results demonstrated that the COR23134 soybean-derived and microbially derived GM-HRA proteins have the expected molecular weight, immunoreactivity, and amino acid sequence, and are not glycosylated. The microbially derived GM-HRA protein was demonstrated to be an appropriate test substance for use in safety studies.

SDS-PAGE Analysis

Samples of the COR23134 soybean-derived GM-HRA protein and the microbially derived GM-HRA protein were analyzed by SDS-PAGE. As expected, the GM-HRA proteins, derived from both COR23134 soybean and the microbial system, migrated as a predominant band consistent with the expected molecular weight (Figure 58).

Additional details regarding SDS-PAGE analytical methods are provided in Appendix H.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	COR23134 Soybean-Derived GM-HRA Protein
4	Microbially Derived GM-HRA Protein (1 µg)

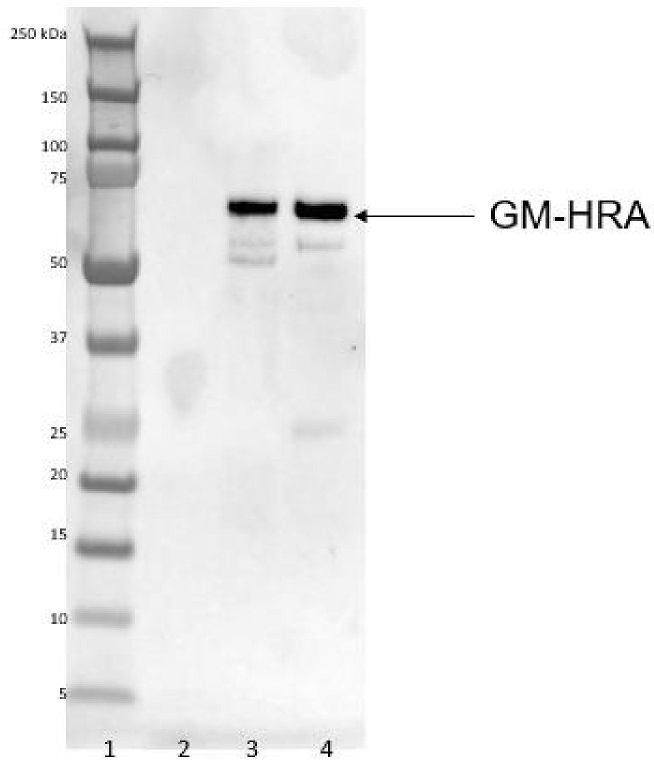
Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 58. SDS-PAGE Analysis of the GM-HRA Protein
Coomassie blue staining of the SDS-PAGE gel demonstrated the protein migrated as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived GM-HRA protein (Lane 3) and the microbially-derived GM-HRA protein (Lane 4).

Western Blot Analysis

Samples of the COR23134 soybean-derived GM-HRA protein and the microbially derived GM-HRA protein were analyzed by Western blot. As expected, the GM-HRA proteins, derived from both COR23134 soybean and the microbial system, are immunoreactive and have the expected molecular weight (Figure 59).

Additional details regarding Western blot analytical methods are provided in Appendix H.



Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	COR23134 Soybean-Derived GM-HRA Protein
4	Microbially Derived GM-HRA Protein (10 ng)

Note: kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

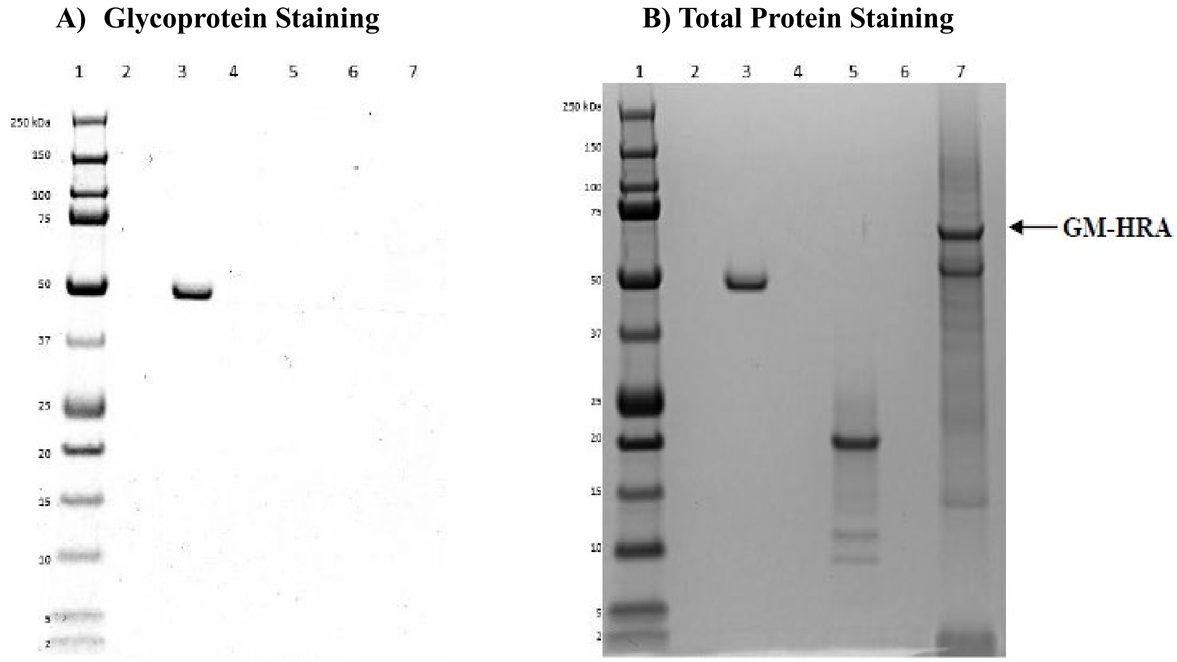
Figure 59. Western Blot Analysis of the GM-HRA Protein
Western blot analysis demonstrated that the GM-HRA protein was immunoreactive to a polyclonal antibody and visible as a predominant band consistent with the expected molecular weight for the COR23134 soybean-derived GM-HRA protein (Lane 3) and the microbially-derived GM-HRA protein (Lane 4).

Protein Glycosylation Analysis

Samples of the COR23134 soybean-derived GM-HRA protein and the microbially derived GM-HRA protein were analyzed by SDS-PAGE followed by the glycoprotein staining for glycosylation analysis. Each gel also included a positive control (horseradish peroxidase) and a negative control (soybean trypsin inhibitor). The gel was first stained using a Pierce Glycoprotein Staining Kit to visualize any glycoproteins, imaged, and then stained with the Coomassie blue reagent to visualize all protein bands.

Glycosylation was determined to be negative for both the COR23134 soybean-derived and microbially derived GM-HRA proteins (Figure 60 and Figure 61, respectively). The horseradish peroxidase positive control was clearly visible as a stained band. The soybean trypsin inhibitor negative control was not stained by the glycoprotein stain.

Additional details regarding glycosylation analytical methods are provided in Appendix H.

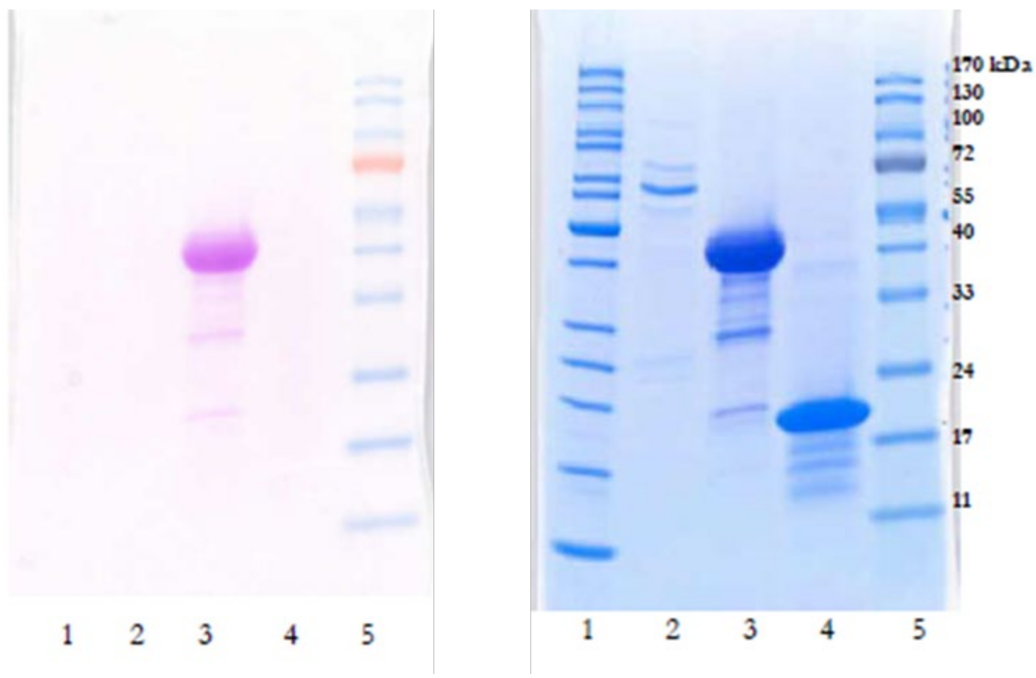


Lane	Sample Identification
1	Pre-stained Protein Molecular Weight Marker[a]
2	1X LDS/DTT Sample Buffer Blank
3	Horseradish Peroxidase Positive Control (1.0 µg)
4	1X LDS/DTT Sample Buffer Blank
5	Soybean Trypsin Inhibitor Negative Control (1.0 µg)
6	1X LDS/DTT Sample Buffer Blank
7	COR23134 Soybean-Derived GM-HRA Protein

Note: The glycoprotein gel was stained with glycoprotein staining reagent. The total protein stain gel was stained with glycoprotein staining reagent followed by staining with Coomassie blue reagent for total proteins. Kilodalton (kDa) and lithium dodecyl sulfate containing dithiothreitol (LDS/DTT).
[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 60. Glycosylation Analysis of the COR23134 Soybean-Derived GM-HRA Protein
A) Glycoprotein staining: Glycosylation was not detected for the COR23134 soybean-derived GM-HRA protein (Lane 7). The horseradish peroxidase positive control was stained (Lane 3), and the soybean trypsin inhibitor negative control was not stained (Lane 5). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the COR23134 soybean-derived GM-HRA protein (Lane 7) and both the positive (Lane 3) and negative (Lane 5) control proteins.

A) Glycoprotein Staining **B) Total Protein Staining**



Lane	Sample Identification
1	Fermentas Protein Molecular Weight Marker[a]
2	Microbially Derived GM-HRA Protein (~1 µg)
3	Horseradish Peroxidase Positive Control (~20 µg)
4	Soybean Trypsin Inhibitor Negative Control (~20 µg)
5	Fermentas Prestained Protein Molecular Weight Marker[a]

[a] Molecular weight markers were included to provide a visual verification of protein migration.

Figure 61. Glycosylation Analysis of the Microbially Derived GM-HRA Protein
A) Glycoprotein staining: Glycosylation was not detected for the microbially derived GM-HRA protein (Lane 2). The horseradish peroxidase positive control was stained (Lane 3), and the soybean trypsin inhibitor negative control was not stained (Lane 4). **B)** Total protein staining: Subsequent Coomassie blue staining of the same gel for total proteins detected the microbially derived GM-HRA protein (Lane 2) and both the positive (Lane 3) and negative (Lane 4) control proteins.

Mass Spectrometry Peptide Mapping Analysis

Samples of the COR23134 soybean-derived GM-HRA protein and the microbially derived GM-HRA protein were analyzed by SDS-PAGE. The gel was stained with Coomassie blue reagent, and the bands containing the GM-HRA protein were excised for each sample.

The excised COR23134 soybean-derived GM-HRA protein bands were digested with trypsin or chymotrypsin. Digested samples were analyzed using liquid chromatography-mass spectrometry (LC-MS). The resulting MS data were used to search and match the peptides from the GM-HRA protein sequence, and the combined sequence coverage was calculated. The combined sequence coverage of the identified tryptic and chymotryptic peptides for the COR23134 soybean-derived GM-HRA protein accounts for 92.2% (557/604) of the amino acid sequence (Figure 62).

The excised microbially derived GM-HRA protein band was digested with trypsin and analyzed by Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS). The sequence coverage of the identified tryptic peptides for the microbially derived GM-HRA protein accounts for 38.3% (232/605) of the amino acid sequence (Figure 63).

Additional details regarding peptide mapping analytical methods are provided in Appendix H.

1 **SISKPPTAAP FTKEAPTTEP FVSRFASGEP RKGADILVEA LERQGVTTVF**
51 **AYPGGASMEI HQALTR**SAAI RNVLPR**HEQG GVFAAEGYAR SSGLPGVCIA**
101 **TSGPGATNLV SGLADALMDS VPVVAITGQV ARRMIGTDAF QETPIVEVSR**
151 **SITKHNYLIL DVDDIPRVVA EAFFVATSGR PGPVLIDIPK DVQQQLAVPN**
201 **WDEPVNLPGY LARLPRPPAE AQLEHIVRLI MEAQKPVLYV GGGSLNSSAE**
251 **LRR**F**VELTGI PVASTLMGLG TFPIGDEYSL QMLGMHGTVY** ANYAVDNSDL
301 LLAF**GVRFDD RVTGKLEAFA SRAKIVHIDI DSAEIGKNKQ AHVSVCADLK**
351 **LALKGINMIL EEKGVEGKFD LGGWREEINV QKHKFPLGYK TFQDAISPQH**
401 **AIEVLDELTN GDAIVSTGVG QHQMWAAQFY K**YKRPR**QWLT SGGLGAMGFG**
451 **LPAAIGAAVA NPGAVVVDID GDGSFIMNVQ ELATIRVENL PVKILLLNNQ**
501 **HLGMVVQLED RFY**KSNR**AHT YLGDPSSESE IFPNMLKFAD ACGIPAAR**VT
551 KKEELRAAIQ R**MLDTPGPYL LDVIVPHQEH VLPMIPSNGS FKDVITEGDG**
601 **RTRY**

Red type	Bold red type indicates soybean-derived GM-HRA peptides identified using LC-MS analysis against the expected GM-HRA protein sequence.
Amino acid residue abbreviations	alanine (A), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

Figure 62. Identified Tryptic and Chymotryptic Peptide Amino Acid Sequence of the COR23134 Soybean-Derived GM-HRA Protein Using LC-MS Analysis

```
GSISKPPTAAPFTKEAPTTEPEVSRFASGEPRKGADILVEALERQGVTTVFAYPGGASME
IHQALTRSAAIRNVLPRHEQGGVFAAEGYARSSGLPGVCIATSGPGATNLVSGLADALMD
SVPVVAITGQVARRMIGTDAFQETPIVEVSRSITKHNYLILDVDDIPRVVAEAFFVATSG
RPGPVLIDIPKDVQQQLAVPNWDEPVNLPGYLARLPRPPAEAQLEHIVRLIMEAQKPVLY
VGGGSLNSSAELRRFVELTGIPVASTLMGLGTFPIGDEYSLQMLGMHGTVYANYAVDNSD
LLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKQAHVSVCADLKLALKGINMI
LEEKGVEGKFDLGGWREEINVQKHKFPLGYKTFQDAISPQHAIEVLDELTNGDAIVSTGV
GQHQMWAAQFYKYKRPRQWLTSGGLGAMGFGLPAAIGAAVANPGAVVVDIDGDGSFIMNV
QELATIRVENLPVKILLLNNQHLGMVVQLEDRFYKSNRAHTYLGDPSSESEIFPNMLKFA
DACGIPAARVTKKEELRAAIQRMLDTPGPYLLDVIVPHQEHVLPMIPSNGSFKDVITEGD
GRTRY
```

Note: Matching peptides are shaded. The N-terminal glycine (G) is an additional amino acid residue remaining from thrombin cleavage of the His-T7 fusion tag used for production and purification of the microbially derived GM-HRA protein.

Figure 63. Identified Tryptic Peptide Amino Acid Sequence of Microbially Derived GM-HRA Protein Using MALDI-MS Analysis

N-Terminal Amino Acid Sequence Analysis

The Edman sequencing analysis of the COR23134 soybean-derived GM-HRA protein sample identified an N-terminal sequence (SISKPPTAAP), matching the amino acid residues 1-10 of the expected mature GM-HRA protein sequence.

The N-terminal peptide for the microbially derived GM-HRA protein was identified by MALD-MS as GSISKPPTAAP from the trypsin digestion, matching the amino acid residues 1-10 of the expected microbially derived GM-HRA protein sequence. This N-terminal glycine (G) identified in the microbially derived GM-HRA protein is an additional amino acid residue remaining from thrombin cleavage of the His-T7 fusion tag used for protein production and purification.

Additional details regarding N-terminal amino acid sequencing analytical methods are provided in Appendix H.

Functional Activity Assay

The functional activity of the microbially derived GM-HRA protein was evaluated by conducting a spectrophotometric assay for enzymatic measurement of acetolactate synthase (ALS) in the presence or absence of an ALS-inhibiting herbicide, chlorsulfuron.

The ALS protein catalyzes the formation of acetolactate from pyruvate. The spectrophotometric assay for determining ALS activity involves an indirect detection of the enzyme product, acetolactate. Following incubation of the enzyme with the substrate (pyruvate), the assay involves the conversion of the end product (acetolactate) to acetoin by decarboxylation with sulfuric acid and high temperature. Acetoin produced is detected by formation of a creatine and α-naphtol complex and measuring the optical density (OD) at 530 nm.

The results demonstrated that the microbially derived GM-HRA protein had equivalent ALS enzymatic activity both in the presence and absence of 100 ng/ml chlorsulfuron inhibitor (Figure 64).

Additional details regarding functional activity analytical methods are provided in Appendix H.

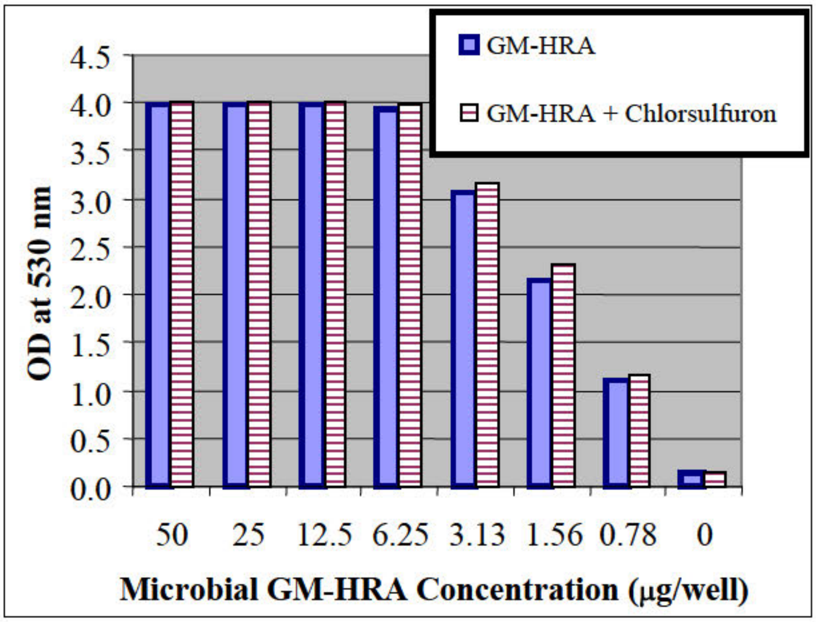


Figure 64. Enzymatic Activity Assay for the Microbially Derived GM-HRA Protein

Allergenicity and Toxicity Analyses of the GM-HRA Protein

A weight-of-evidence approach was applied to determine the allergenic and toxic potential of the GM-HRA protein expressed in COR23134 soybean, including an assessment of the following: the history of safe use of the source organism, a bioinformatic comparison of the amino acid sequence of the GM-HRA protein to known or putative allergen and protein toxin sequences, an evaluation of the heat lability of the GM-HRA protein using a spectrophotometric enzyme activity assay, evaluations of the stability of the GM-HRA protein using *in vitro* gastric and intestinal digestion models, determination of the GM-HRA protein glycosylation status, and an evaluation of acute toxicity in mice following oral exposure to the GM-HRA protein.

Source Organism of the GM-HRA Protein

The GM-HRA protein expressed in COR23134 soybean is encoded by the *gm-hra_1* gene, a modified acetolactate synthase gene, from *Glycine max*.

Soybean is the world's leading oilseed crop with a long history of use (OECD, 2000; OECD, 2012). Historical and geographical evidence suggests that soybeans were first domesticated in the eastern half of China between the 17th and 11th century B.C. Soybeans were first introduced into the United States, now a major producer, in 1765. Today, soybeans are grown as a commercial crop in over 35 countries worldwide (OECD, 2000). Soybeans have a multitude of uses in the human food, animal feed, and industrial sectors, and represent one of the major sources of edible vegetable oil and of proteins for livestock feed use (CFIA, 2021; OECD, 2000).

Bioinformatic Analysis of GM-HRA Homology to Known and Putative Allergens

Assessing newly expressed proteins for potential cross-reactivity with known and putative allergens is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). A bioinformatic assessment of the GM-HRA protein sequence (651 amino acids [aa]) for potential cross-reactivity with allergens was conducted by following established international criteria (Codex Alimentarius Commission, 2009; FAO/WHO, 2001).

Two separate searches for the GM-HRA protein sequence were performed using the Comprehensive Protein Allergen Resource (COMPARE) 2023 database (January 26, 2023 (van Ree et al., 2021)). This peer-reviewed database is a collaborative effort of the Health and Environmental Sciences Institute (HESI) Protein Allergens, Toxins, and Bioinformatics (PATB) Committee and contains 2,631 sequences.

The first search was the sliding 80-mer window search, accomplished with an internally developed Perl script running FASTA v35.04 (Pearson and Lipman, 1988) with an E-score cutoff set to 100. In a sliding window search, each sequentially overlapping 80 aa sub-sequence of the overall GM-HRA protein sequence is used as a query against the COMPARE allergen database sequences. The script examined all alignments generated from the query and reported any possessing > 35% identity over an alignment length of ≥ 80 aa. Additionally, the script rescaled the percent identity to an 80-mer window for any alignments possessing an alignment length shorter than 80 aa; the number of identities in these alignments would be divided by 80, then multiplied by 100, and would report any alignment possessing an adjusted percent identity > 35%.

The second search used EMBOSS fuzzpro v6.6.0 (Rice et al., 2000) to identify any eight or greater contiguous identical amino acid matches between the GM-HRA protein sequence and the COMPARE allergen sequences.

Results of the search of the GM-HRA protein sequence against the COMPARE allergen database sequences found no alignments that were a length of 80 aa or greater with a sequence identity of > 35% and no alignments shorter than 80 aa with a sequence identity > 35% when normalized to an 80-mer window. No contiguous 8-residue exact matches between the GM-HRA protein sequence and the allergen sequences were identified in the second search. Collectively, these data indicate that no allergenicity concern arose from the bioinformatics assessment of the GM-HRA protein.

Bioinformatics evaluation of the GM-HRA protein sequence did not generate biologically relevant amino acid sequence similarities to allergens that are harmful to humans or animals.

Bioinformatic Analysis of GM-HRA Homology to Known and Putative Protein Toxins

Assessing newly expressed proteins for potential sequence similarity with protein toxin is a critical part of the weight-of-evidence approach used to evaluate the safety of these proteins in genetically modified plant products (Codex Alimentarius Commission, 2009). The potential toxicity of the GM-HRA protein was assessed by comparison of its sequence 1) an internal toxin database, and 2) the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein database.

The internal toxin database is a subset of sequences found in UniProtKB/Swiss-Prot (The UniProt Consortium, 2023). UniProtKB/Swiss-Prot is a curated database of non-redundant proteins containing functional information for over 550,000 sequences. To produce the internal toxin database, the proteins in UniProtKB/Swiss Prot are filtered for molecular function by keywords that could imply toxicity or adverse health effects (e.g., toxin, hemagglutinin, vasoactive). The internal toxin database is updated annually and contains 8,858 sequences.

The search between the GM-HRA protein sequence and protein sequences in the internal toxin database was conducted using BLASTP v2.10.0+ with an E-value set to 10^{-4}. No alignments were returned between the GM-HRA protein sequence and any protein sequence in the internal toxin database. Therefore, no toxicity concern arose from the bioinformatics assessment of the GM-HRA protein.

The BLASTP search of the GM-HRA protein against the NCBI nr protein database returned the maximum number of alignment descriptions/sequence alignments, 500/250, all with an E-value of 0. As expected, the BLASTP search returned alignments to various acetolactate synthase proteins. None of the accessions returned by the BLASTP search are proteins known to be toxic to humans or animals.

Bioinformatics evaluation of the GM-HRA protein sequence did not generate biologically relevant amino acid sequence similarities to protein toxins that are harmful to humans or animals.

Thermolability Analysis

Thermal stability of the GM-HRA protein was characterized by determining the activity of the heat-treated GM-HRA protein using the acetolactate synthase (ALS) activity assay as stated previously in the characterization section for GM-HRA in the presence and absence of chlorsulfuron.

The GM-HRA protein samples in individual wells were heated for 15 minutes at a designated temperature ranging from 36-60 °C with 2 °C increments in a gradient thermocycler and analyzed for enzymatic activity. The results demonstrated that the GM-HRA enzyme, either in the absence or presence of chlorsulfuron, was inactivated when incubated for 15 minutes at 50 °C (Figure 65).

Additional details regarding thermolability analytical methods are provided in Appendix H.

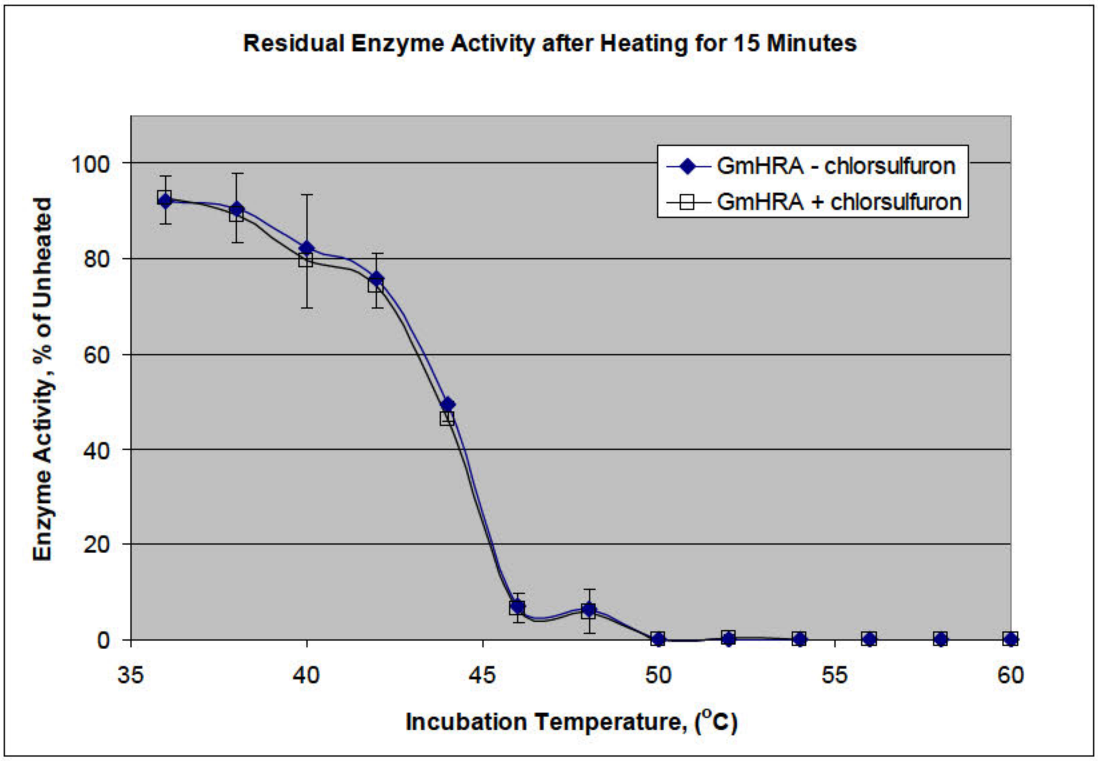


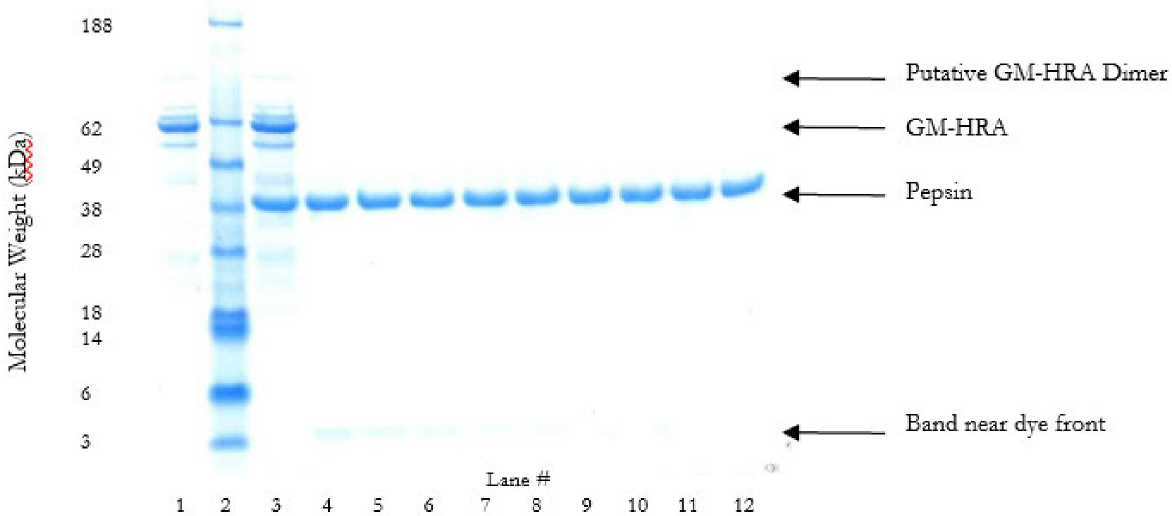
Figure 65. Graph Illustrating the Residual Enzyme Activity versus the Incubation Temperature

Digestibility Analysis with Simulated Gastric Fluid (SGF)

Simulated gastric fluid (SGF) containing pepsin at pH ~1.2 was used to assess the susceptibility of the GM-HRA protein to proteolytic digestion by pepsin *in vitro*. The GM-HRA protein was incubated in SGF for 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. A positive control (bovine serum albumin) and a negative control (β-lactoglobulin) were included in the assay and were incubated in SGF for 0, 1, and 60 minutes. After incubation in SGF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SGF digestibility results showed that the GM-HRA protein migrating at approximately 65 kDa was digested within 0.5 minutes in SGF as demonstrated by both SDS-PAGE and western blot analysis (Figure 66 and Figure 67, respectively).

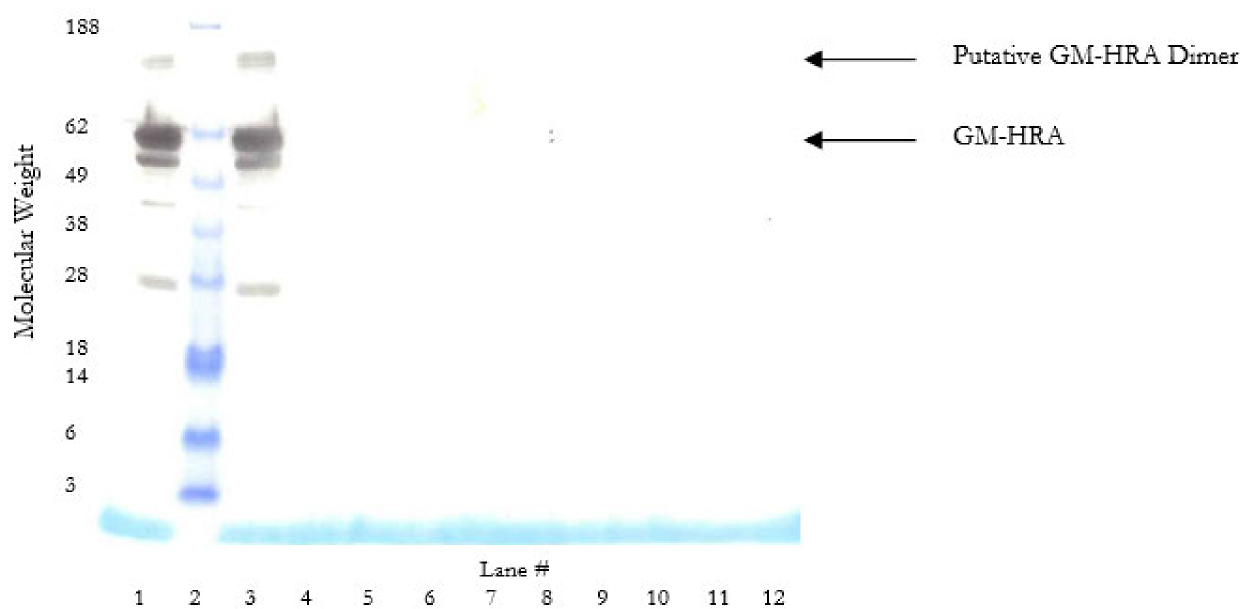
Additional details regarding SGF analytical methods are provided in Appendix H.



Lane	Sample Descriptions
1	GM-HRA protein (~2.3 µg) in water (no SGF), Time 0
2	SeeBlue protein molecular weight marker
3	GM-HRA protein (~2.3 µg) in SGF, Time 0
4	GM-HRA protein (~2.3 µg) in SGF, 0.5 minutes
5	GM-HRA protein (~2.3 µg) in SGF, 1 minutes
6	GM-HRA protein (~2.3 µg) in SGF, 2 minutes
7	GM-HRA protein (~2.3 µg) in SGF, 5 minutes
8	GM-HRA protein (~2.3 µg) in SGF, 10 minutes
9	GM-HRA protein (~2.3 µg) in SGF, 20 minutes
10	GM-HRA protein (~2.3 µg) in SGF, 30 minutes
11	GM-HRA protein (~2.3 µg) in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: simulated gastric fluid (SGF)

Figure 66. SDS-PAGE Analysis of the GM-HRA Protein in Simulated Gastric Fluid Digestion Time Course



Lane	Sample Descriptions
1	GM-HRA protein (~2.3 µg) in water (no SGF), Time 0
2	SeeBlue protein molecular weight marker
3	GM-HRA protein (~2.3 µg) in SGF, Time 0
4	GM-HRA protein (~2.3 µg) in SGF, 0.5 minutes
5	GM-HRA protein (~2.3 µg) in SGF, 1 minutes
6	GM-HRA protein (~2.3 µg) in SGF, 2 minutes
7	GM-HRA protein (~2.3 µg) in SGF, 5 minutes
8	GM-HRA protein (~2.3 µg) in SGF, 10 minutes
9	GM-HRA protein (~2.3 µg) in SGF, 20 minutes
10	GM-HRA protein (~2.3 µg) in SGF, 30 minutes
11	GM-HRA protein (~2.3 µg) in SGF, 60 minutes
12	SGF Control, 60 minutes

Note: simulated gastric fluid (SGF)

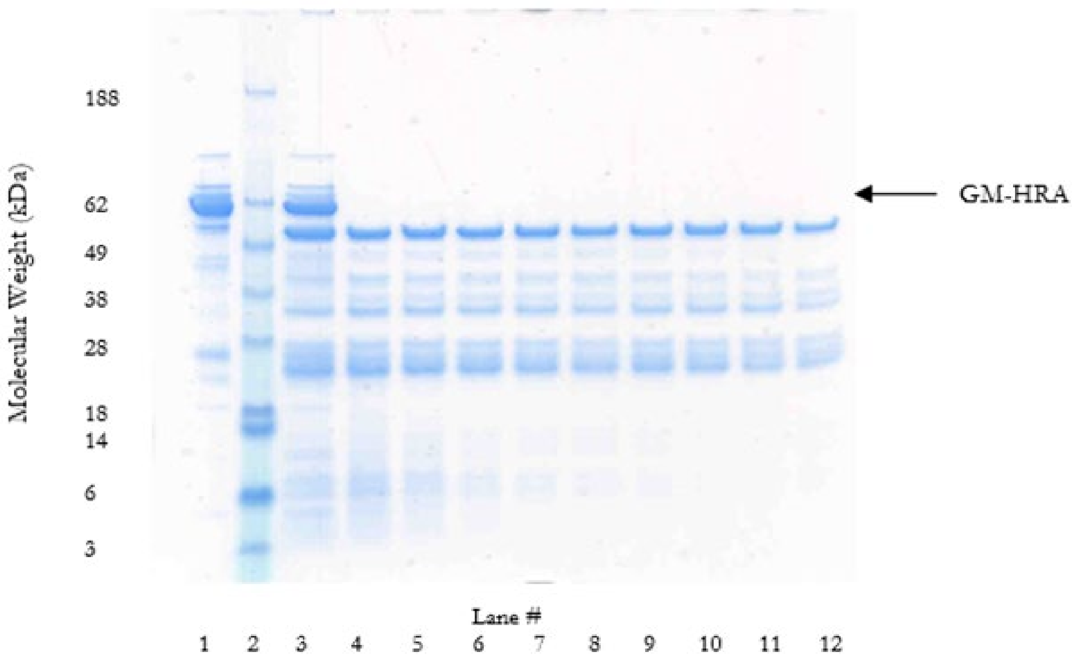
Figure 67. Western Blot Analysis of the GM-HRA Protein in Simulated Gastric Fluid Digestion Time Course

Digestibility Analysis with Simulated Intestinal Fluid (SIF)

Simulated intestinal fluid (SIF) containing pancreatin at ~pH 7.5 was used to assess the susceptibility of the GM-HRA protein to proteolytic digestion by pancreatin in vitro. The GM-HRA protein was incubated in SIF for 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. A positive control (β-lactoglobulin) and a negative control (bovine serum albumin) were included in the assay and were incubated in SIF for 0, 1, and 60 minutes. After incubation in SIF, the samples were analyzed by SDS-PAGE. Coomassie-based stain and western blot were used to detect protein bands.

The SIF digestibility results showed that the GM-HRA protein migrating at approximately 65 kDa was digested within 0.5 minutes in SIF as demonstrated by SDS-PAGE (Figure 68). For western blot analysis, the GM-HRA protein band remained detectable at the 0.5-minute time point in SIF but was not detected at the 1-minute time point (Figure 69).

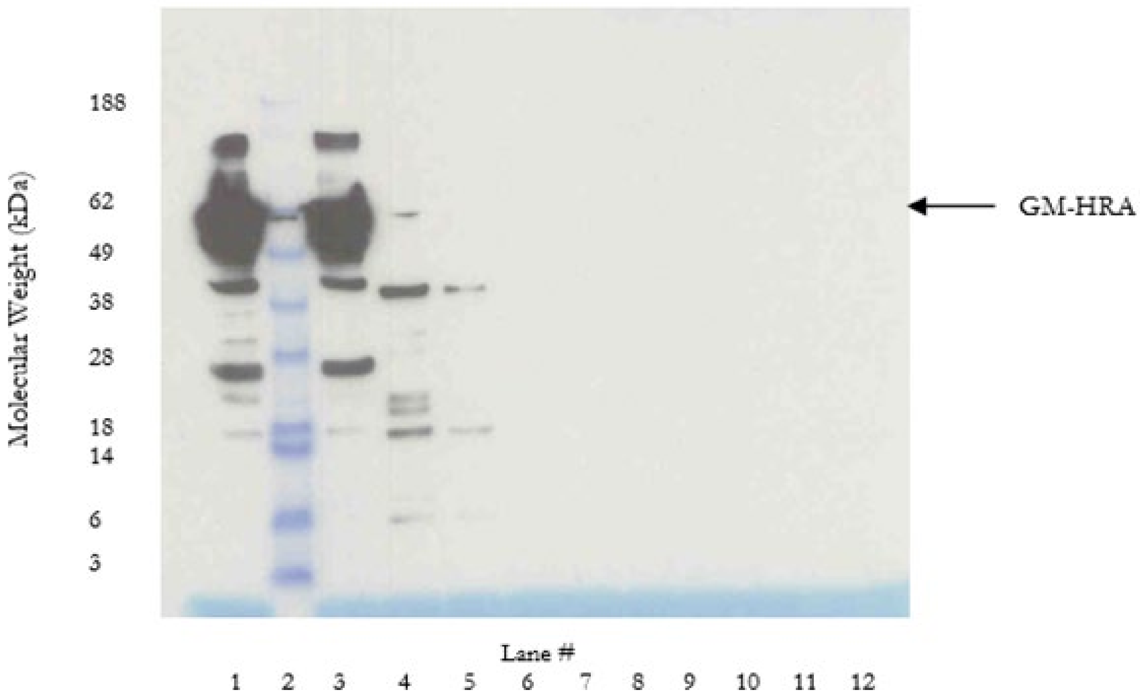
Additional details regarding SIF analytical methods are provided in Appendix H.



Lane	Sample Descriptions
1	GM-HRA protein (~3.3 µg) in water (no SGF), Time 0
2	SeeBlue protein molecular weight marker
3	GM-HRA protein (~3.3 µg) in SIF, Time 0
4	GM-HRA protein (~3.3 µg) in SIF, 0.5 minutes
5	GM-HRA protein (~3.3 µg) in SIF, 1 minutes
6	GM-HRA protein (~3.3 µg) in SIF, 2 minutes
7	GM-HRA protein (~3.3 µg) in SIF, 5 minutes
8	GM-HRA protein (~3.3 µg) in SIF, 10 minutes
9	GM-HRA protein (~3.3 µg) in SIF, 20 minutes
10	GM-HRA protein (~3.3 µg) in SIF, 30 minutes
11	GM-HRA protein (~3.3 µg) in SIF, 60 minutes
12	SGF Control, 60 minutes

Note: simulated intestinal fluid (SIF)

Figure 68. SDS-PAGE Analysis of the GM-HRA Protein in Simulated Intestinal Fluid Digestion Time Course



Lane	Sample Descriptions
1	GM-HRA protein (~3.3 µg) in water (no SIF), Time 0
2	SeeBlue protein molecular weight marker
3	GM-HRA protein (~3.3 µg) in SIF, Time 0
4	GM-HRA protein (~3.3 µg) in SIF, 0.5 minutes
5	GM-HRA protein (~3.3 µg) in SIF, 1 minutes
6	GM-HRA protein (~3.3 µg) in SIF, 2 minutes
7	GM-HRA protein (~3.3 µg) in SIF, 5 minutes
8	GM-HRA protein (~3.3 µg) in SIF, 10 minutes
9	GM-HRA protein (~3.3 µg) in SIF, 20 minutes
10	GM-HRA protein (~3.3 µg) in SIF, 30 minutes
11	GM-HRA protein (~3.3 µg) in SIF, 60 minutes
12	SIF Control, 60 minutes

Note: simulated intestinal fluid (SIF)

Figure 69. Western Blot Analysis of the GM-HRA Protein in Simulated Intestinal Fluid Digestion Time Course

Protein Glycosylation Analysis

As stated previously in the characterization section, the results from glycoprotein staining analysis confirmed the absence of glycosylation for the COR23134 soybean-derived GM-HRA and the microbially derived GM-HRA proteins.

Evaluation of the Acute Toxicity of the GM-HRA Protein

A study was conducted to evaluate the potential acute toxicity of the test substance, GM-HRA, in [redacted] mice following oral exposure at a dose of 2000 mg/kg body weight. This corresponded to a per-animal exposure of at least 436, but less than 582, mg/kg of the microbially derived GM-HRA protein. The GM-HRA protein and bovine serum albumin (BSA) protein were each reconstituted in deionized water. Vehicle control, BSA comparative control, and the GM-HRA test substance formulations were administered in a single oral dose by gavage; the BSA comparative control was administered at an equivalent target dose to that of the GM-HRA protein. The mice were fasted approximately 4.4-4.8 hours prior to dosing.

Body weights were evaluated on test day 0 (prior to fasting and shortly prior to administration of the first dose), 1, 2, 4, 7, and 14. The animals were observed for clinical signs of toxicity while handled before and after fasting, once during the first 30 minutes after dosing, at least 2 more times within 4 hours after dosing, and daily thereafter. Observations for mortality and signs of illness, injury, or abnormal behavior were conducted twice daily. On test day 14, all surviving mice were euthanized and given a gross pathological examination.

All animals survived to scheduled euthanasia. No clinical signs of systemic toxicity or test substance-related body weight losses were observed in any mice. No gross lesions were observed.

Under the conditions of this study, oral exposure via intragastric administration of the GM-HRA protein to male and female mice at 2000 mg/kg body weight did not result in test substance-related clinical signs of toxicity, body weight losses, gross lesions, or mortality.

Conclusion on Safety of the GM-HRA Protein in COR23134 Soybean

In conclusion, protein characterization results via SDS-PAGE, western blot, glycosylation analysis, mass spectrometry peptide mapping analysis, and N-terminal amino acid sequence analysis have demonstrated that the GM-HRA protein derived from COR23134 soybean and the microbial system has the expected molecular weight, immunoreactivity, and amino acid sequence, and is not glycosylated. Characterization of the microbially derived GM-HRA protein demonstrated that it is an appropriate test substance for use in safety studies.

The allergenic potential of the GM-HRA protein was evaluated by assessing the GM-HRA protein source organism and history of safe use, a bioinformatic comparison of the amino acid sequence of the GM-HRA protein with known and putative allergen sequences, an evaluation of the heat lability of the GM-HRA protein using a spectrophotometric enzyme activity assay, evaluations of the stability of the GM-HRA protein using *in vitro* gastric and intestinal digestion models, and determination of the GM-HRA protein glycosylation status. The toxicity potential of the GM-HRA protein was evaluated by a bioinformatic comparison of the GM-HRA amino acid sequence to known and putative protein toxins and by an acute toxicity in mice following oral exposure to the GM-HRA protein.

The bioinformatic comparison of the GM-HRA protein sequence to known and putative allergen and protein toxin sequences showed that the GM-HRA protein is unlikely to be allergenic or toxic for humans or animals. The GM-HRA protein was digested within 0.5 minutes in SGF and 1 minute in SIF. The GM-HRA protein was not glycosylated. The GM-HRA protein, either in the absence or presence of an ALS-inhibitor chlorsulfuron, was inactivated when incubated for 15 minutes at 50 °C. The acute oral toxicity assessment in mice determined the LD_{50} of the GM-HRA protein to be greater than 2000 mg/kg. These data support the conclusion that COR23134 soybean, expressing the GM-HRA protein, is as safe as conventional soybean for the food and feed supply.

Based on this weight of evidence, consumption of the GM-HRA protein from COR23134 soybean is unlikely to cause an adverse effect on humans or animals.

B.3 Other (non-protein substances)

There are no other new substances associated with COR23134 soybean.

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

Not applicable. The GM-HRA enzyme can function in the presence of the ALS-inhibiting class of herbicides, thereby conferring a degree of tolerance to those herbicides. However, the transcript of this gene was used in COR23134 as a selectable marker to identify genetically modified plants during the event development.

B.5 Compositional analyses of the food produced using gene technology

Trait Expression Assessment

The expression levels of the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins were evaluated in COR23134 soybean.

COR23134 soybean plants were grown during the 2022 growing season at six sites in commercial soybean growing regions of the United States and Canada. A randomized complete block design with four blocks was utilized at each site.

The following samples were collected: leaf (V5, R1, and R3 growth stages), flowers (R1-R2 growth stage), root (R3 growth stage), forage (R3 growth stage), and seed (R8 growth stage). Samples were analyzed for the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA protein concentrations using quantitative enzyme linked immunosorbent assay (ELISA) methods.

Concentration results (means, ranges, and standard deviations) are summarized across sites in Table 16 to Table 19 for Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA protein, respectively. Individual sample results below the LLOQ were assigned a value equal to the LLOQ for calculation purposes.

Additional details regarding analytical methods and calculations for trait expression analysis are provided in Appendix I.

Table 16. Across-Site Summary of the Cry1B.34.1 Protein Concentrations in COR23134 Soybean

Tissue (Growth Stage)	ng Cry1B.34.1/mg Tissue Dry Weight				Number of Samples <LLOQ/ Number of Samples Reported
	Mean	Range	Standard Deviation	Sample LLOQ[a]	
COR23134 Soybean					
Leaf (V5)	460	190 - 960	220	0.14	0/24
Leaf (R1)	310	150 - 780	170	0.14	0/24
Leaf (R3)	170	84 - 270	44	0.14	0/24
Flowers (R1 - R2)	260	200 - 320	28	0.28	0/24
Root (R3)	77	15 - 170	39	0.069	0/24
Forage (R3)	150	75 - 180	22	0.069	0/23[b]
Seed (R8)	170	140 - 210	22	0.14	0/24

Note: Growth stages (Pedersen, 2004).
[a] Lower limit of quantification (LLOQ) in ng/mg tissue dry weight.
[b] One forage sample was confirmed negative for the event of interest by polymerase chain reaction (PCR) analysis.

Table 17. Across-Site Summary of the Cry1B.61.1 Protein Concentrations in COR23134 Soybean

Tissue (Growth Stage)	ng Cry1B.61.1/mg Tissue Dry Weight				Number of Samples <LLOQ/ Number of Samples Reported
	Mean	Range	Standard Deviation	Sample LLOQ[a]	
COR23134 Soybean					
Leaf (V5)	400	160 - 720	120	0.28	0/24
Leaf (R1)	490	250 - 1300	220	0.28	0/24
Leaf (R3)	480	260 - 780	130	0.28	0/24
Flowers (R1 - R2)	150	120 - 190	19	0.55	0/24
Root (R3)	0.34[b]	<0.14 - 0.93	0.18[b]	0.14	1/24
Forage (R3)	200	69 - 390	72	0.14	0/23[c]
Seed (R8)	15	11 - 22	2.6	0.28	0/24

Note: Growth stages (Pedersen, 2004).
[a] Lower limit of quantification (LLOQ) in ng/mg tissue dry weight.
[b] One sample result was below the LLOQ. A value equal to the LLOQ value was assigned to that sample to calculate the mean and standard deviation.
[c] One forage sample was confirmed negative for the event of interest by polymerase chain reaction (PCR) analysis.

Table 18. Across-Site Summary of the IPD083Cb Protein Concentrations in COR23134 Soybean

Tissue (Growth Stage)	ng IPD083Cb/mg Tissue Dry Weight				Number of Samples <LLOQ/ Number of Samples Reported
	Mean	Range	Standard Deviation	Sample LLOQ[a]	
COR23134 Soybean					
Leaf (V5)	65	48 - 110	17	1.2	0/24
Leaf (R1)	71	48 - 96	14	1.2	0/24
Leaf (R3)	83	59 - 130	19	1.2	0/24
Flowers (R1-R2)	69	56 - 85	8.7	2.4	0/24
Root (R3)	21	12 - 33	5.6	0.60	0/24
Forage (R3)	59	39 - 84	10	0.60	0/23[b]
Seed (R8)	14	13 - 18	1.4	1.2	0/24

Note: Growth stages (Pedersen, 2004).
[a] Lower limit of quantification (LLOQ) in ng/mg tissue dry weight.
[b] One forage sample was confirmed negative for the event of interest by polymerase chain reaction (PCR) analysis.

Table 19. Across-Site Summary of the GM-HRA Protein Concentrations in COR23134 Soybean

Tissue (Growth Stage)	ng GM-HRA/mg Tissue Dry Weight				Number of Samples <LLOQ/ Number of Samples Reported
	Mean	Range	Standard Deviation	Sample LLOQ[a]	
COR23134 Soybean					
Leaf (V5)	7.6[b]	<2.2 - 17	5.0[b]	2.2	1/24
Leaf (R1)	4.4[b]	<2.2 - 11	2.6[b]	2.2	5/24
Leaf (R3)	2.5[b]	<2.2 - 4.4	0.58[b]	2.2	15/24
Flowers (R1 - R2)	2.3	1.4 - 2.9	0.37	1.1	0/24
Root (R3)	1.0	0.60 - 1.5	0.25	0.27	0/24
Forage (R3)	1.4[b]	<1.1 - 2.2	0.29[b]	1.1	3/23[c]
Seed (R8)	0.73[b]	<0.54 - 1.7	0.28[b]	0.54	8/24

Note: Growth stages (Pedersen, 2004).
[a] Lower limit of quantification (LLOQ) in ng/mg tissue dry weight.
[b] Some, but not all, sample results were below the LLOQ. A value equal to the LLOQ value was assigned to those samples to calculate the mean and standard deviation.
[c] One forage sample was confirmed negative for the event of interest by polymerase chain reaction (PCR) analysis.

Nutrient Composition Assessment

An assessment of the composition of a GM product compared to that of a conventional non-GM comparator with a history of safe use in food and feed is an important part of the weight-of-evidence approach used to evaluate the safety of genetically modified plant products (Codex Alimentarius Commission, 2008; OECD, 1993). Compositional assessments of COR23134 soybean were evaluated in comparison to concurrently grown non-GM, near-isoline soybean (referred to as control soybean) to identify statistical differences, and subsequently were evaluated in the context of biological variation established from multiple sources of composition data from non-GM commercial soybean (referred to as reference soybean).

Forage (R3 growth stage) and seed (R8 growth stage) samples were collected during the 2022 growing season at eight sites in commercial soybean-growing regions of the United States and Canada. A randomized complete block design with four blocks was utilized at each site. Each block included COR23134 soybean, control soybean, and four non-GM commercial soybean reference lines.

The samples were assessed for key nutritional components. Proximate and fiber analytes were assessed in the forage samples (eight analytes total), and seed samples were assessed for proximate, fiber, fatty acid, amino acid, mineral, vitamin, isoflavone, and anti-nutrient analytes. The analytes included in the compositional assessment were selected based on the OECD consensus document on compositional considerations for new varieties of soybean (OECD, 2012). Procedures and methods for nutrient composition analyses of soybean forage and seed were conducted in accordance with the requirements for the U.S. EPA Good Laboratory Practice (GLP) Standards, 40 CFR Part 160. The analytical procedures used were validated methods, with the majority based on methods published by AOAC International, AACC (American Association of Cereal Chemists), and AOCS (American Oil Chemists' Society).

Statistical analyses were conducted to evaluate and compare the nutrient composition of COR23134 soybean and the control soybean. Across-site comparisons were conducted for a total of 82 analytes; 76 analytes were analyzed using mixed model analysis and six analytes did not meet criteria for sufficient quantities of observations above the LLOQ and therefore were not included in the mixed model analyses. Two of those six analytes were subjected to Fisher's exact test, and no statistical analysis was conducted on the remaining four analytes as all data values were below the LLOQ. For a given analyte in the mixed model analysis, if a statistical difference (P-value < 0.05) was observed between COR23134 soybean and the control soybean, the False Discovery Rate (FDR)-adjusted P-value was examined. In cases where the raw P-value indicated a significant difference but the FDR-adjusted P-value was non-significant, it was concluded that the difference was likely a false positive. Additionally, three reference ranges representing the non-GM soybean population with a history of safe use (i.e., tolerance interval, literature range, and in-study reference range) were utilized to evaluate statistical differences in the context of biological variation. If the range of measured values for COR23134 soybean for that analyte fell within at least one of the reference ranges, then this analyte would be considered comparable to conventional soybean.

The outcome of the nutrient composition assessment is provided in Table 20. Nutrient composition analysis results are provided in Table 21 to Table 27. No statistically significant differences were identified between COR23134 soybean and the control soybean in forage for the eight analytes evaluated in the across-site analysis via mixed model. No statistically significant differences were observed between COR23134 soybean and the control soybean for 47 of the 70 analytes in seed that went through across-site analysis via either mixed model analysis or Fisher's exact test. For the remaining 23 analytes in seed (crude protein, ash, carbohydrates, myristic acid [C14:0], palmitic acid [C16:0], heptadecanoic acid [C17:0], heptadecenoic acid [C17:1], linolenic acid [C18:3], arginine, glutamic acid, glycine, isoleucine, leucine, proline, serine, valine, calcium, magnesium, phosphorus, vitamin B5, α-tocopherol, total daidzein equivalent, and total genistein equivalent) evaluated in the across-site analysis, a statistically significant difference (non-adjusted P-value < 0.05), was observed between COR23134 soybean and the control soybean. For each of these analytes the range of values for COR23134 soybean were within one or more of the references ranges (i.e., tolerance interval, literature range, and in-study reference range) representing the non-GM soybean population with a history of safe use, indicating that COR23134 soybean is within the range of normal variation for each of these analytes and the statistical differences are not biologically meaningful. Additionally, for ash, carbohydrates, linolenic acid (C18:3), arginine, glycine, proline, serine, valine, magnesium, phosphorus, vitamin B5, α-tocopherol, and total daidzein equivalent, the FDR-adjusted P-values were non-significant, indicating that the identified statistical differences were likely false positives.

The results of the nutrient composition assessment demonstrate that the nutrient composition of forage and grain derived from COR23134 soybean is comparable to that of conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Additional details regarding methods for nutrient composition and statistical analyses are provided in Appendix J.

Table 20. Outcome of the Nutrient Composition Assessment for COR23134 Soybean

Subgroup	No Statistical Difference Identified	Statistical Difference Identified					Not Included in Statistical Analysis (All Data Values Below the Lower Limit of Quantification)
		All Data Values Within Tolerance Interval	One or More Data Values Outside Tolerance Interval, or Tolerance Interval Not Available			Adjusted P-Value < 0.05	
			All Data Values Within Literature Range	One or More Data Values Outside Literature Range, or Literature Range Not Available			
				All Data Values Within Reference Data Range	One or More Data Values Outside Reference Data Range		
Forage (R3 Growth Stage)							
Proximate and Fiber Composition	Moisture (%) Crude Protein Crude Fat Crude Fiber ADF NDF Ash Carbohydrates	--	--	--	--	--	--
Seed (R8 Growth Stage)							
Proximate and Fiber Composition	Moisture (%) Crude Fat Crude Fiber ADF NDF	Crude Protein Ash Carbohydrates	--	--	--	Crude Protein	--
Fatty Acid Composition	Palmitoleic Acid (C16:1) Heptadecadienoic Acid (C17:2) Stearic Acid (C18:0) Oleic Acid (C18:1) Linoleic Acid (C18:2) Isomer 1 of Nonadecenoic Acid (C19:1,1) Isomer 2 of Nonadecenoic Acid (C19:1,2) Arachidic Acid (C20:0) Eicosenoic Acid (C20:1) Eicosadienoic Acid (C20:2) Heneicosanoic Acid (C21:0) Behenic Acid (C22:0) Tricosanoic Acid (C23:0) Lignoceric Acid (C24:0)	Myristic Acid (C14:0) Heptadecanoic Acid (C17:0) Linolenic Acid (C18:3)	Palmitic Acid (C16:0)	Heptadecenoic Acid (C17:1)	--	Myristic Acid (C14:0) Palmitic Acid (C16:0) Heptadecanoic Acid (C17:0) Heptadecenoic Acid (C17:1)	Lauric Acid (C12:0) Pentadecanoic Acid (C15:0) Pentadecenoic Acid (C15:1) Nonadecanoic Acid (C19:0)

Table 20. Outcome of Nutrient Composition Assessment for COR23134 Soybean (continued)

Subgroup	No Statistical Difference Identified	Statistical Difference Identified					Not Included in Statistical Analysis (All Data Values Below the Lower Limit of Quantification)
		All Data Values Within Tolerance Interval	One or More Data Values Outside Tolerance Interval, or Tolerance Interval Not Available			Adjusted P-Value < 0.05	
			All Data Values Within Literature Range	One or More Data Values Outside Literature Range, or Literature Range Not Available			
				All Data Values Within Reference Data Range	One or More Data Values Outside Reference Data Range		
Seed (R8 Growth Stage)							
Amino Acid Composition	Alanine Aspartic Acid Cystine Histidine Lysine Methionine Phenylalanine Threonine Tryptophan Tyrosine	Arginine Glycine Isoleucine Leucine Proline Serine Valine	Glutamic Acid	--	--	Glutamic Acid Isoleucine Leucine	--
Mineral Composition	Copper Iron Manganese Potassium Sodium Zinc	Calcium Magnesium Phosphorus	--	--	--	Calcium	--
Vitamin Composition	Vitamin B1 (Thiamine) Vitamin B2 (Riboflavin) Vitamin B3 (Niacin) Vitamin B6 (Pyridoxine) Vitamin B9 (Folic Acid) Vitamin K_1	α-Tocopherol	Vitamin B5 (Pantothenic Acid)	--	--	--	--

Table 20. Outcome of Nutrient Composition Assessment for COR23134 Soybean (continued)

Subgroup	No Statistical Difference Identified	Statistical Difference Identified					Not Included in Statistical Analysis (All Data Values Below the Lower Limit of Quantification)
			One or More Data Values Outside Tolerance Interval, or Tolerance Interval Not Available				
				One or More Data Values Outside Literature Range, or Literature Range Not Available			
		All Data Values Within Tolerance Interval	All Data Values Within Literature Range	All Data Values Within Reference Data Range	One or More Data Values Outside Reference Data Range	Adjusted P-Value < 0.05	
Seed (R8 Growth Stage)							
Isoflavone Composition	Total Glycitein Equivalent		Total Daidzein Equivalent	--	Total Genistein Equivalent	Total Genistein Equivalent	--
Anti-Nutrient Composition	Raffinose Stachyose Lectins Phytic Acid Trypsin Inhibitor	--	--	--	--	--	--

Note: Growth stages (Pedersen, 2004).

Proximate and Fiber Assessment of COR23134 Soybean Forage

Proximates and fiber were analyzed in forage derived from COR23134 soybean and control soybean. Results are shown in Table 21. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean.

These results demonstrate that the proximate and fiber composition of forage derived from COR23134 soybean is comparable to conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 21. Proximate and Fiber Results for COR23134 Soybean Forage

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Proximate and Fiber Composition (% Dry Weight or as Indicated)						
Moisture (%)	Mean	87.4	87.7	68.2 - 93.3	32.0 - 89.8	80.6 - 91.0
	Range	80.9 - 91.7	82.2 - 91.1			
	Confidence Interval	85.2 - 89.6	85.5 - 89.8			
	Adjusted P-Value	--	0.640			
	P-Value	--	0.497			
Crude Protein	Mean	22.2	22.2	14.1 - 29.4	9.51 - 46.25	17.5 - 27.0
	Range	18.3 - 25.9	19.8 - 25.2			
	Confidence Interval	21.1 - 23.4	21.0 - 23.3			
	Adjusted P-Value	--	0.942			
	P-Value	--	0.892			
Crude Fat	Mean	8.00	8.07	0.914 - 6.88	0.3 - 20.0	5.54 - 10.6
	Range	6.24 - 10.7	5.59 - 10.4			
	Confidence Interval	7.47 - 8.52	7.55 - 8.59			
	Adjusted P-Value	--	0.879			
	P-Value	--	0.787			
Crude Fiber	Mean	34.0	34.2	15.7 - 41.4	13.59 - 46.20	26.1 - 43.7
	Range	29.4 - 39.8	30.4 - 40.7			
	Confidence Interval	32.5 - 35.4	32.7 - 35.6			
	Adjusted P-Value	--	0.798			
	P-Value	--	0.662			
ADF	Mean	36.8	36.8	19.5 - 50.7	12.845 - 64.100	27.8 - 47.9
	Range	31.0 - 43.6	31.8 - 42.7			
	Confidence Interval	35.0 - 38.6	35.1 - 38.6			
	Adjusted P-Value	--	0.970			
	P-Value	--	0.951			
NDF	Mean	46.8	46.3	25.8 - 60.8	19.26 - 82.00	37.5 - 55.3
	Range	42.1 - 52.8	41.0 - 50.8			
	Confidence Interval	45.4 - 48.2	44.9 - 47.7			
	Adjusted P-Value	--	0.560			
	P-Value	--	0.398			
Ash	Mean	10.3	10.5	6.65 - 16.3	2.866 - 36.600	7.12 - 13.6
	Range	7.74 - 12.7	7.99 - 13.7			
	Confidence Interval	9.20 - 11.4	9.35 - 11.6			
	Adjusted P-Value	--	0.487			
	P-Value	--	0.301			
Carbohydrates	Mean	59.5	59.3	53.5 - 74.5	27.8 - 80.6	51.1 - 66.7
	Range	54.7 - 65.6	53.9 - 63.9			
	Confidence Interval	57.7 - 61.3	57.5 - 61.1			
	Adjusted P-Value	--	0.826			
	P-Value	--	0.710			

Proximate and Fiber Assessment of COR23134 Soybean Seed

Proximates and fiber were analyzed in seed derived from COR23134 soybean and control soybean. Results are shown in Table 22. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean, with a few exceptions. A statistically significant difference (P-value < 0.05) was observed between COR23134 soybean and control soybean for crude protein, ash, and carbohydrates. All individual values for these analytes were within one or more of the reference ranges, indicating that COR23134 soybean is within the range of biological variation for these analytes and the statistical differences are not biologically meaningful. Additionally, for ash and carbohydrates the FDR-adjusted P-values were non-significant, indicating that the identified statistical differences were likely false positives.

These results demonstrate that the proximate and fiber composition of seed derived from COR23134 soybean is comparable to conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 22. Proximate and Fiber Results for COR23134 Soybean Seed

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Proximate and Fiber Composition (% Dry Weight or as Indicated)						
Moisture (%)	Mean	9.50	9.35	5.78 - 32.7	4.7 - 44.5	6.22 - 14.0
	Range	6.03 - 13.4	6.66 - 13.3			
	Confidence Interval	7.78 - 11.2	7.63 - 11.1			
	Adjusted P-Value	--	0.256			
	P-Value	--	0.115			
Crude Protein	Mean	39.1	38.3	32.7 - 45.2	29.51 - 52.1	33.0 - 44.1
	Range	36.6 - 44.0	33.2 - 43.6			
	Confidence Interval	37.6 - 40.6	36.8 - 39.8			
	Adjusted P-Value	--	0.0227†			
	P-Value	--	0.00239*			
Crude Fat	Mean	19.6	19.8	13.9 - 23.8	6.97 - 27.4	16.8 - 22.2
	Range	17.4 - 21.2	17.3 - 21.1			
	Confidence Interval	18.8 - 20.5	19.0 - 20.7			
	Adjusted P-Value	--	0.172			
	P-Value	--	0.0631			
Crude Fiber	Mean	13.8	14.1	6.07 - 18.5	2.7 - 18.50	12.2 - 15.1
	Range	12.9 - 15.4	12.8 - 15.7			
	Confidence Interval	13.4 - 14.3	13.6 - 14.5			
	Adjusted P-Value	--	0.172			
	P-Value	--	0.0635			
ADF	Mean	15.1	15.3	9.10 - 24.5	4.60 - 35.30	12.3 - 18.1
	Range	13.6 - 17.4	14.0 - 17.2			
	Confidence Interval	14.6 - 15.6	14.8 - 15.9			
	Adjusted P-Value	--	0.537			
	P-Value	--	0.360			
NDF	Mean	16.1	16.5	11.2 - 23.5	7.38 - 31.90	14.6 - 19.3
	Range	14.6 - 18.0	14.4 - 17.6			
	Confidence Interval	15.6 - 16.6	16.0 - 17.0			
	Adjusted P-Value	--	0.168			
	P-Value	--	0.0565			
Ash	Mean	4.88	4.81	4.07 - 7.18	3.67 - 10.90	4.29 - 6.05
	Range	4.45 - 5.22	4.55 - 5.02			
	Confidence Interval	4.80 - 4.96	4.73 - 4.89			
	Adjusted P-Value	--	0.0677			
	P-Value	--	0.0152*			
Carbohydrates	Mean	36.5	37.1	29.9 - 43.5	25.2 - 55.8	33.8 - 43.5
	Range	33.8 - 38.2	34.0 - 41.8			
	Confidence Interval	35.6 - 37.3	36.3 - 38.0			
	Adjusted P-Value	--	0.0632			
	P-Value	--	0.0116*			

Note: This table provides results from the mixed model analysis only.

* A statistically significant difference (P-value < 0.05) was observed.

† Adjusted P-value < 0.05 was observed.

Fatty Acid Assessment of COR23134 Soybean Seed

Fatty acids were analyzed in seed derived from COR23134 soybean and control soybean. Results are shown in Table 23. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean, except for five analytes. A statistically significant difference (P-value < 0.05) was observed between COR23134 soybean and control soybean for myristic acid (C14:0), palmitic acid (C16:0), heptadecanoic acid (C17:0), heptadecenoic acid (C17:1), and linolenic acid (C18:3). For linolenic acid (C18:3), the FDR-adjusted P-value was non-significant, indicating that the identified statistical difference was likely a false positive. All individual values for these five analytes were within one or more of the reference ranges, indicating that COR23134 soybean is within the range of biological variation for these analytes and the statistical differences are not biologically meaningful.

These results demonstrate that the fatty acid composition of seed derived from COR23134 soybean is comparable to conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 23. Fatty Acid Results for COR23134 Soybean Seed

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
		Fatty Acid Composition (% Total Fatty Acids)				
Lauric Acid (C12:0)	Mean	<LLOQ[a]	<LLOQ[a]	0 - 0.0930	NQ - 0.134	<LLOQ[a]
	Range	<LLOQ[a]	<LLOQ[a]			
	Confidence Interval	NA	NA			
	Adjusted P-Value	--	NA			
	P-Value	--	NA			
Myristic Acid (C14:0)	Mean	0.0798	0.0763	0.0291 - 0.107	NQ - 0.243	0.0601 - 0.107
	Range	0.0614 - 0.101	0.0659 - 0.103			
	Confidence Interval	0.0727 - 0.0870	0.0691 - 0.0834			
	Adjusted P-Value	--	0.000165[†]			
	P-Value	--	<0.0001[*]			
Pentadecanoic Acid (C15:0)	Mean	<LLOQ[a]	<LLOQ[a]	NC	NQ	<LLOQ[a]
	Range	<LLOQ[a]	<LLOQ[a]			
	Confidence Interval	NA	NA			
	Adjusted P-Value	--	NA			
	P-Value	--	NA			
Pentadecenoic Acid (C15:1)	Mean	<LLOQ[a]	<LLOQ[a]	NC	NQ	<LLOQ[a]
	Range	<LLOQ[a]	<LLOQ[a]			
	Confidence Interval	NA	NA			
	Adjusted P-Value	--	NA			
	P-Value	--	NA			
Palmitic Acid (C16:0)	Mean	11.7	11.5	8.58 - 13.1	8.03 - 15.99	9.66 - 13.5
	Range	10.2 - 13.6	10.1 - 13.4			
	Confidence Interval	11.5 - 11.9	11.3 - 11.7			
	Adjusted P-Value	--	<0.0001[†]			
	P-Value	--	<0.0001[*]			
Palmitoleic Acid (C16:1)	Mean	0.0930	0.0909	0.0545 - 0.204	NQ - 0.247	0.0650 - 0.115
	Range	0.0761 - 0.110	0.0707 - 0.114			
	Confidence Interval	0.0860 - 0.0999	0.0839 - 0.0979			
	Adjusted P-Value	--	0.488			
	P-Value	--	0.308			
Heptadecanoic Acid (C17:0)	Mean	0.0971	0.114	0.0531 - 0.140	NQ - 0.166	0.0778 - 0.135
	Range	0.0835 - 0.129	0.103 - 0.138			
	Confidence Interval	0.0920 - 0.102	0.109 - 0.119			
	Adjusted P-Value	--	0.000941[†]			
	P-Value	--	<0.0001[*]			
Heptadecenoic Acid (C17:1)	Mean	0.0641	0.0701	0 - 0.0954	NQ - 0.088	0.0486 - 0.126
	Range	0.0487 - 0.0922	0.0598 - 0.0980			
	Confidence Interval	0.0579 - 0.0702	0.0639 - 0.0762			
	Adjusted P-Value	--	<0.0001[†]			
	P-Value	--	<0.0001[*]			
Heptadecadienoic Acid (C17:2)	Mean	0.0360	0.0405	0 - 0.0956	NR - NR	<LLOQ[a] - 0.0759
	Range	<LLOQ[a] - 0.0642	<LLOQ[a] - 0.0615			
	Confidence Interval	0.0247 - 0.0472	0.0293 - 0.0518			
	Adjusted P-Value	--	0.256			
	P-Value	--	0.106			
Stearic Acid (C18:0)	Mean	4.34	4.44	2.94 - 5.69	2.68 - 6.74	3.75 - 6.08
	Range	3.25 - 5.61	3.55 - 5.59			
	Confidence Interval	4.19 - 4.48	4.30 - 4.59			
	Adjusted P-Value	--	0.288			
	P-Value	--	0.140			

Table 23. Fatty Acid Results for COR23134 Soybean Seed (continued)

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Oleic Acid (C18:1)	Mean	21.0	20.7	16.0 - 36.4	11.5 - 60.0	17.4 - 26.8
	Range	16.5 - 23.2	16.9 - 23.5			
	Confidence Interval	19.9 - 22.1	19.6 - 21.8			
	Adjusted P-Value	--	0.295			
	P-Value	--	0.148			
Linoleic Acid (C18:2)	Mean	53.3	53.7	42.2 - 58.9	25.0 - 72.5	47.5 - 57.3
	Range	51.4 - 57.5	51.3 - 56.9			
	Confidence Interval	52.4 - 54.2	52.9 - 54.6			
	Adjusted P-Value	--	0.203			
	P-Value	--	0.0776			
Linolenic Acid (C18:3)	Mean	7.74	7.59	1.25 - 11.4	2.9 - 12.84	6.08 - 9.80
	Range	6.72 - 9.55	6.68 - 9.63			
	Confidence Interval	7.37 - 8.12	7.21 - 7.97			
	Adjusted P-Value	--	0.0781			
	P-Value	--	0.0195*			
Nonadecanoic Acid (C19:0)	Mean	<LLOQ[a]	<LLOQ[a]	0 - 0.201	NR - NR	<LLOQ[a]
	Range	<LLOQ[a]	<LLOQ[a]			
	Confidence Interval	NA	NA			
	Adjusted P-Value	--	NA			
	P-Value	--	NA			
Isomer 1 of Nonadecenoic Acid (C19:1,1)	Mean	0.0228	0.0237	NC	NR	<LLOQ[a] - 0.0478
	Range	<LLOQ[a] - 0.0435	<LLOQ[a] - 0.0473			
	Confidence Interval	NA	NA			
	Adjusted P-Value	--	NA			
	P-Value	--	NA			
Isomer 2 of Nonadecenoic Acid (C19:1,2)	Mean	0.0588	0.0633	NC	NR	<LLOQ[a] - 0.136
	Range	<LLOQ[a] - 0.125	0.0410 - 0.104			
	Confidence Interval	0.0493 - 0.0727	0.0524 - 0.0798			
	Adjusted P-Value	--	0.498			
	P-Value	--	0.321			
Arachidic Acid (C20:0)	Mean	0.324	0.329	0.209 - 0.445	NQ - 0.611	0.269 - 0.479
	Range	0.230 - 0.440	0.258 - 0.426			
	Confidence Interval	0.312 - 0.337	0.316 - 0.341			
	Adjusted P-Value	--	0.438			
	P-Value	--	0.259			
Eicosenoic Acid (C20:1)	Mean	0.181	0.181	0.0704 - 0.302	NQ - 0.387	0.139 - 0.228
	Range	0.146 - 0.222	0.149 - 0.230			
	Confidence Interval	0.167 - 0.195	0.167 - 0.195			
	Adjusted P-Value	--	0.803			
	P-Value	--	0.676			
Eicosadienoic Acid (C20:2)	Mean	0.0605	0.0588	0 - 0.0961	NQ - 0.341	0.0425 - 0.101
	Range	<LLOQ[a] - 0.116	0.0471 - 0.0888			
	Confidence Interval	0.0503 - 0.0706	0.0486 - 0.0690			
	Adjusted P-Value	--	0.438			
	P-Value	--	0.257			
Heneicosanoic Acid (C21:0)	Mean	0.0535	0.0509	0 - 0.0690	NR - NR	<LLOQ[a] - 0.222
	Range	0.0415 - 0.162	0.0411 - 0.147			
	Confidence Interval	0.0466 - 0.0646	0.0449 - 0.0603			
	Adjusted P-Value	--	0.430			
	P-Value	--	0.237			

Table 23. Fatty Acid Results for COR23134 Soybean Seed (continued)

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Behenic Acid (C22:0)	Mean	0.355	0.352	0 - 0.502	0.181 - 0.723	0.281 - 0.474
	Range	0.298 - 0.431	0.304 - 0.445			
	Confidence Interval	0.332 - 0.377	0.330 - 0.375			
	Adjusted P-Value	--	0.560			
	P-Value	--	0.420			
Tricosanoic Acid (C23:0)	Mean	0.0857	0.0829	0 - 0.0901	NR - NR	0.0458 - 0.924
	Range	0.0662 - 0.575	0.0656 - 0.722			
	Confidence Interval	0.0748 - 0.104	0.0729 - 0.0987			
	Adjusted P-Value	--	0.381			
	P-Value	--	0.205			
Lignoceric Acid (C24:0)	Mean	0.129	0.131	0 - 0.282	NQ - 0.32	<LLOQ[a] - 0.239
	Range	0.0869 - 0.214	0.0854 - 0.240			
	Confidence Interval	0.102 - 0.156	0.104 - 0.158			
	Adjusted P-Value	--	0.537			
	P-Value	--	0.356			

Note: This table provides results from the mixed model analysis only. Not applicable (NA); mixed model analysis was not performed, or confidence interval was not determined. Not calculated (NC); the tolerance interval calculation was not performed due to insufficient data. Not quantified (NQ); one or more assay values in the published literature references were below the lower limit of quantification (LLOQ) and were not quantified. Analyte ranges were not reported (NR) in the published literature references.

[a] < LLOQ, one or more fatty acid sample values were below the assay LLOQ.

[*] A statistically significant difference (P-value < 0.05) was observed.

[†] Adjusted P-value < 0.05 was observed.

Amino Acid Assessment of COR23134 Soybean Seed

Amino acids were analyzed in seed derived from COR23134 soybean and control soybean. Results are shown in Table 24. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean for 10 analytes. For the remaining eight analytes (arginine, glutamic acid, glycine, isoleucine, leucine, proline, serine, and valine), a statistically significant difference (P-value < 0.05) was observed between COR23134 soybean and control soybean. For arginine, glycine, proline, serine, and valine, the FDR-adjusted P-values were non-significant, indicating that the identified statistical differences were likely false positives. All individual values for these eight analytes were within one or more of the reference ranges, indicating that COR23134 soybean is within the range of biological variation for these analytes and the statistical differences are not biologically meaningful.

These results demonstrate that the amino acid composition of seed derived from COR23134 soybean is comparable to conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 24. Amino Acid Results for COR23134 Soybean Seed

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Amino Acid Composition (% Dry Weight)						
Alanine	Mean	1.71	1.70	1.35 - 2.04	1.15 - 2.35	1.31 - 2.05
	Range	1.37 - 1.97	1.37 - 1.98			
	Confidence Interval	1.59 - 1.83	1.58 - 1.82			
	Adjusted P-Value	--	0.914			
	P-Value	--	0.842			
Arginine	Mean	2.68	2.60	2.08 - 3.53	1.73 - 3.93	2.16 - 3.39
	Range	2.18 - 3.10	2.24 - 3.34			
	Confidence Interval	2.51 - 2.86	2.43 - 2.78			
	Adjusted P-Value	--	0.0643			
	P-Value	--	0.0131*			
Aspartic Acid	Mean	5.00	4.84	3.56 - 5.61	3.13 - 6.83	3.90 - 6.62
	Range	3.92 - 5.84	3.93 - 6.75			
	Confidence Interval	4.70 - 5.30	4.54 - 5.14			
	Adjusted P-Value	--	0.256			
	P-Value	--	0.108			
Cystine	Mean	0.483	0.484	0.383 - 0.741	0.31 - 0.93	0.309 - 0.695
	Range	0.336 - 0.595	0.365 - 0.604			
	Confidence Interval	0.459 - 0.506	0.460 - 0.507			
	Adjusted P-Value	--	0.970			
	P-Value	--	0.944			
Glutamic Acid	Mean	7.85	7.59	5.45 - 8.88	4.35 - 10.90	6.04 - 10.1
	Range	6.03 - 9.05	6.33 - 9.77			
	Confidence Interval	7.38 - 8.31	7.13 - 8.06			
	Adjusted P-Value	--	0.0428†			
	P-Value	--	0.00507*			
Glycine	Mean	1.71	1.66	1.40 - 2.04	1.16 - 2.55	1.32 - 2.06
	Range	1.35 - 1.97	1.43 - 1.99			
	Confidence Interval	1.60 - 1.82	1.55 - 1.77			
	Adjusted P-Value	--	0.162			
	P-Value	--	0.0489*			
Histidine	Mean	1.05	1.05	0.772 - 1.29	0.20 - 1.59	0.722 - 1.27
	Range	0.883 - 1.23	0.882 - 1.26			
	Confidence Interval	0.984 - 1.12	0.983 - 1.12			
	Adjusted P-Value	--	0.970			
	P-Value	--	0.957			
Isoleucine	Mean	1.85	1.80	1.44 - 2.15	1.20 - 2.48	1.45 - 2.17
	Range	1.46 - 2.12	1.56 - 2.11			
	Confidence Interval	1.75 - 1.95	1.70 - 1.90			
	Adjusted P-Value	--	0.0227†			
	P-Value	--	0.00232*			
Leucine	Mean	3.08	3.00	2.45 - 3.53	2.04 - 4.13	2.51 - 3.63
	Range	2.54 - 3.52	2.61 - 3.45			
	Confidence Interval	2.92 - 3.23	2.85 - 3.16			
	Adjusted P-Value	--	0.0441†			
	P-Value	--	0.00581*			

Table 24. Amino Acid Results for COR23134 Soybean Seed (continued)

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Amino Acid Composition (% Dry Weight)						
Lysine	Mean	2.73	2.71	2.10 - 3.14	1.56 - 3.94	2.02 - 3.24
	Range	2.02 - 3.17	2.06 - 3.18			
	Confidence Interval	2.57 - 2.89	2.55 - 2.87			
	Adjusted P-Value	--	0.712			
	P-Value	--	0.562			
Methionine	Mean	0.473	0.479	0.400 - 0.719	0.29 - 1.15	0.327 - 0.639
	Range	0.376 - 0.574	0.349 - 0.597			
	Confidence Interval	0.448 - 0.499	0.453 - 0.505			
	Adjusted P-Value	--	0.775			
	P-Value	--	0.632			
Phenylalanine	Mean	1.94	1.90	1.59 - 2.47	1.40 - 2.73	1.57 - 2.28
	Range	1.63 - 2.28	1.60 - 2.27			
	Confidence Interval	1.82 - 2.06	1.77 - 2.02			
	Adjusted P-Value	--	0.223			
	P-Value	--	0.0879			
Proline	Mean	2.05	1.99	1.58 - 2.42	1.32 - 2.95	1.55 - 2.41
	Range	1.60 - 2.34	1.68 - 2.36			
	Confidence Interval	1.94 - 2.16	1.88 - 2.10			
	Adjusted P-Value	--	0.0643			
	P-Value	--	0.0135*			
Serine	Mean	2.05	2.00	1.60 - 2.48	0.86 - 2.80	1.59 - 2.41
	Range	1.65 - 2.33	1.66 - 2.40			
	Confidence Interval	1.93 - 2.16	1.89 - 2.12			
	Adjusted P-Value	--	0.117			
	P-Value	--	0.0334*			
Threonine	Mean	1.58	1.55	1.28 - 1.90	1.07 - 2.18	1.26 - 1.87
	Range	1.28 - 1.81	1.33 - 1.83			
	Confidence Interval	1.49 - 1.67	1.46 - 1.64			
	Adjusted P-Value	--	0.168			
	P-Value	--	0.0574			
Tryptophan	Mean	0.520	0.516	0.390 - 0.645	0.254 - 0.746	0.423 - 0.608
	Range	0.451 - 0.614	0.464 - 0.582			
	Confidence Interval	0.503 - 0.537	0.499 - 0.534			
	Adjusted P-Value	--	0.608			
	P-Value	--	0.464			
Tyrosine	Mean	1.19	1.16	0.932 - 1.78	0.74 - 2.32	0.937 - 1.48
	Range	0.936 - 1.50	1.00 - 1.49			
	Confidence Interval	1.11 - 1.28	1.08 - 1.25			
	Adjusted P-Value	--	0.256			
	P-Value	--	0.113			
Valine	Mean	1.89	1.85	1.50 - 2.21	1.24 - 2.66	1.49 - 2.22
	Range	1.53 - 2.18	1.59 - 2.15			
	Confidence Interval	1.79 - 2.00	1.74 - 1.95			
	Adjusted P-Value	--	0.0776			
	P-Value	--	0.0184*			

Note: This table provides results from the mixed model analysis only.
* A statistically significant difference (P-value < 0.05) was observed.
† Adjusted P-value < 0.05 was observed.

Mineral Assessment of COR23134 Soybean Seed

Minerals were analyzed in seed derived from COR23134 soybean and control soybean. Results are shown in Table 25. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean, with a few exceptions. A statistically significant difference (P-value < 0.05) was observed between COR23134 soybean and control soybean for calcium, magnesium, and phosphorus. All individual values for these analytes were within one or more of the reference ranges, indicating that COR23134 soybean is within the range of biological variation for these analytes and the statistical difference is not biologically meaningful. The non-significant FDR-adjusted P-values (magnesium and phosphorus) indicate that the differences were likely false positives.

These results demonstrate that the mineral composition of seed derived from COR23134 soybean is comparable to conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 25. Mineral Results for COR23134 Soybean Seed

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Mineral Composition (% Dry Weight)						
Calcium	Mean	0.225	0.261	0.0902 - 0.413	0.09 - 0.49	0.165 - 0.322
	Range	0.159 - 0.281	0.184 - 0.311			
	Confidence Interval	0.201 - 0.249	0.237 - 0.286			
	Adjusted P-Value	--	0.000941†			
	P-Value	--	<0.0001*			
Copper	Mean	0.00116	0.00114	<0.000500[b] - 0.00246	NQ - 0.02	0.000616 - 0.00170
	Range	0.000844 - 0.00160	0.000637 - 0.00159			
	Confidence Interval	0.000970 - 0.00134	0.000949 - 0.00132			
	Adjusted P-Value	--	0.560			
	P-Value	--	0.394			
Iron	Mean	0.00808	0.00801	0.00476 - 0.0391	0.0047 - 0.3780	0.00568 - 0.0948
	Range	0.00590 - 0.0232	0.00585 - 0.0135			
	Confidence Interval	0.00693 - 0.00922	0.00687 - 0.00916			
	Adjusted P-Value	--	0.826			
	P-Value	--	0.728			
Magnesium	Mean	0.236	0.251	0.146 - 0.340	0.09 - 0.40	0.190 - 0.282
	Range	0.199 - 0.287	0.212 - 0.299			
	Confidence Interval	0.220 - 0.252	0.235 - 0.267			
	Adjusted P-Value	--	0.0582			
	P-Value	--	0.00875*			
Manganese	Mean	0.00296	0.00292	0.00183 - 0.00917	0.0014 - 0.0463	0.00197 - 0.00522
	Range	0.00200 - 0.00492	0.00216 - 0.00404			
	Confidence Interval	0.00249 - 0.00343	0.00244 - 0.00339			
	Adjusted P-Value	--	0.560			
	P-Value	--	0.412			
Phosphorus	Mean	0.526	0.505	0.346 - 0.817	0.21 - 0.94	0.375 - 0.651
	Range	0.360 - 0.661	0.394 - 0.615			
	Confidence Interval	0.479 - 0.573	0.458 - 0.553			
	Adjusted P-Value	--	0.103			
	P-Value	--	0.0270*			
Potassium	Mean	1.85	1.81	1.33 - 2.48	0.68 - 2.71	1.41 - 2.09
	Range	1.57 - 2.22	1.53 - 1.99			
	Confidence Interval	1.76 - 1.93	1.73 - 1.90			
	Adjusted P-Value	--	0.283			
	P-Value	--	0.134			
Sodium	Mean	0.00100	0.000979	<LLOQ[a] - 0.0247	NQ - 0.04	<0.000500[a] - 0.00686
	Range	<0.000500[a] - 0.00247	<0.000500[a] - 0.00256			
	Confidence Interval	0.000468 - 0.00153	0.000447 - 0.00151			
	Adjusted P-Value	--	0.917			
	P-Value	--	0.857			
Zinc	Mean	0.00436	0.00430	0.00219 - 0.00953	0.0004 - 0.0125	0.00294 - 0.0299
	Range	0.00310 - 0.00566	0.00311 - 0.00533			
	Confidence Interval	0.00387 - 0.00486	0.00380 - 0.00480			
	Adjusted P-Value	--	0.450			
	P-Value	--	0.273			

Note: This table provides results from the mixed model analysis only. Not quantified (NQ); one or more assay values in the published literature references were below the lower limit of quantification (LLOQ) and were not quantified. Mineral composition is reported as % Dry Weight.

[a] < LLOQ, one or more sample values were below the assay LLOQ.

* A statistically significant difference (P-Value < 0.05) was observed.

† Adjusted P-Value < 0.05 was observed.

Vitamin Assessment of COR23134 Soybean Seed

Vitamins were analyzed in seed derived from COR23134 soybean and control soybean. Results are shown in Table 26. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean, with two exceptions. A statistically significant difference (P-value < 0.05) was observed between COR23134 soybean and control soybean for vitamin B5 and α-tocopherol. All individual values for these analytes were within one or more of the reference ranges, indicating that COR23134 soybean is within the range of biological variation for these analytes and the statistical difference is not biologically meaningful.

These results demonstrate that the vitamin composition of seed derived from COR23134 soybean is comparable to conventional seed represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 26. Vitamin Results for COR23134 Soybean Seed

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Vitamin Composition (mg/kg Dry Weight)						
Vitamin B1 (Thiamine)	Mean	1.04	0.633	0.0663 - 7.17	NQ - 17.8	<0.750[a] - 2.59
	Range	<0.750[a] - 2.35	<0.750[a] - 1.59			
	Confidence Interval	NA	NA			
	Adjusted P-Value	--	NA			
	P-Value	--	NA			
Vitamin B2 (Riboflavin)	Mean	4.46	4.28	2.47 - 7.69	1.70 - 10.87	2.40 - 7.05
	Range	2.59 - 7.16	2.51 - 6.88			
	Confidence Interval	3.84 - 5.08	3.66 - 4.90			
	Adjusted P-Value	--	0.560			
	P-Value	--	0.413			
Vitamin B3 (Niacin)	Mean	47.1	45.3	7.31 - 45.1	10.90 - 75.50	31.9 - 67.0
	Range	34.2 - 55.7	32.5 - 58.1			
	Confidence Interval	45.2 - 48.9	43.5 - 47.1			
	Adjusted P-Value	--	0.283			
	P-Value	--	0.130			
Vitamin B5 (Pantothenic Acid)	Mean	14.2	15.3	3.97 - 21.5	5.05 - 36.2	6.37 - 25.8
	Range	7.60 - 26.5	9.48 - 29.0			
	Confidence Interval	10.2 - 18.2	11.3 - 19.3			
	Adjusted P-Value	--	0.0582			
	P-Value	--	0.00919[*]			
Vitamin B6 (Pyridoxine)	Mean	3.62	3.56	2.81 - 8.98	1.87 - 12.8	2.73 - 5.45
	Range	2.92 - 5.11	2.93 - 4.82			
	Confidence Interval	3.21 - 4.02	3.16 - 3.97			
	Adjusted P-Value	--	0.381			
	P-Value	--	0.201			
Vitamin B9 (Folic Acid)	Mean	7.75	7.54	1.94 - 14.2	1.24 - 22.20	1.82 - 22.0
	Range	2.81 - 15.5	2.16 - 22.3			
	Confidence Interval	6.09 - 9.88	5.92 - 9.60			
	Adjusted P-Value	--	0.904			
	P-Value	--	0.821			
Vitamin K_1	Mean	0.575	0.601	0 - 1.51	NQ - 2.07	0.322 - 1.49
	Range	0.286 - 0.968	0.374 - 1.02			
	Confidence Interval	0.441 - 0.709	0.467 - 0.735			
	Adjusted P-Value	--	0.317			
	P-Value	--	0.163			
α-Tocopherol	Mean	28.5	27.3	<LLOQ[a] - 70.9	NQ - 127.38	8.92 - 43.3
	Range	17.0 - 42.1	16.2 - 39.0			
	Confidence Interval	23.4 - 33.7	22.2 - 32.5			
	Adjusted P-Value	--	0.117			
	P-Value	--	0.0338[*]			

Note: This table provides results from the mixed model analysis only. Not applicable (NA); mixed model analysis was not performed, or confidence interval was not determined. Not quantified (NQ); one or more assay values in the published literature references were below the lower limit of quantification (LLOQ) and were not quantified.

[a] < LLOQ, one or more sample values were below the assay LLOQ.

[*] A statistically significant difference (P-value < 0.05) was observed.

[†] Adjusted P-value < 0.05 was observed.

Isoflavone and Anti-Nutrient Assessment of COR23134 Soybean Seed

Isoflavones and anti-nutrients were analyzed in seed derived from COR23134 soybean and control soybean. Results are shown in Table 27. No statistically significant differences (P-value < 0.05) were observed between COR23134 soybean and control soybean, except for two analytes. A statistically significant difference (P-value < 0.05) was observed between COR23134 soybean and control soybean for total daidzein equivalent and total genistein equivalent. All individual values for these analytes were within one or more of the reference ranges, indicating that COR23134 soybean is within the range of biological variation for these analytes and the statistical differences are not biologically meaningful. The non-significant FDR-adjusted P-value for total daidzein equivalent indicates that the difference was likely a false positive.

These results demonstrate that the isoflavone and anti-nutrient composition of seed derived from COR23134 soybean is comparable to conventional soybean represented by non-GM near-isoline control soybean and non-GM commercial soybean.

Table 27. Isoflavone and Anti-Nutrient Results for COR23134 Soybean Seed

Analyte	Reported Statistics	Control Soybean	COR23134 Soybean	Tolerance Interval	Literature Range	Reference Data Range
Isoflavone Composition (mg/kg Dry Weight)						
Total Daidzein Equivalent	Mean	903	968	<10.0[a] - 1880	60.0 - 3061.20	161 - 1570
	Range	178 - 1640	264 - 1990			
	Confidence Interval	567 - 1440	608 - 1540			
	Adjusted P-Value	--	0.0613			
	P-Value	--	0.0105*			
Total Glycitein Equivalent	Mean	107	101	<10.0[a] - 1390	NQ - 1630.00	36.6 - 364
	Range	52.3 - 235	46.7 - 246			
	Confidence Interval	82.3 - 132	76.2 - 126			
	Adjusted P-Value	--	0.434			
	P-Value	--	0.246			
Total Genistein Equivalent	Mean	1150	1270	<10.0[a] - 2300	35.71 - 2837.20	159 - 2680
	Range	231 - 2280	294 - 2880			
	Confidence Interval	710 - 1870	782 - 2070			
	Adjusted P-Value	--	<0.0001†			
	P-Value	--	<0.0001*			
Anti-Nutrient Composition (% Dry Weight)						
Raffinose	Mean	0.316	0.322	0.136 - 1.41	NQ - 1.8542	<0.200[a] - 1.49
	Range	<0.200[a] - 1.53	<0.200[a] - 1.20			
	Confidence Interval	0.110 - 0.522	0.116 - 0.527			
	Adjusted P-Value	--	0.826			
	P-Value	--	0.720			
Stachyose	Mean	1.24	1.21	1.88 - 5.55	NQ - 6.8900	<0.300[a] - 5.17
	Range	0.376 - 5.27	0.381 - 4.17			
	Confidence Interval	0.315 - 2.16	0.288 - 2.13			
	Adjusted P-Value	--	0.752			
	P-Value	--	0.604			
Lectins (mg/g DW)	Mean	1.76	1.76	1.04 - 7.45	0.7764 - 9.3500	<0.375[a] - 3.84
	Range	0.485 - 3.52	0.768 - 3.17			
	Confidence Interval	1.35 - 2.18	1.35 - 2.18			
	Adjusted P-Value	--	0.999			
	P-Value	--	0.999			
Phytic Acid	Mean	1.36	1.29	0.512 - 2.16	NQ - 2.8600	0.890 - 2.06
	Range	0.923 - 1.77	0.954 - 1.55			
	Confidence Interval	1.22 - 1.49	1.15 - 1.42			
	Adjusted P-Value	--	0.168			
	P-Value	--	0.0551			
Trypsin Inhibitor (TIU/mg DW)	Mean	22.1	22.9	9.19 - 49.9	3.23 - 184	0.525 - 32.8
	Range	9.02 - 30.1	7.47 - 30.9			
	Confidence Interval	19.3 - 24.8	20.1 - 25.6			
	Adjusted P-Value	--	0.539			
	P-Value	--	0.369			

Note: This table provides results from the mixed model analysis only. Not quantified (NQ); one or more assay values in the published literature references were below the lower limit of quantification (LLOQ) and were not quantified.

[a] < LLOQ, one or more sample values were below the assay LLOQ.

* A statistically significant difference (P-value < 0.05) was observed.

† Adjusted P-value < 0.05 was observed.

C. INFORMATION RELATED TO THE NUTRITIONAL IMPACT OF THE FOOD

In section *B.5 Compositional analyses of the food produced using gene technology* the composition of COR23134 soybean was compared with that of a concurrently grown conventional non-GM comparator with a history of safe use in food and feed. The results demonstrated that the nutrient composition of forage and grain derived from COR23134 soybean is comparable to that of conventional soybean represented by non-GM near-isoline soybean and non-GM commercial soybean. Based on these analyses, the seed and forage of COR23134 soybean are comparable to conventional soybean with respect to nutrient composition.

Therefore, no nutritional impact of COR23134 soybean is expected.

D. OTHER INFORMATION

Overall Risk Assessment Conclusion for COR23134 Soybean

This application presents information supporting the safety and nutritional comparability of COR23134 soybean. The molecular characterization analyses conducted on COR23134 soybean demonstrated that the introduced genes are integrated at a single locus, stably inherited across multiple generations, and segregate according to Mendel's law of genetics.

The allergenic and toxic potential of the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins were evaluated and found unlikely to be allergenic or toxic. Based on the weight of evidence, consumption of the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins is unlikely to cause an adverse effect. A composition assessment demonstrated that the nutrient composition of COR23134 soybean forage and grain is comparable to that of conventional soybean, represented by non-genetically modified (non-GM) near-isoline soybean and non-GM commercial soybean.

Overall, data and information contained herein support the conclusion that COR23134 soybean containing the Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins is as safe and nutritious as non-GM soybean for food and feed uses.

REFERENCES

Abbitt SE, inventor, Aug. 8, 2017. SB-UBI terminator sequence for gene expression in plants. United States Patent, Patent No. US 9,725,731 B2

Abbitt SE, Jung R, inventors. Aug. 8, 2017. SB-actin terminator sequence for gene expression in plants. United States Patent, Patent No. US 9,725,729 B2

Abbitt SE, Klein K, Selinger D, inventors. 07 June 2018. Transcriptional Terminators for Gene Expression in Plants. World Intellectual Property Organization, Patent No. WO 2018/102131 A1

AFSI (2022) Crop Composition Database, Version 9.0. Agriculture & Food Systems Institute, https://www.cropcomposition.org/

Albertsen M, Anderson PC, Che P, Glassman KF, Jung R, Zhao Z-Y, inventors. 8 May 2014. Transformed plants having increased beta-carotene levels, increased half-life and bioavailability and methods of producing such. World Intellectual Property Organization, Patent No. WO 2014/070646 A1

Barry J, Gerber RM, Liu L, Lum A, Schepers E, Yalpani N, Zhu G, inventors. Mar. 12, 2019. Insecticidal proteins from plants and methods for their use. United States Patent, Patent No. US 10,227,608 B2

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society B 57: 289-300

Bhyri P, Krishnamurthy N, Narayanan E, Nott A, Sarangi RR, inventors. Aug. 28, 2018. Plant terminator sequences. United States Patent, Patent No. US 10,059,953 B2

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

CFIA (2021) The Biology of *Glycine max* (L.) Merr. (Soybean). Canadian Food Inspection Agency, BIO2021-01

Cheo DL, Titus SA, Byrd DRN, Hartley JL, Temple GF, Brasch MA (2004) Concerted Assembly and Cloning of Multiple DNA Segments Using In Vitro Site-Specific Recombination: Functional Analysis of Multi-Segment Expression Clones. Genome Research 14: 2111-2120

Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Molecular Biology 18: 675-689

Codex Alimentarius Commission (2008) Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants. Codex Alimentarius, CAC/GL 45-2003

Codex Alimentarius Commission (2009) Foods derived from modern biotechnology, Ed 2. Food and Agriculture Organization of the United Nations, World Health Organization, Rome

Cox MM (1988) FLP Site-Specific Recombination System of *Saccharomyces cerevisiae*. In R Kucherlapati, GR Smith, eds, Genetic Recombination. American Society for Microbiology, pp 429-443

Crickmore N, Berry C, Panneerselvam S, Mishra R, Connor TR, Bonning BC (2021) A structure-based nomenclature for *Bacillus thuringiensis* and other bacteria-derived pesticidal proteins. Journal of Invertebrate Pathology 186: 107438

Dale EC, Ow DW (1990) Intra- and intramolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. Gene 91: 79-85

Dong J, Feng Y, Kumar D, Zhang W, Zhu T, Luo M-C, Messing J (2016) Analysis of tandem gene copies in maize chromosomal regions reconstructed from long sequence reads. Proceedings of the National Academy of Sciences 113: 7949-7956

Dummitt B, Micka WS, Chang Y-H (2003) N-Terminal Methionine Removal and Methionine Metabolism in *Saccharomyces cerevisiae*. Journal of Cellular Biochemistry 89: 964-974

EFSA (2010) Scientific Opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. The EFSA Journal 8: 1700

EFSA Panel on Genetically Modified Organisms (GMO) (2011) Guidance for risk assessment of food and feed from genetically modified plants. The EFSA Journal 9: 2150

Elsing MT, Gordon-Kamm W, Perugini LD, Ye H, inventors. 22 September 2016. Methods and Compositions for Accelerated Trait Introgression. World Intellectual Property Organization, Patent No. WO 2016/149351 A1

EU-RL-GMFF (2013) Event-specific Method for the Quantification of Soybean Event DP-305423-1 Using Real-time PCR: Validated Method. European Union Reference Laboratory for Genetically Modified Food and Feed, http://gmo-crl.jrc.ec.europa.eu/summaries/CRLVL0707VP%20Corr2%20EURL%20web.pdf

EU-RL-GMFF (2015) Definition of Minimum Performance Requirements for Analytical Methods of GMO Testing. European Union Reference Laboratory for Genetically Modified Food and Feed, http://gmo-crl.jrc.ec.europa.eu/doc/MPR%20Report%20Application%2020_10_2015.pdf

Ewing B, Green P (1998) Base-Calling of Automated Sequencer Traces Using *Phred*. II. Error Probabilities. Genome Research 8: 186-194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-Calling of Automated Sequencer Traces Using *Phred*. I. Accuracy Assessment. Genome Research 8: 175-185

Falco SC, Li Z, inventors. Nov. 16, 2010. S-adenosyl-L-methionine synthetase promoter and its use in expression of transgenic genes in plants. United States Patent, Patent No. US 7,834,242 B2

FAO/WHO (2001) Evaluation of Allergencity of Genetically Modified Foods: Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology, 22 - 25 January 2001. Food and Agriculture Organization of the United Nations, Rome

Fernández H (2011) Chapter 1 Introduction. In H Fernández, A Kumar, A Revilla, eds, Working with Ferns Issues and Applications. Springer-Verlag New York, pp 1-8

Fling ME, Kopf J, Richards C (1985) Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme, 3" (9)-*0*-nucleotidyltransferase. Nucleic Acids Research 13: 7095-7106

Flynn PJ, Reece RJ (1999) Activation of Transcription by Metabolic Intermediates of the Pyrimidine Biosynthetic Pathway. Molecular and Cellular Biology 19: 882-888

Greger IH, Demarchi F, Giacca M, Proudfoot NJ (1998) Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Research 26: 1294-1300

Hartley JL, Temple GF, Brasch MA (2000) DNA Cloning Using In Vitro Site-Specific Recombination. Genome Research 10: 1788-1795

Hong B, Fisher TL, Sult TS, Maxwell CA, Mickelson JA, Kishino H, Locke MEH (2014) Model-Based Tolerance Intervals Derived from Cumulative Historical Composition Data: Application for Substantial Equivalence Assessment of a Genetically Modified Crop. Journal of Agricultural and Food Chemistry 62: 9916-9926

Horn C, Lau S, Izumi Wilcoxon M, Yamamoto T, Zheng Y, inventors. 19 October 2017. Insecticidal polypeptides having improved activity spectrum and uses thereof. World Intellectual Property Organization, Patent No. WO 2017/180715 A2

Hymowitz T (1970) On the Domestication of the Soybean. Economic Botany 24: 408-421

ISAAA (2023) GM Approval Database. International Service for the Acquisition of Agri-Biotech Applications, https://www.isaaa.org/gmapprovaldatabase/default.asp

Itoh Y, Watson JM, Haas D, Leisinger T (1984) Genetic and Molecular Characterization of the *Pseudomonas* Plasmid pVS1. Plasmid 11: 206-220

Izumi Wilcoxon M, Yamamoto T, inventors. 21 April 2016. Insecticidal polypeptides having improved activity spectrum and uses thereof. World Intellectual Property Organization, Patent No. WO 2016/061197 A9

Jurat-Fuentes JL, Crickmore N (2017) Specificity determinants for Cry insecticidal proteins: Insights from their mode of action. Journal of Invertebrate Pathology 142: 5-10

Katzen F (2007) Gateway® recombinational cloning: a biological operating system. Expert Opinion on Drug Discovery 2: 571-589

Kenward MG, Roger JH (2009) An improved approximation to the precision of fixed effects from restricted maximum likelihood. Computational Statistics & Data Analysis 53: 2583-2595

Kew Science (2020) *Ophioglossum* L. Plants of the World *online*, http://www.plantsoftheworldonline.org/taxon/urn:lsid:ipni.org:names:17408900-1#distribution-map

Kim MY, Van K, Kang YJ, Kim KH, Lee SH (2012) Tracing soybean domestication history: From nucleotide to genome. Breeding Science 61: 445-452

Kim SH, Jung WS, Ahn JK, Chung IM (2005) Analysis of isoflavone concentration and composition in soybean [*Glycine max* (L.)] seeds between the cropping year and storage for 3 years. European Food Research and Technology 220: 207-214

Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. The Plant Journal 10: 165-174

La Paz JL, Pla M, Papazova N, Puigdomenech P, Vicient CM (2010) Stability of the MON 810 transgene in maize. Plant Molecular Biology 74: 563-571

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357-359

Lee KY, Townsend J, Tepperman J, Black M, Chui C-F, Mazur B, Dunsmuir P, Bedbrook J (1988) The molecular basis of sulfonylurea herbicide resistance in tobacco. The EMBO Journal 7: 1241-1248

Lee SJ, Yan W, Ahn JK, Chung IM (2003) Effects of year, site, genotype and their interactions on various soybean isoflavones. Field Crops Research 81: 181-192

Liu L, Schepers E, Lum A, Rice J, Yalpani N, Gerber R, Jiménez-Juárez N, Haile F, Pascual A, Barry J, Qi X, Kassa A, Heckert MJ, Xie W, Ding C, Oral J, Nguyen M, Le J, Procyk L, Diehn SH, Crane VC, Damude H, Pilcher C, Booth R, Liu L, Zhu G, Nowatzki TM, Nelson ME, Lu AL, Wu G (2019) Identification and Evaluations of Novel Insecticidal Proteins from Plants of the Class Polypodiopsida for Crop Protection against Key Lepidopteran Pests. Toxins 11: 383

Marçais G, Kingsford C (2011) A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. Bioinformatics 27: 764-770

Martin M (2011) Cutadapt removes adapter sequences from high-througput sequencing reads. EMBnet.journal 17: 10-12

Mathesius CA, Barnett Jr JF, Cressman RF, Ding J, Carpenter C, Ladics GS, Schmidt J, Layton RJ, Zhang JXQ, Appenzeller LM, Carlson G, Ballou S, Delaney B (2009) Safety assessment of a modified acetolactate synthase protein (GM-HRA) used as a selectable marker in genetically modified soybeans. Regulatory Toxicology and Pharmacology 55: 309-320

Morse WJ (1950) History of Soybean Production. In KS Markley, ed, Soybeans and Soybean Products, Vol 1. Interscience Publishers, New York, pp 3-59

OECD (1993) Safety Evaluation of Foods Derived by Modern Biotechnology: Concepts and Principles. Organisation for Economic Cooperation and Development

OECD (2000) Consensus Document on the Biology of *Glycine max* (L.) Merr. (Soybean). Organisation for Economic Co-operation and Development, ENV/JM/MONO(2000)9

OECD (2006) Section 2 - Soybean (*Glycine Max* (L.) Marr.). In Safety Assessment of Transgenic Organisms- OECD Consensus Documents- Volume 1. Organisation for Economic Co-operation and Development, pp 40-46, https://www.oecd-ilibrary.org/science-and-technology/safety-assessment-of-transgenic-organisms/section-2_9789264095380-5-en

OECD (2012) Revised Consensus Document on Compositional Considerations for New Varieties of Soybean [*Glycine max* (L.) Merr]: Key Food and Feed Nutrients, Anti-nutrients, Toxicants and Allergens. Organisation for Economic Co-operation and Development, ENV/JM/MONO(2012)24

Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A (2005) False discovery rate, sensitivity and sample size for microarray studies. Bioinformatics 21: 3017-3024

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

Pedersen P (2004) Soybean Growth and Development. Iowa State University Extension, PM 1945 (Reviewed December, 2009)

Peralta EG, Hellmiss R, Ream W (1986) *Overdrive*, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. The EMBO Journal 5: 1137-1142

Perkins DN, Pappin DJC, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20: 3551-3567

Proteau G, Sidenberg D, Sadowski P (1986) The minimal duplex DNA sequence required for site-specific recombination promoted by the FLP protein of yeast *in vitro*. Nucleic Acids Research 14: 4787-4802

Rastogi S, Pandey MM, Rawat AKS (2018) Ethnopharmacological uses, phytochemistry and pharmacology of genus *Adiantum*: A comprehensive review. Journal of Ethnopharmacology 215: 101-119

Rice DA, inventor, Jan. 23, 2001. Maize Promoters. United States Patent, Patent No. US 6,177,611 B1

Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. Trends in Genetics 16: 276-277

Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH (1998) *Bacillus thuringiensis* and Its Pesticidal Crystal Proteins. Microbiology and Molecular Biology Reviews 62: 775-806

Sedivy EJ, Wu F, Hanzawa Y (2017) Soybean domestication: the origin, genetic architecture and molecular bases. New Phytologist 214: 539-553

Seguin P, Zheng W, Smith DL, Deng W (2004) Isoflavone content of soybean cultivars grown in eastern Canada. Journal of the Science of Food and Agriculture 84: 1327-1332

Shearwin KE, Callen BP, Egan JB (2005) Transcriptional interference – a crash course. Trends in Genetics 21: 339-345

Shelton A (2012) Bacteria. Biological Control: A Guide to Natural Enemies in North America, http://www.biocontrol.entomology.cornell.edu/pathogens/bacteria.html

Sherman F, Stewart JW, Tsunasawa S (1985) Methionine or Not Methionine at the Beginning of a Protein. Bioessays 3: 27-31

Sidorenko L, Bevan SA, Larsen CM, Anthony GI, Robinson AE, Yerkes CN, inventors. Dec. 31, 2020. Compositions and methods for expressing transgenes using regulatory elements from chlorophyll binding Ab genes. United States Patent, Patent No. US 2020/0407742 A1

Sikorski RS, Hieter P (1989) A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in *Saccharomyces cerevisiae*. Genetics 122: 19-27

Simmons CR, Herman RA (2023) Non-seed plants are emerging gene sources for agriculture and insect control proteins. The Plant Journal 116: 23-37

Soberón M, Monnerat R, Bravo A (2016) Mode of Action of Cry Toxins from *Bacillus thuringiensis* and Resistance Mechanisms. In Microbial Toxins, pp 1-13

Spelman RJ, Bovenhuis H (1998) Moving from QTL experimental results to the utilization of QTL in breeding programmes. Animal Genetics 29: 77-84

Taylor NB, Fuchs RL, MacDonald J, Shariff AR, Padgette SR (1999) Compositional Analysis of Glyphosate-Tolerant Soybeans Treated with Glyphosate. Journal of Agricultural and Food Chemistry 47: 4469-4473

Tetreau G, Andreeva EA, Banneville A-S, De Zitter E, Colletier J-P (2021) How Does *Bacillus thuringiensis* Crystallize Such a Large Diversity of Toxins? Toxins 13: 443

The UniProt Consortium (2023) UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research 51: D523-D531

Thorpe HM, Smith MCM (1998) *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proceedings of the National Academy of Sciences of the United States of America 95: 5505-5510

Trieu-Cuot P, Courvalin P (1983) Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'5"-aminoglycoside phosphotransferase type III. Gene 23: 331-341

US-EPA (1998) R.E.D. Facts: *Bacillus thuringiensis*. United States Environmental Protection Agency, EPA-738-F-98-001

US-EPA (2001) Overview. In Biopesticides Registration Action Document: *Bt* Plant-Incorporated Protectants. United States Environmental Protection Agency, pp I1-I27, https://www3.epa.gov/pesticides/chem_search/reg_actions/pip/bt_brad2/1-overview.pdf

USDA-FAS (2024) Soybean 2023 World Production: 396,850 (1000 MT). United States Department of Agriculture, Foreign Agricultural Service, https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=2222000

USDA-NRCS (2023) *Adiantum trapeziforme* L., diamond maidenhair. United States Department of Agriculture - Natrual Resources Conservation Service, https://plants.usda.gov/home/plantProfile?symbol=ADTR4

van Ree R, Sapiter Ballerda D, Berin MC, Beuf L, Chang A, Gadermaier G, Guevera PA, Hoffmann-Sommergruber K, Islamovic E, Koski L, Kough J, Ladics GS, McClain S, McKillop KA, Mitchell-Ryan S, Narrod CA, Pereira Mouriès L, Pettit S, Poulsen LK, Silvanovich A, Song P, Teuber SS, Bowman C (2021) The COMPARE Database: A Public Resource for Allergen Identification, Adapted for Continuous Improvement. Frontiers in Allergy 2: 700533

Waigmann E, Gomes A, Lanzoni A, Perry JN (2013) Editorial: New Commission Implementing Regulation on Risk Assessment of GM plant applications: novel elements and challenges. The EFSA Journal 11: e11121

Wang J, Chu S, Zhang H, Zhu Y, Cheng H, Yu D (2016) Development and application of a novel genome-wide SNP array reveals domestication history in soybean. Scientific Reports 6: 20728

Weber N, Halpin C, Hannah LC, Jez JM, Kough J, Parrott W (2012) Editor's Choice: Crop Genome Plasticity and Its Relevance to Food and Feed Safety of Genetically Engineered Breeding Stacks. Plant Physiology 160: 1842-1853

Westfall PH, Tobias RD, Rom D, Wolfinger RD, Hochberg Y (1999) Concepts and Basic Methods for Multiple Comparisons and Tests. In Multiple Comparisons and Multiple Tests: Using SAS. SAS Institute Inc., Cary, NC, pp 13-40

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. Gene 33: 103-119

Zastrow-Hayes GM, Lin H, Sigmund AL, Hoffman JL, Alarcon CM, Hayes KR, Richmond TA, Jeddeloh JA, May GD, Beatty MK (2015) Southern-by-Sequencing: A Robust Screening Approach for Molecular Characterization of Genetically Modified Crops. The Plant Genome 8: 1-15

STUDY INDEX

[redacted] (2022) “IPD083Cb Protein: Acute Oral Toxicity Study in [redacted] Mice” Corteva Agriscience Study ID: PHI-2021-205

[redacted] (2023) “Cry1B.61.1 Protein: Acute Oral Toxicity Study in [redacted] Mice” Corteva Agriscience Study ID: PHI-2022-098

[redacted] (2023) “Comparison of the Amino Acid Sequence of the Cry1B.34.1 Protein to the Amino Acid Sequences of Allergens and Protein Toxins” Corteva Agriscience Study ID: PHI-R127-Y23/205

[redacted]. (2023) “Comparison of the Amino Acid Sequence of the Cry1B.61.1 Protein to the Amino Acid Sequences of Allergens and Protein Toxins” Corteva Agriscience Study ID: PHI-R126-Y23/205

[redacted]. (2023) “Comparison of the Amino Acid Sequence of the GM-HRA Protein to the Amino Acid Sequences of Allergens and Protein Toxins” Corteva Agriscience Study ID: PHI-R124-Y23/205

[redacted]. (2023) “Comparison of the Amino Acid Sequence of the IPD083Cb Protein to the Amino Acid Sequences of Allergens and Protein Toxins” Corteva Agriscience Study ID: PHI-R125-Y23/205

[redacted] (2021) “Cry1B.34: Acute Oral Toxicity Study in Mice” Corteva Agriscience Study ID: PHI-2020-188 (submitted with the A1281 application)

[redacted]. (2006) “GM-HRA: Acute Oral Toxicity Study in Mice” Corteva Agriscience Study ID: PHI-2006-008 (submitted with the A1018 application)

[redacted]. (2023) “Reading Frame Analysis at the Insertion Site of Soybean Event COR-23134-4” Corteva Agriscience Study ID: PHI-R100-Y23/225

[redacted] (2023) “Nutrient Composition of a Soybean Line Containing Event COR-23134-4” Corteva Agriscience Study ID: PHI-2022-014/020

APPENDIX A. METHODS FOR SOUTHERN-BY-SEQUENCING ANALYSIS

Test and Control Substances

The test substance is COR-23134-4 contained within the soybean seed. The presence or absence of the COR23134 insert was confirmed by event-specific qualitative real-time PCR analysis. Four plants containing the event (transgenic plants) and six plants not containing the event (null segregant plants) from the segregating T1 generation were used for SbS analysis.

The control substances are 1) non-genetically modified (non-GM) soybean that has the same genetic background (93Y21) as COR23134 soybean (the absence of the COR23134 insert was confirmed by event-specific qualitative real-time PCR analysis) and 2) [redacted] plasmid DNA.

DNA Extraction and Quantitation

Genomic DNA isolation was performed by Pioneer Hi-Bred International, Inc. Genomics Technologies (hereafter referred to as Genomics Technologies). Genomic DNA was extracted from leaf tissue of ten segregating plants of the T1 generation of COR23134 and two control soybean plants (one used as a control and one used as a positive control with [redacted] plasmid DNA diluted in the genomic DNA). The tissue was pulverized in tubes containing grinding beads using a Geno/Grinder (SPEX CertiPrep), and the genomic DNA was isolated using a Sbeadex DNA extraction kit (Biosearch Technologies). Following extraction, the DNA was quantified on a Lunatic (Unchained Labs) and visualized on the TapeStation 4200 (Agilent) to determine the DNA quality.

Southern-by-Sequencing

SbS (performed by Genomics Technologies) utilizes probe-based sequence capture, NGS techniques, and bioinformatics procedures to capture, sequence, and identify inserted DNA within the soybean genome (Zastrow-Hayes *et al.*, 2015). By compiling a large number of unique sequencing reads and aligning them against the linearized transformation plasmid sequence and control soybean genome, unique junctions due to inserted DNA are identified in the bioinformatics analysis. This information is used to determine the number and organization of insertions within the plant genome and confirm the absence of plasmid backbone sequences.

Genomic DNA isolated from ten plants of the T1 generation of COR23134 soybean was analyzed by SbS to determine the insertion copy number and organization and to confirm the absence of plasmid backbone or other unintended plasmid sequences. SbS was also performed on one control soybean DNA sample and a positive control sample ([redacted] DNA diluted in control soybean DNA) to confirm that the assay could reliably detect plasmid fragments within the genomic DNA.

The following processes were performed by Genomics Technologies using standard methods and were based on the procedures described in Zastrow-Hayes et al. (2015).

Capture Probe Design and Synthesis

Biotinylated probes (approximately 120 nucleotides long) for hybridization to plasmid sequences were designed and synthesized by [redacted]. The probe set was designed to target all sequences within the [redacted] plasmid.

Sequencing Library Construction

NGS libraries were constructed for DNA samples from individual soybean plants, including segregating plants from the T1 generation of COR23134 soybean, a control soybean plant, and the positive control sample [redacted] plasmid DNA diluted in control soybean DNA). Genomic DNA isolated as described above was sheared using the Covaris E220 focused-ultrasonicator to an average fragment size of 400 base pairs (bp). Sheared DNA was end-repaired, A-tailed, and ligated to NEXTFLEX Unique Dual Index adapters (PerkinElmer) following the KAPA HTP Library Preparation Kit (Roche) manufacturer's recommended protocol so that samples would be indexed to enable identification after sequencing. The DNA fragment libraries were amplified by PCR for eight cycles prior to the capture process. Amplified libraries were analyzed using a TapeStation 4200, diluted to 6 ng/µl with nuclease-free water, and pooled.

Probe Hybridization and Sequence Enrichment

A double capture procedure was used to capture and enrich DNA fragments that contained sequences homologous to the capture probes. The genomic DNA libraries described above were mixed with hybridization buffer and blocking oligonucleotides corresponding to the adapter sequences and denatured. Following denaturation, the biotinylated probes were added to the genomic DNA library and incubated at 47 °C for 16 hours. Streptavidin beads were added to the hybridization mix to bind DNA fragments that were associated with the probes. Bound fragments were washed and eluted, PCR-amplified for five cycles, and purified using KAPA HyperPure Beads (Roche). The enriched DNA libraries underwent a second capture reaction using the same conditions to further enrich the sequences targeted by the probes. This was followed by PCR amplification for 16 cycles and purification as described above. The final double-enriched libraries were quantified and diluted to 2 nM for sequencing.

Next Generation Sequencing on Illumina Platform

Following sequence capture, the libraries were submitted for NGS (Illumina MiSeq) to a total read depth of at least 300x for the captured sequences. The sequencing reads were trimmed [redacted] (Ewing and Green, 1998; Ewing *et al.*, 1998) and assigned to the corresponding individual plant based on the indexed adapters. A complete sequence set from each plant is referred to as "AllReads" for bioinformatics analysis of that plant.

Quality Assurance of Sequncing Reads

The adapter sequences were trimmed from the NGS sequence using Cutadapt, v2.10 (Martin, 2011). Further analysis to eliminate sequencing errors used JELLYFISH, version 2.2.10 (Marçais and Kingsford, 2011), to [redacted] within

"AllReads" as described in Zastrow-Hayes et al. (2015). This set of sequences was used for further bioinformatics analysis and is referred to as "CleanReads." Identical sequencing reads were combined into non-redundant read groups (referred to as "Non-redundantReads") while retaining abundance information for each group and were used for further analysis, as described in Zastrow-Hayes et al. (2015).

Aligning Reads

Each set of "Non-redundantReads" was aligned to the plasmid sequence, including the plasmid backbone sequence, using Bowtie2, version 2.3.4.2 (Langmead and Salzberg, 2012), [redacted] Remaining "Non-redundantReads" were aligned to the control soybean genome using Bowtie2, version 2.3.4.2, [redacted].

Junction Detection

Following removal of "Non-redundantReads" with alignments to the control soybean genome or plasmid sequence identified during the quality assurance phase, the remaining "Non-redundantReads" were aligned to the entire plasmid sequence using Bowtie2 version 2.3.4.2 with the soft-trimming feature enabled. Chimeric reads contain sequence that is non-contiguous with the plasmid sequence from the alignment, such as plasmid-genome junctions or rearrangements of the plasmid. These chimeric reads are referred to as junction reads or junctions [redacted]. This identifier (referred to as a 30_20 mer) includes 20 bp of sequence from [redacted] and 30 bp of sequence adjacent to the plasmid-derived 20 bp within a sequencing read. The adjacent 30 bp either did not align to the plasmid contiguously to the known 20 bp or aligned to the control soybean genome. When the 20 bp from the plasmid and the adjacent 30 bp were identified as a 30_20 mer, they indicated the junction shown by the chimeric read. [redacted] The total number of sequencing reads (referred to as "TotalSupportingReads") for each unique junction was retained for filtering. [redacted]

Junction Identification

Variations between the soybean endogenous elements used in plasmid [redacted] (Table 1; Figures 1 and 2) and the soybean genome may result in the identification of junctions that are due to these differences. To detect these endogenous junctions, a control soybean genomic DNA library was captured and sequenced in the same manner. The 30_20 mers of the endogenous junctions detected in the control sample were used to filter the same endogenous junctions in the COR23134 soybean samples, so that the only junctions remaining in the COR23134 samples are due to actual insertions derived from [redacted]

Data QC

The transgenic and null segregant samples were compared to the control sample and a quality check was performed. If regions of the plasmid backbone contain low to medium sequencing coverage compared to the control sample and no junctions were identified, the data was reviewed a second time. If no junctions were identified for these reads, there is no insertion of the plasmid into the genome.

SbS Results

Results for the control soybean, positive control, one transgenic COR23134 soybean plant (plant ID 437164754), and a representative null segregant (negative) plant (plant ID 437164750) are presented in the main body of this document (see section A.3 (c) *Molecular characterisation*).

Remaining transgenic COR23134 soybean plant results from SbS analysis are presented in Figures A1 to A3 below:

A. Alignment of NGS Reads to [redacted] T-DNA Region

[redacted]

B. Alignment of NGS Reads to [redacted]

[redacted]

Figure A.1. SbS Analysis for Transgenic COR23134 Soybean (Plant ID 437164755)
The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] T-DNA or [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in each plasmid derived from the soybean genome (identified by numbers; Table 3), while tan bars indicate genetic elements derived from other sources. **A)** SbS results for transgenic COR23134 soybean aligned against the [redacted] T-DNA region intended for insertion [redacted] bp; Figure 2) indicating that this plant contains the insertion. Arrows in the Junctions panel indicate the two plasmid-genome junctions (black arrows) and one plasmid-plasmid junction (red arrow) identified by SbS; the numbers below the arrows refer to the bp location of the junction relative to the [redacted] T-DNA (Figure 2). The insertion comprises bp [redacted] of the [redacted] T-DNA shown in Figure 2. The presence of two plasmid-genome junctions (Junctions [redacted]) demonstrates the presence of a single insertion in the COR23134 soybean genome. One plasmid-plasmid junction (Junction [redacted] indicates the location of a 21-bp deletion ([redacted] bp) identified in all plants containing the COR23134 insertion. **B)** SbS results for transgenic COR23134 soybean aligned against the plasmid [redacted] sequence [redacted] bp; Figure 1). Coverage was obtained for the elements between the Right and Left Borders transferred into COR23134 soybean; however, for clarity the junctions identified in panel A are not shown in this view. The absence of any other junctions to the [redacted] sequence shows that there are no additional insertions or backbone sequence present in COR23134 soybean.

A. Alignment of NGS Reads to [redacted] T-DNA Region

[redacted]

B. Alignment of NGS Reads to [redacted]

[redacted]

Figure A.2. SbS Analysis for Transgenic COR23134 Soybean (Plant ID 437164756)
The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] T-DNA or [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in each plasmid derived from the soybean genome (identified by numbers; Table 3), while tan bars indicate genetic elements derived from other sources. **A)** SbS results for transgenic COR23134 soybean aligned against the [redacted] T-DNA region intended for insertion [redacted] bp; Figure 2) indicating that this plant contains the insertion. Arrows in the Junctions panel indicate the two plasmid-genome junctions (black arrows) and one plasmid-plasmid junction (red arrow) identified by SbS; the numbers below the arrows refer to the bp location of the junction relative to the [redacted] T-DNA (Figure 2). The insertion comprises bp [redacted] of the [redacted] T-DNA shown in Figure 2. The presence of two plasmid-genome junctions (Junctions [redacted]) demonstrates the presence of a single insertion in the COR23134 soybean genome. One plasmid-plasmid junction (Junction [redacted]) indicates the location of a 21-bp deletion [redacted] bp) identified in all plants containing the COR23134 insertion. **B)** SbS results for transgenic COR23134 soybean aligned against the plasmid [redacted] sequence ([redacted] bp; Figure 1). Coverage was obtained for the elements between the Right and Left Borders transferred into COR23134 soybean; however, for clarity the junctions identified in panel A are not shown in this view. The absence of any other junctions to the [redacted] sequence shows that there are no additional insertions or backbone sequence present in COR23134 soybean.

A. Alignment of NGS Reads to [redacted] T-DNA Region

[redacted]

B. Alignment of NGS Reads to [redacted]

[redacted]

Figure A.3. SbS Analysis for Transgenic COR23134 Soybean (Plant ID 437164757)
The blue coverage graph shows the number of individual NGS reads aligned at each point on the [redacted] T-DNA or [redacted] plasmid using a logarithmic scale at the middle of the graph. Green bars above the coverage graph indicate endogenous genetic elements in each plasmid derived from the soybean genome (identified by numbers; Table 3), while tan bars indicate genetic elements derived from other sources. **A)** SbS results for transgenic COR23134 soybean aligned against the [redacted] T-DNA region intended for insertion [redacted] bp; Figure 2) indicating that this plant contains the insertion. Arrows in the Junctions panel indicate the two plasmid-genome junctions (black arrows) and one plasmid-plasmid junction (red arrow) identified by SbS; the numbers below the arrows refer to the bp location of the junction relative to the [redacted] T-DNA (Figure 2). The insertion comprises bp [redacted] of the [redacted] T-DNA shown in Figure 2. The presence of two plasmid-genome junctions (Junctions [redacted]) demonstrates the presence of a single insertion in the COR23134 soybean genome. One plasmid-plasmid junction (Junction [redacted] indicates the location of a 21-bp deletion ([redacted] bp) identified in all plants containing the COR23134 insertion. **B)** SbS results for transgenic COR23134 soybean aligned against the plasmid [redacted] sequence [redacted] bp; Figure 1). Coverage was obtained for the elements between the Right and Left Borders transferred into COR23134 soybean; however, for clarity the junctions identified in panel A are not shown in this view. The absence of any other junctions to the [redacted] sequence shows that there are no additional insertions or backbone sequence present in COR23134 soybean.

APPENDIX B. METHODS FOR SOUTHERN BLOT ANALYSIS

Test, Control, and Reference Substances

The test substance in the study is COR-23134-4 contained within the soybean seed. Seeds containing the test substance were from COR23134 soybean of the T1, T2, T3, T4, and T5 generations. The control substance was defined as seed from a non-genetically modified (non-GM) soybean line (93Y21). 93Y21 soybean has the same genetic background to the test substance; however, it does not contain the COR23134 soybean insertion.

Plasmid DNA of [redacted] that was used for *Agrobacterium*-mediated transformation to produce COR23134 soybean was defined as a reference substance. This plasmid was used as a positive control for Southern analysis to verify probe hybridization. The *cry1B.34.1*, *cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* probes used in the study were derived from plasmid [redacted].

DNA molecular weight markers for gel electrophoresis and Southern blot analysis were obtained from commercial vendors and were used as reference substances to determine approximate molecular weights of DNA fragments. For Southern analysis, DNA Molecular Weight Marker III and VII, Digoxigenin (DIG)-labeled (Roche), were used as size standards for hybridizing fragments.

Sample Collection, Handling, Identification and Storage

Seed from each of the five generations of COR23134 soybean and the control soybean were planted in a controlled environment at Pioneer (Johnston, Iowa, USA). Fresh leaf tissue samples from test and control soybean were harvested, stored frozen (≤ -50 °C freezer unit), and then lyophilized. Lyophilized tissue samples were shipped to Regulatory Sciences, Multi Crop Research Center, Pioneer Hi-Bred Private Limited at Hyderabad, at ambient temperature. Upon arrival, samples were stored frozen (≤ -50 °C freezer unit) until processing.

DNA Extraction and Quantification

Genomic DNA was isolated from leaf tissue of one plant each from five generations (T1, T2, T3, T4, and T5) of COR23134 soybean and one plant from the 93Y21 control soybean.

Lyophilized leaf samples were pulverized with steel beads in tubes using a paint shaker (AGS Transact Technology Ltd.). Care was taken to ensure leaf samples were ground sufficiently for DNA isolation. Genomic DNA was isolated using a high salt extraction buffer (2.0 M sodium chloride, 100 mM Tris-Hydrochloride pH 8.0, 50 mM sodium salt of EDTA, 3% β-mercaptoethanol (v/v), and 100 mM sodium metabisulphite) and sequentially precipitated using potassium acetate and isopropyl alcohol. Extracted DNA was treated with Ribonuclease A (RNase A), purified using phenol/chloroform/isoamyl alcohol (25:24:1), and precipitated using sodium acetate and chilled ethanol. The purified DNA was quantified using Quant-iT™ PicoGreen® reagent (Molecular Probes, Invitrogen) and visualized on a 1% agarose gel to check the quality of the isolated DNA.

<u>Digestion of DNA and Electrophoretic Separation</u>

Genomic DNA (10 µg) isolated from both COR23134 soybean and control soybean leaves was digested with the restriction enzyme *Bst*1107 I (Thermo Fisher Scientific). Plasmid [redacted] DNA was added to additional control soybean DNA samples at a level equivalent to one plasmid copy per genomic copy and digested in the same manner. Following digestion with the restriction enzyme, the fragments produced were electrophoretically separated on a 0.9% agarose gel. After electrophoresis, the gel was stained using GelRed (Biotium Inc.) and documented by photographing the gel under UV illumination (BioRad Gel doc XR^+ System).

<u>Southern Transfer</u>

The DNA fragments separated on the agarose gel were denatured *in situ*, transferred to a nylon membrane (GE Healthcare, LC) using vacuum blotter (BioRad) and fixed to the membrane by UV crosslinking (UV Stratalinker, UVP).

<u>Probe Labeling and Southern Blot Hybridization</u>

DNA probes specific to the *cry1B.34.1*, *cry1B.61.1*, *ipd083Cb*, and *gm-hra_1* gene elements (Figure 2) were labeled by incorporation of Digoxigenin (DIG) labeled nucleotide [DIG-11-dUTP] into the fragments by PCR labeling method. Detailed descriptions of these probes are provided in Table 1.

Labeled probes were hybridized to the DNA on the nylon membrane for detection of the specific genomic DNA fragments. DNA Molecular Weight Marker III and VII, Digoxigenin (DIG)-labeled (Roche), were used for visualization as the fragment size standards on the blot.

<u>Detection of Hybridized Probes</u>

After overnight hybridization, the membrane was washed and processed using the DIG Wash and Block Buffer Set (Roche). DIG-labeled DNA standards and single stranded DIG-labeled probes hybridized to DNA bound to the nylon membrane were visualized using CDP-Star Chemiluminescent Nucleic Acid Detection System. Blots were exposed for one or more time points to detect hybridized fragments and to visualize molecular weight standards. Images were captured by the Syngene G-Box Chemi XX6 (Syngene, Inc.). Detected bands were documented for each probe.

<u>Stripping of Probes and Subsequent Hybridization</u>

Following hybridization and detection, membranes were stripped of DIG-labeled probe to prepare the blots for subsequent re-hybridization to a different probe. Membranes were rinsed briefly in distilled and de-ionized water and then stripped in a solution of 0.2 N NaOH and 0.1% SDS at 37 °C with constant shaking. The membranes were then rinsed in 2x Saline sodium citrate and either used directly for subsequent hybridizations or stored for later use. The alkali-based stripping procedure effectively removed probes labeled with alkali-labile DIG used in these experiments.

APPENDIX C. METHODS FOR MULTI-GENERATION SEGREGATION ANALYSIS

Greenhouse Experimental Design

Six generations of COR23134 soybean (T1, T2, T3, T4, T5, and T6) were evaluated using polymerase chain reaction (PCR) analyses and herbicide-tolerance testing to confirm Mendelian inheritance of genotype and phenotype.

Planting, Thinning, and Plant Selection

Seeds were planted in separate pots (one seed per pot). All seeds were grown in a controlled environment under suitable conditions for producing soybean plants.

For T1 generation plants intended for genotypic analysis and tissue production, 25 and 15 seeds were planted over two plantings, respectively. After copy number confirmation via qPCR, T1 generation plants intended for tissue production were thinned by selecting a total of three to four plants confirmed positive and six to seven plants confirmed negative over both plantings.

For T2 and T3 generation plants intended for genotypic analysis and tissue production, 100 seeds were planted per entry. After copy number confirmation via qPCR, T2 and T3 generation plants intended for tissue production for molecular analysis were thinned by selecting three to four plants confirmed positive and three to five plants confirmed negative per entry. T2 and T3 generation plants intended for tissue production for ELISA analysis were thinned by selecting five plants confirmed positive per entry.

For T4 and T5 generation plants intended for genotypic and phenotypic analysis and tissue production, 200 total seeds were planted per entry over two plantings (100 seeds per entry per planting). After trait confirmation via PCR, T4 and T5 generation plants from the first planting intended for tissue production were thinned by selecting seven to eight plants confirmed positive. All seeds from the second planting for the T4 and T5 generations that successfully germinated and developed into plants were maintained and used for phenotypic analysis.

For T6 generation plants intended for genotypic and phenotypic analysis, 100 seeds were planted. All seeds for the T6 generation that successfully germinated and developed into plants were maintained and used for phenotypic analysis.

For control soybean plants intended for genotypic and/or phenotypic analysis and tissue production, 15, 100, and 100 seeds were planted over three plantings, respectively. After trait confirmation of the absence of the COR23134 soybean event via PCR, control soybean plants from the first planting intended for tissue production were thinned by selecting eight to nine plants confirmed negative. All seeds from the second planting for the control soybean that successfully germinated and developed into plants were maintained and used for phenotypic analysis. All control soybean seeds planted for the third planting that successfully germinated and developed into plants were maintained and used for phenotypic analysis.

Genotypic and Phenotypic Analyses

Qualitative and Quantitative PCR Sample Collection and Genotypic Analyses

Prior to planting, seed chip samples were collected from seeds representing all six generations and the control soybean. Using a nail clipper, a small 'chip' of the seed containing seed coat and cotyledon was removed from each seed and placed into individual wells of 96-well collection plates. Each seed (now with a chip removed) was placed into a corresponding well within 96-well bubble packs. Each plate and well and corresponding bubble pack were uniquely labeled to allow a given seed chip sample to be tracked back to the originating seed. The seed chip samples were used for genotypic trait confirmation of the target event by qualitative endpoint PCR.

For the T1, T2, and T3 generations, leaf punch samples were collected (one leaf punch per plant) at the VC or V1 growth stage. Individual plants and corresponding leaf punch samples were uniquely labeled to allow a given sample to be tracked back to the originating plant and/or seed chip. The leaf punch samples were used for determination of the copy number of the COR23134 soybean event and the following genes of interest by qPCR: *gm-hra_1*, *cry*1B.34.1, *cry*1B.61.1, and *ipd083Cb*.

Herbicide Application and Evaluation

At the V3 growth stage, diclosulam was applied to the T4, T5, and T6 generations of COR23134 soybean. The spray mixture consisted of Strongarm containing 0.84 pounds of diclosulam per pound (0.84 kg diclosulam per kg) and Methylated Seed Oil adjuvant (MSO Concentrate with Leci-Tech) at a rate of 1.2% volume per volume. No other adjuvants or additives were included in the spray mixture. Strongarm was applied at a target rate of 0.71 ounces per acre (49.74 grams per hectare) with a target spray volume of approximately 20 gallons per acre (187 liters per hectare) using a spray chamber to simulate a broadcast (over-the-top) application. Actual application rates were within +/-10% of the target herbicide application rate.

Nine days after herbicide application, each plant was visually evaluated for herbicide tolerance in which presence of herbicide injury corresponded to an herbicide-susceptible phenotype and absence of herbicide injury corresponded to an herbicide-tolerant phenotype.

Statistical Analysis

A chi-square test was performed at the 0.05 significance level on the segregation results of T1, T2, and T3 generations of COR23134 soybean. A chi-square test was performed separately for each generation to compare the observed segregation ratio to the expected segregation ratio (3:1 for T1, T2, and T3). A chi-square test was not performed for the T4, T5, and T6 generations of COR23134 soybean as all plants were identified as positive as expected for a homozygous generation. Statistical analyses were conducted using SAS software, Version 9.4 (SAS Institute Inc.).

APPENDIX D. METHODS FOR SANGER SEQUENCING ANALYSIS

Test and Control Substances

The test substance in the study was defined COR-23134-4 contained within the soybean seed. The control substance was defined as seed from a non-genetically modified (non-GM) soybean line, 93Y21, that has a similar genetic background to the test substance but does not contain the COR23134 soybean insertion.

DNA Extraction and Quantification

Genomic DNA (gDNA) from COR23134 soybean and non-GM control soybean was extracted from finely ground fresh leaf tissue using the Synergy 2.0 Plant DNA Extraction Kit (OPS Diagnostics). For COR23134 soybean, approximately 500 grams of leaf tissue from each of eight plants was pooled and used for extraction. For the non-GM control soybean, 500 mg of leaf tissue from one plant was used for extraction.

Tissue samples were extracted based on the manufacturer's instructions with minor modifications. Briefly, each sample was homogenized with kit-supplied homogenization buffer and grinding matrix and centrifuged. Additional homogenization buffer was added to the pelleted cell debris, heated, and centrifuged. The supernatants were combined and treated with RNase A. Cold isopropanol was added to each sample. The mixture was loaded onto Silica Spin columns, washed twice with cold 70% ethanol, and eluted in 1X TE, pH 8. The quality of the extracted gDNA was assessed by agarose gel electrophoresis and visualized under ultraviolet light, and the concentration was determined using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific).

Polymerase Chain Reaction (PCR) Amplification of the Insert and Flanking Genomic Regions in COR23134 Soybean

All PCR and sequencing primers were designed based on the sequences of [redacted] and the soybean genome. PCR primers were designed to amplify seven overlapping PCR fragments (A, B, C, D, E, F, and G) spanning the insert and its 5' and 3' flanking genomic regions. PCR Fragments A and G contain sequence from the insert and either the 5' or 3' flanking genomic region, whereas PCR Fragments B, C, D, E, and F contain sequence from [redacted] M13 forward and reverse primers and multiple internal sequencing primers were used for Sanger sequencing. All primers were synthesized by Integrated DNA Technologies.

Two independent PCR reactions were performed for each fragment using gDNA extracted from COR23134 soybean. All PCR fragments were generated using 160 ng of gDNA as a template with each primer at a final concentration of 0.4 µM in a 50-µl reaction. Phusion High Fidelity PCR Master Mix with GC Buffer (Thermo Fisher Scientific) was used to amplify all PCR fragments, except for Fragment G, which used Phusion Green HSII High Fidelity PCR Master Mix (Thermo Fisher Scientific). Fragment F also used 1 M betaine in the PCR reaction. PCR conditions optimized to yield targeted PCR products are detailed in Table D.1. The gDNA extracted from non-GM control soybean and a no-template control (NTC) served as negative controls. All PCR products were confirmed as a dominant band at the expected size by agarose gel electrophoresis and visualized under ultraviolet light.

Table D.1. PCR Fragment Amplification Conditions for COR23134 Soybean

Cycles	A	B	C	D	E	F	G
1×	98 ºC 2'	98 ºC 2'	98 ºC 2'	98 ºC 2'	98 ºC 2'	98 ºC 2'	98 ºC 2'
30×[a]	98 ºC 15" 65 ºC 10" 72 °C 2'	98 ºC 15" 65 ºC 10" 72 °C 2'	98 ºC 15" 65 ºC 10" 72 °C 2'	98 ºC 15" 65 ºC 10" 72 °C 2'	98 ºC 15" 65 ºC 10" 72 °C 2'	98 ºC 30" 65 ºC 10" 68 °C 2'	98 ºC 15" 65 ºC 10" 72 °C 3'
1×	72 ºC 10' 4 °C ∞	72 ºC 10' 4 °C ∞	72 ºC 10' 4 °C ∞	72 ºC 10' 4 °C ∞	72 ºC 10' 4 °C ∞	72 ºC 10' 4 °C ∞	72 ºC 10' 4 °C ∞

Note: Time is indicated in minutes (') and seconds (").
[a] All PCR fragments used 30 cycles, except Fragment G which used 35 cycles.

Cloning of PCR Products

PCR products from two independent reactions of each fragment were separately cloned into a pCR BluntII-TOPO vector using a Zero Blunt TOPO PCR Cloning Kit (Invitrogen). At least five individual colonies from each transformation were inoculated for liquid bacterial culture, and subsequently the plasmid DNA was isolated from each culture using QIAprep Spin Miniprep Kit (Qiagen). The presence of the PCR insert within the plasmid DNA was confirmed by restriction enzyme digestion, and plasmids containing the PCR insert were quantified using a NanoDrop 2000 spectrophotometer.

Sanger DNA Sequencing

Six plasmids (three from each of the two independent PCR reactions) for each PCR fragment, except Fragments E and F (ten plasmids; five from each of the two independent PCR reactions), were sequenced in both forward and reverse directions to cover every nucleotide by Sanger sequencing (Eurofins Genomics; Louisville, KY, USA). Sequencher 5.4.6 (Gene Codes Corporation) was used to analyze and assemble the sequences using default parameters. Low-quality data determined by the analysis software and cloning vector sequence were trimmed from the 5' and 3' ends of each trace file prior to assembly. All sequencing reads were manually reviewed, and any ambiguous nucleotides were visually verified from the original chromatograms and compared with the sequencing reads from the other plasmids to make a final base call.

Sequencing reads from the six (ten for Fragments E and F) plasmids were used to determine the consensus sequence for each PCR fragment. The consensus sequences of all seven overlapping fragments were combined to determine the sequence for COR23134 soybean, and the determined sequence was compared with the sequence of the [redacted] T-DNA.

APPENDIX E. METHODS FOR CHARACTERIZATION OF THE CRY1B.34.1 PROTEIN

<u>Test Materials</u>

COR23134 Soybean-Derived Cry1B.34.1 Protein

Cry1B.34.1 protein was isolated from whole plant tissue derived from COR23134 soybean. The whole plant tissue was collected at the V5 growth stage (the stage when the leaflets on the sixth leaf node have unrolled (Pedersen, 2004)) of development from plants grown at a Pioneer owned field location (Johnston, IA, USA). The tissue was lyophilized, homogenized, and stored frozen at -80°C. The Cry1B.34.1 protein was extracted from lyophilized soybean tissue by homogenization with a Waring blender vessel using phosphate-buffered saline containing polysorbate 20 (PBST) extraction buffer. The sample extract was then clarified by centrifugation and filtration. The filtered extract was purified by immunoaffinity chromatography. The immunoaffinity columns were prepared by coupling a Cry1B.34.1 polyclonal antibody (R11955; Pioneer) to AminoLink Plus Coupling Resin (Thermo Scientific). The Cry1B.34.1 protein sample was eluted off the column using IgG Elution buffer (Thermo Scientific). Select elutions were collected and concentrated using a centrifugal concentrator (30K Vivaspin; Sartorius) to a volume of approximately 500 µl. The concentrated sample was buffer exchanged using 50 mM Tris buffer, pH 8, and then concentrated again to a volume of approximately 300 µl.

Following extraction, purification, and concentration, the final volume in the concentrator was estimated, NuPAGE LDS Sample Buffer (Life Technologies) was added at 25% along with 10% NuPAGE Reducing Agent containing DTT (Life Technologies) to the concentrated sample in the concentrator. The sample in the concentrator was transferred to a microcentrifuge tube, heat-treated, and stored frozen (-20 °C freezer unit) until use in characterization of the Cry1B.34.1 protein.

Microbially Derived Cry1B.34.1 Protein

The Cry1B.34.1 protein was produced by [redacted] for Pioneer. The protein was expressed in an *Escherichia coli* protein expression system and then purified using immobilized metal affinity chromatography. Tangential flow filtration was used to change the buffer to 50 mM ammonium bicarbonate. After lyophilization and mixing, a lot number was assigned.

<u>SDS-PAGE Analysis</u>

The purified COR23134 soybean-derived Cry1B.34.1 protein sample was allowed to thaw, diluted as applicable for the sensitivity of the assay, heated, and then loaded into 4-12% Bis-Tris gels along with pre-stained protein molecular weight markers (Precision Plus Protein Dual Xtra Standards). For applicable SDS PAGE and western blot analysis, the Cry1B.34.1 protein reference substance was diluted in 1X LDS/DTT, heated, and loaded into the gels to 1 µg for SDS PAGE and 10 ng for western blot analysis. Electrophoresis was conducted using a pre-cast gel electrophoresis system with MES running buffer at a constant 200 volts for 35 minutes.

Upon completion of electrophoresis, the gels were either prepared for protein staining or protein transfer to a membrane for sequencing or western blot analysis.

For Coomassie staining, following electrophoresis, the gel was washed with water and stained with GelCode Blue Stain Reagent (Thermo Scientific). Following staining, the gel was de stained with water until the gel background was clear. Protein bands were stained on the gels and the gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

Following SDS-PAGE as described above, the resulting gel was assembled into a nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in PBST containing 5% weight/volume (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with a Cry1B.34.1 monoclonal mouse antibody 10C2.D4.H10 (Pioneer Hi-Bred International, Inc.) diluted 1:5000 in PBST containing 1% w/v non-fat dry milk. Following primary antibody incubation, the membrane was washed with PBST. The membrane was incubated with a secondary antibody (anti-mouse IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:10,000 in PBST containing 1% w/v non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP (Bio Rad) imaging system.

Protein Glycosylation Analysis

A Pierce Glycoprotein Staining Kit (Thermo Fisher) was used to determine whether the COR23134 soybean-derived Cry1B.34.1 protein was glycosylated. The purified soybean-derived Cry1B.34.1 protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor) were run by SDS PAGE as described above.

Following electrophoresis, the gel was washed with water, fixed with 50% methanol, and washed with 3% acetic acid. The gel was then incubated with oxidizing solution and washed with 3% acetic acid. The gel was incubated with glycoprotein staining reagent and then incubated in a reducing reagent. The gel was then washed with 3% acetic acid followed by water. Glycoproteins were detected as stained bands on the gel.

Following glycoprotein detection, the image of the gel was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system. The same gel was then stained with GelCode Blue Stain Reagent (Thermo Scientific) followed by washes with water to visualize all protein bands. The image of the GelCode-stained gel was then captured electronically.

Peptide Mapping by Mass Spectrometry

Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described above, the COR23134 soybean-derived Cry1B.34.1 protein was loaded onto a gel on two separate days. The first day, the COR23134 soy-derived Cry1B.34.1 protein was loaded onto a gel in each of two lanes and the second day in each of four lanes. Following SDS PAGE, Coomassie staining, and gel imaging using the methods as described above, bands containing the soybean derived Cry1B.34.1 protein were excised from a gel and stored frozen (20 °C freezer unit). The protein in each gel slice was reduced with DTT, alkylated with iodoacetamide, and then subsequently digested with trypsin or chymotrypsin. For the analysis day with four bands, two bands were combined for each digestion and the combined digests were concentrated to about 50% of the original volume under a stream of nitrogen prior to analysis. The digested samples were separated on a nanoACQUITY UPLC (Waters Corporation) fitted with a Peptide BEH C18 300 Å 1.7 µm column (75 µm x 100 mm; Waters Corporation) by gradient elution. Eluent from the column was directed into an electrospray source, operating in positive ion mode, on a TripleTOF 5600+ hybrid quadrupole-TOF mass spectrometer (AB Sciex; currently Sciex). The resulting mass spectrometry (MS) data were processed using MSConvert to produce a peak list. The peak list was used to perform an MS/MS ion search (Mascot Software version 2.8.0) and match peptides from the expected Cry1B.34.1 protein sequence (Perkins *et al.*, 1999). The following search parameters were used: peptide and fragment mass tolerance, ± 0.1 Da; fixed modifications, cysteine carbamidomethyl; variable modifications, methionine oxidation; and maximum missed cleavages, 1 for trypsin and 2 for chymotrypsin. The Mascot-generated peptide ion score threshold was > 13 which indicates identity or extensive homology ($P < 0.05$). The combined sequence coverage was calculated with GPMAW version 12.11.0.

N-Terminal Amino Acid Sequencing Analysis

For N-terminal amino acid sequence analysis, COR23134 soybean-derived Cry1B.34.1 protein was loaded on to a gel. Following SDS-PAGE using the methods as described above the resulting gel was incubated in cathode buffer (60 mM Tris, 40 mM CAPS, 0.075% SDS, pH 9.6). An Immobilon-P PVDF membrane (Millipore) was wetted in 100% methanol, followed by immersion in anode buffer (60 mM Tris, 40 mM CAPS, 15% methanol, pH 9.6). A Trans-Blot SD Semi Dry Electrophoretic Transfer Cell system (Bio-Rad) was used to transfer proteins from the gel to the membrane. Following protein transfer, the membrane was washed with water, stained with GelCode Blue Stain Reagent (Thermo Scientific), and then destained with water to visualize the Cry1B.34.1 protein band. A band containing the Cry1B.34.1 protein was excised and stored frozen (-20 °C freezer unit). The band was analyzed using a Shimadzu PPSQ-51A sequencer. Ten cycles of Edman sequencing were performed. During each cycle, the N terminal amino acid was sequentially derivatized with phenylisothiocyanate (PITC), cleaved with trifluoracetic acid, and converted to PTH amino acid which was identified through chromatography. LabSolutions Software was used to automatically identify the N-terminal sequence.

Bioactivity Bioassay

The biological activity of Cry1B.34.1 protein was evaluated by conducting a 7-day bioassay using *Spodoptera frugiperda* (fall armyworm; Lepidoptera: Noctuidae), a species sensitive to the Cry1B.34.1 protein. Eggs were obtained from [redacted] and their identity was recorded by study personnel.

The bioassay utilized a generalized randomized block design containing four blocks. Each block consisted of a 12-well bioassay plate and contained five replicates from each treatment for a target of 20 individuals per treatment. *S. frugiperda* larvae were exposed via oral ingestion to one of the following two treatments:

- Treatment 1: Buffer Control Diet (containing 10 mM CAPS buffer)
- Treatment 2: Test Diet (targeting 15 ng Cry1B.34.1 protein per mg wet diet)

On each day of diet preparation (Day 0 and Day 4), the Cry1B.34.1 protein test substance was solubilized in 10 mM CAPS buffer, pH 10.5 (referred to as buffer), to a target concentration of 2 mg per ml. The solubilized test substance was diluted in buffer to achieve the concentration in the test dosing solution (20 ng/µl). The buffer control dosing solution consisted of 10 mM CAPS buffer, pH 10.5. Dosing solutions were maintained chilled (on wet ice) until use. The carrier for the *S. frugiperda* bioassay consisted of an artificial Stonefly Heliothis diet. On each day of diet preparation, each dosing solution was mixed with carrier in a 3:1 ratio (i.e., 3 ml of dosing solution to 1 g of carrier), generating Treatments 1 and 2.

S. frugiperda eggs were incubated in an environmental chamber until the eggs hatched. Neonates were used in the bioassay within 24 hours of hatching. On Day 0 of the bioassay, approximately 300 µl (i.e., 1 g of wet diet equated to 1 ml of wet diet) of freshly prepared diets were dispensed into individual wells of the bioassay plate and one *S. frugiperda* neonate was placed in each well containing diet. Each bioassay plate was sealed with heat-sealing film and ventilated with a small hole over each well. The bioassay was conducted in an environmental chamber set at 25 °C, 65% relative humidity, and continuous dark. On Day 4, new bioassay plates were prepared with fresh diet as described for Day 0 with the exception that 600 µl of freshly prepared diet were dispensed per well, surviving organisms were transferred to the new plates, and the plates were placed in the environmental chambers. After 7 days, the bioassay was complete, mortality was assessed, and surviving organisms were individually weighed.

The bioassay acceptability criterion indicated the bioassay may be repeated if the combined number of dead and missing organisms exceeded 20% for the buffer control diet (Treatment 1) group. The bioassay met the acceptability criterion with 0% dead and missing organisms.

Control of bias in the bioactivity assay was achieved through the use of a control diet and the random allocation of treatments within each block.

Thermolability Analysis

The test substance consisted of Cry1B.34 protein solubilized from a lyophilized powder (lot number PCF-0042). The carrier consisted of Stonefly Heliothis diet. The buffer control dosing solution used to prepare Treatment 1 consisted of 10 mM CAPS. The bulk dosing solution used to prepare Treatments 2-6 consisted of aliquots of the test substance diluted in 10 mM CAPS to achieve the targeted concentration in Treatments 2-6. The dosing solution aliquots used to prepare Treatments 3-6 were incubated for 30-35 minutes at several targeted temperatures.

The test system was *Spodoptera frugiperda* (fall armyworm; Lepidoptera: Noctuidae). The test system was chosen because *S. frugiperda* is an insect sensitive to the Cry1B.34 protein. *S. frugiperda* larvae were exposed via oral ingestion to one of the following six treatments:

- Treatment 1: Buffer Control Diet containing 10 mM CAPS
- Treatment 2: Control Diet containing the unheated Cry1B.34 protein dosing solution

- Treatment 3: Test Diet containing the Cry1B.34 protein dosing solution incubated at 25 °C
- Treatment 4: Test Diet containing the Cry1B.34 protein dosing solution incubated at 50 °C
- Treatment 5: Test Diet containing the Cry1B.34 protein dosing solution incubated at 75 °C
- Treatment 6: Test Diet containing the Cry1B.34 protein dosing solution incubated at 95 °C

The unheated control diet and each test diet contained a targeted concentration of 25 ng Cry1B.34 protein per mg diet wet weight. Treatments were arranged in a generalized randomized block design with a total of 10 blocks. Each block consisted of a 12-well bioassay plate and contained 2 replicates from each treatment. On Day 0, each treatment was provided to a target of 20 *S. frugiperda* individuals. The bioassay was conducted in an environmental chamber set at 25 °C, 65% relative humidity, and continuous dark. Larvae were refed on Day 4. After 7 days, the bioassay was complete, final mortality was assessed, and surviving organisms were individually weighed.

The bioassay acceptability criteria were as follows: The bioassay may be terminated and repeated if the combined number of dead and missing organisms is greater than 20% for the buffer control diet (Treatment 1) group. The bioassay may be terminated and repeated if the mortality of the unheated control diet (Treatment 2) group does not exceed 80%. The *S. frugiperda* bioassay met both acceptability criteria. An enzyme linked immunosorbent assay (ELISA) was used to assess the homogeneity of the Cry1B.34 protein in Treatment 2 and concentration of the Cry1B.34 protein dosing solutions. The absence of Cry1B.34 protein in the buffer control dosing solutions was also assessed. Bias in the *S. frugiperda* bioassay was controlled through the randomization of treatments within blocks and the use of one or more control diets. Bias in the characterization portion of the study was controlled through the use of replicate testing and appropriate assay controls.

On each day of diet preparation, dosing solutions for Treatments 1-6 were prepared. Each dosing solution was mixed with carrier in a 3:1 ratio (i.e., 3 ml of dosing solution to 1 g of carrier), generating Treatments 1-6. Dosing solutions were maintained chilled (in a refrigerator set at 4 °C or on wet ice) until use. *S. frugiperda* eggs were incubated in an environmental chamber until the eggs hatched. *S. frugiperda* neonates were used in the bioassay within 24 hours of hatching.

On Day 0, approximately 300 µl (i.e., 1 g of wet diet equated to 1 ml of wet diet) of freshly prepared diets were dispensed into wells of the bioassay plates. One *S. frugiperda* neonate was placed in each well containing diet. Each bioassay plate was sealed with heat-sealing film, a small hole was poked over each well to allow for ventilation, and the plates were placed in an environmental chamber. On Day 4, new bioassay plates were prepared with fresh diet as described for Day 0, with the exception that 600 µl was dispensed per well. Living *S. frugiperda* larvae were transferred to the new plates, missing or dead larvae were recorded, and the freshly prepared plates were placed in the environmental chamber. After 7 days, the bioassay was complete, mortality was assessed, and surviving larvae were individually weighed.

Statistical Analysis

Statistical analyses of data were conducted using SAS software, Version 9.4 (SAS Institute, Inc.). The response variable of interest was mortality. Statistical comparisons were made between mortality of *S. frugiperda* fed diet containing heat-treated Cry1B.34 protein (Treatments 3, 4, 5 and 6) and that of *S. frugiperda* fed the unheated Cry1B.34 protein control diet (Treatment 2).

Statistical analysis was conducted using Fisher's exact test to determine if the mortality rate of *S. frugiperda* fed diets containing the heat-treated Cry1B.34 protein (m_T) was lower than the mortality rate of those fed the unheated Cry1B.34 protein control diet (m_C). The corresponding hypothesis test was

$$H_0: m_T - m_C = 0 \quad vs. \quad H_a: m_T - m_C < 0.$$

A significant difference was established if the P-value was < 0.05. SAS PROC MULTTEST was utilized to conduct the Fisher's exact test.

Digestibility in Simulated Gastric Fluid (SGF)

Test and control solutions were prepared as follows:

- The gastric control solution was prepared fresh on the day of use and was comprised of 0.2% weight per volume (w/v) NaCl in 0.7% volume per volume (v/v) HCl, with a pH of ~1.2.
- The pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use by dissolving pepsin (Sigma-Aldrich) into gastric control solution. The SGF was prepared so that the pepsin to protein ratio of the final digestion mixture was 10 units of pepsin per µg of test substance.
- The test substance consisted of Cry1B.34 protein solubilized from a lyophilized powder (PCF-0042).
- To prepare the stock solutions for each of the control proteins (BSA and β-lactoglobulin), a 5.0-mg sub-sample of powder was weighed into an individual tube for each protein and solubilized by adding 1 ml of 10 mM CAPS buffer to a target protein concentration of 5 mg/ml.
- The final concentration of protein and pepsin in the control digestion mixtures was 0.25 mg/ml Cry1B.34 protein or control protein and 2500 units/ml pepsin.

SGF solution (1900 µl) was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to the addition of 100 µl of Cry1B.34 protein test substance at Time 0. The digestion reaction mixture was mixed constantly using a stir bar and a submersible magnetic stirrer.

A 120-µl sub-sample of the Cry1B.34 protein digestion reaction mixture was removed from the vial at the following analytical time points (± 10 seconds): 0.5, 1, 2, 5, 10, 20, and 60 minutes. The sub-samples were inactivated by adding them to pre-labeled tubes containing 139 µl of a pre-mixed sample stop solution (consisting of 48 µl of 200 mM sodium carbonate, 65 µl NuPAGE 4X LDS sample buffer, and 26 µl NuPAGE 10X sample reducing agent) and heating to 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

To prepare control digestion samples at 1 and 60 minutes, a 114-µl sample of the appropriate digestion solution (see table) was pre-warmed in a 37 °C water bath for 2-5 minutes prior to adding 6 µl of the Cry1B.34 protein test substance, control protein stock solution, or 10 mM CAPS buffer. The tubes were incubated in the water bath for the allotted time and then inactivated by mixing with 139 µl of the pre-mixed sample stop solution.

The Time 0 control reaction mixtures were prepared by first neutralizing 114 µl of the appropriate digestion solution (see table) with 139 µl of the pre-mixed sample stop solution, and then adding 6 µl of the Cry1B.34 protein test substance, control protein stock solution, or 10 mM CAPS buffer to the appropriate tube and mixing.

Control digestion samples included in the SGF assay are provided in Table E.1. Following digestion and inactivation, all control reaction mixtures were heated at 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

Table E.1. Control Samples for Simulated Gastric Fluid (SGF) Digestibility Analysis

Protein	Digestion Solution	Digestion Time (min)		
		0	1	60
None (10 mM CAPS buffer) (SGF Control)	SGF	X	--	X
BSA	SGF	X	X	X
β-lactoglobulin	SGF	X	X	X
Cry1B.34	SGF	X	--	--
Cry1B.34	None (10 mM CAPS buffer)	X	--	X
Cry1B.34	Gastric Control Solution (No Pepsin)	--	--	X

SDS-PAGE Analysis

The Cry1B.34 protein digestion time-course samples and control digestion samples were removed from frozen storage, heated at 90-100 °C for 5 minutes, and loaded (20 µl/well) into 4-12% Bis-Tris gels for SDS-PAGE analysis. To demonstrate the sensitivity of the SDS-PAGE gel and western blot analyses, an aliquot of the Cry1B.34 protein in SGF (Time 0) sample was loaded into the gel at a 1:20 dilution (116 ng Cry1B.34 protein) for protein staining, and at a 1:200 dilution (11.6 ng Cry1B.34 protein) for the western blot. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into the gels to provide a visual estimate of molecular weight. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gels were removed from the gel cassettes for use in protein staining or western blot analyses. For protein staining, the gels were washed with water three times for 5 minutes each and stained with GelCode Blue Stain Reagent for 61 minutes. Following staining, the gels were destained with water four times for a minimum of 5 minutes each until the gel background was clear. Proteins were stained as blue-colored bands on the gels. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

The Cry1B.34 protein digestion time-course samples were also analyzed by western blot. Following SDS-PAGE, the gel intended for western blot analysis was assembled into a nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane for 7 minutes with a pre-set program (P3).

Following protein transfer, the membrane was blocked in phosphate-buffered saline containing polysorbate 20 (PBST) containing 5% (w/v) non-fat dry milk for 46 minutes at ambient laboratory temperature. Before and after the blocking step, the membrane was washed with PBST three times for at least 5 minute each to reduce the background. The blocked membrane was incubated for 45 minutes at ambient laboratory temperature with a Cry1B.34 polyclonal antibody R11957 (Pioneer Hi-Bred International, Inc.) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. Following primary antibody incubation, the membrane was washed in PBST three times for 5 minutes each. The membrane was incubated for 48 minutes at ambient laboratory temperature with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. The membrane was washed in PBST four times for 5 minutes each. The membrane remained in PBST prior to incubating with a chemiluminescent substrate for 5 minutes. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP imaging system.

<u>Digestibility in Simulated Intestinal Fluid (SIF)</u>

Test and control solutions were prepared as follows:

- The pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by dissolving 26.4 mg of pancreatin (Sigma-Aldrich) into 5 ml of intestinal control solution (I-Con 1X buffer) to a final concentration of 0.5% weight per volume (w/v) pancreatin and 50 mM KH_2PO_4, pH 7.5.
- The test substance consisted of Cry1B.34 protein solubilized from a lyophilized powder (lot number PCF-0042).
- To prepare the stock solutions for each of the control proteins (BSA and β-lactoglobulin), a 5.0-mg sub-sample of powder was weighed into an individual tube for each control and solubilized by adding 1.0 ml of 10 mM CAPS buffer (to a target protein concentration of 5.0 mg/ml).
- The final concentration of the protein and pancreatin in the SIF reaction mixture was 0.25 mg/ml Cry1B.34 protein and 0.5% (w/v) pancreatin.

SIF solution (1900 µl) was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to the addition of 100 µl of Cry1B.34 protein test substance at Time 0. The digestion reaction mixture was mixed constantly using a stir bar and a submersible magnetic stirrer.

A 120-µl sub-sample of the Cry1B.34 protein digestion reaction mixture was removed from the vial at the following analytical time points (± 10 seconds): 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. The sub-samples were inactivated by adding them to pre-labeled tubes containing 64 µl of pre-mixed sample solution (consisting of 46 µl NuPAGE 4X LDS Sample Buffer and 18 µl

NuPAGE 10X Sample Reducing Agent) and heating to 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

To prepare control digestion samples at 1 and 60 minutes, a 114-µl sample of the appropriate digestion solution was pre-warmed in a 37 °C water bath for 2-5 minutes prior to adding 6 µl of the Cry1B.34 protein test substance, control protein stock solution, or 10 mM CAPS buffer. The tubes were incubated in the water bath for the allotted time and then inactivated by mixing with 64 µl of the pre-mixed sample solution.

The Time 0 control reaction mixtures were prepared by first neutralizing 114 µl of the appropriate digestion solution with 64 µl of the pre-mixed sample solution, and then adding 6 µl of the Cry1B.34 protein test substance, protein stock solution, or 10 mM CAPS buffer to the appropriate tube and mixing.

Following digestion and inactivation, all control reaction mixtures were heated at 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

Control digestion samples included in the SIF assay are provided in Table E.2.

Table E.2. Control Samples for Simulated Intestinal Fluid (SIF) Digestibility Analysis

Protein	Digestion Solution	Digestion Time (min)		
		0	1	60
None (10 mM CAPS buffer), {SIF Control}	SIF	X	--	X
BSA	SIF	X	X	X
β-lactoglobulin	SIF	X	X	X
Cry1B.34	SIF	X	--	--
Cry1B.34	10 mM CAPS	X	--	X
Cry1B.34	Intestinal Control Solution (No Pancreatin)	--	--	X

SDS-PAGE Analysis

The Cry1B.34 protein digestion time-course samples and control samples were removed from frozen storage, heated at 90-100 °C for 5 minutes, and loaded (10 µl/well) into 4-12% Bis-Tris gels for SDS-PAGE analysis. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into each gel to provide a visual estimate of molecular weight. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gels were removed from the gel cassettes for use in protein staining or western blot analyses. For protein staining, the gels were washed three times for 5 minutes each with water and stained with GelCode Blue Stain Reagent for 60 minutes. Following staining, the gels were destained with water four times for a minimum of 5 minutes each until the gel background was clear. Proteins were stained as blue-colored bands on the gels. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

The Cry1B.34 protein digestion time-course samples were also analyzed by western blot. Following SDS-PAGE, one of the resulting gels was assembled into a nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane for 7 minutes with a pre-set program (P3).

Following protein transfer, the membrane was blocked in phosphate-buffered saline containing polysorbate 20 (PBST) containing 5% (w/v) non-fat dry milk for 60 minutes at ambient laboratory temperature. Before and after the blocking step, the membrane was washed with PBST three times for 1 minute each to reduce the background. The blocked membrane was incubated for 60 minutes at ambient laboratory temperature with an Cry1B.34 polyclonal antibody R11956 (Pioneer Hi-Bred International, Inc.) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. Following primary antibody incubation, the membrane was washed in PBST four times for 5 minutes each. The membrane was incubated for 60 minutes at ambient laboratory temperature with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. The membrane was washed in PBST four times for at least 5 minutes each. The membrane remained in PBST prior to incubating with a chemiluminescent substrate for 5 minutes. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP imaging system.

<u>Sequential Digestibility Analysis with Simulated Gastric Fluid (SGF) and Simulated Intestinal Fluid (SIF)</u>

Test solutions were prepared as follows:

- A concentrated (i.e., 2X) pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use by solubilizing pepsin (Sigma-Aldrich) in a previously prepared 2X gastric control solution. The final concentration of gastric control solution in SGF was 0.2% weight per volume (w/v) NaCl and 0.7% volume per volume (v/v) HCl; pH ~1.2. The SGF was prepared so that the pepsin to protein ratio of the final digestion mixture was 10 units of pepsin per µg of test protein.

- A concentrated (i.e., 2.5X) pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by solubilizing pancreatin (Sigma-Aldrich) in 2.5X intestinal control solution (2.5X I-Con). The final concentration of intestinal control solution in SIF was 50 mM KH_2PO_4, with a pH of ~7.5. Pancreatin content in SIF was adjusted so that there was approximately 0.5% (w/v) pancreatin in the final digestion reaction mixture.

- The pre-mixed sample stop solutions used to inactivate samples were prepared fresh on the day of use. The solution for SGF reactions was prepared by mixing 1200 µl of 200 mM Na2CO3, 1625 µl NuPAGE 4X LDS Sample Buffer, and 650 µl NuPAGE 10X Sample Reducing Agent. The solution for SIF reactions was prepared by mixing 1150 µl NuPAGE 4X LDS Sample Buffer and 450 µl NuPAGE 10X Sample Reducing Agent.

- The test substance consisted of Cry1B.34 protein solubilized from a lyophilized powder (lot number PCF-0042).

In Vitro Pepsin Digestion

Cry1B.34 Protein in SGF 10 Minutes Sample for Sequential Digestion: An aliquot (1 ml) of the 2X SGF solution and 800 µl water were dispensed into a 7-ml glass vial and pre-warmed in the 37 °C water bath for 2-5 minutes prior to addition of 200 µl of the Cry1B.34 protein test substance. The SGF digestion reaction mixture was incubated and mixed constantly using a stir bar and submersible stir plate for 10 minutes (± 10 seconds) after adding the Cry1B.34 protein test substance. At the end of the time period, a 1.5-ml sample of the Cry1B.34 SGF digestion reaction mixture was transferred to a separate vial and inactivated by neutralization with 0.3 ml of 0.5 N NaOH. This sample was used for the sequential SIF digestion.

Cry1B.34 Protein in SGF 10 Minutes: A 120-µl control sample (Cry1B.34 in SGF 10 minutes) was taken out from the SGF digestion reaction mixture at the end of 10 minutes (± 10 seconds) and inactivated by neutralization with 139 µl of pre-mixed SGF sample stop solution. The neutralized sample was heated for 5 minutes at 90-100 °C prior to frozen storage (-20 °C freezer unit).

Cry1B.34 Protein in SGF Time 0: A control sample (Cry1B.34 in SGF Time 0) was prepared by first inactivating 60 µl of 2X SGF and 49 µl water in 139 µl of pre-mixed SGF sample stop solution and then adding 12 µl of Cry1B.34 protein test substance to the neutralized SGF. The neutralized sample was heated for 5 minutes at 90-100 °C prior to frozen storage (-20 °C freezer unit).

SGF-Only 10 Minutes Incubation: An SGF-only control sample without Cry1B.34 protein test substance (SGF Control 10 minute) was prepared by mixing 60 µl 2X SGF and 49 µl water in a tube and pre-warming at 37 °C for 2-5 minutes. Following the addition of 12 µl of 10 mM CAPS buffer, the tube was incubated in a 37 °C water bath for 10 minutes (± 10 seconds). After incubation, the sample was inactivated by neutralization with 139 µl of pre-mixed SGF sample stop solution. The neutralized sample was heated for 5 minutes at 90-100 °C prior to frozen storage (-20 °C freezer unit).

Sequential Pancreatin Digestion

Cry1B.34 Protein in SGF 10 Minutes, SIF 0.5-30 Minutes: For the sequential SIF digestion time course, a 1.2-ml sample of the neutralized Cry1B.34 SGF digestion reaction mixture was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to addition of 800 µl 2.5X SIF solution. The SIF digestion reaction mixture was mixed constantly using a stir bar and a submersible stir plate.

A 120-µl sub-sample of the SIF digestion reaction mixture was removed from the vial at each of the following analytical time points (± 10 seconds): 0.5, 1, 2, 5, 10, 20 and 30 minutes. Each sub-sample was neutralized by adding it to a pre-labeled tube containing 64 µl of pre-mixed SIF sample stop solution. The neutralized samples were inactivated by heating at 90-100 °C for 5 minutes.

Cry1B.34 Protein in SGF 10 Minutes, SIF Time 0: An SIF control sample (Cry1B.34 10 minutes SGF Time 0 SIF) was prepared by mixing 48 µl 2.5X SIF with 64 µl of pre-mixed SIF sample stop solution and then heating for 5 minutes at 90-100 °C. A sub-sample (72 µl) of the neutralized Cry1B.34 SGF digestion reaction mixture was added to the heat-inactivated SIF control sample and then heated again for 5 minutes at 90-100 °C.

After neutralization and heating, all SIF reaction samples were stored frozen (-20 °C freezer unit).

SDS-PAGE Analysis

The digestion samples were removed from frozen storage, heated at 90-100 °C for 5 minutes, and loaded (10 µl/well) into a 4-12% Bis-Tris gel for SDS-PAGE analysis. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into the gel to provide a visual estimate of molecular weight. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gel was removed from the gel cassette and washed three times for 5 minutes each with water and stained with GelCode Blue Stain Reagent for 60 minutes. Following staining, the gel was destained with water four times for a minimum of 3 minutes each until the gel background was clear. Proteins were stained as blue-colored bands on the gel. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

APPENDIX F. METHODS FOR CHARACTERIZATION OF THE CRY1B.61.1 PROTEIN

Test Materials

COR23134 Soybean-Derived Cry1B.61.1 Protein

Cry1B.61.1 protein was isolated from whole plant tissue derived from COR23134 soybean. The whole plant tissue was collected at the V5 growth stage (the stage when the leaflets on the sixth leaf node have unrolled; Pedersen, 2004) of development from plants grown at a Pioneer owned field location (Johnston, IA, USA). The tissue was lyophilized, homogenized, and stored frozen at -80 °C. The Cry1B.61.1 protein was extracted from lyophilized soybean tissue by homogenization in a pre-chilled Waring blender vessel using phosphate-buffered saline containing polysorbate 20 (PBST) extraction buffer. The sample extract was then clarified by centrifugation and filtration. The filtered extract was purified by immunoaffinity chromatography. The immunoaffinity columns were prepared by coupling a monoclonal antibody (21C2.E8.H2.D10; Pioneer) to AminoLink Plus Coupling Resin (Thermo Scientific). The Cry1B.61.1 protein sample was eluted off the column using IgG Elution buffer (Thermo Scientific). Select elutions were collected and concentrated using a centrifugal concentrator (30K Vivaspin; Sartorius) to a volume of approximately 500 µl. The concentrated sample was buffer exchanged using 50 mM Tris buffer, pH 8, and then concentrated again to a volume of approximately 350 µl.

Following extraction, purification, and concentration, the final volume in the concentrator was estimated, NuPAGE LDS Sample Buffer (Life Technologies) was added at 25% along with 10% NuPAGE Reducing Agent containing DTT (Life Technologies) to the concentrated sample in the concentrator. The sample in the concentrator was transferred to a microcentrifuge tube, heat-treated, and stored frozen (-20 °C freezer unit).

Microbially Derived Cry1B.61.1 Protein

In order to have sufficient amounts of the purified Cry1B.61.1 protein for the multiple studies required to assess its safety, the Cry1B.61.1 protein was produced at Pioneer Hi-Bred International, Inc. using a microbial expression system. The protein was expressed in an *Escherichia coli* protein expression system as a fusion protein with a C-terminal 6XHis tag and was purified using Immobilized Metal Affinity Chromatography. Following purification, the protein was concentrated, and the buffer was changed to 10 mM CAPS, pH 11.0. After lyophilization and mixing, a lot number was assigned.

SDS-PAGE Analysis

The purified COR23134 soybean-derived Cry1B.61.1 protein sample was allowed to thaw, diluted as applicable for the sensitivity of the assay, heated, and then loaded into 4-12% Bis-Tris gels along with pre-stained protein molecular weight markers (Precision Plus Protein Dual Xtra Standards). For applicable SDS PAGE and western blot analysis, the Cry1B.61.1 protein reference substance was allowed to thaw, diluted in 1X LDS/DTT, heated, and loaded into the gels to 1 µg for SDS PAGE and 10 ng for western blot analysis. Electrophoresis was conducted using a pre-cast gel electrophoresis system with MES running buffer at a constant 200 volts for 35 minutes.

Upon completion of electrophoresis, the gels were either prepared for protein staining or protein transfer to a membrane for sequencing or western blot analysis.

For Coomassie staining, following electrophoresis, the gel was washed with water and stained with GelCode Blue Stain Reagent (Thermo Scientific). Following staining, the gel was de stained with water until the gel background was clear. Protein bands were stained on the gels and the gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

Following SDS-PAGE as described above and diluted as applicable for the sensitivity of the assay, the resulting gel was assembled into a nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in PBST containing 5% weight/volume (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with a polyclonal rabbit antibody 12000 (Pioneer Hi-Bred International, Inc.) diluted 1:5000 in PBST containing 1% w/v non-fat dry milk. Following primary antibody incubation, the membrane was washed with PBST. The membrane was incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:10,000 in PBST containing 1% w/v non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP (Bio Rad) imaging system.

Protein Glycosylation Analysis

COR23134 Soybean -Derived Cry1B.61.1 Protein

A Pierce Glycoprotein Staining Kit (Thermo Fisher) was used to determine whether the COR23134 soybean-derived Cry1B.61.1 protein was glycosylated. The purified soybean-derived Cry1B.61.1 protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor), were run by SDS PAGE as described in Methods Section B. For glycosylation staining, the soybean-derived Cry1B.61.1 protein was loaded on to the gel at approximately the same concentration as the positive and negative control proteins (1 μg).

Following electrophoresis, the gel was washed with water, fixed with 50% methanol, and washed with 3% acetic acid. The gel was then incubated with oxidizing solution and washed with 3% acetic acid. The gel was incubated with glycoprotein staining reagent and then incubated in a reducing reagent. The gel was then washed with 3% acetic acid followed by water. Glycoproteins were detected as stained bands on the gel.

Following glycoprotein detection, the image of the gel was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system. The same gel was then stained with GelCode Blue Stain Reagent (Thermo Scientific) followed by washes with water to visualize all protein bands. The image of the GelCode-stained gel was then captured electronically.

Microbially Derived Cry1B.61.1 Protein

A Pierce Glycoprotein Staining Kit was used to determine whether the Cry1B.61.1 protein was glycosylated. For glycosylation staining, 1 µg of Cry1B.61.1 protein was loaded on to the gel. The Cry1B.61.1 protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor), were run by SDS-PAGE as described in Methods Section B.

Following electrophoresis, the gel was washed with water twice for 5 minutes each wash, fixed with 50% methanol for 30-35 minutes, and washed twice with 3% acetic acid for 10-15 minutes each wash. The gel was then incubated with oxidizing solution for 15-20 minutes and washed three times with 3% acetic acid for 5-7 minutes each wash. The gel was incubated with glycoprotein staining reagent for 15 20 minutes and then incubated in a reducing reagent for 5-7 minutes. The gel was then washed three times with 3% acetic acid and once in water for 5 minutes each wash. Glycoproteins were detected as bands stained a magenta color on the gel.

Following glycoprotein detection, the image of the gel was captured electronically using a ChemiDoc MP imaging system. The same gel was then stained with GelCode Blue stain reagent for 60 minutes followed by three washes with water for at least 5 minutes each to visualize all protein bands. The image of the GelCode-stained gel was then captured electronically.

<u>Mass Spectrometry Peptide Mapping Analysis</u>

COR23134 Soybean-Derived Cry1B.61.1 Protein

Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described above, the COR23134 soybean-derived Cry1B.61.1 protein was loaded onto a gel in each of two lanes. Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described in Methods Section B and C, bands containing the soybean derived Cry1B.61.1 protein were excised from a gel and stored frozen (-20 °C freezer unit). The protein in each gel slice was reduced with DTT, alkylated with iodoacetamide, and then subsequently digested with trypsin or chymotrypsin. The digested samples were separated on a nanoACQUITY UPLC (Waters Corporation) fitted with a Peptide BEH C18 300 Å 1.7 µm column (75 µm x 100 mm; Waters Corporation) by gradient elution. Eluent from the column was directed into an electrospray source, operating in positive ion mode, on a TripleTOF 5600+ hybrid quadrupole-TOF mass spectrometer (AB Sciex; currently Sciex). The resulting mass spectrometry (MS) data were processed using MSConvert to produce a peak list. The peak list was used to perform an MS/MS ion search (Mascot Software version 2.8.0) and match peptides from the expected Cry1B.61.1 protein sequence (Perkins *et al.*, 1999). The following search parameters were used: peptide and fragment mass tolerance, ± 0.1 Da; fixed modifications, cysteine carbamidomethyl; variable modifications, methionine oxidation; and maximum missed cleavages, 1 for trypsin and 2 for chymotrypsin. The Mascot-generated peptide ion score threshold was > 13 which indicates identity or extensive homology ($P < 0.05$). The combined sequence coverage was calculated with GPMAW version 12.11.0.

Microbially Derived Cry1B.61.1 Protein

For mass spectrometry sequencing analyses, 4 µg of Cry1B.61.1 protein was loaded onto a gel in each of three lanes. Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described in Methods Section B, Cry1B.61.1 protein bands were excised from a gel and stored frozen (-20 °C freezer unit). The protein in two of the gel slices was reduced with DTT, alkylated with iodoacetamide, and then subsequently digested with trypsin or chymotrypsin. The digested samples were separated on a nanoACQUITY UPLC (Waters Corporation) fitted with a Peptide BEH C18, 300 Å, 1.7 µm column (75 µm x 100 mm; Waters Corporation) by gradient elution. Eluent from the column was directed into an electrospray source, operating in positive ion mode, on a TripleTOF 5600+ hybrid quadrupole TOF mass spectrometer (AB Sciex; currently Sciex). The resulting MS data were processed using MSConverter to produce a peak list. The peak list was used to perform an MS/MS ion search (Mascot Software version 2.8.0) and match peptides from the expected Cry1B.61.1 protein sequence (Perkins *et al.*, 1999). The following search parameters were used: peptide and fragment mass tolerance, ± 0.1 Da; fixed modifications, cysteine carbamidomethyl; variable modifications, methionine oxidation, acetyl (protein N terminal); and maximum missed cleavages, 1 for trypsin and 2 for chymotrypsin. The Mascot-generated peptide ion score threshold was > 13, which indicates identity or extensive homology ($p < 0.05$). The combined sequence coverage was calculated with GPMAW version 12.11.0.

N-Terminal Amino Acid Sequence Analysis

COR23134 Soybean-Derived Cry1B.61.1 Protein

The N-terminal amino acid sequence was also determined through peptide mapping as it was determined with previous Cry1B.61.1 protein lots that the N terminus was blocked and proteins with a blocked N-terminal residue are unable to be directly sequenced by Edman degradation.

Microbially Derived Cry1B.61.1 Protein

For N-terminal amino acid sequence analyses, 6 µg of Cry1B.61.1 protein was loaded on to a gel in each of three lanes. Following SDS-PAGE, using the methods as described in Methods Section B, the resulting gel was incubated in cathode buffer (60 mM Tris, 40 mM CAPS, 0.075% SDS, pH 9.6) for 10 20 minutes. An Immobilon-PSQ PVDF membrane was wetted in 100% methanol for 1 minute, followed by immersion in anode buffer (60 mM Tris, 40 mM CAPS, 15% methanol, pH 9.6) for 10-20 minutes. A Trans-Blot SD Semi Dry Electrophoretic Transfer Cell system was used to transfer proteins from the gel to the membrane at 10 V for 60 minutes. Following protein transfer, the membrane was washed with water three times for 5 minutes each, stained with GelCode Blue stain reagent for 5 minutes, and then destained with water to visualize Cry1B.61.1 protein bands. The bands containing the Cry1B.61.1 protein were excised and stored frozen (20 °C freezer unit) and one band was analyzed using a Shimadzu PPSQ-51A sequencer. Ten cycles of Edman sequencing were performed. During each cycle, the N terminal amino acid was sequentially derivatized with phenylisothiocyanate (PITC), cleaved with trifluoracetic acid, and converted to PTH amino acid which was identified through chromatography. LabSolutions Software was used to identify the N terminal sequence.

Bioactivity Bioassay

The biological activity of Cry1B.61.1 protein was evaluated by conducting a 7-day bioassay using *Chrysodeixis includens* (soybean looper; Lepidoptera: Noctuidae), a species sensitive to the

Cry1B.61.1 protein. Eggs were obtained from Pioneer and their identity was recorded by study personnel.

The bioassay utilized a generalized randomized block design containing 10 blocks. Each block consisted of a 12-well bioassay plate and contained two replicates from each treatment for a target of 20 individuals per treatment. *C. includens* larvae were exposed via oral ingestion to one of the following four treatments:

- Treatment 1: Buffer Control Diet (containing 10 mM CAPS buffer)
- Treatment 2: Test Diet (targeting 5 ng Cry1B.61.1 protein per mg wet diet)
- Treatment 3: Test Diet (targeting 50 ng Cry1B.61.1 protein per mg wet diet)
- Treatment 4: Test Diet (targeting 500 ng Cry1B.61.1 protein per mg wet diet)

On each day of diet preparation (Day 0 and Day 4), the Cry1B.61.1 protein test substance was solubilized in 10 mM CAPS buffer, pH 10.5 (referred to as buffer), to a target concentration of 1 mg per ml. The solubilized test substance was diluted in buffer to achieve the concentrations in the test dosing solutions (6.7 ng/µl, 66.7 ng/µl, and 666.7 ng/µl, for Treatments 2, 3, and 4, respectively). The buffer control dosing solution consisted of 10 mM CAPS buffer, pH 10.5. Dosing solutions were maintained chilled (on wet ice) until use. The carrier for the *C. includens* bioassay consisted of an artificial Stonefly Heliothis diet. On each day of diet preparation, each dosing solution was mixed with carrier in a 3:1 ratio (i.e., 3 ml of dosing solution to 1 g of carrier), generating Treatments 1-4.

C. includens eggs were incubated in an environmental chamber until the eggs hatched. Neonates were used in the bioassay within 24 hours of hatching. On Day 0 of the bioassay, approximately 300 µl (i.e., 1 g of wet diet equated to 1 ml of wet diet) of freshly prepared diets were dispensed into individual wells of the bioassay plate and one *C. includens* neonate was placed in each well containing diet. Each bioassay plate was sealed with heat-sealing film and ventilated with a small hole over each well. The bioassay was conducted in an environmental chamber set at 25 °C, 65% relative humidity, and continuous dark. On Day 4, new bioassay plates were prepared with fresh diet as described for Day 0, surviving organisms were transferred to the new plates, and the plates were placed in the environmental chambers. After 7 days, the bioassay was complete, mortality was assessed, and surviving organisms were individually weighed. Only wells that contained one organism were included in the total number of observed individuals; organisms recorded as missing from a well were excluded from data analysis.

The bioassay acceptability criterion indicated the bioassay may be repeated if the combined number of dead and missing organisms exceeded 20% for the buffer control diet (Treatment 1) group. The bioassay exceeded the acceptability criterion with 30% dead and missing organisms; however, it was not repeated as the purpose of the bioassay (i.e., demonstrating the biological activity of the test substance) was met.

Control of bias in the bioactivity assay was achieved through the use of a control diet and the random allocation of treatments within each block.

Thermolability Analysis

The test substance consisted of Cry1B.61.1 protein solubilized from a lyophilized powder (lot number PCF-0062). The carrier consisted of Stonefly Heliothis diet. The buffer control dosing solution used to prepare Treatment 1 consisted of chilled 10 mM CAPS. The dosing solutions used to prepare Treatments 2-6 consisted of the test substance diluted in 10 mM CAPS buffer to achieve the concentration in each treatment. The dosing solutions used to prepare Treatments 3-6 were incubated for 30-35 minutes at various temperatures.

The test system was *Chrysodeixis includens* (soybean looper; Lepidoptera: Noctuidae). The test system was chosen because *C. includens* is an insect sensitive to Cry1B.61.1 protein. *C. includens* larvae were exposed via oral ingestion to one of the following six treatments:

- Treatment 1: Buffer Control Diet containing a dosing solution of 10 mM CAPS buffer
- Treatment 2: Control Diet containing the unheated Cry1B.61.1 protein dosing solution
- Treatment 3: Test Diet containing the Cry1B.61.1 protein dosing solution incubated at 25 °C
- Treatment 4: Test Diet containing the Cry1B.61.1 protein dosing solution incubated at 50 °C
- Treatment 5: Test Diet containing the Cry1B.61.1 protein dosing solution incubated at 75 °C
- Treatment 6: Test Diet containing the Cry1B.61.1 protein dosing solution incubated at 95 °C

The control diet containing the unheated test dosing solution and each test diet contained a targeted concentration of 100 ng Cry1B.61.1 protein per mg diet wet weight. Treatments were arranged in a generalized randomized block design with a total of 10 blocks. Each block consisted of a 12-well bioassay plate and contained two replicates from each treatment. Each treatment was provided to a target of 20 *C. includens* individuals. The bioassay was conducted in an environmental chamber set at 25 °C, 65% relative humidity, and continuous dark. Larvae were refed on Day 4. After 7 days, the bioassay was complete, mortality was assessed, and surviving organisms were individually weighed.

The bioassay acceptability criteria indicated the bioassay may be terminated and repeated if: the combined number of dead and missing organisms exceeds 20% for the buffer control diet (Treatment 1) group and the mortality rate does not exceed 80% in the unheated control (Treatment 2) group. The *C. includens* bioassay met the acceptability criteria (0% dead and 0% missing in Treatment 1; 100% mortality in Treatment 2). An enzyme linked immunosorbent assay (ELISA) was used to verify the homogeneity of the Cry1B.61.1 protein in Treatment 2 and the concentration of the Cry1B.61.1 protein dosing solutions used to prepare Treatments 2-6. The absence of Cry1B.61.1 protein in one buffer control dosing solution preparation was also verified. Bias in the *C. includens* bioassay was controlled through the randomization of treatments within blocks and the use of one or more control diets. Bias in the characterization portion of the study was controlled through the use of replicate testing and appropriate assay controls.

The buffer control and Cry1B.61.1 protein dosing solutions were prepared and characterized. Dosing solutions were prepared on each day of feeding for the *C. includens* bioassay and maintained chilled (in a refrigerator set at 4 °C or on wet ice) when not under heat treatment. Each dosing solution was mixed with carrier in a 3:1 ratio (i.e., 3 ml of dosing solution to 1 g of carrier), generating Treatments 1-6. *C. includens* eggs were incubated in an environmental chamber and neonates were used in the bioassay within 24 hours of hatching.

On Day 0, approximately 300 µl (i.e., 1 g of wet diet equated to 1 ml of wet diet) of freshly prepared diets were dispensed into wells of the bioassay plates. One neonate was placed in each well containing diet and each bioassay plate was sealed with heat-sealing film and ventilated with a small hole over each well. The plates were placed in an environmental chamber. On Day 4, new bioassay plates were prepared with fresh diet as described for Day 0, living larvae were transferred to the new plates, missing or dead larvae were recorded, and the freshly prepared plates were placed in the environmental chamber. After 7 days, the bioassay was complete, mortality was assessed, and surviving larvae were individually weighed.

Statistical Analysis

Statistical analyses of data were conducted using SAS software, Version 9.4 (SAS Institute Inc.). The response variable of interest was mortality.

Statistical comparisons for mortality were made between *C. includens* fed diets containing heated Cry1B.61.1 protein (Treatments 3-6) and those fed a diet containing unheated Cry1B.61.1 protein (Treatment 2). Statistical analysis was conducted using Fisher's exact test to determine if the mortality rate of *C. includens* fed diets containing the heated Cry1B.61.1 protein (m_T) was less than the mortality rate of those fed the buffer control diet with unheated Cry1B.61.1 protein diet (m_C). The corresponding hypothesis test was:

$$H_0: m_T - m_C = 0 \quad vs. \quad H_a: m_T - m_C < 0$$

A significant difference was established if the P-value was < 0.05. SAS PROC MULTTEST was utilized to conduct the Fisher's exact test.

<u>Digestibility in Simulated Gastric Fluid (SGF)</u>

Test and control solutions were prepared as follows:

- The gastric control solution was prepared fresh on the day of use and was comprised of 0.2% weight per volume (w/v) NaCl in 0.7% volume per volume (v/v) HCl, with a pH of ~1.2.
- The pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use by dissolving pepsin (Sigma-Aldrich) into gastric control solution. The SGF was prepared so that the pepsin to protein ratio of the final digestion mixture was approximately 10 units of pepsin per µg of test or control protein.
- The test substance consisted of Cry1B.61.1 protein solubilized from a lyophilized powder (lot number PCF-0067).
- To prepare the solutions for each of the control proteins (BSA and β-lactoglobulin), a 5.0-mg sub-sample of powder was weighed into an individual tube for each control and solubilized by adding 1.0 ml of buffer to a target protein concentration of 5.0 mg/ml.
- The final concentration of the protein and pepsin in the SGF reaction mixture was 0.25 mg/ml Cry1B.61.1 protein and 2500 units/ml pepsin.

SGF (1900 µl) was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2 5 minutes prior to the addition of 100 µl of Cry1B.61.1 protein solution at Time 0. The digestion reaction mixture was mixed constantly using a stir plate.

A 120-µl sub-sample of the Cry1B.61.1 protein digestion reaction mixture was removed from the vial at the following analytical time points (± 10 seconds): 0.25, 1, 2, 5, 10, 30, and 60 minutes. The sub samples were inactivated by neutralization with 139 µl of pre-mixed sample stop solution and heating to 90-100 °C for 5 minutes prior to SDS-PAGE analysis and subsequent frozen storage (-20 °C freezer unit).

To prepare control digestion samples at 1 and 60 minutes, a 114-µl sample of the appropriate digestion solution (see table) was pre-warmed in a 37 °C water bath for 2-5 minutes prior to adding 6 µl of the Cry1B.61.1 protein solution, control protein solutions, or buffer. The tubes were incubated in the water bath for the allotted time and then inactivated by mixing with 139 µl of the pre-mixed sample solution.

The Time 0 control reaction mixtures were prepared by first neutralizing 114 µl of the appropriate digestion solution (see table) with 139 µl of the pre-mixed sample solution. A 6-µl sample of the appropriate Cry1B.61.1 protein solution, control protein solutions, or buffer was then added, and the mixture was heated at 90-100 °C for 5 minutes. The final concentration of the protein and pepsin in the control digestion samples was 0.25 mg/ml Cry1B.61.1 protein or control proteins and 2500 units/ml pepsin.

Following digestion and inactivation, all control digestion samples were analyzed by SDS-PAGE analysis and then stored frozen (-20 °C freezer unit).

Control digestion samples included in the SGF assay are provided in Table F.1.

Table F.1. Control Samples for Simulated Gastric Fluid (SGF) Digestibility Analysis

Protein	Digestion Solution	Digestion Time (min)		
		0	1	60
None (SGF Control – Buffer)	SGF	X	--	X
BSA	SGF	X	X	X
β-lactoglobulin	SGF	X	X	X
Cry1B.61.1	SGF	X	--	--
Cry1B.61.1	None (Ultrapure Water)	X	--	X
Cry1B.61.1	Gastric Control Solution (No Pepsin)	--	--	X

SDS-PAGE Analysis

The Cry1B.61.1 protein digestion time-course samples and control samples were heated at 90-100 °C for 5 minutes and loaded (10 µl/well) into 4-12% Bis-Tris gels for SDS PAGE analysis. To demonstrate the sensitivity of the SDS-PAGE gel and western blot analyses, an aliquot of the Cry1B.61.1 protein in SGF (Time 0) sample was loaded into the gel at a 1:20 dilution for protein staining, and at a 1:50 dilution for the western blot. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into each gel. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gels were removed from the gel cassettes for use in protein staining or western blot analyses. For protein staining, the gels were washed with water and stained with GelCode Blue Stain Reagent. Following staining, the gels were destained with water until the gel background was clear. Protein bands were stained on the gels. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

The Cry1B.61.1 protein digestion time-course samples were also analyzed by western blot. Following SDS-PAGE, one of the resulting gels was assembled into a mini nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in phosphate-buffered saline containing polysorbate 20 (PBST) containing 5% (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with a polyclonal antibody R12000 (Pioneer Hi-Bred International, Inc.) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. Following primary antibody incubation, the membrane was washed in PBST and incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP imaging system.

Digestibility in Simulated Intestinal Fluid (SIF)

Test and control solutions were prepared as follows:

- The pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by dissolving 26.3 mg of pancreatin (Sigma-Aldrich) into 5 ml of intestinal control solution (I-Con 1X buffer) to a final concentration of 0.5% weight per volume (w/v) pancreatin and 50 mM KH2PO4, pH 7.5.
- The test substance consisted of purified Cry1B.61.1 protein in the form of a lyophilized powder (lot number PCF-0067).
- To prepare the solutions for each of the control proteins (BSA and β-lactoglobulin), a 5.0-mg sub sample of powder was weighed into an individual tube for each control and solubilized by adding 1.0 ml of CAPS buffer, pH 10.5, to a target protein concentration of 5 mg/ml.
- The final concentration of 0.5% weight per volume (w/v) pancreatin and 50 mM KH2PO4, pH 7.5.

SIF (1900 µl) was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2 5 minutes prior to the addition of 100 µl of Cry1B.61.1 protein solution at Time 0. The digestion reaction mixture was mixed constantly using a stir plate.

A 120-µl sub-sample of the Cry1B.61.1 protein digestion reaction mixture was removed from the vial at the following analytical time points (± 10 seconds): 0.25, 1, 2, 5, 10, 30, and 60 minutes.

The sub samples were inactivated by adding them to pre-labeled tubes containing 64 µl of pre- mixed sample solution and heating to 90 100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

To prepare control digestion samples at 1 and 60 minutes, a 114-µl sample of the appropriate digestion solution (see table below) was pre-warmed in a 37 °C water bath for 2-5 minutes prior to adding 6 µl of the Cry1B.61.1 protein solution, control protein solution, or buffer. The tubes were incubated in the water bath for the allotted time and then inactivated by mixing with 64 µl of the pre-mixed sample solution.

The Time 0 control reaction mixtures were prepared by first neutralizing 114 µl of the appropriate digestion solution (see table) with 64 µl of the pre-mixed sample solution, and then heating at 90-100 °C for 5 minutes. A 6-µl sample of the appropriate Cry1B.61.1 protein solution, control protein solution, or buffer was then added, and the mixture was heated again at 90-100 °C for 5 minutes. The final concentration of the protein and pancreatin in the control digestion samples was 0.25 mg/ml Cry1B.61.1 protein or control protein and 0.5% (w/v) pancreatin.

Following digestion and inactivation, all control reaction mixtures were heated at 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

Control digestion samples included in the SIF assay are provided in Table F.2.

Table F.2. Control Samples for Simulated Intestinal Fluid (SIF) Digestibility Analysis

Protein	Digestion Solution	Digestion Time (min)		
		0	1	60
None (SIF Control – Buffer)	SIF	X	--	X
BSA	SIF	X	X	X
β-lactoglobulin	SIF	X	X	X
Cry1B.61.1	SIF	X	--	--
Cry1B.61.1	None (Ultrapure Water)	X	--	X
Cry1B.61.1	Intestinal Control Solution (No Pancreatin)	--	--	X

SDS-PAGE Analysis

The Cry1B.61.1 protein digestion time-course samples and control samples were heated at 90-100 °C for 5 minutes and loaded (10 µl/well) into 4-12% Bis-Tris gels for SDS PAGE analysis. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into each gel. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gels were removed from the gel cassettes for use in protein staining or western blot analyses. For protein staining, the gels were washed with water and stained with GelCode Blue Stain Reagent. Following staining, the gels were destained with water until the gel background was clear. Protein bands were stained on the gels. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

The Cry1B.61.1 protein digestion time-course samples were also analyzed by western blot. Following SDS PAGE, one of the resulting gels was assembled into a mini nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in phosphate-buffered saline containing polysorbate 20 (PBST) containing 5% (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with a polyclonal antibody R12000 (Pioneer Hi Bred International, Inc.) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. Following primary antibody incubation, the membrane was washed in PBST and incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP imaging system.

<u>Sequential Digestibility Analysis with Simulated Gastric Fluid (SGF) and Simulated Intestinal Fluid (SIF)</u>

Test solutions were prepared as follows:

- A 2X concentrated pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use by solubilizing pepsin (Sigma-Aldrich) in a previously prepared 2X gastric control solution. The final concentration of gastric control solution in the final digestion mixture was 0.2% weight per volume (w/v) NaCl and 0.7% volume per volume (v/v) HCl; pH ~1.2. The SGF was prepared so that the pepsin to protein ratio of the final digestion mixture was approximately 8.6 units of pepsin per µg of test substance.
- A 2.5X concentrated pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by solubilizing pancreatin (Sigma-Aldrich) in 2.5X intestinal control solution. The final concentration of intestinal control solution in SIF was 50 mM KH2PO4, pH ~7.5. Pancreatin content in SIF was adjusted so that there was approximately 0.5% (w/v) pancreatin in the final digestion reaction mixture.
- The pre-mixed sample solutions used to inactivate samples were prepared fresh on the day of use. The stop solution for SGF reactions was prepared by mixing 1200 µl of 200 mM Na2CO3, 1625 µl of NuPAGE 4X LDS Sample Buffer, and 650 µl NuPAGE 10X Sample Reducing Agent containing DTT. The SIF sample solution for SIF reactions was prepared by mixing 1150 µl NuPAGE 4X LDS Sample Buffer and 450 µl NuPAGE 10X Sample Reducing Agent containing DTT.
- The test substance consisted of purified Cry1B.61.1 protein in the form of a lyophilized powder (lot number PCF-0067).

In Vitro Pepsin Digestion

Cry1B.61.1 Protein in SGF 10 Minutes Sample for Sequential Digestion: A 1-ml aliquot of the 2X SGF solution was dispensed into a 7-ml glass vial and pre-warmed in the 37 °C water bath for 2-5 minutes prior to addition of 800 µl ultrapure water and 200 µl of the Cry1B.61.1 protein solution at Time Zero. The SGF digestion reaction mixture was incubated and mixed constantly for 5 minutes (± 10 seconds) after adding the Cry1B.61.1 protein. At the end of the time period, a 1.5-ml sample of the Cry1B.61.1 SGF digestion reaction mixture was transferred to a separate vial and inactivated by neutralization with 0.3 ml of 0.5 N NaOH. This sample was used for the sequential SIF digestion.

Cry1B.61.1 Protein in SGF 10 Minutes: A sample (120 µl) was collected from the Cry1B.61.1 SGF digestion reaction mixture at the end of 5 minutes (± 10 seconds) and inactivated by neutralization with 139 µl of pre-mixed SGF sample stop solution. The neutralized sample was heated for 5 minutes at 90-100 °C prior to frozen storage (-20 °C freezer unit). This sample served as an SGF-only sample without sequential digestion.

Cry1B.61.1 Protein in SGF Time 0: A control sample (Cry1B.61.1 protein in SGF) was prepared by first neutralizing a 60 µl sample of 2X SGF and 49 µl ultrapure water with 139 µl of SGF sample stop solution and then adding 12 µl of Cry1B.61.1 protein. The neutralized sample was then heated at 90 100 °C for 5 minutes.

SGF-Only 5 Minutes Incubation: An SGF only control sample (without Cry1B.61.1 protein) was prepared by mixing 60 µl of 2X SGF with 49 µl ultrapure water and then pre-warming in a 37 °C water bath for 2-5 minutes. After pre-warming, 12 µl of buffer was added and the SGF solution was incubated in the water bath for 5 minutes. At the end of the time period, 139 µl of SGF sample stop solution was added and the neutralized sample was then heated at 90-100 °C for 5 minutes.

Sequential Pancreatin Digestion

For the sequential SIF digestion time-course, a 1.2-ml sample of the NaOH-neutralized Cry1B.61.1 SGF digestion reaction mixture was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to addition of 800 µl of 2.5X SIF solution. The SIF digestion reaction mixture was mixed constantly using a stir plate. A 120-µl sub sample of the SIF digestion reaction mixture was removed from the vial at each of the following analytical time points (± 10 seconds): 0.25, 1, 2, 5, 10, 20 and 30 minutes. Each sub sample was inactivated by adding it to a pre-labeled tube containing 64 µl of pre-mixed SIF sample solution and heating at 90-100 °C for 5 minutes.

After heating, all SIF reaction samples were stored frozen (-20 °C freezer unit).

SDS-PAGE Analysis

The Cry1B.61.1 protein digestion time-course samples and control samples were removed from frozen storage, heated at 90-100 °C for 5 minutes, and loaded (10 µl/well) into a 4-12% Bis-Tris gel for SDS PAGE analysis. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into the gel. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gel was removed from the gel cassette for use in protein staining. For protein staining, the gel was washed with water and stained with GelCode Blue Stain Reagent. Following staining, the gel was destained with water until the gel background was clear. Protein bands were stained on the gel. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

APPENDIX G. METHODS FOR CHARACTERIZATION OF THE IPD083CB PROTEIN

Test Materials

COR23134 Soybean-Derived IPD083Cb Protein

The IPD083Cb protein isolated from whole plant tissue derived from COR23134 soybean. The whole plant tissue was collected at the V5 growth stage (the stage when the leaflets on the sixth leaf node have unrolled; Pedersen, 2004) of development from plants grown at a Pioneer owned field location (Johnston, IA, USA). The IPD083Cb protein was extracted from lyophilized soybean tissue by homogenization in a pre-chilled Waring blender vessel using phosphate-buffered saline containing polysorbate 20 (PBST) extraction buffer. The sample extract was then clarified by centrifugation and filtration. The filtered extract was purified by immunoaffinity chromatography. The immunoaffinity columns were prepared by coupling an IPD083Cb monoclonal antibody (19F4.G1.F6; Pioneer) to AminoLink Plus Coupling Resin (Thermo Scientific). The IPD083Cb protein sample was eluted off the column using IgG Elution buffer (Thermo Scientific). Select elutions were collected and concentrated using a centrifugal concentrator (30K Vivaspin Turbo 4; Sartorius) to a volume of approximately 500 µl. The concentrated sample was buffer exchanged using 50 mM Tris buffer, pH 8, and then concentrated again to a volume of approximately 225 µl.

Following extraction, purification, and concentration, the final volume in the concentrator was estimated, NuPAGE LDS Sample Buffer (Life Technologies) was added at 25% along with 10% NuPAGE Reducing Agent containing DTT (Life Technologies) to the concentrated sample in the concentrator. The sample in the concentrator was transferred to a microcentrifuge tube, heat-treated, and stored frozen (-20 °C freezer unit).

Tobacco-Expressed* IPD083Cb *Protein

In order to have sufficient amounts of the purified IPD083Cb protein for the multiple studies required to assess its safety, the IPD083Cb protein was produced by iBio CDMO for Pioneer using a *Nicotiana benthamiana* protein expression system. The protein was first purified by immobilized Metal Affinity Chromatography (IMAC), followed by anion exchange chromatography, then concentrated and buffer exchanged. After lyophilization and mixing, a lot number was assigned.

SDS-PAGE Analysis

The purified COR23134 soybean-derived IPD083Cb protein sample was diluted as applicable for the sensitivity of the assay, heated, and then loaded into 4-12% Bis-Tris gels along with pre-stained protein molecular weight markers (Precision Plus Protein Dual Xtra Standards). For SDS PAGE and western blot analysis, as applicable, the IPD083Cb protein reference substance was diluted in 1X LDS/DTT, heated, and loaded into the gels to 1 µg for SDS PAGE and 10 ng for western blot analysis. Electrophoresis was conducted using a pre-cast gel electrophoresis system with MES running buffer at a constant 200 volts for 35 minutes.

Upon completion of electrophoresis, the gels were either prepared for protein staining or protein transfer to a membrane for sequencing or western blot analysis.

For Coomassie staining, following electrophoresis, the gel was washed with water and stained with GelCode Blue Stain Reagent (Thermo Scientific). Following staining, the gel was de stained with water until the gel background was clear. Protein bands were stained on the gels and the gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

Following SDS-PAGE as described above, the resulting gel was assembled into a nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in PBST containing 5% weight/volume (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with an IPD083Cb polyclonal rabbit antibody R3373 (Pioneer Hi-Bred International, Inc.) diluted 1:5000 in PBST containing 1% w/v non-fat dry milk. Following primary antibody incubation, the membrane was washed. The membrane was incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:10,000 in PBST containing 1% w/v non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP (Bio Rad) imaging system.

Protein Glycosylation Analysis

COR23134 Soybean-Derived IPD083Cb Protein

A Pierce Glycoprotein Staining Kit (Thermo Fisher) was used to determine whether the COR23134 soybean-derived IPD083Cb protein was glycosylated. The purified soybean-derived IPD083Cb protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor), were run by SDS PAGE as described above. For glycosylation staining, the soybean-derived IPD083Cb protein was loaded on to the gel at approximately the same concentration as the positive and negative control proteins (1 μg).

Following electrophoresis, the gel was washed with water, fixed with 50% methanol, and washed with 3% acetic acid. The gel was then incubated with oxidizing solution and washed with 3% acetic acid. The gel was incubated with glycoprotein staining reagent and then incubated in a reducing reagent. The gel was then washed with 3% acetic acid followed by water. Glycoproteins were detected as stained bands on the gel.

Following glycoprotein detection, the image of the gel was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system. The same gel was then stained with GelCode Blue Stain Reagent (Thermo Scientific) followed by washes with water to visualize all protein bands. The image of the GelCode-stained gel was then captured electronically.

Tobacco-Expressed* IPD083Cb *Protein

A Pierce Glycoprotein Staining Kit was used to determine whether the IPD083Cb protein was glycosylated. For glycosylation staining, 1 µg of IPD083Cb protein was loaded on to the gel. The IPD083Cb protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor) were run by SDS-PAGE as described in Methods Section B.

Following electrophoresis, the gel was washed with water twice for 5 minutes each wash, fixed with 50% methanol for 30-35 minutes, and washed twice with 3% acetic acid for 10-15 minutes each wash. The gel was then incubated with oxidizing solution for 15-20 minutes and washed three times with 3% acetic acid for 5-7 minutes each wash. The gel was incubated with glycoprotein staining reagent for 15-20 minutes and then incubated in a reducing reagent for 5-7 minutes. The gel was then washed once with 3% acetic acid and once in water for 5 minutes each wash. Glycoproteins were detected as bands stained a magenta color on the gel.

Following glycoprotein detection, the image of the gel was captured electronically using a ChemiDoc MP imaging system. The same gel was then stained with GelCode Blue stain reagent for 60 minutes followed by three washes with water for 5 minutes each wash to visualize all protein bands. The image of the GelCode stained gel was then captured electronically.

Mass Spectrometry Peptide Mapping Analysis

COR23134 Soybean-Derived IPD083Cb Protein

For mass spectrometry sequencing analyses, the COR23134 soybean-derived IPD083Cb protein was loaded onto a gel in each of two lanes. Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described above, bands containing the soybean-derived IPD083Cb protein were excised from a gel and stored frozen (-20 °C freezer unit). The protein in each gel slice was reduced with DTT, alkylated with iodoacetamide, and then subsequently digested with trypsin or chymotrypsin. The digested samples were separated on a nanoACQUITY UPLC (Waters Corporation) fitted with a Peptide BEH C18 300 Å 1.7 µm column (75 µm x 100 mm; Waters Corporation) by gradient elution. Eluent from the column was directed into an electrospray source, operating in positive ion mode, on a TripleTOF 5600+ hybrid quadrupole-TOF mass spectrometer (AB Sciex; currently Sciex). The resulting mass spectrometry (MS) data were processed using MSConverter to produce a peak list. The peak list was used to perform an MS/MS ion search (Mascot Software version 2.8.0) and match peptides from the expected IPD083Cb protein sequence (Perkins *et al.*, 1999). The following search parameters were used: peptide and fragment mass tolerance, ± 0.1 Da; fixed modifications, cysteine carbamidomethyl; variable modifications, acetyl (protein N-terminal), methionine oxidation; and maximum missed cleavages, 1 for trypsin and 2 for chymotrypsin. The Mascot-generated peptide ion score threshold was > 13 which indicates identity or extensive homology ($P < 0.05$). The combined sequence coverage was calculated with GPMAW version 12.11.0.

Tobacco-Expressed* IPD083Cb *Protein

For mass spectrometry sequencing analyses, 4 µg of IPD083Cb protein was loaded onto a gel in each of three lanes. Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described in Methods Section B, protein bands at the expected molecular weight of

IPD083Cb protein were excised from a gel and stored frozen (-20 °C freezer unit). The protein in two of the gel slices was reduced with DTT, alkylated with iodoacetamide, and then subsequently digested with trypsin or chymotrypsin. The digested samples were separated on a nanoACQUITY UPLC (Waters Corporation) fitted with a Peptide BEH C18 300 Å 1.7 µm column (75 µm x 100 mm; Waters Corporation) by gradient elution. Eluent from the column was directed into an electrospray source, operating in positive ion mode, on a TripleTOF 5600+ hybrid quadrupole-TOF mass spectrometer (AB Sciex; currently Sciex). The resulting MS data were processed using MS Data Converter to produce a peak list. The peak list was used to perform an MS/MS ion search (Mascot Software version 2.8.0) and match peptides from the expected IPD083Cb protein sequence (Perkins *et al.*, 1999). The following search parameters were used: peptide and fragment mass tolerance, ± 0.1 Da; fixed modifications, cysteine carbamidomethyl; variable modifications, methionine oxidation, acetyl (protein N-terminal); and maximum missed cleavages, 1 for trypsin and 2 for chymotrypsin. The Mascot-generated peptide ion score threshold was > 13, which indicates identity or extensive homology ($p < 0.05$). The combined sequence coverage was calculated with GPMAW version 12.10.0.

<u>N-Terminal Amino Acid Sequencing Analysis</u>

COR23134 Soybean-Derived IPD083Cb Protein

For N-terminal amino acid sequence analysis, the COR23134 soybean-derived IPD083Cb protein was loaded onto a gel. Following SDS-PAGE using the methods as described above, the resulting gel was incubated in cathode buffer (60 mM Tris, 40 mM CAPS, 0.075% SDS, pH 9.6). An Immobilon-P PVDF membrane (Millipore) was wetted in 100% methanol, followed by immersion in anode buffer (60 mM Tris, 40 mM CAPS, 15% methanol, pH 9.6). A Trans-Blot SD Semi Dry Electrophoretic Transfer Cell system (Bio-Rad) was used to transfer proteins from the gel to the membrane. Following protein transfer, the membrane was washed with water, stained with GelCode Blue Stain Reagent (Thermo Scientific), and then destained with water to visualize the IPD083Cb protein band which was excised and stored frozen (-20 °C freezer unit). The band was analyzed using a Shimadzu PPSQ-51A sequencer. Ten cycles of Edman sequencing were performed. No results were obtained due to N-terminal acetylation. The N-terminal amino acid sequence was also determined through peptide mapping as it was determined that the N-terminus is acetylated, rendering Edman degradation unsuccessful.

Tobacco-Expressed* IPD083Cb *Protein

The N-terminal amino acid sequence was also determined through peptide mapping as it was determined with previous IPD083Cb protein lots that the N-terminus is blocked, rendering Edman degradation unsuccessful.

<u>Bioactivity Bioassay</u>

The biological activity of IPD083Cb protein was evaluated by conducting a 7-day bioassay using *Chrysodeixis includens* (soybean looper; Lepidoptera: Noctuidae), a species sensitive to the IPD083Cb protein. Eggs were obtained from [redacted] and their identity was recorded by study personnel.

The bioassay was conducted in a 48-well bioassay plate with the plate considered a single block. Treatments were randomized by column with a total of 24 replicates per treatment. *C. includens* larvae were exposed via oral ingestion to one of the following two treatments prepared by surface application of dosing solutions to the targeted protein concentration per cm2 diet:

- Treatment 1: Buffer Control Diet (containing a dosing solution of 25 mM Tris-HC1, pH 8, 100 mM NaCl, 0.01% PS-80; referred to as buffer)
- Treatment 2: Test Diet (targeting 50 µg IPD083Cb protein per cm^2)

The carrier for the *C. includens* bioassay consisted of an agar-based artificial diet prepared by Pioneer by blending dry Southland pre-mix diet and other diet ingredients with molten agar. The blended diet was stirred on a hot plate until ready to dispense. Prior to the beginning of the bioassay, approximately 500 µl of freshly prepared agar diet was dispensed into wells of the bioassay plates and allowed to cool and solidify. Plates were stored refrigerated (~ 4 °C) until use.

An IPD083Cb stock solution was prepared by solubilizing 5.0 mg of test substance in 1.788 ml of chilled buffer to a nominal concentration of 1.6 mg/ml. The stock solution was diluted in buffer to prepare the test dosing solution at a targeted concentration of 1500 ng/µl. The buffer control dosing solution consisted of the buffer used to dilute the test substance. The IPD083Cb stock solution and the test and buffer control dosing solutions were prepared fresh on each day of diet distribution (Day 0 and Day 4). Dosing solutions were maintained chilled (on wet ice) until use.

C. includens eggs were incubated in an environmental chamber until the eggs hatched. Neonates were used in the bioassay within 24 hours of hatching. On Day 0 of the bioassay, dosing solutions were dispensed topically by treatment to assigned wells using a surface application (25 µl per well) and plates were dried under a hood. One *C. includens* neonate was placed in each well containing diet, each bioassay plate was sealed with heat-sealing film, and a small hole was poked over each well to allow for ventilation. The bioassay was conducted in an environmental chamber set at 25 °C, 65% relative humidity, and continuous dark. On Day 4, new bioassay plates were prepared with fresh diet as described for Day 0, organisms were transferred to corresponding wells of the new plates, and the plates were sealed with a hole for ventilation and placed in the environmental chamber. After 7 days, the bioassay was complete, mortality was assessed, and surviving larvae were individually weighed.

The bioassay acceptability criterion indicated the bioassay may be repeated if the combined number of dead and missing organisms exceeds 20% for the buffer control diet (Treatment 1) group. The bioassay met the acceptability criteria (16.7% combined dead and missing).

Thermolability Analysis

The test substance consisted of IPD083Cb protein solubilized in study from a lyophilized powder (lot number PCF-0061A). The carrier consisted of an agar-based artificial diet. The bioassay control dosing solutions used to prepare Treatment 1 consisted of chilled ultrapure (American Society for Testing and Materials [ASTM] Type 1) water. The bulk dosing solutions used to prepare Treatments 2-6 consisted of aliquots of the test substance diluted in chilled ultrapure water to achieve the concentration in each treatment. The dosing solutions used to prepare Treatments 3-6 were incubated for 30-35 minutes at various temperatures.

The test system was *Anticarsia gemmatalis* (velvetbean caterpillar; Lepidoptera: Erebidae). The test system was chosen because *A. gemmatalis* is an insect sensitive to IPD083Cb protein. *A. gemmatalis* larvae were exposed via oral ingestion to one of the following six treatments prepared by surface application of dosing solutions to agar-based artificial diet:

- Treatment 1: Bioassay Control Diet with a dosing solution of ultrapure water
- Treatment 2: Control Diet with the unheated IPD083Cb protein dosing solution
- Treatment 3: Test Diet with the IPD083Cb protein dosing solution incubated at 25 °C
- Treatment 4: Test Diet with the IPD083Cb protein dosing solution incubated at 50 °C
- Treatment 5: Test Diet with the IPD083Cb protein dosing solution incubated at 75 °C
- Treatment 6: Test Diet with the IPD083Cb protein dosing solution incubated at 95 °C

Treatments 2-6 were prepared to a targeted protein concentration of 50 µg IPD083Cb protein per cm2 diet. Incubations for the heated test dosing solutions used to prepare Treatments 3-6 were performed for 30-35 minutes. Treatments were arranged in a generalized randomized block design with a total of four blocks. Each block consisted of a 48-well bioassay plate and contained six replicates randomized by column from each treatment. Each treatment was provided to a target of 24 *A. gemmatalis* individuals. The bioassay was conducted in an environmental chamber set at 25 °C, 65% relative humidity, and continuous dark. Larvae were refed on Day 4. After 7 days, the bioassay was complete, mortality was assessed, and surviving organisms were individually weighed.

The bioassay acceptability criteria indicated the bioassay may be terminated and repeated if: the combined number of dead and missing organisms exceeds 20% for the bioassay control (Treatment 1) group and the mortality rate does not exceed 80% in the unheated control (Treatment 2) group. The *A. gemmatalis* bioassay met the acceptability criteria (0% dead and 0% missing in Treatment 1; 100% mortality in Treatment 2). Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) and densitometry analysis were used to determine the concentration of the IPD083Cb protein test substance and verify the concentration of IPD083Cb protein in the bulk dosing solutions used to prepare Treatments 2-6. Bias in the A. gemmatalis bioassay was controlled through the randomization of treatments within blocks and the use of one or more control diets.

On each day of diet preparation, dosing solutions for Treatments 1-6 were prepared and heat-treated as applicable. Dosing solutions were maintained chilled (in a refrigerator set at 4 °C or on wet ice) until use. *A. gemmatalis* eggs were incubated in an environmental chamber until the eggs hatched. *A. gemmatalis* neonates were used in the bioassay within 24 hours of hatching.

On Day 0 of the bioassay, dosing solutions were dispensed by treatment to assigned wells using a surface application (25 µl per well) and plates were dried under a hood. One A. gemmatalis neonate was placed in each well containing a treatment. Each plate was sealed with heat-sealing film and ventilated with a small hole over each well. The plates were placed in an environmental chamber. On Day 4, new bioassay plates were removed from refrigerated storage and prepared with surface applications of the dosing solutions as described for Day 0, living organisms were transferred to the new plates, missing or dead organisms were recorded, and the freshly prepared plates were placed in the environmental chamber. After 7 days, the bioassay was complete, mortality was assessed, and surviving larvae were individually weighed.

Statistical Analysis

Statistical analyses of data were conducted using SAS software, Version 9.4 (SAS Institute Inc.). The response variable of interest was mortality. Statistical comparisons were made between A. gemmatalis fed diets containing heated IPD083Cb protein (Treatments 3-6) and those fed a diet containing unheated IPD083Cb protein (Treatment 2) for mortality. Statistical analysis was conducted using Fisher's exact test to determine if the mortality rate of A. gemmatalis fed diets containing the heated IPD083Cb protein (m_T) was less than the mortality rate of those fed the control diet with unheated IPD083Cb protein diet (m_C). The corresponding hypothesis test was:

$$H_0: m_T - m_C = 0 \quad vs. \quad H_a: m_T - m_C < 0.$$

A significant difference was established if the P-value was < 0.05. SAS PROC MULTTEST was utilized to conduct the Fisher's exact test.

Digestibility in Simulated Gastric Fluid (SGF)

Test and control solutions were prepared as follows:

- The gastric control solution was prepared fresh on the day of use and was comprised of 0.2% weight per volume (w/v) NaCl in 0.7% volume per volume (v/v) HCl, with a pH of ~1.2.
- The pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use by dissolving pepsin (Sigma-Aldrich) into gastric control solution. The SGF was prepared so that the pepsin to protein ratio of the final digestion mixture was approximately 10 units of pepsin per µg of test or control protein.
- The test substance consisted of IPD083Cb protein solubilized from a lyophilized powder (lot number PCF-0061A).
- To prepare the solutions for each of the control proteins (BSA and β-lactoglobulin), a 5.0-mg sub sample of powder was weighed into an individual tube for each control and solubilized by adding 1.0 ml of ultrapure water to a target protein concentration of 5 mg/ml.
- The final concentration of the protein and pepsin in the SGF reaction mixture was 0.25 mg/ml IPD083Cb protein and 2500 units/ml pepsin.

SGF solution (1900 µl) was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to the addition of 100 µl of IPD083Cb protein solution at Time 0. The digestion reaction mixture was mixed constantly using a stir bar and a submersible magnetic stirrer.

A 120-µl sub-sample of the IPD083Cb protein digestion reaction mixture was removed from the vial at the following analytical time points (± 10 seconds): 0.25, 1, 2, 5, 10, 30, and 60 minutes. The sub samples were inactivated by neutralization with 139 µl of pre-mixed sample stop solution and heating to 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

To prepare control digestion samples at 1 and 60 minutes, a 114-µl sample of the appropriate digestion solution (see table) was pre-warmed in a 37 °C water bath for 2-5 minutes prior to adding

6 µl of the IPD083Cb protein solution, control protein solution, or ultrapure water. The tubes were incubated in the water bath for the allotted time and then inactivated by mixing with 139 µl of the pre-mixed sample solution and heating at 90-100 °C for 5 minutes.

The Time 0 control reaction mixtures were prepared by first neutralizing 114 µl of the appropriate digestion solution (see table) with 139 µl of the pre-mixed sample solution, and then heating at 90 100 °C for 5 minutes. A 6-µl sample of the appropriate IPD083Cb protein solution, control protein solution, or ultrapure water was then added, and the mixture was heated again at 90 100 °C for 5 minutes.

Control digestion samples included in the SGF assay are provided in Table E.1. Following digestion and inactivation, all control reaction mixtures were heated at 90-100 °C for 5 minutes prior to frozen storage (-20 °C freezer unit).

Control digestion samples included in the SGF assay are provided in Table G.1.

Table G.1. Control Samples for Simulated Gastric Fluid (SGF) Digestibility Analysis

Protein	Digestion Solution	Digestion Time (min)		
		0	1	60
None (SGF Control – Ultrapure water)	SGF	X	--	X
BSA	SGF	X	X	X
β-lactoglobulin	SGF	X	X	X
IPD083Cb	SGF	X	--	--
IPD083Cb	None (Ultrapure Water)	X	--	X
IPD083Cb	Gastric Control Solution (No Pepsin)	--	--	X

SDS-PAGE Analysis

The IPD083Cb protein digestion time-course samples and control samples were removed from frozen storage, heated at 90-100 °C for 5 minutes, and loaded (10 µl/well) into 4-12% Bis-Tris gels for SDS PAGE analysis. To demonstrate the sensitivity of the SDS-PAGE gel and western blot analyses, an aliquot of the IPD083Cb protein in SGF (Time 0) sample was loaded into the gel at a 1:20 dilution for protein staining, and at a 1:100 dilution for the western blot. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into each gel. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gels were removed from the gel cassettes for use in protein staining or western blot analyses. For protein staining, the gels were washed with water and stained with GelCode Blue Stain Reagent. Following staining, the gels were destained with water until the gel background was clear. Protein bands were stained on the gels. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

The IPD083Cb protein digestion time-course samples were also analyzed by western blot. Following SDS PAGE, one of the resulting gels was assembled into a mini nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in phosphate-buffered saline containing polysorbate 20 (PBST) containing 5% (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with an IPD083Cb polyclonal antibody 3373 (Pioneer Hi Bred International, Inc.) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. Following primary antibody incubation, the membrane was washed in PBST and incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP imaging system.

<u>Digestibility in Simulated Intestinal Fluid (SIF)</u>

Test and control solutions were prepared as follows:

- The pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by dissolving 52.6 mg of pancreatin (Sigma-Aldrich) into 5 ml of intestinal control solution (I-Con 1X buffer) to a final concentration of 0.5% weight per volume (w/v) pancreatin and 50 mM KH2PO4, pH 7.5.
- To prepare the solutions for each of the control proteins (BSA and β-lactoglobulin), a 5.0-mg sub sample of powder was weighed into an individual tube for each control and solubilized by adding 1.0 ml of ultrapure water to a target protein concentration of 5 mg/ml.
- The test substance consisted of IPD083Cb protein solubilized from a lyophilized powder (lot number PCF-0061A).
- The final concentration of the protein and pancreatin in the SIF reaction mixture was 0.25 mg/ml IPD083Cb protein and 1% (w/v) pancreatin.

SIF solution (1900 µl) was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to the addition of 100 µl of IPD083Cb protein solution at Time 0. The digestion reaction mixture was mixed constantly using a stir bar and a submersible magnetic stirrer.

A 120-µl sub-sample of the IPD083Cb protein digestion reaction mixture was removed from the vial at the following analytical time points (± 10 seconds): 0.25, 1, 2, 5, 10, 30, and 60 minutes. The sub samples were inactivated by adding them to pre-labeled tubes containing 64 µl of pre-mixed sample solution and heating to 90-100 °C for 5 minutes prior to SDS-PAGE analysis and subsequent frozen storage (-20 °C freezer unit).

To prepare control digestion samples at 1 and 60 minutes, a 114-µl sample of the appropriate digestion solution (see table below) was pre-warmed in a 37 °C water bath for 2-5 minutes prior to adding 6 µl of the IPD083Cb protein solution, control protein solution, or ultrapure water. The tubes were incubated in the water bath for the allotted time and then inactivated by mixing with 64 µl of the pre-mixed sample solution.

The Time 0 control reaction mixtures were prepared by first neutralizing 114 µl of the appropriate digestion solution (see table below) with 64 µl of the pre-mixed sample solution, and then heating at 90-100 °C for 5 minutes. A 6-µl sample of the appropriate IPD083Cb protein solution, control protein solution, or ultrapure water was then added, and the mixture was heated again at 90-100 °C for 5 minutes.

Following digestion and inactivation, all control digestion samples were analyzed by SDS-PAGE analysis and then stored frozen (-20 °C freezer unit).

Control digestion samples included in the SIF assay are provided in Table G.2.

Table G.2. Control Samples for Simulated Intestinal Fluid (SIF) Digestibility Analysis

Protein	Digestion Solution	Digestion Time (min)		
		0	1	60
None (SIF Control-Ultrapure water)	SIF	X	--	X
BSA	SIF	X	X	X
β-lactoglobulin	SIF	X	X	X
IPD083Cb	SIF	X	--	--
IPD083Cb	None (Ultrapure water)	X	--	X
IPD083Cb	Intestinal Control Solution (No Pancreatin)	--	--	X

SDS-PAGE Analysis

The IPD083Cb protein digestion time-course samples and control samples were heated at 90-100 °C for 5 minutes and loaded (10 µl/well) into 4-12% Bis-Tris gels for SDS PAGE analysis. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into each gel. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gels were removed from the gel cassettes for use in protein staining or western blot analyses. For protein staining, the gels were washed with water and stained with GelCode Blue Stain Reagent. Following staining, the gels were destained with water until the gel background was clear. Protein bands were stained on the gels. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

The IPD083Cb protein digestion time-course samples were also analyzed by western blot. Following SDS PAGE, one of the resulting gels was assembled into a mini nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in phosphate-buffered saline containing polysorbate 20 (PBST) containing 5% (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with an IPD083Cb polyclonal antibody 3373 (Pioneer Hi Bred International, Inc.) diluted 1:100,000 in PBST containing 1% (w/v) non-fat dry milk. Following primary antibody incubation, the membrane was washed in PBST and incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:100,000 in

PBST containing 1% (w/v) non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP imaging system.

Sequential Digestibility Analysis with Simulated Gastric Fluid (SGF) and Simulated Intestinal Fluid (SIF)

Test solutions were prepared as follows:

- A 2X concentrated pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use by solubilizing pepsin (Sigma-Aldrich) in a previously prepared 2X gastric control solution. The final concentration of gastric control solution in the final digestion mixture was 0.2% weight per volume (w/v) NaCl and 0.7% volume per volume (v/v) HCl; pH ~1.2. The SGF was prepared so that the pepsin to protein ratio of the final digestion mixture was approximately 8.6 units of pepsin per µg of test substance.
- A 2.5X concentrated pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by solubilizing pancreatin (Sigma-Aldrich) in 2.5X intestinal control solution. The final concentration of intestinal control solution in SIF was 50 mM KH2PO4, pH ~7.5. Pancreatin content in SIF was adjusted so that there was approximately 0.5% (w/v) pancreatin in the final digestion reaction mixture.
- The pre-mixed sample solutions used to inactivate samples were prepared fresh on the day of use. The stop solution for SGF reactions was prepared by mixing 1200 µl of 200 mM Na2CO3, 1625 µl of NuPAGE 4X LDS Sample Buffer, and 650 µl NuPAGE 10X Sample Reducing Agent containing DTT. The sample solution for SIF reactions was prepared by mixing 1150 µl NuPAGE 4X LDS Sample Buffer and 450 µl NuPAGE 10X Sample Reducing Agent containing DTT.

In Vitro Pepsin Digestion

IPD083Cb Protein in SGF 10 Minutes Sample for Sequential Digestion: A 1-ml aliquot of the 2X SGF solution was dispensed into a 7-ml glass vial and pre-warmed in the 37 °C water bath for 2-5 minutes prior to addition of 800 µl ultrapure water and 200 µl of the IPD083Cb protein solution. The SGF digestion reaction mixture was incubated and mixed constantly using a stir bar and submersible stir plate for 10 minutes (± 10 seconds) after adding the IPD083Cb protein. At the end of the time period, a 1.5-ml sample of the IPD083Cb SGF digestion reaction mixture was transferred to a separate vial and inactivated by neutralization with 0.3 ml of 0.5 N NaOH. This sample was used for the sequential SIF digestion.

IPD083Cb Protein in SGF 10 Minutes: A sample (120 µl) was collected from the IPD083Cb SGF digestion reaction mixture at the end of 10 minutes (± 10 seconds) and inactivated by neutralization with 139 µl of pre-mixed SGF sample stop solution. The neutralized sample was heated for 5 minutes at 90-100 °C prior to frozen storage (-20 °C freezer unit).

IPD083Cb Protein in SGF Time 0: A control sample was prepared by mixing 48 µl of 2.5X SIF with 64 µl of pre-mixed SIF sample solution and then heating for 5 minutes at 90-100 °C. A 72-µl sub-sample of the NaOH-neutralized IPD083Cb SGF digestion reaction mixture was added to the heat-inactivated SIF control sample and then heated again for 5 minutes at 90 100 °C. After inactivation and heating, all SIF reaction samples were stored frozen (-20 °C freezer unit).

SGF-Only 10 Minutes Incubation: An SGF only control sample (without IPD083Cb protein) was prepared by mixing 60 µl of 2X SGF with 49 µl ultrapure water and then pre-warming in a 37 °C water bath for 2-5 minutes. After pre-warming, 12 µl of ultrapure water was added and the SGF solution was incubated in the water bath for 10 minutes. At the end of the time period, 139 µl of SGF sample stop solution was added and the neutralized sample was then heated at 90 100 °C for 5 minutes.

Sequential Pancreatin Digestion

For the sequential SIF digestion time-course, a 1.2-ml sample of the NaOH-neutralized IPD083Cb SGF digestion reaction mixture was dispensed into a 7-ml glass vial and placed in a 37 °C water bath for 2-5 minutes prior to addition of 800 µl of 2.5X SIF solution. The SIF digestion reaction mixture was mixed constantly using a stir bar and a submersible stir plate. A 120-µl sub sample of the SIF digestion reaction mixture was removed from the vial at each of the following analytical time points (± 10 seconds): 0.25, 1, 2, 5, 10, 20, and 30 minutes. Each sub sample was inactivated by adding it to a pre-labeled tube containing 64 µl of pre-mixed SIF sample solution and heating at 90-100 °C for 5 minutes. After heating, all SIF reaction samples were stored frozen (-20 °C freezer unit).

SDS-PAGE Analysis

The IPD083Cb protein digestion time-course samples and control samples were removed from frozen storage, heated at 90-100 °C for 5 minutes, and loaded (10 µl/well) into a 4-12% Bis-Tris gel for SDS PAGE analysis. Pre-stained protein molecular weight markers (Precision Plus Dual Xtra Standards) were also loaded into the gel. Electrophoresis was conducted using a pre-cast gel electrophoresis system with 1X MES SDS running buffer at a constant 200 volts (V) for 35 minutes.

Upon completion of electrophoresis, the gel was removed from the gel cassette for use in protein staining. For protein staining, the gel was washed with water and stained with GelCode Blue Stain Reagent. Following staining, the gel was destained with water until the gel background was clear. Protein bands were stained on the gel. The gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

APPENDIX H. METHODS FOR CHARACTERIZATION OF THE GM-HRA PROTEIN

<u>Test Materials</u>

COR23134 Soybean-Derived GM-HRA Protein

The GM-HRA protein was isolated from whole plant tissue derived from COR23134 soybean. The whole plant tissue was collected at the V5 growth stage (the stage when the leaflets on the sixth leaf node have unrolled (Pedersen, 2004)) of development from plants grown at a Pioneer owned field location (Johnston, IA, USA). The tissue was lyophilized, homogenized, and stored frozen (-80 °C freezer unit). The GM-HRA protein was extracted from lyophilized soybean tissue by homogenization in a pre-chilled Waring blender vessel using an extraction buffer comprised of 50 mM Tris-HCl, pH 7.5, 5 mM sodium pyruvate, 10 mM FAD, 1 mM EDTA, 5% glycerol, 5 mM magnesium chloride, and 50 mM sodium chloride. Protease inhibitors and 20% (based on tissue weight) polyvinylpolypyrrolidone were also added. The sample extract was then clarified by centrifugation and filtration. The filtered extract was purified by immunoaffinity chromatography. The immunoaffinity column was prepared by coupling a GM-HRA monoclonal antibody (16A3.D11; Pioneer) to AminoLink Plus Coupling Resin (Thermo Scientific). The GM-HRA protein sample was eluted off the column using IgG Elution buffer (Thermo Scientific). Select elutions were collected and concentrated using a centrifugal concentrator (30K Vivaspin; Sartorius) to a volume of approximately 500 µl. The concentrated sample was buffer exchanged using 50 mM Tris buffer, pH 8, and then concentrated again to a volume of approximately 100 µl.

Following extraction, purification, and concentration, the final volume in the concentrator was estimated, NuPAGE LDS Sample Buffer (Life Technologies) was added at 25% along with 10% NuPAGE Reducing Agent containing DTT (Life Technologies) to the concentrated sample in the concentrator. The sample in the concentrator was transferred to a microcentrifuge tube, heat-treated, and stored frozen (-20 °C freezer unit) until use in characterization of the GM-HRA protein.

Microbially Derived GM-HRA Protein

In order to have sufficient amounts of the purified GM-HRA protein for the multiple studies required to assess its safety, the GM-HRA protein was produced at Aldveron using a microbial expression system. The protein was expressed in an *E. coli* BL21 (DE3) RIPL as a fusion protein containing a His-T7 tag protein expression system and purified using an immobilized metal affinity column. The His-T7 tag was cleaved from the affinity purified protein with thrombin and diafiltration was used to remove the cleaved tag and thrombin. Thrombin cleavage resulted in one additional N-terminal amino acid residue, glycine, on the microbial GM-HRA protein which is not found in the plant expressed mature GM-HRA protein. The purified microbial GM-HRA protein was dialyzed into 100 mM ammonium bicarbonate pH 7.5 and then lyophilized. After lyophilization and mixing, a lot number was assigned.

SDS-PAGE Analysis

The purified COR23134 soybean-derived GM-HRA protein sample was allowed to thaw, diluted as applicable for the sensitivity of the assay, heated, and then loaded into 4-12% Bis-Tris gels along with pre-stained protein molecular weight markers (Precision Plus Protein Dual Xtra Standards). For applicable SDS PAGE and western blot analysis, the GM-HRA protein reference substance was diluted in 1X LDS/DTT, heated, and loaded into the gels to 1 µg for SDS PAGE and 10 ng for western blot analysis. Electrophoresis was conducted using a pre-cast gel electrophoresis system with MES running buffer at a constant 200 volts for 35 minutes.

Upon completion of electrophoresis, the gels were either prepared for protein staining or protein transfer to a membrane for sequencing or western blot analysis.

For Coomassie staining, following electrophoresis, the gel was washed with water and stained with GelCode Blue Stain Reagent (Thermo Scientific). Following staining, the gel was de stained with water until the gel background was clear. Protein bands were stained on the gels and the gel image was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system.

Western Blot Analysis

Following SDS-PAGE as described above, the resulting gel was assembled into a nitrocellulose (NC) iBlot Gel Transfer Stack. An iBlot Gel Transfer Device was used to transfer proteins from the gel to the NC membrane.

Following protein transfer, the membrane was blocked in PBST containing 5% weight/volume (w/v) non-fat dry milk. Before and after the blocking step, the membrane was washed with PBST to reduce the background. The blocked membrane was incubated with a GM-HRA polyclonal rabbit antibody R9222 (Pioneer Hi-Bred International, Inc.) diluted 1:5000 in PBST containing 1% w/v non-fat dry milk. Following primary antibody incubation, the membrane was washed with PBST. The membrane was incubated with a secondary antibody (anti-rabbit IgG, horseradish peroxidase conjugate; Promega Corporation) diluted 1:10,000 in PBST containing 1% w/v non-fat dry milk. The membrane was washed and remained in PBST prior to incubating with a chemiluminescent substrate. The chemiluminescent signal and the pre-stained markers were detected and captured using a ChemiDoc MP (Bio Rad) imaging system.

Protein Glycosylation Analysis

COR23134 soybean-Derived GM-HRA Protein

A Pierce Glycoprotein Staining Kit (Thermo Fisher) was used to determine whether the COR23134 soybean-derived GM-HRA protein was glycosylated. The purified soybean-derived GM-HRA protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor), were run by SDS PAGE as described above. For glycosylation staining, soybean-derived GM-HRA protein was loaded on to the gel at approximately the same concentration as the positive and negative control proteins (1.0 µg).

Following electrophoresis, the gel was washed with water, fixed with 50% methanol, and washed with 3% acetic acid. The gel was then incubated with oxidizing solution and washed with 3% acetic acid. The gel was incubated with glycoprotein staining reagent and then incubated in a reducing reagent. The gel was then washed with 3% acetic acid followed by water. Glycoproteins were detected as stained bands on the gel.

Following glycoprotein detection, the image of the gel was captured electronically using a ChemiDoc MP (Bio-Rad) imaging system. The same gel was then stained with GelCode Blue Stain Reagent (Thermo Scientific) followed by washes with water to visualize all protein bands. The image of the GelCode-stained gel was then captured electronically.

Microbially Derived GM-HRA Protein

A GelCode glycoprotein staining kit (Pierce Biotechnology, Inc.) was used according to the manufacturer's instructions to determine whether the microbial GM-HRA protein was glycosylated. The microbial GM-HRA protein, a positive control protein (horseradish peroxidase), and a negative control protein (soybean trypsin inhibitor) were separated by electrophoresis on SDS-PAGE as described in Section B. The microbial GM-HRA protein and the control proteins were loaded at concentrations of ~1 ug/lane and ~20 ug/lane, respectively. Following electrophoresis, the gel was fixed with 50% methanol for approximately 30 minutes and washed with 3% acetic acid. The gel was then incubated with oxidizing solution for approximately 15 minutes and washed three times with 3% acetic acid. The gel was incubated with GelCode glycoprotein staining reagent (Pierce Biotechnology, Inc.) for 15 minutes and then treated with reducing reagent. Next the gel was extensively washed with 3% acetic acid and deionized water. Glycoproteins were detected as magenta-colored bands on the gel. Following glycoprotein detection, the gel was scanned and the image captured electronically. The same gel was stained with Coomassie Blue as described in Section B to visualize the total protein content of all protein bands.

Peptide Mapping by Mass Spectrometry

COR23134 soybean-Derived GM-HRA Protein

For mass spectrometry sequencing analyses, the COR23134 soybean-derived GM-HRA protein was loaded onto a gel in each of two lanes. Following SDS-PAGE, Coomassie staining, and gel imaging using the methods as described above, bands containing the soybean derived GM HRA protein were excised from a gel and stored frozen (-20 °C freezer unit). The protein in each gel slice was reduced with DTT, alkylated with iodoacetamide, and then subsequently digested with trypsin or chymotrypsin. The digested samples were separated on a nanoACQUITY UPLC (Waters Corporation) fitted with a Peptide BEH C18 300 Å 1.7 µm column (75 µm x 100 mm; Waters Corporation) by gradient elution. Eluent from the column was directed into an electrospray source, operating in positive ion mode, on a TripleTOF 5600+ hybrid quadrupole-TOF mass spectrometer (AB Sciex; currently Sciex). The resulting mass spectrometry (MS) data were processed using MSConvert to produce a peak list. The peak list was used to perform an MS/MS ion search (Mascot Software version 2.8.0) and match peptides from the expected GM-HRA protein sequence (Pedersen, 2004). The following search parameters were used: peptide and fragment mass tolerance, ± 0.1 Da; fixed modifications, cysteine carbamidomethyl; variable modifications, methionine oxidation; and maximum missed cleavages, 1 for trypsin and 2 for chymotrypsin. The Mascot-generated peptide ion score threshold was > 13 which indicates identity or extensive homology ($P < 0.05$). The combined sequence coverage was calculated with GPMAW version 14.02.

Microbially Derived GM-HRA Protein

Following electrophoresis, the microbial GM-HRA protein band was visualized by staining with Coomassie Blue and the band was then excised from the gel. The gel slice was placed in a labeled tube and shipped overnight on dry ice to the Keck Biotechnology Resource Laboratory (Yale University, 300 George Street, Box 201, New Haven, CT 06511, USA) for trypsin digestion and MALDI-MS analysis. The protein in the gel slice was digested with trypsin for 18 hours at 37°C. An aliquot of the digest was analyzed by MALDI-MS on a Waters MALDI-L/R spectrometer (Waters Corporation) in reflectron mode of operation. Detected peptide peaks were considered a match if the observed experimental mass was within 100 parts per million (ppm) of the theoretical mass of peptides derived from trypsin cleavage of the mature GM-HRA protein. Allowances were also made for the following expected modifications: oxidation of methionine or tryptophan residues (observed value is 15.995 Da greater than the theoretical value) and modification of cysteine residues by acrylamide free radicals during SDS-PAGE (observed value is 71.037 Da greater than the theoretical value).

N-Terminal Amino Acid Sequencing Analysis

COR23134 Soybean-Derived GM-HRA Protein

For N-terminal amino acid sequence analysis, COR23134 soybean-derived GM-HRA protein was loaded onto a gel. Following SDS-PAGE using the methods as described above, the resulting gel was incubated in cathode buffer (60 mM Tris, 40 mM CAPS, 0.075% SDS, pH 9.6). An Immobilon-P PVDF membrane (Millipore) was wetted in 100% methanol, followed by immersion in anode buffer (60 mM Tris, 40 mM CAPS, 15% methanol, pH 9.6). A Trans-Blot SD Semi Dry Electrophoretic Transfer Cell system (Bio-Rad) was used to transfer proteins from the gel to the membrane. Following protein transfer, the membrane was washed with water, stained with GelCode Blue Stain Reagent (Thermo Scientific), and then destained with water to visualize the GM-HRA protein band. A band containing GM-HRA protein was excised and stored frozen (-20 °C freezer unit). The band was analyzed using a Shimadzu PPSQ-51A sequencer. Ten cycles of Edman sequencing were performed. During each cycle, the N terminal amino acid was sequentially derivatized with phenylisothiocyanate (PITC), cleaved with trifluoracetic acid, and converted to PTH amino acid which was identified through chromatography. LabSolutions Software was used to automatically identify the N-terminal sequence with manual adjustments as necessary.

Microbially Derived GM-HRA Protein

The microbial GM-HRA protein sample was separated by SDS-PAGE and electrophoretically transferred to a PVDF membrane as described in Section C. After transfer, the PVDF membrane was stained with Ponceau S solution (Sigma-Aldrich, 3050 Spruce Street, St. Louis, MO 63103, USA) to visualize the GM-HRA protein band. The resulting band was excised and shipped to the Keck Biotechnology Resource Laboratory (Yale University, 300 George Street, Box 201, New

Haven, CT 06511, USA) for Edman N-terminal amino acid sequencing using the Procise 494 cLC analyzer (Applied Biosystems, Inc., 850 Lincoln Centre Drive, Foster City, CA 94404, USA) equipped with an online high performance liquid chromatography (HPLC) system.

Bioactivity Bioassay

The bioactivity of the GM-HRA protein was verified by a spectrophotometric assay for enzymatic activity measurement of acetolactate synthase (ALS).

The enzyme acetolactate synthase (ALS), also known as acetohydroxyacid synthase (AHAS) (EC 2.2.1.6, formerly EC 4.1.3.18), catalyzes the first common step in the biosynthesis of the essential branched-chain amino acids isoleucine, leucine, and valine. Two reactions are catalyzed by ALS enzymes, the conversion of two molecules of pyruvate to form acetolactate leading to the synthesis of leucine and valine and the condensation of pyruvate with 2-ketobutyrate to form 2-acetohydroxybutyrate in the pathway to isoleucine.

The spectrophotometric assay for detecting ALS activity involves an indirect detection of the enzyme product, acetolactate. Following incubation of the enzyme with the substrate (pyruvate), the assay involves the conversion of the end product (acetolactate) to acetoin by decarboxylation with sulfuric acid and high temperature. Acetoin produced is detected by a modified Westerfield method by formation of a creatine and α-naphtol complex and measuring the O.D. at 530 nm.

The microbial GM-HRA lyophilized powder was resuspended in 2 mM phosphate buffer pH 7.4, 177.4 mM NaCl, 0.54 mM KCl, 10% glycerol, 0.5 mM thiamine pyrophosphate (TPP), 2 mM flavin adenine dinucleotide (FAD), 0.1 mM pyruvate, 0.5 mM MgCl2 and then diluted in extraction buffer [0.1 M phosphate buffer pH 7.5, 10% glycerol, 0.5 mM TPP, 20 µM FAD, and 2 mM MgCl2]. An acetoin standard curve was prepared in 0.1 M phosphate buffer, pH 7.5 and triplicate 100 µl aliquots of each protein sample and standard dilution was dispensed into a 96 well plate. Then 5 µl of either 100 mM phosphate buffer or the ALS inhibitor, chlorsulfuron (5µg/ml) was added to wells as applicable. The enzymatic production of acetolactate from pyruvate was initiated by the addition of 10 µl of a 1.1 M pyruvate solution to each well, followed by incubation at 37 °C for 1 hour.

The acetolactate was then converted to acetion by adding 5 µl of 2 N H2SO4 to each well, followed by incubation at 60 °C for 15 minutes. The plates were allowed to cool at room temperature for 15 additional minutes prior to the detection step.

Indirect detection of the enzymatic reaction was done by adding 50 µl of creatine/naphtol solution per well, followed by incubation at 60 °C for 15 minutes. Plates were allowed to cool at room temperature for an additional 15 minutes and then read at 530 nm using a SpectraMax Model 190 spectrophotometer (Molecular Devices Corporation). The relative amount of product produced by the enzyme sample was interpolated from the standard curve.

Thermolability Analysis

The test substance consisted of the GM-HRA lyophilized protein (lot # PCF-0008) powder. The protein was first dissolved in deionized water at 3 milligrams (mg) powder per milliliter (ml) and diluted with an equal volume of 2X ALS buffer (4.0 mM magnesium chloride, 1.0 mM thiamine pyrophosphate, 40 µM flavin adenine dinucleotide, 20% glycerol, 200 mM phosphate buffer pH

7.5). The suspension was then centrifuged in a microcentrifuge at 13,000 rpm for approximately 5 minutes. The supernatant was removed, and the soluble protein concentration was determined using Coomassie Plus Protein Assay Reagent (Pierce Biotechnology, Inc.) according to the manufacturer's instructions in which bovine serum albumin (BSA) was used as the standard and absorbance was measured at 595 nanometers. The concentration of this solution was then adjusted to a final concentration of 100 micrograms (μg) of protein per ml with 1X ALS buffer.

The GM-HRA samples in individual wells were heated for 15 minutes at a designated temperature ranging from 36-60 °C with 2 °C increments in a gradient thermocycler and analyzed for activity by dispensing three 100 μl replicates of each non-heated and heat-treated GM-HRA protein sample into a 96 well plate. Either 5 μl of 100 mM phosphate buffer or chlorsulfuron (5 μg/ml) was added to each well as applicable. Enzyme production of acetolactate from pyruvate was initiated by the addition of 10 μl of a 1.1 M pyruvate solution to each well, followed by incubation at 37 °C for 60 minutes. The acetolactate was then converted to acetoin by adding 5 μl of 2 N H2SO4 to each well, followed by incubation at 60 °C for 15 minutes. The plates were allowed to cool at room temperature for 15 additional minutes. Indirect detection of the enzymatic reaction was performed by adding 50 μl of creatine/napthol solution per well, followed by incubation at 60 °C for 15 minutes. Plates were allowed to cool at room temperature for 15 minutes and then read at 530 nm using a spectrophotometer. The relative amount of product produced by the enzyme sample was interpolated from the acetoin standard curve.

Enzymatic Activity

The OD values obtained from the spectrophotometer readings were converted to nM of acetoin by interpolation of the standard curve (see Figure 1). The standard curve yielded the following equation:

$OD_{530\ nm}$ = (slope of the acetoin standard curve)(concentration of acetoin (nM)) + $OD_{530\ nm}$ at 0 nM of acetoin

$OD_{530\ nm}$ = (0.0343)(concentration of acetoin (nM)) + 0.1469 or

nM acetoin = ($OD_{530\ nm}$ - 0.1469) / (0.0343)

For example, at an OD_{530nm} value of 1.531:

nM Acetoin = (1.531 - 0.1469) / 0.0343 = 40.353 nM Acetoin

The nM of Acetoin were converted into a rate using the following equation:

nM acetoin mg total protein^{-1} minute^{-1} = (nM Acetoin * 1000 μg/mg) / (μg total protein per well) (time of 37°C incubation in minutes)

From the previous example:

Rate = (40.353 nM * 1000 ug/mg) / (2 μg * 60 minutes) = 336.3 nM acetoin mg total protein^{-1} minute^{-1}

The amount of residual GM-HRA enzymatic activity for each treatment temperature was then calculated as follows. An average rate value of the three unheated samples was determined and the average rate value for a given heated sample was compared to the average rate of the unheated samples using the following equation:

Percent residual enzymatic activity = (average rate of acetoin production for a heated sample / average rate of unheated sample) * 100

For example (heated at 36°C, analyzed in the absence of chlorsulfuron):

= (306 / 333) *100 = 91.9%

Limit of Detection (LOD)

The average OD_{530} and standard deviation (SD) was calculated for the 0 nM acetoin level of the standard curve. The LOD was calculated as follows:

LOD = (average OD_{530} at 0 nM acetoin) + (3)(SD of 0 nM acetoin)

For example, with an average OD_{530} of 0.138 and a SD of 0.0045

LOD = 0.138 + (3 * 0.0045) = 0.152

The OD_{530} values below the LOD were considered equal to zero and therefore, the rate was assigned a value of zero.

<u>Digestibility in Simulated Gastric Fluid (SGF)</u>

Test and control solutions were prepared as follows:

- The gastric control solution was prepared as per US Pharmacopoeia with final concentrations of 0.084 N HCl and 35 mM NaCl at pH 1.2.
- The pepsin digestion solution, referred to as simulated gastric fluid (SGF), was prepared fresh on the day of use and was adjusted to approximately 10 units/µg of the test and control proteins in the final digestion mixture based on stated pepsin activity of 3,370 units per mg of protein, or approximately 0.74 mg/ml pepsin.
- The test substance consisted of GM-HRA protein solubilized from a lyophilized powder (lot number PCF-0008).
- The stock solutions for each of the control proteins (BSA and β-lactoglobulin), were dissolved in water to a concentration of 5 mg/ml.
- The final concentrations in the SGF reaction mixture were 0.25 mg/ml GM-HRA protein in 0.084 N HCl

SGF solution, 1900 microliters (µl), was dispensed into a 4 ml glass screw top vial and placed in a 37°C water bath for approximately two minutes prior to addition of 100 µl of GM-HRA protein stock solution. The reaction mixture was constantly mixed using a stir bar and submersible stir plate.

A 120 µl sub-sample of the digestion reaction was neutralized (stopped) by mixing with 48 µl of 200 mM Na2CO3 (pH 11.0), 65 µl NuPAGE LDS 4X Sample Buffer and 26 µl of NuPAGE4 Sample Reducing Agent 10X containing 0.5 M dithiothreitol, in a tube at each of the following analytical time points (+/- 10 seconds): 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. Time zero for each SGF reaction mixture was prepared by first neutralizing a sample of SGF and then adding test protein. Stopped samples were kept on ice.

To prepare control digestion samples at prepared at approximately 60 minutes, SGF alone (no test protein) prepared at time zero and approximately 60 minutes, and gastric control solution containing GM-HRA (no pepsin) prepared at approximately 60 minutes. The BSA and β-Lactoglobulin controls were prepared at the following time points 0, 1 (+/- 10 seconds), and approximately 60 minutes. A 120 µl aliquot of the control solutions were neutralized (stopped) by mixing with 48 µl of 200 mM Na2CO3 (pH 11.0), 65 µl NuPAGE5 LDS 4X Sample Buffer (Invitrogen Corporation), and 26 µl of NuPAGE5 Sample Reducing Agent 10X containing 0.5 M dithiothreitol. Stopped samples were kept on ice until preparation for electrophoresis.

Control digestion samples included in the SGF assay are provided in Table H.1.

Table H.1. Control Samples for Simulated Gastric Fluid (SGF) Digestibility Analysis

Protein	Pepsin Resistance Determined by SDS-PAGE	Pepsin Resistance Determined by Western Blot	Approximate Molecular Weight (kDa)
GM-HRA	<30 seconds	<30 seconds	~65
BSA (positive control)	<60 seconds	NA[1]	~66
β-Lactoglobulin (negative control)	>60 minutes	NA	~18

[1] NA- Not applicable, the control samples were not analyzed by Western blotting procedure.

SDS-PAGE Analysis

SDS-PAGE was performed using NuPAGE Novex 4-12% Bis-Tris Gels (Invitrogen Corporation) with 12 wells. Protein samples were prepared for electrophoresis by heating for approximately ten minutes at 70°C. Samples were loaded at 20 µl/well and SeeBlue5 Pre-stained Standard (Invitrogen Corporation) was loaded at 13 µl/well. Electrophoresis was conducted with NuPAGE5 MES SDS running buffer (Invitrogen Corporation) at a constant 200 volts for approximately 30 minutes. Once complete, gels were removed from the cassette and stained or used for Western blot analysis.

After electrophoresis, gels were removed from the gel cassette and washed 3 times for at least 5 minutes each wash with de-ionized water prior to staining in RAPIDStain (G-Biosciences) for approximately 45 minutes. Gels were destained using de-ionized water. Destained gels were evaluated for apparent disappearance of the GM-HRA protein.

Western Blot Analysis

The SGF digestion reaction time course samples for the GM-HRA protein were analyzed by Western blot. After SDS-PAGE, gels were removed from the gel cassette and electrophoretically transferred to a polyvinylidene fluoride (PVDF) membrane (Invitrogen, # LC2002) for approximately 1 hour using a XCell II5 Blot Module (Invitrogen, #EI5091) with NuPAGE5 Transfer Buffer (Invitrogen, #NP0006) and a constant voltage of 35 V.

Following transfer, the PVDF membrane blot was blocked with a phosphate-buffered saline solution with Tween -20 (PBST: 8.1 mM phosphate buffer, pH 7.4, 137 mM sodium chloride, 2.7

mM potassium chloride, and 0.05% Tween7-20) containing 5% low dry milk solution followed by a 30 minute incubation in a 1:10,000 dilution of a primary anti-GM-HRA antibody (affinity purified rabbit polyclonal; R8001) in 1% low fat milk/PBST solution. The blot was washed with PBST at least three times, for approximately 10 minutes each wash and then incubated for approximately 30 minutes with a 1:10,000 dilution of the secondary Donkey anti-Rabbit IgG antibody conjugated to Horseradish Peroxidase (Promega U.S.). The unbound secondary antibody-HRP conjugate was removed by another 3 washes of at least 10 minutes each in PBST.

Digestibility in Simulated Intestinal Fluid (SIF)

Test and control solutions were prepared as follows:

- The pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by dissolving in intestinal control solution at a final concentration of 1% w/v with a final concentration of 50 mM KH2PO4, pH of 7.5.
- The test substance consisted of GM-HRA protein solubilized from a lyophilized powder (lot number PCF-0008).
- The stock solutions for each of the control proteins (BSA and β-lactoglobulin), were dissolved in water to a concentration of 5 mg/ml.
- The final nominal concentration of GM-HRA was 0.25 mg/ml in the SIF (50 mM KH2PO4 and 1% w/v pancreatin, pH 7.5) reaction mixture.

SIF solution (1900 µl) was dispensed into a 4 ml glass screw top vial and placed in a 37 °C water bath for approximately two minutes prior to addition of 100 µl of GM-HRA protein stock solution. The solution was constantly mixed using a stir bar and submersible stir plate.

A 120 µl sub-sample of the digestion reaction was inactivated by mixing with 46 µl of NuPAGE LDS 4X Sample Buffer (Invitrogen Corporation), and 18 µl of NuPAGE4 Sample Reducing Agent 10X (Invitrogen Corporation) containing 0.5 M dithiothreitol, in separate tubes at each of the following analytical time points (+/- 10 seconds): 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes. Inactivated samples were heated for 10 minutes at 70°C. Time zero for each SIF reaction mixture was prepared by first inactivating a sample of SIF, heating at 70°C for approximately 10 minutes, and then adding 6 µl of the test protein. All inactivated samples were kept on ice prior to electrophoresis.

To prepare control digestion samples at approximately 60 minutes and/or time zero, SIF alone (no test protein) inactivated at time zero and approximately 60 minutes, and intestinal control solution containing GM-HRA protein (no pancreatin) inactivated at approximately 60 minutes. The BSA and β-Lactoglobulin controls were inactivated at time points 0, 1 (+/- 10 seconds), and approximately 60 minutes. A 120 µl aliquot of the control solutions were inactivated by mixing with 46 µl of NuPAGE4 LDS 4X Sample Buffer (Invitrogen Corporation), and 18 µl of NuPAGE Sample Reducing Agent 10X containing 0.5 M dithiothreitol. Inactivated samples were heated for 10 minutes at 70 °C and then kept on ice prior to electrophoresis.

Control digestion samples included in the SIF assay are provided in Table H.2.

Table H.2. Control Samples for Simulated Intestinal Fluid (SIF) Digestibility Analysis

Protein	Digestion Time Determined by SDS-PAGE (minutes)	Digestion Time Determined by Western Blot (minutes)	Approximate Molecular Weight (kDa)
GM-HRA	<0.5	<1	~65
BSA (negative control)	>60	NA[1]	~66
β-Lactoglobulin (positive control)	<60	NA	~18

[1] Control samples did not go through the western blotting procedures.

SDS-PAGE Analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis was conducted on two independent SIF digestion reactions, the first on May 4, 2006 and the second on June 6, 2006. SDS-PAGE was performed using NuPAGE5 Novex 4-12% Bis-Tris Gels. Samples were loaded at a volume of 20 µl/well and SeeBlue5 Pre-stained Standard was loaded at 13 µl/well. Electrophoresis was conducted using a Novex XCell II5 Mini-Cell electrophoresis unit with NuPAGE5 MES SDS running buffer at a constant 200 volts for approximately 30 minutes. Once complete, gels were removed from the gel cassette and washed three times for at least five minutes each with deionized water prior to staining in RAPIDstain for at least 1 hour. The gel was destained with deionized water and a digital image was generated by scanning with a Hewlett-Packard Scanjet 6200C scanner and capturing the images as a JPEG file using Hewlett-Packard PrecisionScan Pro software.

Western Blot Analysis

Western blot analysis of the SIF digestion reaction samples generated on June 6, 2006 was conducted in order to clearly differentiate between the GM-HRA protein and the pancreatin protein bands. SDS-PAGE of the GM-HRA SIF time course samples from June 6, 2006 was conducted using duplicate gels. One gel was stained in RAPIDstain7 as previously described and the other was used for western blot preparation. The proteins in the unstained gel were electrophoretically transferred to a polyvinylidene fluoride (PVDF) membrane using a Novex XCell II5 Mini-Cell blot module with NuPAGE5 transfer buffer (Invitrogen Corporation) run at a constant 30 volts for approximately 1 hour. The membrane containing transferred proteins (blot) was blocked with a 5% low fat milk in phosphate buffered saline solution containing 0.05% Tween -20 (PBST), followed by an approximately 60-minute incubation in a 1:10,000 dilution of primary rabbit anti-GM-HRA polyclonal antibody (R8001) in 1 % low fat milk/PBST solution. The blot was washed with PBST at least three times for approximately ten minutes each and incubated for approximately 45 minutes with a 1:10,000 dilution of the secondary donkey anti-rabbit IgG antibody conjugated to horseradish peroxidase. Unbound secondary antibody-HRP conjugate was removed with at least three washes of approximately ten minutes with PBST. The antibody binding to GM-HRA protein bands was detected using an ECL chemiluminescent detection kit according to the manufacturer's instructions, followed by exposure to film. The film was developed in a M35A X-Omat developer. The film was then evaluated for apparent disappearance of the GM-HRA protein.

APPENDIX I. METHODS FOR TRAIT EXPRESSION ANALYSES

Field Trial Experimental Design

A multi-site field trial was conducted during the 2022 growing season at six sites in commercial soybean-growing regions of the United States (one site in Iowa, Illinois, Indiana, Nebraska, and Pennsylvania) and Canada (one site in Ontario). A randomized complete block design with four blocks was utilized at each site.

Sample Collection

The following samples were collected: leaf (V5, R1, and R3 growth stages), flowers (R1-R2 growth stage), root (R3 growth stage), forage (R3 growth stage), and seed (R8 growth stage). Growth stage descriptions are provided in Table I.1. One sample per plot was collected for each tissue at the applicable growth stages. All samples were collected from impartially selected, healthy, representative plants to minimize potential bias.

Table I.1. Soybean Growth Stage Descriptions

Growth Stage	Description
VE	*Emergence* - Occurs when plants first emerge from the soil.
VC	*Cotyledon* - Occurs when the unifoliate leaves have unfolded.
V1	*First Node* - Occurs when the leaflets on the second leaf node have unrolled.
V2	*Second Node* - Occurs when the leaflets on the third leaf node have unrolled.
V3	*Third Node* - Occurs when the leaflets on the fourth leaf node have unrolled.
V4	*Fourth Node* - Occurs when the leaflets on the fifth leaf node have unrolled.
V5	*Fifth Node* - Occurs when the leaflets on the sixth leaf node have unrolled.
R1	*Beginning Bloom* - Occurs when plants have one open flower at any node on the main stem.
R2	*Full Bloom* - Occurs when plants have an open flower at one of the two uppermost nodes.
R3	*Beginning Pod* - Occurs when a pod is 5 millimeters long at one of the four uppermost nodes.
R4	*Full Pod* - Occurs when a pod is 2 centimeters long at one of the four uppermost nodes.
R5	*Beginning Seed* - Occurs when a pod at one of the four uppermost nodes has seed 3 millimeters long.
R6	*Full Seed* - Occurs when a pod is filled to capacity with green seed at one of the four uppermost nodes.
R7	*Beginning Maturity* - Occurs when one pod on the main stem has reached mature pod color.
R8	*Full Maturity* - Occurs when 95% of the pods have reached their mature pod color.

Note: Growth stages (Pedersen, 2004).

Samples were collected as follows:

- Each V5 growth stage leaf sample was obtained by pruning the youngest, trifoliate leaf with fully developed leaflets collected from each of six plants. Each R1 growth stage leaf sample was obtained by pruning three of the youngest trifoliate leaves with fully developed leaflets collected from each of two plants. Each R3 growth stage leaf sample was obtained by pruning the youngest trifoliate leaves with fully developed leaflets from one individual plant to fill the sample container. Each growth stage leaf tissue sample was collected into a pre-labeled vial.

- Each flower sample was obtained by collecting 60-70 freshly opened flowers from a minimum of 10 individual plants. Each flower sample was collected into a pre-labeled vial.
- Each root sample was obtained by removing the plant from the ground after cutting a circle in the soil 9-12 in. (23-30 cm) in diameter and depth. The roots were thoroughly cleaned with water and removed from the plant. Root tissue was cut into sections of approximately 0.5 in. (1.25 cm) or less in length and collected into a pre-labeled vial.
- Each forage sample was obtained by cutting the plant approximately 2-4 in. (5-10 cm) above the soil surface line. The plant tissue (aerial portion, without roots) was placed into a pre-labeled, plastic-lined, cloth bag.
- Each seed sample was obtained by threshing all seeds from an individual plant. A sub-sample of 20 seeds was collected into a pre-labeled vial.

<u>Sample Processing, Shipping, and Storage</u>

Each sample was placed on dry ice within 10 minutes of collection in the field and transferred to frozen storage (≤ -10 °C) until shipment. Expressed trait protein samples were then shipped frozen to Pioneer Hi-Bred International, Inc. for processing. Upon arrival, samples were stored frozen (-20 °C freezer unit). All forage samples were coarsely homogenized prior to lyophilization. All samples were lyophilized under vacuum until dry. Following lyophilization, all samples were finely homogenized and stored frozen until analysis.

<u>Protein Concentration Determination</u>

The concentrations of Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA proteins were determined using quantitative ELISA methods that have been internally validated to demonstrate method suitability.

Processed tissue sub-samples were weighed at the following target weights: 5 mg for flowers, 10 mg for leaf and seed, and 20 mg for root and forage. Sub-samples were stored in a -20 ºC freezer unit until analysis. Each sample analyzed for Cry1B.34.1, Cry1B.61.1, and IPD083Cb was extracted with 0.60 ml of chilled phosphate-buffered saline containing 0.05% polysorbate 20 (PBST). Each leaf, root, forage, and seed sample analyzed for GM-HRA was extracted with 0.60 ml of chilled buffer which was comprised of PBST with 0.2% CHAPS. Flower samples analyzed for GM-HRA were extracted with 0.60 ml of chilled buffer comprised of PBST with 0.2% CHAPS and 5% StabilZyme Select. Extracted samples were centrifuged, and then supernatants were removed and prepared for analysis.

ELISA Methods

ELISA methods were performed as follows:

- <u>Cry1B.34.1 Protein ELISA Method:</u> Prior to analysis, samples were diluted as applicable in PBST. Standards (typically analyzed in triplicate wells) and diluted samples (typically analyzed in duplicate wells) were incubated in a plate pre-coated with monoclonal antibody

5B7.D9.H2.G4 (Pioneer). Following incubation, unbound substances were washed from the plate. A different monoclonal antibody 10C2.D4.H10 (Pioneer) conjugated to the enzyme horseradish peroxidase (HRP) was added to the plate and incubated. Unbound substances were washed from the plate. Detection of the bound Cry1B.34.1 antibody complex was accomplished by the addition of substrate, which generated a colored product in the presence of HRP. The reaction was stopped with an acid solution and the optical density (OD) of each well was determined using a plate reader.

- Cry1B.61.1 Protein ELISA Method: Prior to analysis, samples were diluted as applicable in PBST. Standards (typically analyzed in triplicate wells) and diluted samples (typically analyzed in duplicate wells) were incubated in a plate pre-coated with monoclonal antibody 14B5.D7.D6 (Pioneer). Following incubation, unbound substances were washed from the plate. A different monoclonal antibody 21C2.E8.H2.D10 (Pioneer) conjugated to the enzyme HRP was added to the plate and incubated. Unbound substances were washed from the plate. Detection of the bound Cry1B.61.1-antibody complex was accomplished by the addition of substrate, which generated a colored product in the presence of HRP. The reaction was stopped with an acid solution and the OD of each well was determined using a plate reader.
- IPD083Cb Protein ELISA Method: Prior to analysis, samples were diluted as applicable in PBST. Standards (typically analyzed in triplicate wells) and diluted samples (typically analyzed in duplicate wells) were incubated in a plate pre-coated with monoclonal antibody 19F4.G1.F6 (Pioneer). Following incubation, unbound substances were washed from the plate. A different monoclonal antibody 5E6.F5.D5 (Pioneer) conjugated to the enzyme HRP was added to the plate and incubated. Unbound substances were washed from the plate. Detection of the bound IPD083Cb antibody complex was accomplished by the addition of substrate, which generated a colored product in the presence of HRP. The reaction was stopped with an acid solution and the OD of each well was determined using a plate reader.
- GM-HRA Protein ELISA Method: Prior to analysis, samples were diluted as applicable in PBST containing 0.2% CHAPS and 5% StabilZyme Select. Standards (typically analyzed in triplicate wells) and diluted samples (typically analyzed in duplicate wells) were incubated in a plate precoated with polyclonal antibody R9222 (Pioneer). Following incubation, unbound substances were washed from the plate and the bound GM-HRA protein was incubated with monoclonal antibody 14G5.H4 (Pioneer) conjugated to the enzyme HRP. Unbound substances were washed from the plate. Detection of the bound GM-HRA-antibody complex was accomplished by the addition of substrate, which generated a colored product in the presence of HRP. The reaction was stopped with an acid solution and the OD of each well was determined using a plate reader.

Calculations for Determining Cry1B.34.1, Cry1B.61.1, IPD083Cb, and GM-HRA Protein Concentrations

SoftMax Pro GxP Version 7.0.3 (Molecular Devices) microplate data software was used to perform the calculations required to convert the OD values obtained for each set of sample wells to a protein concentration value.

A standard curve was included on each ELISA plate. The equation for the standard curve was derived by the software, which used a quadratic fit to relate the OD values obtained for each set of standard wells to the respective standard concentration (ng/ml).

The sample concentration values were adjusted for a dilution factor expressed as 1:N by multiplying the interpolated concentration by N.

Adjusted Concentration = Interpolated Sample Concentration x Dilution Factor

Adjusted sample concentration values obtained from SoftMax Pro GxP software were converted from ng/ml to ng/mg sample weight as follows:

$$\text{Sample Concentration (ng protein/mg sample weight)} = \text{Sample Concentration (ng/ml)} \times \frac{\text{Extraction Buffer Volume (ml)}}{\text{Sample Target Weight (mg)}}$$

The reportable assay lower limit of quantification (LLOQ) in ng/ml was calculated as follows:

Reportable Assay LLOQ (ng/ml) = (lowest standard concentration - 10%) x minimum dilution

The LLOQ, in ng/mg sample weight, was calculated as follows:

$$\text{LLOQ} = \text{Reportable Assay LLOQ (ng/ml)} \times \frac{\text{Extraction Buffer Volume (ml)}}{\text{Sample Target Weight (mg)}}$$

Trait Confirmation

To confirm sample identity, event-specific polymerase chain reaction (PCR) analyses were performed for samples with unexpected ELISA results. If a given test sample was confirmed as not containing the event of interest, the protein results were excluded from reporting.

Statistical Analysis

Statistical analysis of the protein concentration results consisted of the calculations of means, ranges, and standard deviations. Individual sample results below the LLOQ were assigned a value equal to the LLOQ for calculation purposes.

APPENDIX J. METHODS FOR NUTRIENT COMPOSITION ANALYSIS

Field Trial Experimental Design

A multi-site field trial was conducted during the 2022 growing season at eight sites in commercial soybean-growing regions of United States (one site in Illinois, Indiana, Missouri, Nebraska, and Pennsylvania, and two sites in Iowa) and Canada (one site in Ontario). Each block included COR23134 soybean, control soybean, and four of the following reference soybean lines: 92M35, 92B63, 92M72, BK291, P29T50, BK310, BK317, BK331N, P33T60, BK340, 93Y41, P34A50, P35A41, BK360, BK361, 93M62, BK370, and 93B82. A randomized complete block design with four blocks was utilized at each site. A quizalofop and fomesafen herbicide treatment was applied to all soybean plots (control, reference, and COR23134) soybean.

Sample Collection

One forage sample (R3 growth stage) and one seed sample (R8 growth stage) were collected from each plot. Each forage sample (containing at least 600 g) was obtained by cutting the aerial portion of the plants approximately 1 in. (2.5 cm) above the soil surface. The plants were chopped into sections of 6 in. (15 cm) or less in length and collected in a pre-labeled, plastic-lined, cloth bag. Each seed sample (containing at least 400 g) was obtained. The seed was collected into a large pre labeled, plastic, resealable bag and then placed into a pre-labeled, plastic lined, cloth bag.

All samples were collected from impartially selected, healthy, representative plants. Reference soybean and control soybean samples were collected prior to the collection of COR23134 soybean samples to minimize the potential for contamination. Each sample was uniquely labeled with a sample identification number and barcode for sample tracking, and is traceable by site, entry, block, tissue, and growth stage. Samples were placed in chilled storage (e.g., coolers with wet ice, artificial ice, or dry ice), transferred to a freezer (≤ -10 °C), or placed on dry ice within 3 hours of collection. Samples from each site were shipped frozen to [redacted] for nutrient composition analyses.

Nutrient Composition Analyses

The forage and seed samples were analyzed at [redacted]. Experimental bias was controlled through the use of the same sample identification numbers assigned to the originally collected samples, through the use of pre-set data acceptability criteria, sample randomization prior to homogenization, and through the arrangement of samples for analyses without consideration of sample identity. The following nutrient composition analytes were determined:

- *Forage proximate and fiber composition:* moisture, crude protein, crude fat, crude fiber, acid detergent fiber (ADF), neutral detergent fiber (NDF), ash, and carbohydrates
- *Seed proximate and fiber composition*: moisture, crude protein, crude fat, crude fiber, acid detergent fiber (ADF), neutral detergent fiber (NDF), ash, and carbohydrates
- *Seed fatty acid composition*: lauric acid (C12:0), myristic acid (C14:0), pentadecanoic acid (C15:0), pentadecanoic acid (C15:1), palmitic acid (C16:0), palmitoleic acid (C16:1), heptadecanoic acid (C17:0), heptadecenoic acid (C17:1), heptadecadienoic acid (C17:2),

stearic acid (C18:0), oleic acid (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), nonadecanoic acid (C19:0), Isomer 1 of nonadecanoic acid (C19:1,1), Isomer 2 of nonadecanoic acid (C19:1,2), arachidic acid (C20:0), eicosenoic acid (C20:1), eicosadienoic acid (C20:2), heneicosanoic acid (C21:0), behenic acid (C22:0), tricosanoic acid (C23:0) and lignoceric acid (C24:0)

- *Seed amino acid composition*: alanine, arginine, aspartic acid, cystine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine
- *Seed mineral composition*: calcium, copper, iron, magnesium, manganese, phosphorus, potassium, sodium, and zinc
- *Seed vitamin composition*: vitamin B1 (thiamine), vitamin B2 (riboflavin), vitamin B3 (niacin), vitamin B5 (pantothenic acid), vitamin B6 (pyridoxine), vitamin B9 (folic acid), vitamin K_1, α-tocopherol
- *Seed isoflavone composition*: total daidzein equivalent, total genistein equivalent, and total glycitein equivalent
- *Seed anti-nutrient composition:* raffinose, stachyose, lectins, phytic acid and trypsin inhibitor

Nutrient composition analytical methods and procedures are summarized in Table J.1.

Table J.1. Methods for Compositional Analysis of COR23134 Soybean

Nutritional Analyte	Method
Moisture Forage and Seed	The analytical procedure for moisture determination was based on a method published by the Association of Official Analytical Chemists (AOAC) International. Samples were analyzed to determine the percentage of moisture by gravimetric measurement of weight loss after drying in a forced air oven (forage) and vacuum oven (seed). The moisture content and dry matter of the soybean forage and seed was determined.
Ash Forage and Seed	The analytical procedure for ash determination in soybean forage and seed was based on a method published by the AOAC International. Samples were analyzed to determine the percentage of ash by gravimetric measurement of the weight loss after ignition in a muffle furnace.
Crude Protein Forage and Seed	The analytical procedure for crude protein determination in soybean forage and seed utilized an automated Kjeldahl technique based on a method provided by the manufacturer of the titrator unit (Foss-Tecator) and the AOAC International. Ground samples were digested in the presence of a catalyst. The digestate was then distilled and titrated with a Foss-Tecator Kjeltec Analyzer Unit.
Crude Fat Forage and Seed	The analytical procedure for crude fat determination in soybean forage and seed was based on methods provided by the American Oil Chemists' Society (AOCS) and the manufacturer of the hydrolysis and extraction apparatus (Ankom Technology). Forage samples were hydrolyzed with 3N hydrochloric acid at 90 ºC for 80 minutes. The hydrolysates were extracted with a petroleum ether/ethyl ether/ethyl alcohol solution at 90 ºC for 60 minutes. Seed samples were extracted with petroleum ether using the ANKOM ^{XT15} extraction system at 90 ºC for 70 minutes. After extraction, the samples were oven dried, and the crude fat content was determined gravimetrically.
Crude Fiber Forage and Seed	The analytical procedure for crude fiber determination in soybean forage and seed was based on methods provided by the manufacturer of the extraction apparatus (Ankom Technology), the AOAC International, and the AOCS. Samples were analyzed to determine the percentage of crude fiber by digestion and solubilization of other materials present. After rinsing with reverse osmosis (RO) water, the remaining residue was dried and weighed to determine the crude fiber content.

Table J.1. Methods for Compositional Analysis of COR23134 Soybean (continued)

Nutritional Analyte	Method
Neutral Detergent Fiber Forage and Seed	The analytical procedure for neutral detergent fiber (NDF) determination in soybean forage and seed was based on a method provided by the manufacturer of the extraction apparatus (Ankom Technology), the AOAC International, and the *Journal of AOAC International*. Samples were analyzed to determine the percentage of NDF by digesting with a neutral detergent solution, sodium sulfite and alpha amylase. After rinsing with RO water, the remaining residue was dried and weighed to determine the NDF content.
Acid Detergent Fiber Forage and Seed	The analytical procedure for acid detergent fiber (ADF) determination in soybean forage and seed was based on a method provided by the manufacturer of the extraction apparatus (Ankom Technology) and the AOAC International. Samples were analyzed to determine the percentage of ADF by digesting with an acid detergent solution. After rinsing with RO water, the remaining residue was dried and weighed to determine the ADF content.
Carbohydrates Forage and Seed	The carbohydrate content in soybean forage and seed on a dry weight basis was calculated using a formula obtained from the United States Department of Agriculture "*Energy Value of Foods*," in which the percent dry weight of crude protein, crude fat, and ash was subtracted from 100%.
Tryptophan Seed	The analytical procedure for tryptophan determination in soybean seed was based on an established lithium hydroxide hydrolysis procedure using reverse phase ultra-performance liquid chromatography (UPLC) with ultraviolet (UV) detection published by the *Journal of Micronutrient Analysis*.
Cystine and Methionine Seed	The analytical procedure for cystine and methionine determination in soybean seed was based on methods obtained from the Waters Corporation, the AOAC International, and the *Journal of Chromatography A*. The procedure converts cystine to cysteic acid and methionine to methionine sulfone, to the 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate derivatives which were then analyzed by reverse phase UPLC with UV detection.
Additional Amino Acids Seed	In addition to tryptophan, cystine, and methionine, 15 other amino acids were also determined. The analytical procedure for analysis of these amino acids in soybean seed was based on methods obtained from the Waters Corporation and the *Journal of Chromatography A*. The procedure converts the free acids, after acid hydrolysis, to the 6-aminoquinoyl-N-hydroxysuccinimidyl carbamate derivatives, which were analyzed by reverse phase UPLC with UV detection.
Fatty Acids Seed	The analytical procedure for determination of fatty acids in soybean seed was based on methods published by the AOAC International and the AOCS. The procedure converts the free acids, after microwave assisted ether extraction and base hydrolysis, to the fatty acid methyl ester (FAME) derivatives, which were analyzed by gas chromatography with flame ionization detection (GC/FID). Results are reported as percent total fatty acids but presented in the raw data as percent fresh weight.

Table J.1. Methods for Compositional Analysis of COR23134 Soybean (continued)

Nutritional Analyte	Method
Thiamine (Vitamin B_1) Seed	The analytical procedure for the determination of thiamine in soybean seed was based on a method published by the American Association of Cereal Chemists (AACC). The samples were extracted with 10% acetic acid/4.3% trichloroacetic acid solution. A 100-fold dilution was performed then the samples were analyzed by reverse phase high performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS).
Riboflavin (Vitamin B2) Seed	The analytical procedure for determination of riboflavin in soybean seed was based on a method validated by r-biopharm. Riboflavin was extracted from the matrix using pH adjusted samples and heating in the autoclave. After heating, filtering, and diluting, the extract was assayed microbiologically with the organism *Lactobacillus rhamnosus* using the VitaFast® Riboflavin assay kit from r-biopharm. The turbidity (optical density) produced by the organism is measured using a spectrophotometer set to measure absorbance at 630 nm. The riboflavin concentration is then calculated within the r-biopharm software (RIDASOFT) using a 4-parameter logistic curve
Tocopherols Seed	The analytical procedure for the determination of tocopherols in soybean seed was based on methods from the *Journal of the American Oil Chemists' Society and Analytical Sciences*. Alpha tocopherol was extracted with hot hexane and the extracts were analyzed by normal phase UPLC with fluorescence detection.
Trypsin Inhibitor Seed	The analytical procedure for the determination of trypsin inhibitor in soybean seed was based on a method published by the AOCS. Trypsin inhibitor was extracted with sodium hydroxide. Benzoyl-DL-arginine-p-nitroanilide hydrochloride (BAPNA) was added to the sample dilutions to act as a chromogenic substrate for trypsin. Trypsin was added to the sample dilutions and reacted with trypsin inhibitor in the extracts. The amount of trypsin activity present in the reaction was measured using a spectrophotometer, and the amount of trypsin inhibitor was calculated based on the inhibition of trypsin activity from the sample extracts.
Phytic Acid Seed	The analytical procedure for the determination of phytic acid in soybean seed was based on a method published by the AOAC International. The samples were analyzed to determine the amount of phytic acid by extracting the phytic acid with dilute hydrochloric acid (HCl) and isolating it using an aminopropyl silica solid phase extraction column. Once isolated and eluted, the phytic acid was analyzed for elemental phosphorus by inductively coupled plasma optical emission spectroscopy (ICP-OES). The phytic acid content was then calculated from the phosphorus concentration.
Minerals Seed	The analytical procedure for the determination of minerals in soybean seed was based on methods published by the AOAC International and CEM Corporation. The soybean seed minerals determined were calcium, copper, iron, magnesium, manganese, phosphorus, potassium, sodium, and zinc. The samples were digested in a microwave-based digestion system and the digestate was diluted using DI water. Both the diluted and undiluted portions were analyzed by ICP-OES.

Table J.1. Methods for Compositional Analysis of COR23134 Soybean (continued)

Nutritional Analyte	Method
Oligosaccharide Seed	The analytical procedure for the determination of raffinose and stachyose contents in soybean seed were determined based on a method published by the AACC. Samples were first extracted with methanol/DI water followed by an extraction with chloroform. After removal of the chloroform, an aliquot of the extract was evaporated to dryness. The sample residues were dissolved in 50:50 acetonitrile (ACN): DI water and assayed for raffinose and stachyose by reverse phase HPLC with refractive index (RI) detection.
Vitamin B9 (Folic Acid) Seed	The analytical procedure for the determination of folic acid in soybean seed was based on a method published by AACC. Samples were hydrolyzed and digested by protease and amylase enzymes to release the folates from the soybean seed. A conjugase enzyme was used to convert the naturally occurring folylpolyglutamates to folyldiglutamates. An aliquot of the extracted folates was mixed with a folate and folic acid free microbiological growth medium. The mixture was inoculated with *Lactobacillus casei*. The total folate content was determined by measuring the turbidity of the *Lactobacillus casei* growth response in the sample and comparing it to the turbidity of the growth response in folic acid standards using a spectrophotometer set to measure absorbance at 600 nm.
Vitamin B_3 (Niacin) Seed	The analytical procedure for the determination of niacin in soybean seed was based on a method published by the AACC. Niacin was extracted from the sample by adding DI water and autoclaving. A tube array was prepared using three different dilutions of the samples. This tube array was inoculated with *Lactobacillus plantarum* and allowed to incubate for 18-22 hours. After incubation, the bacterial growth was determined using a spectrophotometer set to measure absorbance at 660 nm. The absorbance readings were compared to a standard curve generated using known concentrations of nicotinic acid.
Vitamin B_6 (Pyridoxine) Seed	The analytical procedure for the determination of pyridoxine in soybean seed was based on a method validated from r-biopahrm. Pyridoxine content was determined using a microbiological assay. After heating, filtering, and diluting the extract, vitamin B_6 is assayed microbiologically using the organism *Saccharomyces cerevisiae* using the VitaFast® Vitamin B_6 assay kit from r-biopharm. The growth of *Saccharomyces cerevisiae* is proportional to the amount of vitamin B_6 in the extract. The turbidity (optical density) produced by the organism is measured using a spectrophotometer set to measure absorbance at 630 nm. The Peridoxin Hydrochloride concentration is then calculated within r-biopharm software (RIDASOFT) using a 4-parameter logistic curve).

Table J.1. Methods for Compositional Analysis of COR23134 Soybean (continued)

Nutritional Analyte	Method
Vitamin B_5 (Pantothenic Acid) Seed	The analytical procedure for the determination of pantothenic acid in soybean seed was based on a method from the AOAC International. Pantothenic acid content was determined using a microbiological assay. Pantothenic acid was extracted from the sample by an acetic acid buffer solution, consisting of acetic acid adjusted to a pH of 5.65 with sodium hydroxide, and autoclaving the samples. A tube array was prepared using three different dilutions of the samples. This tube array was inoculated with *Lactobacillus plantarum* and allowed to incubate for 18-22 hours. After incubation, the bacterial growth was determined using a spectrophotometer set to measure absorbance at 660 nm. The absorbance readings were compared to a standard curve generated using known concentrations of D-pantothenic acid hemicalcium salt.
Lectins Seed	The analytical procedure for the determination of lectin in soybean seed was based on a method from the *Journal of American Oil Chemists*. Lectins were extracted using a phosphate buffered saline solution with Tween and shaking in a Geno/Grinder. After centrifuging and diluting, samples were loaded onto a 96-well microtiter plate coated with a-GaINAc-PAA. The plates were further processed with a-GaINAc-PAA-biotin and Neutravidin-HRP, and then developed with TMB substrate. Absorbance of microtiter plates is measured at 450 nm using a microtiter plate reader. The lectin content was then calculated from the absorbance value.
Isoflavones Seed	The analytical procedure for the determination of isoflavones in soybean seed was based on methods from AACC and AOAC International. Isoflavones were extracted from the samples using methanol: DI water, hydrolyzed with sodium hydroxide, and neutralized with acetic acid. Samples were then assayed by reverse phase UPLC with UV detection.
Vitamin K_1 Seed	The analytical procedure for the determination of vitamin K_1 in soybean seed was based on methods published in the *Journal of AOAC International, European Food Research and Technology*, the AOAC, the United States Pharmacopeia (USP), and the *Journal of Agricultural and Food Chemistry*. Samples were extracted with a mixture of dimethyl sulfoxide and hexane and cleaned up on a silica solid-phase extraction cartridge. Following elution from the SPE cartridge, the sample solution was evaporated to dryness under a stream of nitrogen and the residue was re-dissolved in isopropyl alcohol. The alcohol solutions were analyzed by HPLC-MS/MS.

Statistical Analysis of Nutrient Composition Data

Prior to statistical analysis, the data were processed as follows:

- *LLOQ Sample Values*: For statistical analysis, nutrient composition values reported as below the assay lower limit of quantification (LLOQ) were each assigned a value equal to half the LLOQ

  Conversion of fatty acid assay values: The raw data for all fatty acid analytes were provided by [redacted] in units of percent fresh weight (%FW). Any fatty acid values below the %FW LLOQ were set to half the LLOQ value, and then all assay values were converted to units of % total fatty acids for statistical analyses.

  For a given sample, the conversion to units of % total fatty acids was performed by dividing each fatty acid analyte value (%FW) by the total fresh weight of all fatty acids for that sample; for analyte values below the LLOQ, the half LLOQ value was used as the analyte value. Half LLOQ values were also included in the total fresh weight summations. After the conversion, a fixed LLOQ value was not available for a given individual fatty acid analyte on the % total fatty acids basis. If the assay value of an individual analyte was below the LLOQ for a given sample, half of the LLOQ value was used in computing the total. The total was considered below the LLOQ only when all the individual analytes contributing to its calculation were below the LLOQ.

Statistical analyses were conducted using SAS software, Version 9.4 (SAS Institute, Inc.). The following rules were implemented for each analyte:

- If both COR23134 soybean and the control soybean had < 50% of samples below the LLOQ, then an across-site mixed model analysis was conducted. In addition, if both soybean lines had at least two samples at a given site above the LLOQ, then an individual-site mixed model analysis was conducted.
- If, either COR23134 soybean or the control soybean had ≥ 50% samples below the LLOQ, but not both entries had 100% of samples below the LLOQ across sites, then Fisher's exact test was conducted. The Fisher's exact test assessed whether there was a significant difference (P value < 0.05) in the proportion of samples below the LLOQ between these two soybean lines across sites. Individual-site analyses were not performed.
- If both COR23134 soybean and the control soybean had 100% of samples below the LLOQ, then statistical analyses was not performed (Table J.2).

Statistical Model for Across-Site Analysis

For a given analyte, data were analyzed using the following linear mixed model:

$$y_{ijk} = \mu_i + \ell_j + r_{k(j)} + (\mu\ell)_{ij} + \varepsilon_{ijk} \qquad \text{Model 1}$$

$\ell_j \sim$ iid $N(0, \sigma^2_{Site})$, $r_{k(j)} \sim$ iid $N(0, \sigma^2_{Rep})$, $(\mu\ell)_{ij} \sim$ iid $N(0, \sigma^2_{Ent\times Site})$, and $\varepsilon_{ijk} \sim$ iid $N(0, \sigma^2_{Error})$

where μ_i denotes the mean of the i^{th} entry (fixed effect), ℓ_j denotes the effect of the j^{th} site (random effect), $r_{k(j)}$ denotes the effect of the k^{th} block within the j^{th} site (random effect), $(\mu\ell)_{ij}$ denotes the interaction between the entries and sites (random effect), and ε_{ijk} denotes the effect of the plot assigned the i^{th} entry in the k^{th} block of the j^{th} site (random effect or residual). Notation ~ iid $N(0, \sigma^2{}_a)$ indicates random variables that are identically independently distributed (iid) as normal with zero mean and variance $\sigma^2{}_a$. Subscript a represents the corresponding source of variation.

The residual maximum likelihood estimation procedure was utilized to generate estimates of variance components and entry means across sites. The estimated means are known as empirical best linear unbiased estimators (hereafter referred to as LS-Means). The statistical comparison was conducted by testing for a difference in LS-Means between COR23134 soybean and the control soybean. The approximated degrees of freedom for the statistical test were derived using the Kenward-Roger method (Kenward and Roger, 2009). A significant difference was identified if a P-value was < 0.05.

For each analyte, goodness-of-fit of the model was assessed in terms of meeting distributional assumptions of normally, independently distributed errors with homogeneous variance. Deviations from assumptions were addressed using an appropriate transformation or allowing for heterogeneous error variance among sites.

Statistical Model for Individual-Site Analysis

For a given analyte, individual sites were analyzed separately using the following linear mixed model:

$$y_{ik} = \mu_i + r_k + \varepsilon_{ik} \qquad \text{Model 2}$$

$r_k \sim$ iid $N(0, \sigma^2{}_{Rep})$ and $\varepsilon_{ik} \sim$ iid $N(0, \sigma^2{}_{Error})$

where μ_i denotes the mean of the i^{th} entry (fixed effect), r_k denotes the effect of the k^{th} block (random effect), and ε_{ik} denotes the residual for the observation obtained from the plot assigned to the i^{th} entry in the k^{th} block (random effect or residual).

The residual maximum likelihood estimation procedure was used to generate estimates of variance components and entry means (LS-Means). The statistical comparison was conducted by testing for a difference in LS-Means between COR23134 soybean and the control soybean. The approximated degrees of freedom for the statistical test were derived using the Kenward-Roger method. The same transformations applied during across-site analysis were also utilized for individual-site analyses.

False Discovery Rate Adjustment

The false discovery rate (FDR) method (Benjamini and Hochberg, 1995; Westfall et al., 1999) was used to control for false positive outcomes across all analytes analyzed using linear mixed models. A false positive outcome occurs if the difference in means between two entries is declared significant, when in fact the two means are not different. Since its introduction in the mid-1990s, the FDR approach has been widely employed across a number of scientific disciplines, including genomics, ecology, medicine, plant breeding, epidemiology, dairy science, and signal/image processing (e.g., Pawitan et al., 2005; Spelman and Bovenhuis, 1998). In the FDR method, the false discovery rate is held at 5% across comparisons of multiple analytes via an adjustment to the P-value and is not inflated by the number of analytes in the comparison. The FDR adjustment of raw P-values was conducted separately for the across-site analysis and each of the individual-site analyses.

Interpretation of Statistical Results

For a given analyte, when a statistically significant difference (P-value from mixed model analysis < 0.05, or Fisher's exact test P-value < 0.05) was identified in the across-site analysis, the respective range of individual values from COR23134 soybean was compared to a tolerance interval. Tolerance intervals are expected to contain at least 99% of the values for corresponding analytes of the conventional soybean population with a 95% confidence level (Hong *et al.*, 2014). The tolerance intervals were derived from Pioneer and Dow Agroscience proprietary accumulated data from non-GM soybean lines, which were grown in commercial soybean-growing region in the United States, Canada, Chile, Brazil, and Argentina in growing seasons ranging from 2005 to 2016. The combined data represent 81 commercial soybean lines and 175 unique environments. The selected commercial soybean lines represent the non-GM soybean population with a history of safe use, and the selected environments (site and year combinations) represent soybean growth under a wide range of environmental conditions (i.e., soil type, temperature, precipitation, and irrigation) and soybean maturity group zones.

If the range of COR23134 soybean contained individual values outside the tolerance interval, it was then compared to the respective literature range obtained from published literature: AFSI (2022); Kim *et al.* (2005); Lee *et al.* (2003); Morse (1950); OECD (2012); Seguin *et al.* (2004); Taylor *et al.* (1999). Literature ranges complement tolerance intervals in that they are composed of non-proprietary data from additional non-GM commercial soybean lines and growing environments, which are not included in Pioneer's proprietary database.

If the range of COR23134 soybean contained individual values outside the literature range, it was then compared to the respective in-study reference range comprised of all individual values across-sites from all non-GM reference soybean lines grown in this study. In-study reference data ranges complement tolerance intervals and literature ranges in that they provide additional context of natural variation specific to the current study.

In cases when a raw P-value indicated a significant difference, but the FDR adjusted P-value was > 0.05, it was concluded that the difference was likely a false positive.

Table J.2. Number of Sample Values Below the Lower Limit of Quantification

Analyte	Number of Samples Below the LLOQ		Fisher's Exact Test P-Value
	Control Soybean (n=32)	COR23134 Soybean (n=32)	
Fatty Acid Composition (% Total Fatty Acids)			
Lauric Acid (C12:0)	32	32	--
Pentadecanoic Acid (C15:0)	32	32	--
Pentadecenoic Acid (C15:1)	32	32	--
Heptadecadienoic Acid (C17:2)[a]	11	7	--
Nonadecanoic Acid (C19:0)	32	32	--
Isomer 1 of Nonadecenoic Acid (C19:1,1)	25	22	0.572
Isomer 2 of Nonadecenoic Acid (C19:1,2)[a]	1	0	--
Eicosadienoic Acid (C20:2)[a]	1	0	--
Mineral Composition (% Dry Weight)			
Sodium[a]	14	12	--
Vitamin Composition (mg/kg Dry Weight)			
Vitamin B1 (Thiamine)	13	20	0.133
Anti-Nutrient Composition (% Dry Weight)			
Raffinose[a]	12	10	--

[a] This analyte had < 50% of sample values below the lower limit of quantification (LLOQ) in each soybean line and was subjected to the mixed model analysis.