Petition for the Determination of Nonregulated Status for Herbicide Tolerant 356043 Soybean

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

Submitting Company:

Pioneer Hi-Bred International, Inc.
7100 NW 62nd Avenue
PO Box 1014
Johnston, IA 50131-1014

Submitted by:

Tracy A. Rood, Registration Manager
Pioneer Hi-Bred International, Inc.
7250 NW 62nd Ave.
PO Box 552
Johnston, IA 50131
Telephone: 515-270-4036
Fax: 515-334-4478

Prepared by:

Tracy A. Rood (lead author), Natalie Weber, Annie Tang Gutsche, Padma Commuri, and Maria Fedorova

OECD Unique Identifier: DP-356Ø43-5

September 27, 2006

Updated November 22, 2006

NO CBI

Release of Information

Pioneer is submitting the information in this assessment for review by the USDA as part of the regulatory process. By submitting this information, Pioneer does not authorize its release to any third party except to the extent it is requested under the Freedom of Information Act (FOIA), 5 U.S.C., Section 552; USDA complies with the provisions of FOIA and USDA's implementation regulations (7 CFR Part 1.4); and this information is responsive to the specific request. Except in accordance with the Freedom of Information Act, Pioneer does not authorize the release, publication or other distribution of this information (including website posting) without Pioneer's prior notice and consent.

Certification

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

Tracy A. Rood
Registration Manager
Pioneer Hi-Bred International, Inc.
7250 NW 62nd Ave.
PO Box 552
Johnston, IA 50131
Telephone: 515-270-4036
Fax: 515-334-4478

Summary

Pioneer Hi-Bred International, Inc. (Pioneer) is submitting a Petition for Determination of Nonregulated Status for herbicide tolerant soybean event DP-356Ø43-5, hereafter referred to as 356043 soybean. Event DP-356Ø43-5 was developed by Pioneer, a DuPont Company. Pioneer requests a determination from USDA - Animal and Plant Health Inspection Service (APHIS) that 356043 soybean and any crosses of this line with nonregulated soybean lines no longer be considered regulated articles under 7 CFR 340.

Pioneer has developed a transgenic soybean product that provides tolerance to two different classes of herbicides: glyphosate and acetolactate synthase (ALS)-inhibiting herbicides. The dual herbicide tolerance of 356043 soybean will enable growers to choose an optimal combination of these herbicides to best manage their individual weed populations. The availability of 356043 soybean will enable growers to proactively manage weed populations while avoiding adverse population shifts of troublesome weeds or the development of resistance. Herbicide tolerant 356043 soybean will be marketed in the U.S. under the brand name Optimum™ GAT™.[1]

The 356043 soybean plants have been genetically modified to express the GAT4601 (glyphosate acetyltransferase) and GM-HRA (modified version of a soybean acetolactate synthase) proteins. The GAT4601 protein, encoded by the *gat4601* gene, confers tolerance to glyphosate-containing herbicides by acetylating glyphosate and thereby rendering it non-phytotoxic. The GM-HRA protein, encoded by the *gm-hra* gene, confers tolerance to the ALS-inhibiting class of herbicides.

The *gat4601* gene is based on the sequences of three *gat* genes from the common soil bacterium *Bacillus licheniformis*. *Bacillus licheniformis* is widespread in the environment; therefore, animals and humans are regularly exposed without adverse consequences to this organism and its components, such as the glyphosate acetyltransferase (GAT) protein. GAT proteins are members of the GCN 5-related family of N-acetyltransferases (also known as the GNAT family). The GNAT superfamily is one of the largest enzyme superfamilies recognized to date with over 10,000 representatives from plants, animals and microbes. The GAT4601 protein is 84% homologous to each of the three native GAT proteins from which it was derived. In 356043 soybean, the expression of the *gat4601* gene is driven by the constitutive synthetic core promoter SCP1.

The GM-HRA protein from soybean is a modified version of the soybean GM-ALS protein. GM-ALS is involved in branched chain amino acid (leucine, isoleucine and valine) biosynthesis in the plastid. The herbicide tolerant *gm-hra* gene was made by isolating the herbicide sensitive soybean *gm-als* gene and introducing two specific amino acid changes, known to confer herbicide tolerance to tobacco ALS. The GM-HRA protein is >99% homologous to the native GM-ALS protein from which it was derived. In 356043 soybean, the expression of the *gm-hra* gene is driven by the soybean constitutive S-adenosyl-L-methionine synthetase (SAMS) promoter.

The allergenic potential of the GAT4601 and GM-HRA proteins was assessed using a step-wise, weight-of-the evidence approach utilizing guidance from the Codex Alimentarius Commission, 2003. Bioinformatic analyses revealed no identities to known

---

[1] Optimum™ and GAT™ are trademarks of Pioneer Hi-Bred International, Inc.

or putative protein allergens or toxins for either the GAT4601 or GM-HRA amino acid sequences. Both the GAT4601 and GM-HRA proteins were non-glycosylated and heat labile. Both proteins hydrolyzed rapidly (within 30 seconds) in simulated gastric fluid. In simulated intestinal fluid the GAT4601 protein hydrolyzed within 2 minutes and the GM-HRA protein hydrolyzed within 30 seconds. There was no evidence of acute toxicity in mice for either GAT4601 or GM-HRA at doses of 1680 or 582 mg protein per kg of body weight, respectively. Based on the GAT4601 and GM-HRA protein levels in 356043 soybean, exposure levels would be well below the tested doses. These data support the food and feed safety of the GAT4601 and GM-HRA proteins. An Early Food Safety Evaluation comprehensive safety data package for the GAT4601 protein was submitted to FDA on June 16, 2006.

Transformation, via microprojectile bombardment, of somatic embryos with a 5361 base pair DNA fragment containing the *gat4601* and *gm-hra* expression cassettes resulted in the generation of 356043 soybean. Molecular characterization of 356043 soybean by Southern blot analysis confirmed that a single, intact fragment was inserted into the soybean genome to produce 356043 soybean. A single copy of each of the genetic elements of the *gat4601* and *gm-hra* expression cassettes was present, and the integrity of the inserted fragment was demonstrated in three different generations, confirming stability during traditional breeding procedures. Southern blot analysis also verified the absence of unwanted DNA such as plasmid backbone sequences in 356043 soybean. Segregation data for five generations confirmed Mendelian inheritance of the *gat4601* and *gm-hra* genes.

356043 soybean has been field tested since 2003 in the major soybean growing regions of the continental United States as well as Hawaii. All field tests have occurred under field permits granted by USDA - APHIS. Comprehensive agronomic performance and ecological observation assessments for 356043 soybean were conducted in replicated, multisite field studies at 16 total North American locations over the 2004 and 2005 growing seasons. Characteristics such as emergence, seedling vigor, plant height, lodging, days to maturity, shattering, seed weight, yield, disease incidence and insect damage were measured. Seed germination and dormancy data were also collected in laboratory experiments. All field trials of 356043 soybean were observed for opportunistic disease or insect biotic stressors and for normal phenotypic characteristics. Analysis of agronomic and ecological data showed no biologically meaningful differences between 356043 soybean and control soybean lines, indicating no plant pest characteristics for 356043 soybean. Likewise, assessment of the ecological data detected no biologically significant differences between 356043 soybean and control soybean lines indicative of a selective advantage that would result in increased weed potential for 356043 soybean. These data support a conclusion of agronomic and phenotypic comparability of 356043 soybean to conventional soybean varieties with respect to the lack of increased weed or plant pest potential.

Extensive nutrient composition analysis of forage and grain was conducted to compare the compositon of 356043 soybean compared to that of a non-transgenic near isoline and four conventional soybean varieties. In total, data from 72 different analytical components (72 in grain and five of those in forage) are presented. Compositional analysis of 356043 soybean was used to evaluate any changes in the levels of key nutrients, isoflavones, or antinutrients. Along with the agronomic data included in this petition, compositional analysis is a general indicator that 356043 soybean will not exhibit unexpected or unintended effects with respect to plant pest risk.

Based on the compositional evaluation, the grain and forage of 356043 soybean were considered to be comparable to conventional soybean. An increase in two minor fatty acids, heptadecanoic acid (C17:0) and heptadecenoic acid (C17:1), was detected, but

this is not unexpected as expression of the GM-HRA protein likely results in a slight shift in availability of the GM-HRA enzyme substrates, pyruvate and 2-ketobutyrate. These two compounds are also substrates for the enzyme complex that initiates oil biosynthesis. Levels of C17:0 and C17:1 fatty acids are very low in 356043 soybean (0.3% and 0.2% of the total fatty acids, respectively). In addition to also being present in conventional soybean varieties, C17:0 is present in vegetable oils, butter and meat. C17:1 is present in beef, cheese and olive oil. Levels of two acetylated amino acids, N-acetylaspartate and N-acetylglutamate, were also elevated. GAT proteins are known to acetylate certain amino acids under specific *in vitro* conditions. Levels of N-acetylaspartate and N-acetylglutamate are low in 356043 soybean (together less than 0.15% of the total amino acids). In addition to also being found in conventional soybean varieties, these two analytes are components of commonly consumed food such as eggs, chicken, turkey, beef, vegetable bouillon, mushrooms, yeast, and soy sauce.

A 42-day broiler study confirmed the nutritional wholesomeness and comparability of poultry diets made from 356043 and control soybean. Based on the food and feed safety assessment of 356043 soybean, we conclude there will be no significant adverse effects to animal or human health. A detailed assessment of the food and feed safety and nutritional value of 356043 soybean will be submitted to FDA.

Information presented herein demonstrates that 356043 soybean exhibits no plant pathogenic properties and is no more likely to become a weed of agriculture or be invasive of natural habitats than conventional soybean varieties. GAT4601 and GM-HRA proteins do not exhibit properties of toxins or allergens and are therefore unlikely to have an impact on threatened or endangered species, beneficial organisms, animals or humans. Agronomic practices for 356043 soybean will be unchanged from those for existing soybean varieties with the exception of additional herbicide options for weed control and integrated weed management.

Therefore, Pioneer requests that APHIS grant the request for a determination of nonregulated status for 356043 soybean and any crosses of this line with nonregulated soybean lines.

Abbreviations, Acronyms and Definitions

~	approximately
ADF	acid detergent fiber
ALS	acetolactate synthase
APHIS	Animal and Plant Health inspection Service
ATCC	American Type Culture Collection
B. licheniformis	*Bacillus licheniformis*
B. subtilis	*Bacillus subtilis*
BAR	phosphinothricin acetyltransferase from *Streptomyces hygroscopicus*
bp	base pair
CaMV	Cauliflower Mosaic Virus
Da	dalton
DIG	digoxygenin
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
E score	expectation score
ELISA	enzyme linked immunosorbent assay
EPA	Environmental Protection Agency
EPSPS	enolpyruvylshikimate-3-phosphate synthase
ESMS	electrospray ionization mass spectroscopy
FDA	Food and Drug Administration
FDR	false discovery rate
GAT	glyphosate N-acetyltransferase
GAT4601	specific GAT protein
gat4601	specific *gat* gene
Gly	glyphosate
gm-als	native soybean acetolactate synthase gene
GM-ALS	native soybean acetolactate synthase protein
gm-hra	modified version of soybean acetolactate synthase gene
GM-HRA	modified version of soybean acetolactate synthase protein
HRP	horseradish peroxidase
IgG	immunoglobulin G
ILSI	International Life Sciences Institute
kb	kilobase pair
kDa	kilodalton
LLOQ	lower limit of quantitation

LOD	limit of detection
MALDI-MS	matrix assisted laser desorption ionization mass spectrometry
NAA	N-acetylaspartate
NAG	N-acetylglutamate
NCBI	National Center for Biotechnology Information
NDF	neutral detergent fiber
OD	optical density
OECD	Organisation for Economic Co-operation and Development
PAT	phosphinothricin acetyltransferase from *Streptomyces viridochomogenes*
PCR	polymerase chain reaction
pinII	proteinase inhibitor II
ppm	parts per million
qPCR	quantitative polymerase chain reaction
RNA	ribonucleic acid
SAMS	S-adenosyl-L-methionine synthetase
SCP1	synthetic core promoter
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SU	sulfonylurea
TMV	Tobacco Mosaic Virus
USDA	United States Department of Agriculture
UTR	untranslated region

*Abbreviations of units of measurement and of physical and chemical quantities are done according to the standard format described in Instructions to Authors in the Journal of Biological Chemistry (http://www.jbc.org/).

Table of Contents

List of Tables

List of Figures

I. Rationale for the Development of 356043 Soybean

I-A. Basis for the Request for a Determination of Nonregulated Status under 7 CFR Part 340.6

The Animal and Plant Health Inspection Service (APHIS) of the U.S. Department of Agriculture (USDA) has responsibility, under the Plant Protection Act (7 U.S.C. 7701-7772) and the Plant Quarantine Act (7 U.S.C. 151-167), to prevent the introduction or dissemination of plant pests into or within the United States. Part 340 regulates introduction of organisms altered or produced through genetic engineering which are plant pests or for which there is a reason to believe are plant pests. The APHIS regulations at 7 CFR 430.6 provide that an applicant may petition APHIS to evaluate submitted data on the genetically engineered crop to determine that a regulated article does not present a plant pest risk and therefore should no longer be regulated.

Pioneer Hi-Bred International, Inc. is submitting data for genetically engineered herbicide tolerant 356043 soybean and requests a determination from APHIS that event DP-356Ø43-5 and crosses of this event with nonregulated soybean lines no longer be considered regulated articles under 7 CFR 340.

I-B. Benefits of 356043 Soybean

The commercialization of herbicide tolerant 356043 soybean is expected to have a beneficial impact on weed control practices, providing growers with another tool to address their weed control needs. The availability of 356043 soybean will enable growers to choose an optimal combination of glyphosate, ALS-inhibiting herbicides, and other herbicides to best manage their individual weed populations. The availability of 356043 soybean will enable growers to proactively manage weed populations while avoiding adverse population shifts of troublesome weeds or the development of resistance.

Glyphosate has proven to be a popular herbicide available in soybeans, and combinations of glyphosate with ALS-inhibiting herbicides or other herbicides with different modes of action and inherent crop tolerance will enable more effective management of weed populations. Glyphosate and the ALS-inhibiting herbicide family are complementary in spectrum. Glyphosate is a post-emergence herbicide used to control weeds after emergence, whereas ALS inhibitors can effectively control weeds when applied post-emergence as well as pre-emergence (prior to weed emergence). Both chemical families are broad-spectrum herbicides and effective in controlling grasses and annual broadleaf weeds. Both herbicides have excellent environmental profiles and low mammalian toxicity. Glyphosate has no residual soil activity, thus it does not control weeds prior to their emergence, whereas the ALS inhibitors can be applied to the soil for pre-emergence activity as well as to emerged weeds in a post-emergence application. Finally, the ALS-inhibiting herbicides are typically applied at very low usage rates that are in the ounces per acre or less range. Therefore, the availability of 356043 soybean with tolerance to both glyphosate and ALS-inhibiting herbicides will be beneficial for growers and the environment.

I-C. Submissions to Other Regulatory Agencies

An Early Food Safety Evaluation for the GAT4601 protein was submitted to FDA on June 16, 2006. A safety and nutritional assessment for feed and food derived from 356043 soybean will be submitted to FDA in 4th quarter, 2006. Submission of a tolerance petition and supporting residue data to EPA to amend the soybean tolerance expression for glyphosate to include N-acetylglyphosate is targeted for 4th quarter, 2006, as are submissions to Health Canada and CFIA for food, feed and environmental approvals.

II. The Biology of Soybean

Refer to the OECD Consensus Document on the Biology of *Glycine max* (L.) Merr. (Soybean), 2000, for information pertaining to the following aspects of soybean biology:

- general description, including taxonomy and morphology and use as a crop plant;
- agronomic practices;
- centers of origin;
- reproductive biology;
- cultivated *Glycine max* as a volunteer weed;
- ability to cross inter-species/genus, introgression into relatives, and interactions with other organisms;
- summary of the ecology of *Glycine max*.

Characterization of the Recipient Soybean Cultivar

The publicly available cultivar, Jack, was used as the recipient line for the generation of 356043 soybean. The variety was originally developed at the Illinois Agricultural Experimental Station and commercially released in 1989 (Nickell *et al*., 1990). Jack is classified as maturity group II and is best adapted to approximately 40 to 42 degrees of Northern latitude. It has white flowers, gray pubescence, brown pods at maturity, and seeds with dull yellow coat and yellow hila. Jack was developed and released because of its resistance to soybean cyst nematode (Races 3 and 4) and higher yield when compared with cultivars of similar maturity. It is susceptible to phytophthora rot (Races 1, 4, and 7).

Jack is extensively used in soybean transformation because of its high embryogenic capacity (Stewart *et al*., 1996; Santarem *et al*., 1998; Yan *et al*., 2000). Somatic embryos can be induced from immature cotyledons, proliferated, and maintained in liquid medium until transformation.

III. Method of Development of 356043 Soybean

III-A. Description of the Transformation System

A linear DNA fragment PHP20163A (5.4 kb, containing *gat4601* and *gm-hra* gene cassettes) was used for transformation to generate 356043 soybean. Refer to section IV for the detailed description of the DNA fragment. The fragment was obtained from the plasmid PHP20163 by digestion with *Asc* I and *Not* I restriction enzymes and purified using agarose gel electrophoresis.

The transgenic 356043 soybean was generated using the Biolistics PDS-1000/He particle gun, manufactured by Bio-Rad (Hercules, CA), essentially as described by Klein *et al*. (1987). The targets for transformation were clumps of secondary somatic embryos derived from explants from small, immature soybean seeds of the cultivar Jack. The secondary somatic embryos were excised from immature explants, transferred to a liquid soybean culture maintenance medium, and subcultured at regular intervals until prepared for bombardment.

Soybean somatic embryogenic cultures were used in transformation experiments two to four months after initiation. On the day of transformation, microscopic gold particles were coated with the purified fragment PHP20163A DNA and accelerated into the embryogenic soybean cultures. Only PHP20163A DNA was used, and no additional DNA (*e.g.,* carrier DNA) was used in the transformation process.

Following transformation, the soybean tissue was transferred to flasks of fresh liquid culture maintenance medium for recovery. After seven days, the liquid culture medium was changed to culture maintenance medium supplemented with chlorsulfuron as the selection agent. Chlorsulfuron belongs to a family of ALS-inhibiting herbicides, and therefore only soybean cells that had stably inherited the *gm-hra* transgene continued to grow.

After several weeks in the culture maintenance medium supplemented with chlorsulfuron, small islands of healthy, chlorsulfuron-tolerant green tissue became visible and started to grow out of pieces of dying somatic embryogenic tissue. Green embryogenic clumps were excised from associated pieces of dying or dead tissue and received regular changes of fresh liquid selection medium until the start of the regeneration process. Embryogenic tissue samples were analyzed to confirm the presence of the *gat4601* and *gm-hra* transgenes by Southern blot hybridization. T0 plants were regenerated and transferred to the greenhouse for seed production. Refer to Figure 1 for a schematic diagram of the development process for 356043 soybean and Figure 2 for a breeding diagram.

III-B. Selection of Comparators for 356043 Soybean

To ensure the accurate assessment of the impact of transgene insertion on various characteristics of 356043 soybean, a proper selection of comparator plants is important. Two types of soybean lines—control soybean and reference soybean, were used as comparators for 356043 soybean (Figure 2).

The control plants should have a genetic background similar to that of 356043 soybean but lack the transgenic insert. In most cases, the variety Jack was used as the control. Jack is the recipient variety that was transformed to generate 356043 soybean and is a variety that is homozygous at nearly all loci. For this reason, the T0 plants and all

subsequent selfed generations derived post-transformation were essentially genetically identical to Jack soybean with the exception of the transgenic insert DNA.

For the molecular analysis, a Pioneer elite variety (“Elite 1”) was also employed as a non-transgenic control for the F3 generation. Elite 1 was one of the elite varieties used in breeding crosses to generate F3 lines. In some agronomic experiments, a null segregant of 356043 soybean was used. The null segregant is also an appropriate comparator because it is essentially genetically identical to 356043 soybean with the exception of the transgenic insert DNA.

For the nutritent compositional assessment and poultry study, nontransgenic commercial soybean reference varieties were used to help determine the normal variation seen in soybean and to develop the statistical tolerance intervals.

Figure 1. Schematic of the Development of 356043 Soybean

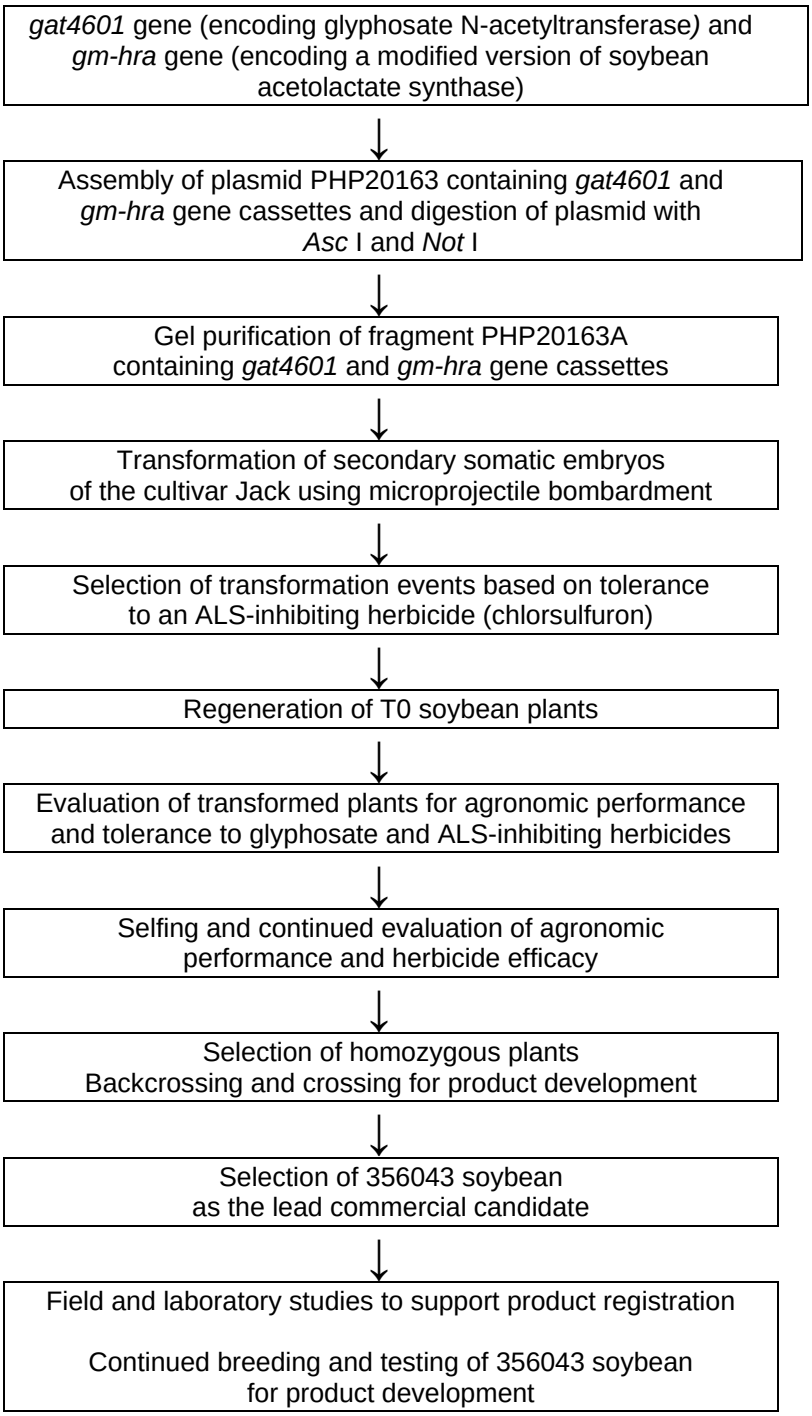
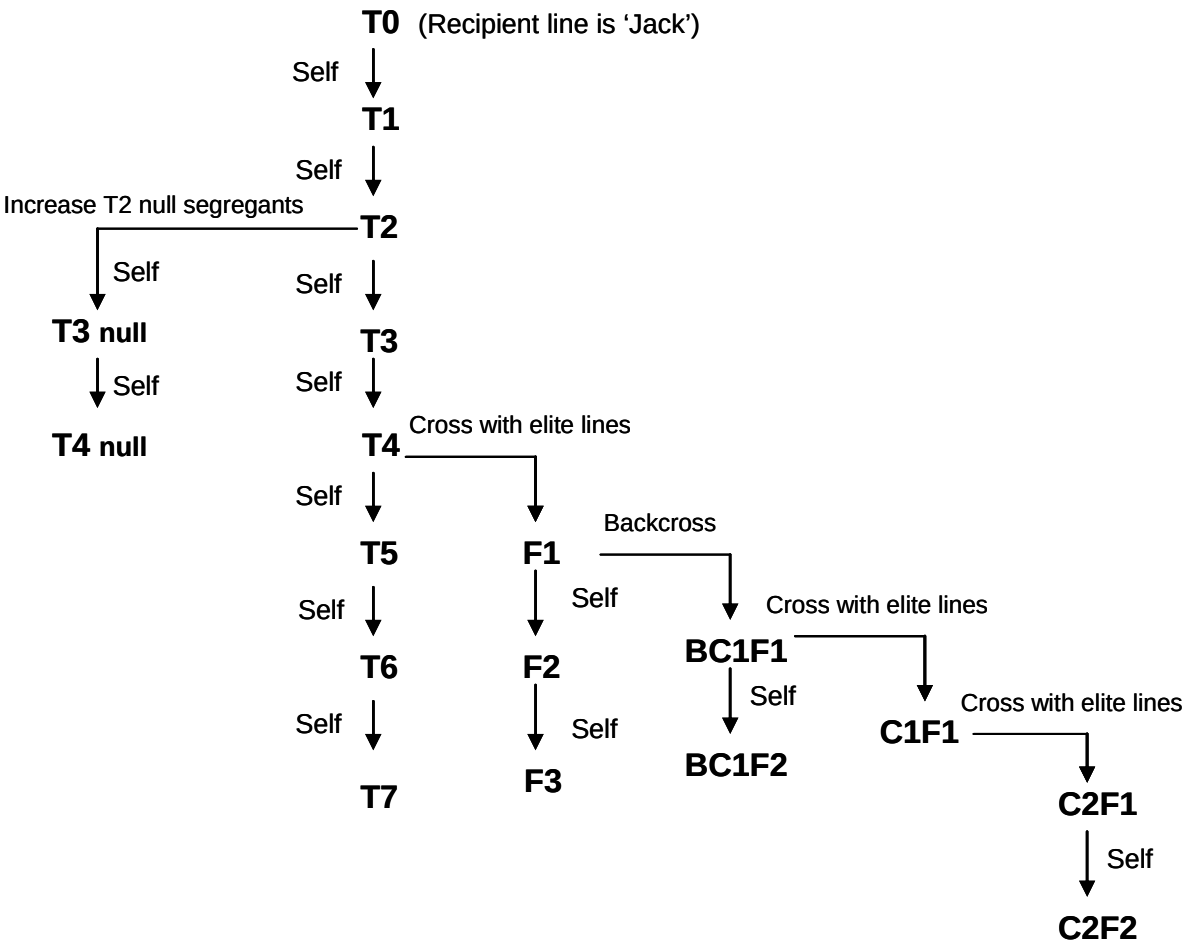


Figure 2. Breeding Diagram for 356043 Soybean and Generations Used for Analyses



Analysis	Relevant Section of Petition	356043 Soybean Generation Used	Control Used
Molecular	V-A, B, C and D	T4, T5 and F3	Jack and Elite 1*
Inheritance	V-E	T1, F2, F3, BC1F2 and C2F2	Not applicable
Levels of transgenic proteins	VI-C	T5	Jack
Germination / dormancy	VII-A	T7	Jack
Field agronomics	VII-B	T3	T3 null segregants
		T4	T4 null segregants
		T5	Jack
Compositional assessment	VIII	T5	Jack and four commercial soybean varieties
Poultry study	VIII-D	T7	Jack and three commercial soybean varieties

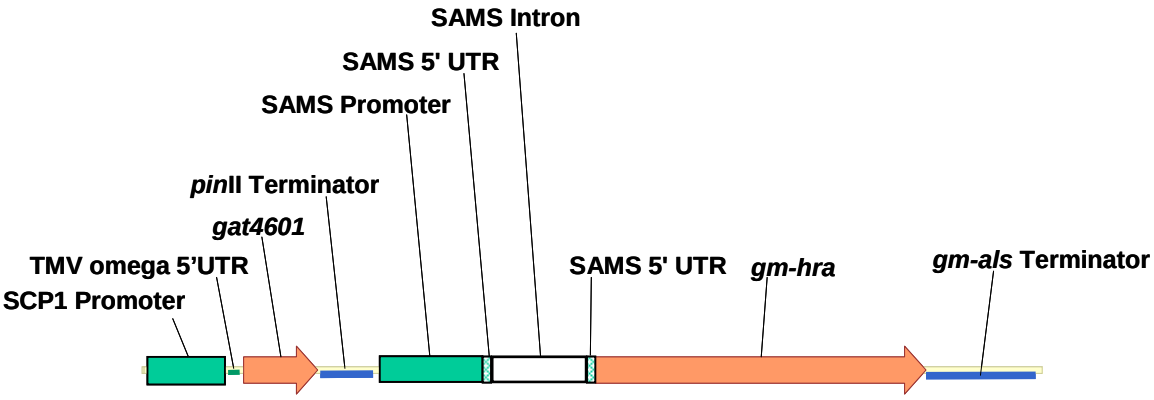
*****Elite 1 is one of several elite soybean varieties crossed with T4 plants to yield F1 plants.

IV. Donor Genes and Regulatory Sequences

IV-A. DNA Fragment PHP20163A Used in Transformation

356043 soybean was produced by particle bombardment with DNA fragment PHP20163A (Figure 3). The source plasmid PHP20163 is represented schematically in Figure 4. A summary of the elements and their position on the source plasmid, PHP20163, is given in Table 1.

Figure 3. Map of the Fragment PHP20163A



Schematic diagram of the fragment used for particle bombardment with the *gat4601* gene cassette and *gm-hra* gene cassette elements indicated. Length of the fragment is 5361 base pairs.

Figure 4. Plasmid Map of PHP20163

Schematic diagram of plasmid PHP20163 with the location of genes and regulatory elements indicated. Plasmid size is 7953 base pairs. PHP20163A (Figure 3) was isolated from this plasmid by a *Not* I and *Asc* I double digestion. Enzyme sites for *Not* I and *Asc* I are indicated at base pair position 1 and 5362, respectively.

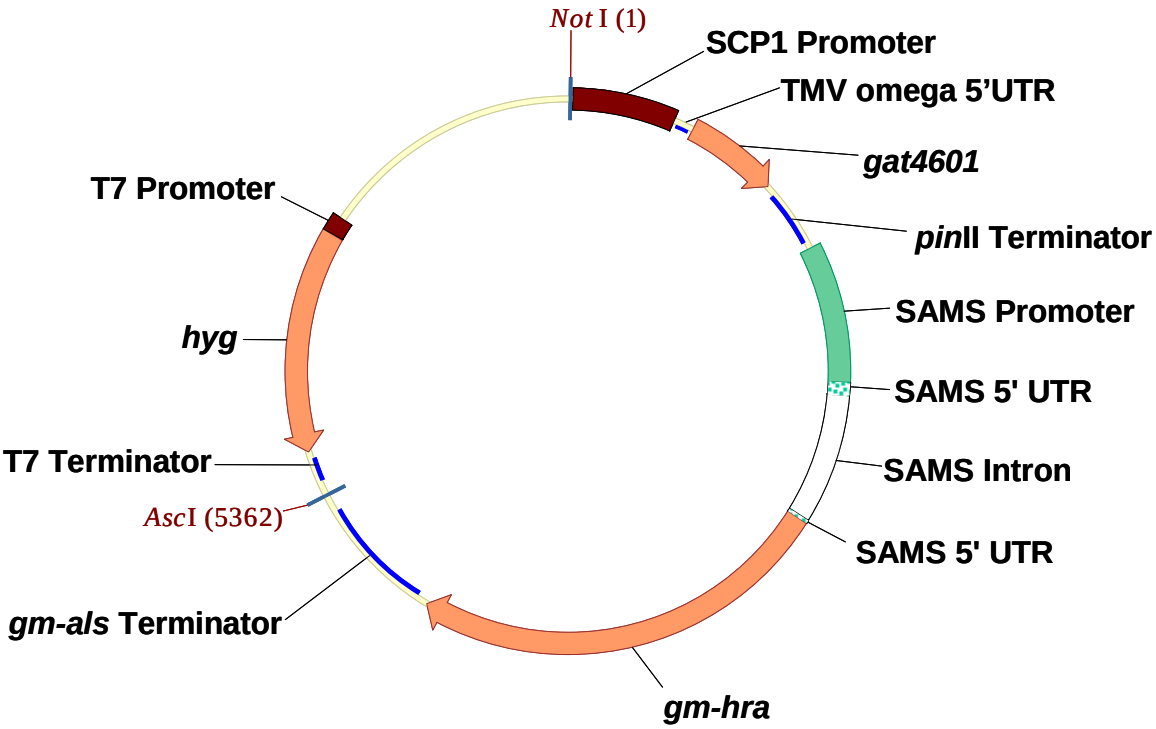


Table 1. Description of Genetic Elements in Fragment PHP20163A

Location on Plasmid PHP20163 (base pair position)	Genetic Element	Size (base pairs)	Description
1 to 16	polylinker region	16	Region required for cloning genetic elements
17 to 502	SCP1 promoter	486	Constitutive synthetic promoter comprising a portion of the CaMV 35S promoter (Odell *et al.*, 1985) and the Rsyn7-Syn II Core consensus promoter (Bowen *et al.*, 2000 and 2003).
503 to 504	polylinker region	2	Region required for cloning genetic elements
505 to 571	TMV omega 5'-UTR	67	An element derived from the Tobacco Mosaic Virus omega 5' untranslated leader that enhances translation (Gallie and Walbot, 1992).
572 to 596	polylinker region	25	Region required for cloning genetic elements
597 to 1037	*gat4601* gene	441	Synthetic glyphosate N-acetyltransferase (*gat*) gene (Castle *et al.*, 2004).
1038 to 1053	polylinker region	16	Region required for cloning genetic elements
1054 to 1369	*pin*II terminator	316	Terminator region from *Solanum tuberosum* proteinase inhibitor II (*pin*II) gene (Keil *et al.,* 1986; An *et al.*, 1989).
1370 to 1385	polylinker region	16	Region required for cloning genetic elements
1386 to 2030	SAMS promoter	645	Promoter of the S-adenosyl-L-methionine synthetase (SAMS) gene from soybean (Falco and Li, 2003).
2031 to 2089	SAMS 5'-UTR	59	5' untranslated region of the SAMS gene from soybean (Falco and Li, 2003).
2090 to 2680	SAMS intron	591	Intron within the 5'-untranslated region of the SAMS gene from soybean (Falco and Li, 2003).
2681 to 2696	SAMS 5'-UTR	16	5' untranslated region (UTR) of the SAMS gene from soybean (Falco and Li, 2003).
2697 to 4667	*gm-hra* gene	1971	Modified version of the acetolactate synthase gene from soybean with 15 additional nucleotides on the 5' end (2697 to 2711) derived from the 5' untranslated region of the *gm-als* gene and two nucleotide changes within the coding sequence.
4668 to 5318	*gm-als* terminator	651	Native terminator from the soybean acetolactate synthase gene.
5319 to 5361	polylinker region	43	Region required for cloning genetic elements

IV-B. Identity and Source of Genetic Material in Fragment PHP20163A

B1. The *gat4601* Gene Cassette

Fragment PHP20163A contains two gene cassettes. The first cassette contains the synthetic glyphosate N-acetyltransferase (*gat4601*) gene that encodes the GAT4601 protein. The *gat* genes were isolated from three strains of *Bacillus licheniformis*, and the *gat4601* sequence was generated by functional optimization of these genes using the gene shuffling process to enhance the acetylation activity of the GAT enzyme (see Section VI-A for additional details). The GAT4601 protein acetylates glyphosate, rendering it non-phytotoxic to plants. The insertion of the *gat4601* gene in soybean plants confers tolerance to glyphosate-containing herbicides. The *gat4601* gene encodes a protein of 146 amino acids that has a molecular weight of approximately 17 kDa (Castle *et al.*, 2004).

B. licheniformis, the source organism for the *gat4601* gene, is used for the production of a number of enzymes such as proteases and amylases that have wide application in the detergent industry. *B. licheniformis* has been used in the United States, Canada, and Europe in the fermentation industry for production of food enzymes (*e.g.,* alpha-amylase, cyclodextrin glycosyltransferase, hemicellulase, proteases, and pullulanase; Rey *et al.,* 2004). All *B. licheniformis* cultures available from the American Type Culture Collection (ATCC) are classified as Biosafety Level 1. Items that the ATCC classifies at Biosafety Level 1 have no known history of causing disease in humans or animals based on their assessment of potential risk using U.S. Public Health Service guidelines, with assistance provided by ATCC scientific advisory committees.

Expression of the *gat4601* gene is controlled by two regulatory elements upstream of the *gat4601* coding region. Furthest upstream is the SCP1 promoter that is a synthetic constitutive promoter comprising a portion of the CaMV 35S promoter (Odell *et al.*, 1985) and the Rsyn7-Syn II Core synthetic consensus promoter (Bowen *et al.*, 2000 and 2003). Downstream from this element is the omega 5' untranslated leader of the Tobacco Mosaic Virus (TMV omega 5'-UTR) for enhancing translation (Gallie and Walbot, 1992). The terminator for the *gat4601* gene is the 3' terminator sequence from the proteinase inhibitor II gene of *Solanum tuberosum* (*pin*II terminator) (Keil *et al*., 1986; An *et al.*, 1989).

B2. The *gm-hra* Gene Cassette

The second cassette of fragment PHP20163A contains *gm-hra*, a modified version of the endogenous soybean acetolactate synthase gene (*gm-als*). The *gm-hra* gene encodes the GM-HRA protein, which has two amino acid residues modified in comparison to the wild-type ALS protein and five additional amino acids at the N-terminus derived from the translation of 15 nucleotides from the 5' untranslated region of the soybean *gm-als* gene. This gene encodes a resistant form of acetolactate synthase that confers tolerance to the ALS-inhibiting class of herbicides. The GM-HRA protein is comprised of 656 amino acids and has a molecular weight of approximately 71 kDa.

The expression of the *gm-hra* gene is controlled by the promoter from an S-adenosyl-L-methionine synthetase (SAMS) gene from soybean (Falco and Li, 2003). This consists of a constitutive promoter and an intron that interrupts the SAMS 5' untranslated region (Falco and Li, 2003). The terminator for the *gm-hra* gene is the native soybean acetolactate synthase terminator (*gm-als* terminator).

V. Genetic Characterization of 356043 Soybean

V-A. Molecular Analysis Overview

To characterize the DNA insertion in 356043 soybean, Southern blot analysis was conducted. Individual plants of the T4 generation were analyzed to determine the copy number of each of the genetic elements inserted into 356043 soybean and to verify that the integrity of the PHP20163A fragment was maintained upon integration. The analysis confirmed a single, intact PHP20163A fragment has been inserted into the soybean genome to produce 356043 soybean (Section V-B).

This analysis also confirmed that the DNA insertion remained stable during traditional soybean breeding procedures. Southern blot analysis was conducted on two selfed generations, T4 and T5, and verified that the insertion remained intact and stably integrated as demonstrated by identical hybridization patterns in the two generations. In addition to the selfed generations, the F3 generation was also analyzed by Southern blot analysis, confirming the same stable, event-specific hybridization pattern as exhibited by the T4 and T5 generations. These results confirmed the stability of the insertion in 356043 soybean across multiple breeding generations (Section V-C).

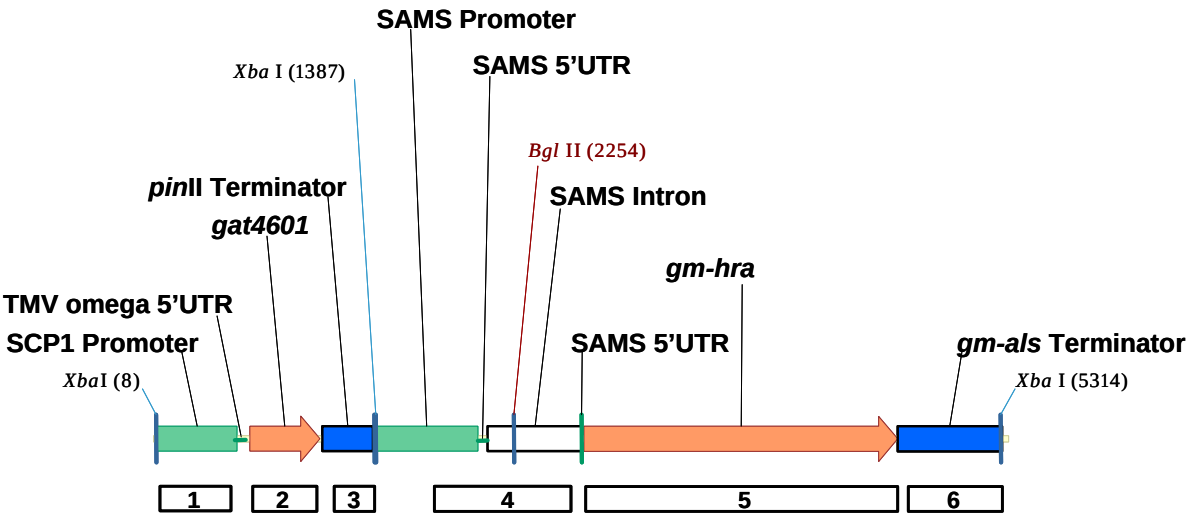
Both the T4 and T5 generations were analyzed to confirm the absence of plasmid sequence from PHP20163 outside of the transformation fragment PHP20163A, *i.e.* the plasmid backbone sequence removed prior to transformation. The results verified the absence of backbone sequences in 356043 soybean (Section V-D).

Genomic DNA from leaf material of the Jack soybean variety was used as a negative control for all Southern blot analyses. Genomic DNA from an elite soybean variety (Elite 1) was included as an additional negative control for analysis of the F3 generation. Plasmid PHP20163 was used as a positive control for probe hybridization and to verify fragment sizes internal to the transformation fragment PHP20163A. All probes used for the analysis are indicated on the schematic maps of PHP20163A and PHP20163 (Figures 5 and 6, respectively) and outlined in Table 2.

Based on these analyses, a schematic map of the insertion region in 356043 soybean was determined and is presented in Figure 7.

Detailed descriptions of the methods used for Southern blot analysis are further described in Appendix 1. The explanations of the T4, T5 and F3 soybean generations can be found in Figure 2.

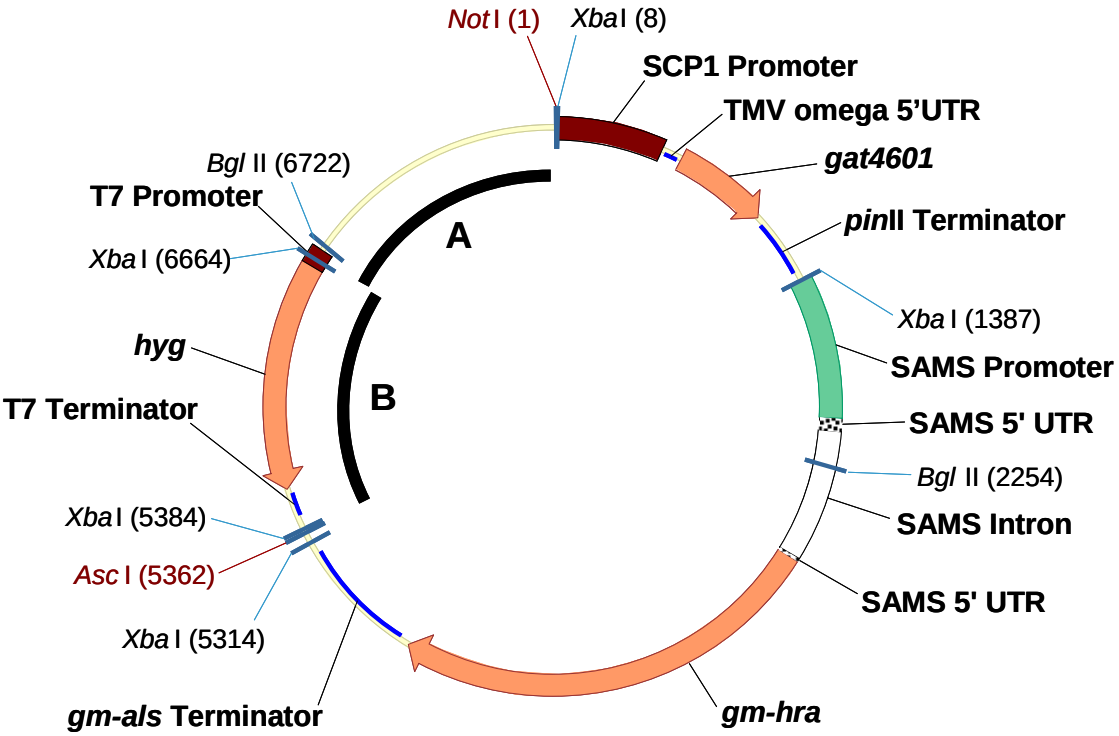
Figure 5. Map of Fragment PHP20163A with Genetic Element Probes Indicated



Schematic map of fragment PHP20163A indicating location of the genetic elements contained in the two gene expression cassettes and base pair positions for *Bgl* II and *Xba* I restriction enzyme sites. The *Not* I and *Asc* I restriction enzyme sites are lost upon excision of this fragment from PHP20163. The total fragment size is 5361 base pairs. Approximate locations of the probes used are shown as numbered boxes below the fragment and are identified below. Additional details on these probes are provided in Table 2.

Number	Probe Identification
1	SCP1 promoter probe
2	*gat4601* probe
3	*pin*II terminator probe
4	SAMS probe
5	*gm-hra* probe
6	*gm-als* terminator probe

Figure 6. Plasmid Map of PHP20163 with Backbone Probes Indicated



Schematic plasmid map of PHP20163 indicating the location of genetic elements and base pair positions for restriction enzyme sites for *Not* I, *Bgl* II, *Xba* I, and *Asc* I. The *Not* I-*Asc* I fragment of this plasmid was isolated (PHP20163A; map, Figure 5) and used for transformation to produce 356043 soybean. The *Xba* I site located at bp 1387 contains a Dam methylase recognition site and is resistant to digestion by *Xba* I if the plasmid is prepared from a Dam^+ strain. The total plasmid size is 7953 base pairs. Backbone probes are indicated schematically as lines within the plasmid diagram and are identified below. Additional details on these probes are provided in Table 2.

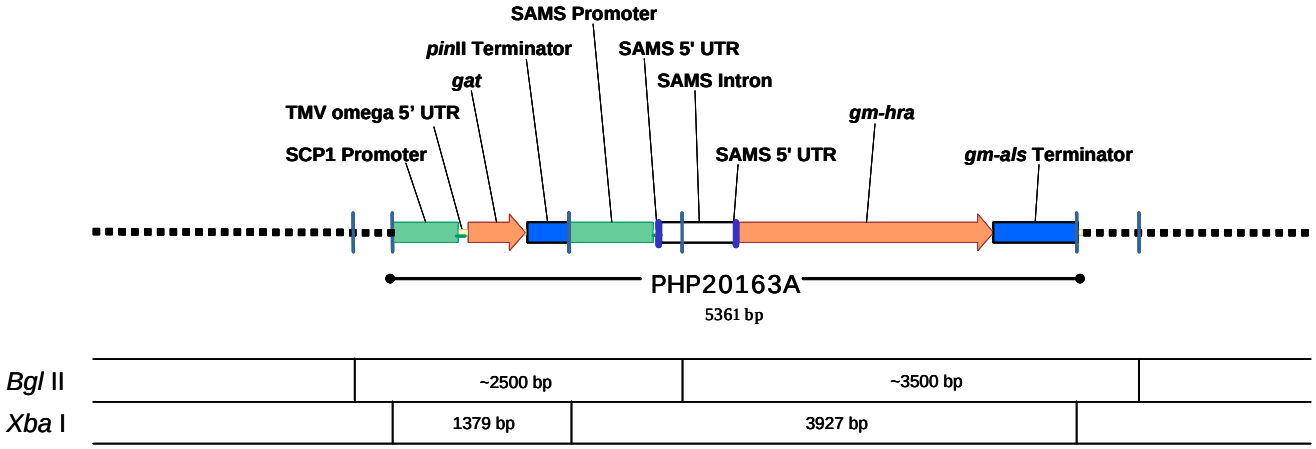
Letter	Backbone Probe Identification
A	*backbone 20163* probe
B	*hyg 20163* probe

Table 2. Description of DNA Probes Used for Southern Blot Hybridization

Probe Name	Genetic Element	Figure *Probe*	Position on PHP20163A (bp to bp)	Position on PHP20163 (bp to bp)	Length (bp)
SCP1 promoter	SCP1 promoter	Figure 5 *probe 1*	12 to 479	12 to 479	486
gat4601	*gat4601* gene	Figure 5 *probe 2*	597 to 1012	597 to 1012	416
*pin*II terminator	*pin*II terminator	Figure 5 *probe 3*	1107 to 1340	1107 to 1340	234
SAMS[1]	SAMS promoter and intron elements	Figure 5 *probe 4*	1702 to 2146 2147 to 2638	1702 to 2146 2147 to 2638	445 492
gm-hra[1]	*gm-hra* gene	Figure 5 *probe 5*	2700 to 3629 3635 to 4664	2700 to 3629 3635 to 4664	930 1030
gm-als terminator	*gm-als* terminator	Figure 5 *probe 6*	4670 to 5318	4670 to 5318	649
backbone 20163[1]	plasmid backbone of PHP20163	Figure 6 *probe A*	N/A[2]	7426-7953 6664-7415	528 752
hyg 20163[1]	hygromycin resistance gene of PHP20163	Figure 6 *probe B*	N/A	6096-6618 5388-6090	523 703

[1]Two non-overlapping segments were generated for this probe and were combined for hybridization.
[2]Not Applicable; these are not present on the PHP20163A fragment.

Figure 7. Schematic Map of the Insertion in 356043 Soybean



Schematic map of the transgene insertion in 356043 soybean based on Southern blot analysis. The flanking soybean genome is represented by the horizontal dotted line. A single, intact copy of the PHP20163A fragment integrated into the soybean genome. *Bgl* II and *Xba* I restriction enzyme sites are indicated with the sizes of observed fragments on Southern blots shown below the map in base pairs (bp).

V-B. Transgene Copy Number and Insertion Integrity

The integration pattern of the insertion in 356043 soybean was investigated with *Bgl* II digestion to determine copy number and with *Xba* I digestion to determine insertion integrity. Southern blots were hybridized to several probes to confirm copy number and integrity of each genetic element. SCP1 promoter, *gat4601*, and *pin*II terminator probes were used to characterize the *gat4601* cassette (Table 2 and Figure 5). SAMS, *gm-hra*, and *gm-als* terminator probes were used to characterize the *gm-hra* cassette (Table 2 and Figure 5).

B1. Transgene Copy Number

The *Bgl* II digest provides information about number of copies integrated into the genome of 356043 soybean as there is a single restriction enzyme site in the PHP20163A fragment at base pair (bp) position 2254 (Figure 5) and additional sites outside the fragment in the soybean genome. Hybridization with the probes from each cassette, except for the SAMS probe, would indicate the number of copies of each element found in 356043 soybean based on the number of hybridizing bands (*e.g.* one hybridizing band indicates one copy of the element). For the SAMS probe, since the *Bgl* II site is located within the probe region, two hybridizing bands would be expected for every one copy of the element. Predicted and observed fragment sizes for 356043 soybean with *Bgl* II are given in Table 3 for the *gat4601* cassette and in Table 4 for the *gm-hra* cassette.

Based on the Southern blot analyses as discussed below, it was determined that a single, intact PHP20163A fragment has been inserted into the genome of 356043 soybean as diagramed in the insertion map (Figure 7).

Table 3. Predicted and Observed Hybridizing Bands on Southern Blots with *gat4601* Cassette Probes

Probe	Restriction Enzyme	Figure(s)	Predicted Fragment Size from PHP20163A[1] (bp)	Predicted Fragment Size from PHP20163[2] (bp)	Observed Fragment Size in 356043 soybean[3] (bp)
SCP1 promoter	*Bgl* II	8	>2300[4]	3485	~2500
gat4601	*Bgl* II	8, 14, 16	>2300[4]	3485	~2500
*pin*II terminator	*Bgl* II	8	>2300[4]	3485	~2500
SCP1 promoter	*Xba* I	10	1379	1379[5]	1379[6]
gat4601	*Xba* I	10, 11	1379	1379[5]	1379[6]
*pin*II terminator	*Xba* I	10	1379	1379[5]	1379[6]

[1] Predicted fragment sizes for 356043 soybean are based on the map of PHP20163A as shown in Figure 5.
[2] Predicted fragment sizes for hybridization to samples containing the plasmid positive control are based on the PHP20163 plasmid map as shown in Figure 6.
[3] Observed fragment sizes are considered approximate from these analyses and are based on the indicated sizes of the DIG-labeled DNA Molecular Weight Marker VII fragments on the Southern blots. Due to incorporation of DIG (digoxygenin) molecules for visualization, the marker fragments typically run approximately 5-10% larger than their actual indicated molecular weight.
[4] Minimum fragment size predicted based on an intact insertion of PHP20163A (Figure 5).
[5] Predicted hybridizing *Xba* I fragment size from plasmid PHP20163 grown in a Dam^- strain. For plasmid grown in a Dam^+ strain, the hybridizing *Xba* I fragment size is predicted to be 5307 bp.
[6] Observed fragment is equal to the predicted size based on blot presented in Figure 12 showing comparison to plasmid PHP20163 grown in a Dam^- strain.

Table 4. Predicted and Observed Hybridizing Bands on Southern Blots with *gm-hra* Cassette Probes

Probe	Restriction Enzyme	Figure(s)	Predicted Fragment Size from PHP20163A[1] (bp)	Predicted Fragment Size from PHP20163[2] (bp)	Observed Fragment Size in 356043 Soybean[3] (bp)
SAMS	*Bgl* II	9	>3100[4] >2300[4]	4468 3485	~3500 ~2500 ~6900* ~4700*
gm-hra	*Bgl* II	9, 15, 17	>3100[4]	4468	~3500 >8600* >8600*
gm-als terminator	*Bgl* II	9	>3100[4]	4468	~3500 >8600*
SAMS	*Xba* I	12	3927	3927[5]	3927[6] >8600*
gm-hra	*Xba* I	12, 13	3927	3927[5]	3927[6] >8600* ~8600* ~7400* ~6800* ~5800* ~4500*
gm-als terminator	*Xba* I	12	3927	3927[5]	3927[6] ~5800*

Note: An asterisk (*) and gray shading indicates the designated band is due to probe hybridization to endogenous soybean genome sequences, as can be determined by the presence of the same band in all lanes, both 356043 soybean and control. Certain endogenous bands may be difficult to discern on a printed copy but are visible on the original film and electronic version.

[1] Predicted fragment sizes for 356043 soybean are based on the map of PHP20163A as shown in Figure 5.
[2] Predicted fragment sizes for hybridization in samples containing the plasmid positive control are based on the PHP20163 plasmid map as shown in Figure 6.
[3] Observed fragment sizes are considered approximate from these analyses and are based on the indicated sizes of the DIG-labeled DNA Molecular Weight Marker VII fragments on the Southern blots. Due to incorporation of DIG molecules for visualization, the marker fragments typically run approximately 5-10% larger than their actual indicated molecular weight.
[4] Minimum fragment size predicted based on an intact insertion of PHP20163A (Figure 5).
[5] Predicted hybridizing *Xba* I fragment size from plasmid PHP20163 grown in a Dam^- strain. For plasmid grown in a Dam^+ strain, the hybridizing *Xba* I fragment size is predicted to be 5307 bp.
[6] Observed fragment is equal to the predicted size based on blot presented in Figure 13 showing comparison to plasmid PHP20163 grown in a Dam^- strain.

gat4601 Cassette:

A single copy of all the elements of the *gat4601* cassette was inserted into 356043 soybean. SCP1 promoter, *gat4601*, and *pin*II terminator probes were hybridized to *Bgl* II-digested genomic DNA from individual 356043 soybean plants of the T4 generation (Table 3, Figure 8). Each of the probes hybridized to the same single fragment of approximately 2500 base pairs (bp) (Table 3, Figure 8), indicating the expected arrangement of genetic elements on the fragment inserted in 356043 soybean.

gm-hra Cassette:

Likewise, a single copy of all the elements of the *gm-hra* cassette was inserted into 356043 soybean. The elements comprising this cassette - the SAMS promoter region, *gm-hra* gene, and *gm-als* terminator - were used as probes to determine the number of copies inserted. The probes of this cassette are homologous to elements endogenous to the soybean genome and therefore each probe hybridized to bands in control soybean samples. The hybridizing bands in 356043 soybean from the endogenous soybean genome are indicated by asterisks in the shaded boxes of Table 4 and were determined by their presence in control soybean samples and are thus not associated with the insertion.

The SAMS probe hybridized to one band of approximately 2500 bp and a second band of approximately 3500 bp in 356043 soybean (Table 4, Figure 9, SAMS probe). Two bands would be expected to indicate one insertion of this element with this probe as the *Bgl* II site is located within the SAMS region of PHP20163A (Figure 5) and the results indicated one copy of the element. The 2500 bp fragment was determined to be the same fragment containing the *gat4601* cassette as described above and as shown in Figure 8. The intensity of the 3500 bp band was noticeably less than that of the 2500 bp band (Figure 9, SAMS probe) due to the location of the restriction site and the specific sequence of the probe regions in each fragment (Figure 5). The region hybridizing to the probe on the 3500 bp fragment is shorter and has a lower percentage of guanine and cytosine bases and, therefore, will hybridize less efficiently than the hybridizing region on the 2500 bp fragment. This differential hybridization intensity was also reflected in the hybridization pattern of the equivalent regions in plasmid PHP20163. The 4468 bp plasmid fragment hybridized less efficiently than the 3485 bp plasmid fragment with the SAMS probe (Figure 9, SAMS probe, Lane 6).

The *gm-hra* and *gm-als* terminator probes, the final two probes from the *gm-hra* cassette, hybridized to the same 3500 bp fragment as the SAMS probe (Table 4, Figure 9, *gm-hra* and *gm-als* terminator probes). The hybridization of all three probes to the same 3500 bp fragment and of the SAMS probe to the 2500 bp fragment confirmed the expected arrangement of the genetic elements in the DNA insertion in 356043 soybean.

B2. Insertion Integrity

Xba I digestion was used to verify that the *gat4601* and *gm-hra* cassettes were complete and intact in 356043 soybean as there are three sites in the PHP20163A fragment (base pair positions 8, 1387, and 5314) which precisely flank each gene expression cassette (Figure 5). Hybridization with the probes of the *gat4601* and *gm-hra* cassettes confirmed that all the elements were found on the appropriate internal fragments containing the cassette. Expected and observed fragment sizes with *Xba* I are given in Table 3 for the *gat4601* cassette and Table 4 for the *gm-hra* cassette.

The SCP1 promoter, *gat4601*, and *pin*II terminator probes each hybridized to the expected internal band of 1379 bp (Table 3, Figure 10) and the size was confirmed by

additional hybridizations as described in the section below and shown in Figure 11. Because these probes hybridized to the same internal fragment of the predicted size, the *gat4601* cassette in 356043 soybean was determined to be intact and all elements of the cassette were confirmed on this fragment.

The SAMS, *gm-hra*, and *gm-als* terminator probes each hybridized to the expected internal band of 3927 bp band (Table 4, Figure 12) and the size was confirmed by additional hybridization described in the section below and shown in Figure 13. Because these probes hybridized to the same fragment, the *gm-hra* cassette in 356043 soybean was determined to be intact and all elements were present.

Clarification of *Xba* I-Digested Plasmid PHP20163 Hybridization Results

For the blots in Figures 10 and 12, plasmid PHP20163 was prepared from a strain of *E. coli* that expresses a DNA methylase (a Dam^+ strain). This plasmid did not produce the expected bands when digested with *Xba* I and probed with the *gat4601* and *gm-hra* cassette probes. A band of approximately 5300 bp was observed in lanes containing plasmid PHP20163 for all probes (Figures 10 and 12, Lane 6) instead of the predicted 1379 bp and 3927 bp size bands (Tables 3 and 4). Based on the plasmid sequence, it was determined that the *Xba* I site at bp position 1387 of PHP20163 (Figure 6) overlaps a Dam methylation recognition sequence ($^{5'}$GATC$^{3'}$). The final adenine in this site (recognition sequence $^{5'}$TCTAG<u>A</u>$^{3'}$) is expected to be methylated, thus blocking digestion by *Xba* I.

The inability for *Xba* I enzyme to cut at this site affected the size prediction from fragment PHP20163A in 356043 soybean (Tables 3 and 4, Figure 5). In order to confirm the size in 356043 soybean and the plasmid result, plasmid PHP20163 was prepared from a strain of *E. coli* lacking Dam methylase (Dam^- strain) and was compared to the plasmid PHP20163 analyzed in Figures 10 and 12. Southern hybridization results presented in Figures 11 and 13 show the plasmid comparison alongside samples of the T4 and T5 generations of 356043 soybean digested with *Xba* I. The blot probed with *gat4601* and *gm-hra* probes demonstrated that plasmid PHP20163 prepared from the Dam^- strain digested as expected by *Xba* I and produced the predicted size bands of 1379 bp and 3927 bp (Lanes 1 and 2 in Figure 11 for the *gat4601* probe and Figure 13 for the *gm-hra* probe), while the original plasmid from the Dam^+ strain again produced a band of approximately 5300 bp for both probes (Figures 11 and 13, Lanes 18 and 19). These results confirmed that the central *Xba* I site at position 1387 was blocked from digestion due to Dam methylation. Furthermore, the hybridizing bands in 356043 soybean in Figures 11 and 13 were of the equivalent size as those in the unmethylated plasmid PHP20163 from the Dam^- strain (Figures 11 and 13, Lanes 1 and 2) confirming that 356043 soybean contained a complete and intact insertion.

V-C. Stability of the Insertion Across Generations

As discussed above, the *Bgl* II restriction enzyme has a single site (bp position 2254) located within the PHP20163A fragment (Figure 5) and will generate a unique event-specific hybridization pattern for 356043 soybean when hybridized to the *gat4601* and *gm-hra* probes. This analysis would confirm event stability across generations as changes to the insertion structure in 356043 soybean would be detected. As discussed in the section above, a band of approximately 2500 bp would be expected with the *gat4601* probe to confirm stability across generations (Table 3). Likewise, for the *gm-hra* probe, a band of approximately 3500 bp would be expected to confirm stability across generations (Table 4).

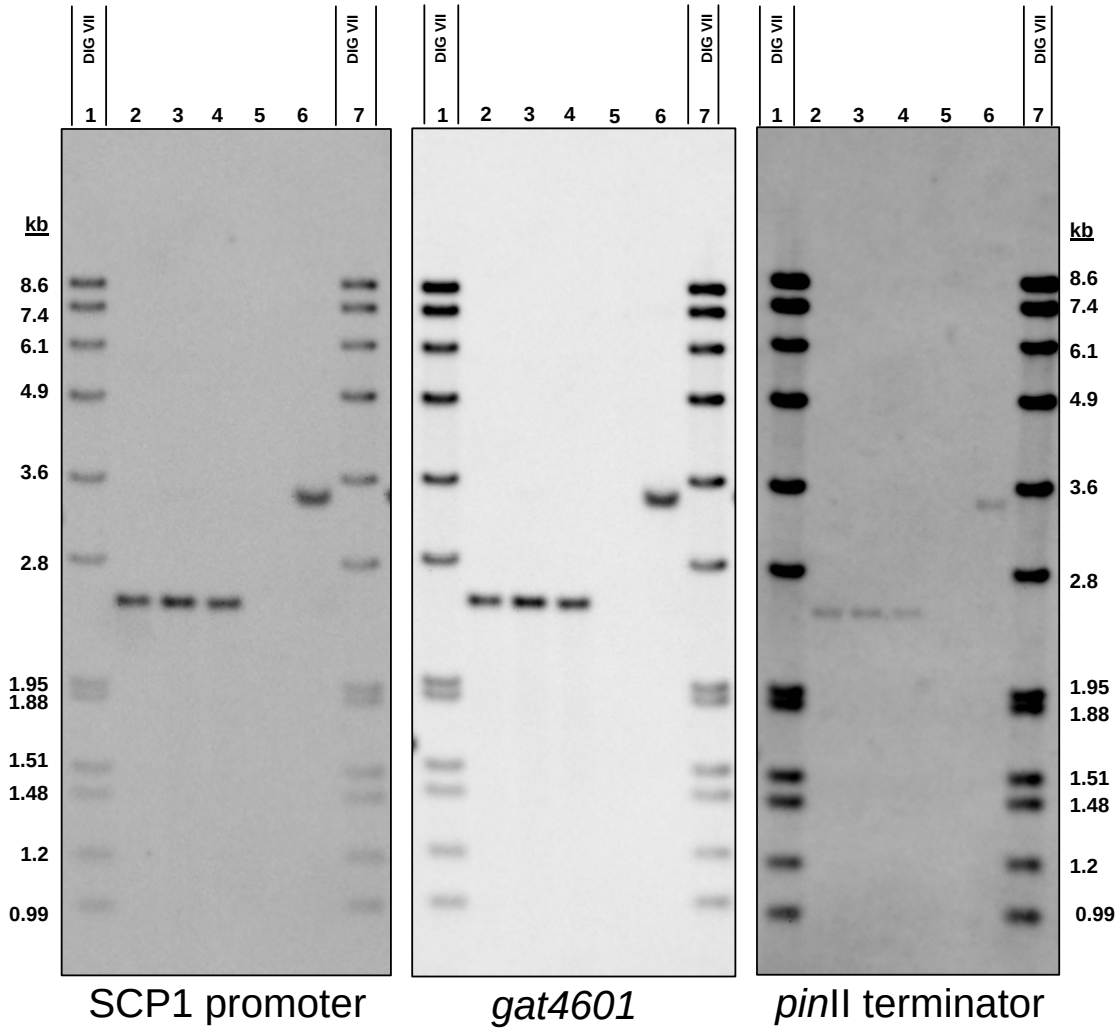
Genomic DNA of T4 and T5 generations of 356043 soybean was digested with *Bgl* II and hybridized to the *gat4601* and *gm-hra* probes to confirm stability across generations (Figures 14 and 15, respectively). A band of approximately 2500 bp specific to 356043 soybean hybridized to the *gat4601* probe in both the T4 and T5 generations (Table 3, Figure 14). With the *gm-hra* probe, a single band of approximately 3500 bp specific to 356043 soybean was present in both generations (Table 4, Figure 15). In addition to the 3500 bp band, the *gm-hra* probe also hybridized to additional bands that were determined to be endogenous to the soybean genome since these bands were present in both 356043 soybean and control soybean plants (Figure 15). Hybridization results from both the *gat4601* and *gm-hra* probes confirmed that the insertion of the PHP20163A DNA fragment in 356043 soybean remained stable across the selfed T4 and T5 generations.

Southern blot analysis of the F3 generation of 356043 soybean was also conducted. A total of 77 individual plants were analyzed and a representative Southern blot from the population is presented in Figures 16 and 17. Genomic DNA of the F3 generation was digested with *Bgl* II and hybridized to the *gat4601* and *gm-hra* probes. A band of approximately 2500 bp was observed with the *gat4601* probe (Figure 16) and a single band of approximately 3500 bp specific to 356043 soybean was observed with the *gm-hra* probe (Figure 17). As with the previous analysis conducted, the *gm-hra* probe hybridized to additional bands in 356043 soybean and control samples which were due to endogenous sequences within the soybean genome (Figure 17). Hybridization results from both the *gat4601* and *gm-hra* probes were consistent with the results from the T4 and T5 generations described above and confirmed the stability of inheritance of the insertion during traditional soybean breeding.

V-D. Absence of Plasmid Backbone DNA

The *backbone 20163* and *hyg 20163* probes were designed to hybridize to areas of plasmid PHP20163 outside of the transformation fragment (Figure 6, probes A and B, respectively) and were hybridized to *Xba* I-digested genomic DNA to confirm absence of these sections of the plasmid in 356043 soybean. Figures 18 and 19 show the results of hybridization with the *backbone 20163* and *hyg 20163* probes, respectively. Neither of the two backbone probes hybridized to 356043 soybean samples (Figures 18 and 19, Lanes 6 through 14), confirming the absence of these sequences.

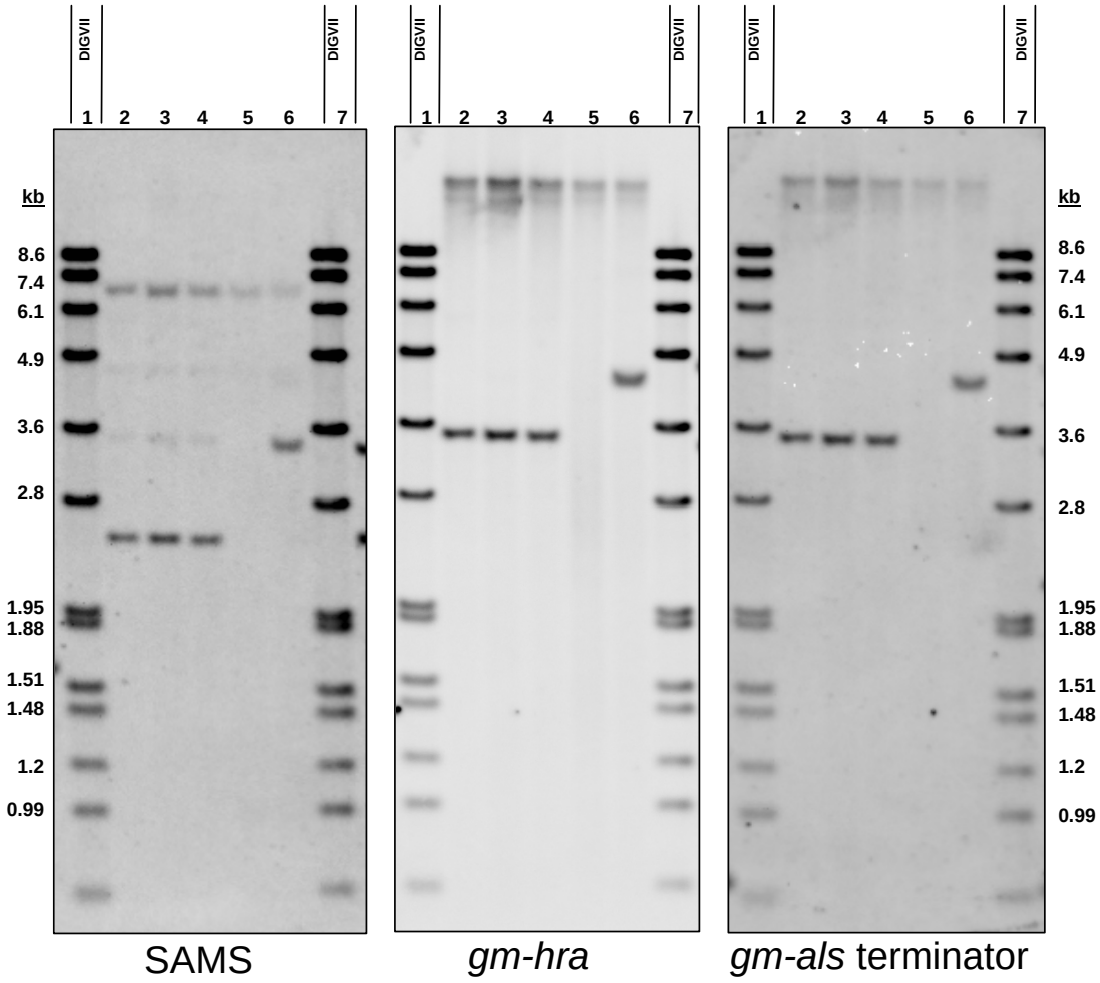
Figure 8. Southern Blot Analysis: *gat4601* Cassette Probes and *Bgl* II Digest



Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T4 generation) and of control soybean (Jack) was digested with *Bgl* II and hybridized to the probes from the *gat4601* cassette. Probes used are indicated below each panel. Approximately 2 µg of genomic DNA was digested and loaded per lane. Lane 6 included plasmid PHP20163 at approximately one gene copy and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Lane	*Sample*
1	DIG VII
2	356043 soybean/ plant 7
3	356043 soybean/ plant 8
4	356043 soybean/ plant 9
5	control (Jack)
6	control (Jack) + 1 copy of PHP20163
7	DIG VII

Figure 9. Southern Blot Analysis: *gm-hra* Cassette Probes and *Bgl* II Digest

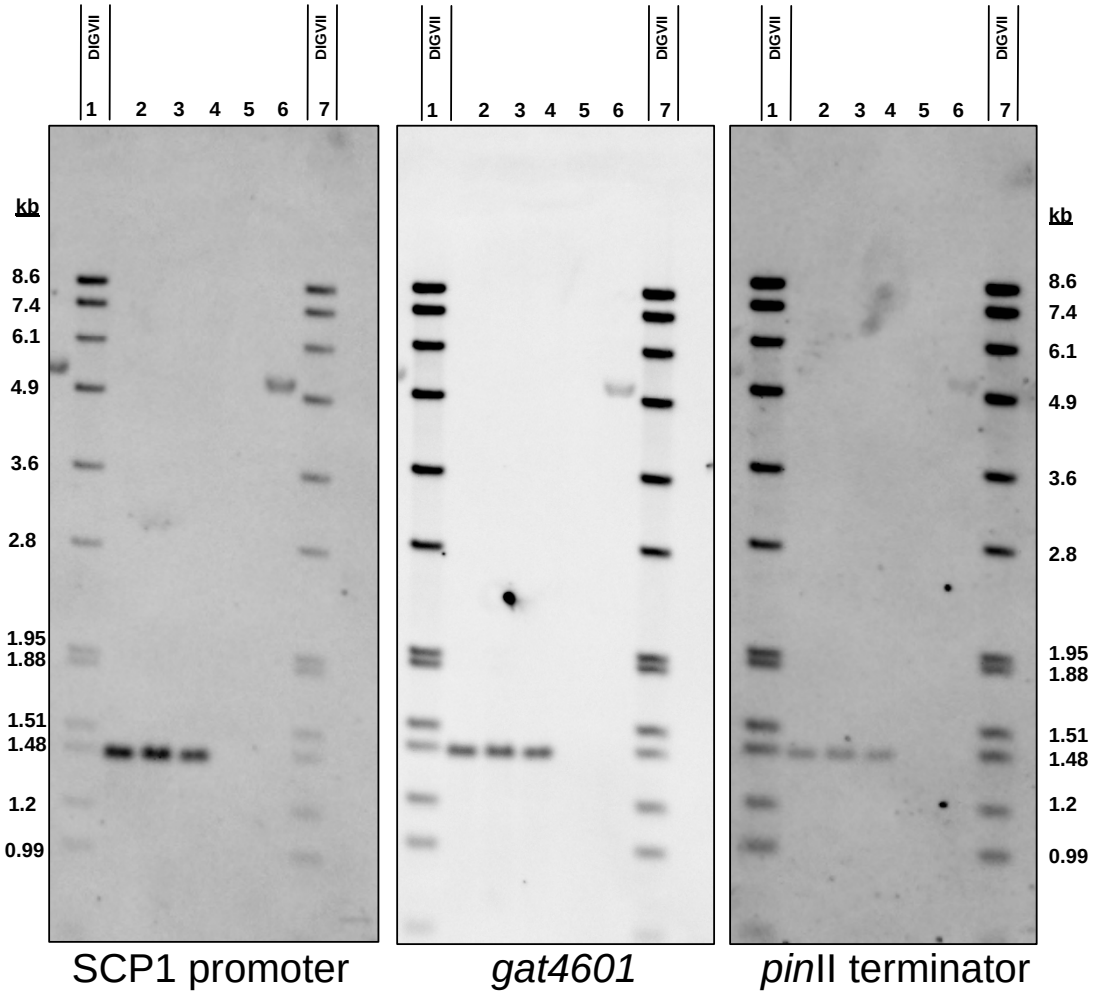


Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T4 generation) and of control soybean (Jack) was digested with *Bgl* II and hybridized to the probes from the gm-hra cassette. Probes used are indicated below each panel. Approximately 2 µg of genomic DNA was digested and loaded per lane. Lane 6 included plasmid PHP20163 at approximately one gene copy and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Note: Hybridizing bands due to endogenous sequences are visible in both 356043 soybean and control soybean lanes. Certain endogenous bands may be difficult to discern on a printed copy but are visible on the original film and electronic version.

Lane	*Sample*
1	DIG VII
2	356043 soybean/ plant 7
3	356043 soybean/ plant 8
4	356043 soybean/ plant 9
5	control (Jack)
6	control (Jack) + 1 copy of PHP20163
7	DIG VII

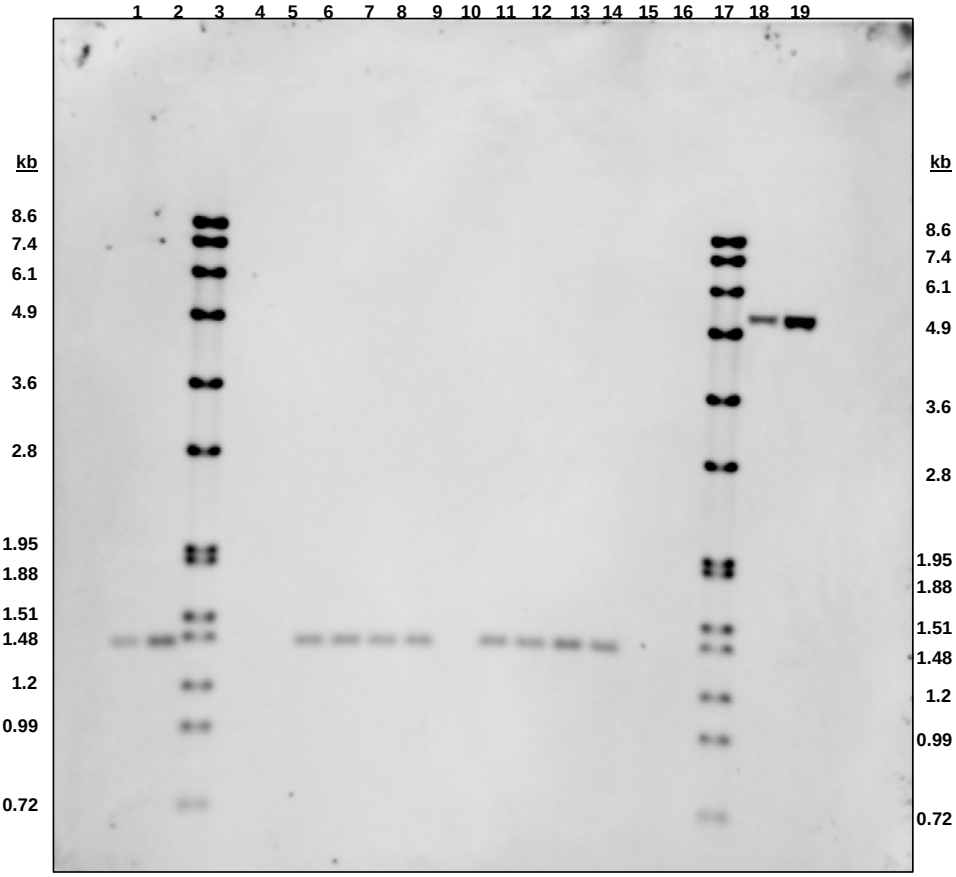
Figure 10. Southern Blot Analysis: *gat4601* Cassette Probes and *Xba* I Digest



Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T4 generation) and of control soybean (Jack) was digested with *Xba* I and hybridized to the probes from the *gat4601* cassette. Probes used are indicated below each panel. Approximately 2 µg of genomic DNA was digested and loaded per lane. Lane 6 included plasmid PHP20163 prepared from a Dam^+ strain at approximately one gene copy and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Lane	*Sample*
1	DIG VII
2	356043 soybean/ plant 7
3	356043 soybean/ plant 8
4	356043 soybean/ plant 9
5	control (Jack)
6	control (Jack) + 1 copy of PHP20163 (Dam^+)
7	DIG VII

Figure 11. Southern Blot Analysis of T4 and T5 Generations: *gat4601* Probe and *Xba* I Digest

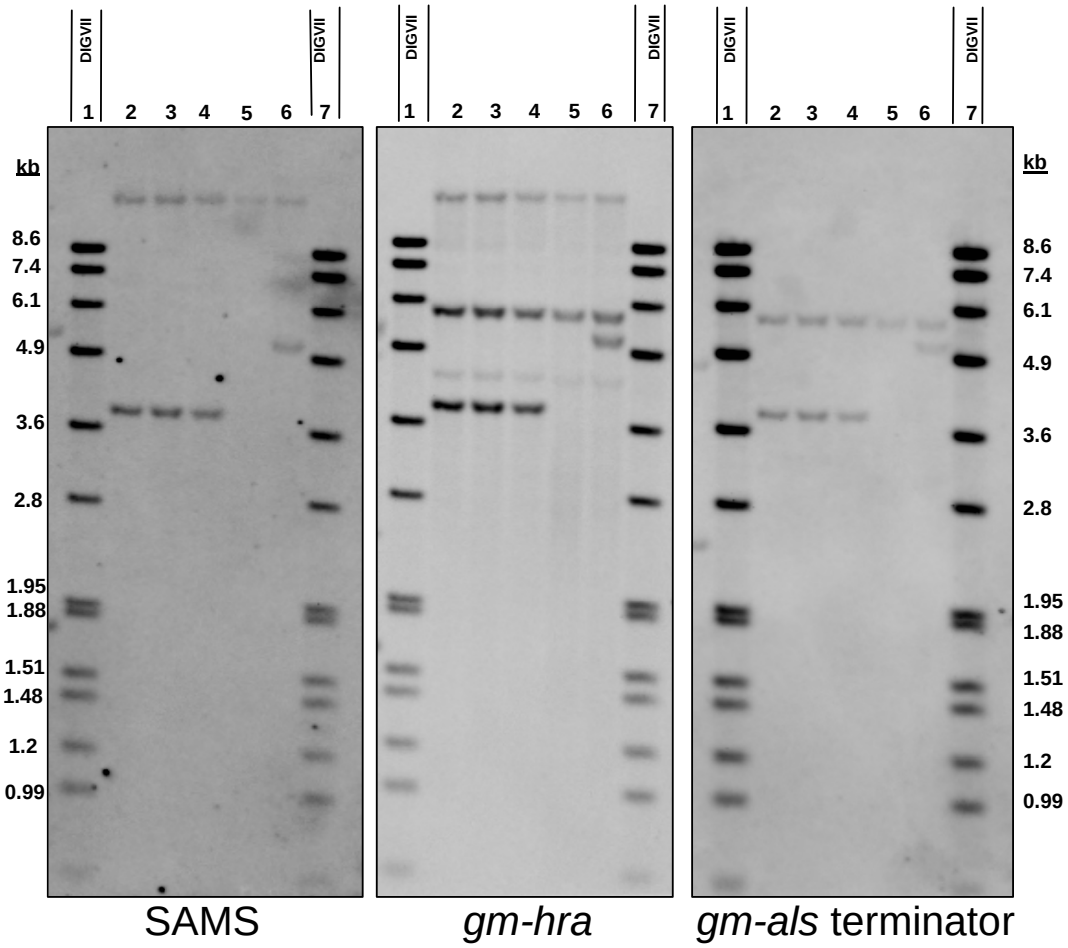


Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T5 and T4 generations) and of control soybean (Jack) was digested with *Xba* I and probed with the *gat4601* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 prepared from a Dam^- strain (lanes 1 and 2) or prepared from a Dam^+ strain (lanes 18 and 19) at the indicated approximate gene copy number and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Lane	Sample
1	Control (Jack) + 1 copy of PHP20163 (Dam^-)
2	Control (Jack) + 3 copies of PHP20163 (Dam^-)
3	DIG VII
4	blank
5	Control (Jack)
6	356043 soybean/ plant 1 (T5 generation)
7	356043 soybean/ plant 3 (T5 generation)
8	356043 soybean/ plant 5 (T5 generation)
9	356043 soybean/ plant 8 (T5 generation)
10	Blank

Lane	Sample
11	356043 soybean/ plant 16 (T4 generation)
12	356043 soybean/ plant 17 (T4 generation)
13	356043 soybean/ plant 18 (T4 generation)
14	356043 soybean/ plant 19 (T4 generation)
15	Control (Jack)
16	Blank
17	DIG VII
18	Control (Jack) + 1 copy of PHP20163 (Dam^+)
19	Control (Jack) + 3 copies of PHP20163 (Dam^+)

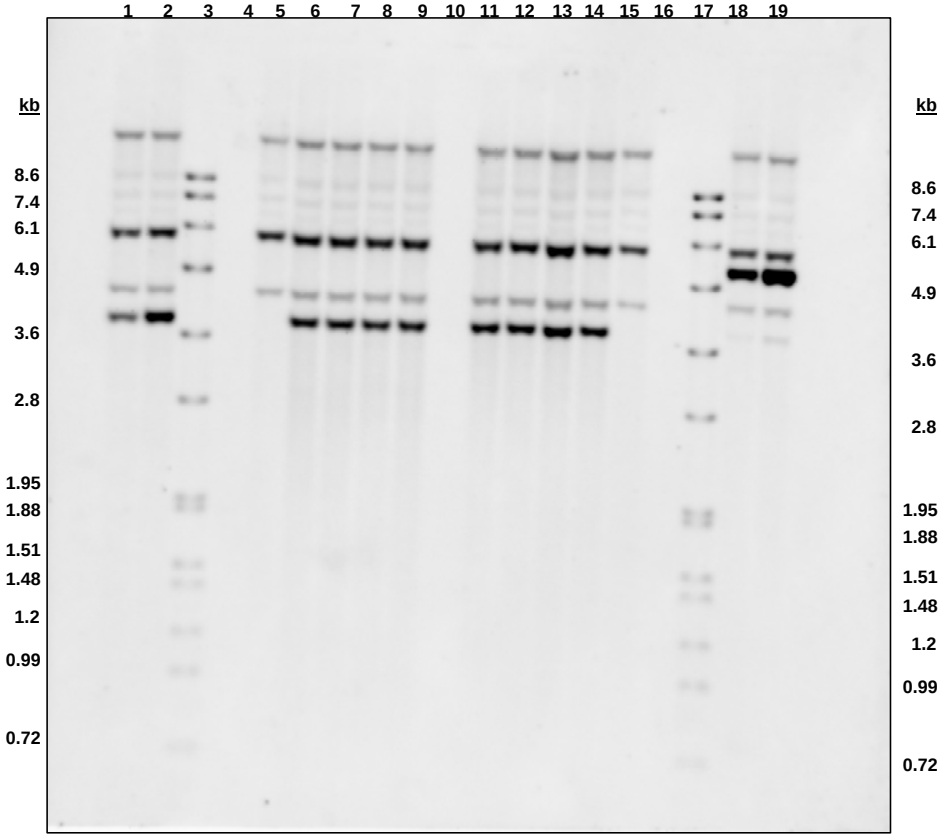
Figure 12. Southern Blot Analysis: *gm-hra* Cassette Probes and *Xba* I Digest



Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T4 generation) and of control soybean (Jack) was digested with *Xba* I and hybridized to the probes from the *gm-hra* cassette. Probes used are indicated below each panel. Approximately 2 µg of genomic DNA was digested and loaded per lane. Lane 6 included plasmid PHP20163 prepared from a Dam^+ strain at approximately one gene copy and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).
Note: Hybridizing bands due to endogenous sequences are visible in both 356043 soybean and control soybean lanes. Certain endogenous bands may be difficult to discern on a printed copy but are visible on the original film and electronic version.

Lane	*Sample*
1	DIG VII
2	356043 soybean/ plant 7
3	356043 soybean/ plant 8
4	356043 soybean/ plant 9
5	control (Jack)
6	control (Jack) + 1 copy of PHP20163 (Dam^+)
7	DIG VII

Figure 13. Southern Blot Analysis of T4 and T5 Generations: *gm-hra* Probe and *Xba* I Digest

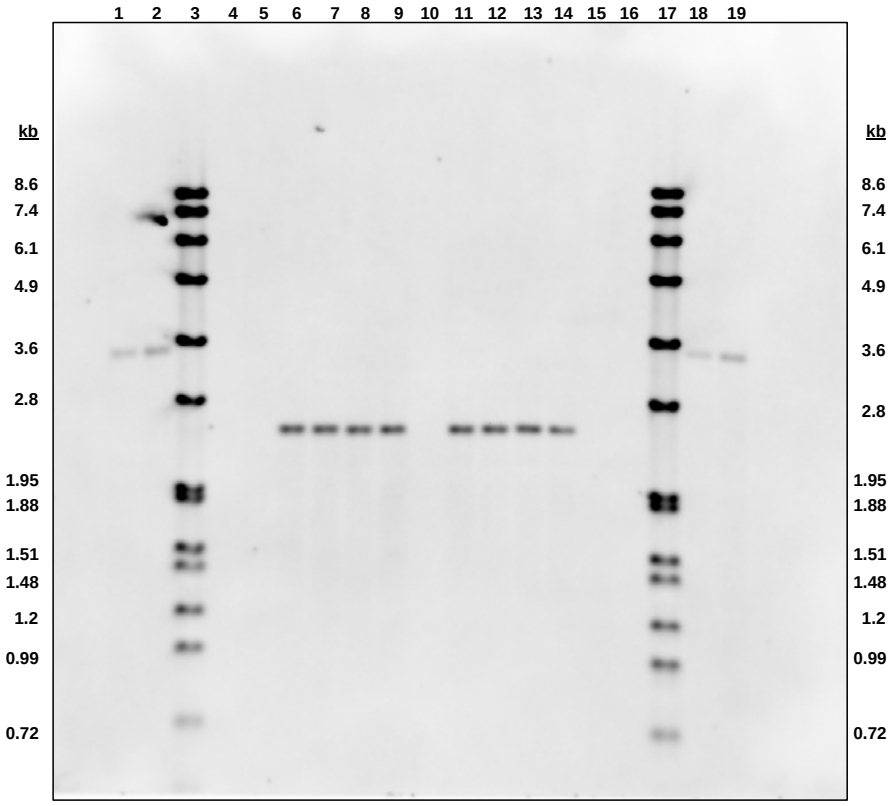


Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T5 and T4 generation) and of control soybean (Jack) was digested with *Xba* I and probed with the *gm-hra* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 prepared from a Dam^- strain (lanes 1 and 2) or prepared from a Dam^+ strain (lanes 18 and 19) at the indicated approximate gene copy number and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).
Note: Hybridizing bands due to endogenous sequences are visible in both 356043 soybean and control soybean lanes. Certain endogenous bands may be difficult to discern on a printed copy but are visible on the original film and electronic version.

Lane	Sample
1	Control (Jack) + 1 copy of PHP20163 (Dam^-)
2	Control (Jack) + 3 copies of PHP20163 (Dam^-)
3	DIG VII
4	blank
5	Control (Jack)
6	356043 soybean/ plant 1 (T5 generation)
7	356043 soybean/ plant 3 (T5 generation)
8	356043 soybean/ plant 5 (T5 generation)
9	356043 soybean/ plant 8 (T5 generation)
10	Blank

Lane	Sample
11	356043 soybean/ plant 16 (T4 generation)
12	356043 soybean/ plant 17 (T4 generation)
13	356043 soybean/ plant 18 (T4 generation)
14	356043 soybean/ plant 19 (T4 generation)
15	Control (Jack)
16	Blank
17	DIG VII
18	Control (Jack) + 1 copy of PHP20163 (Dam^+)
19	Control (Jack) + 3 copies of PHP20163 (Dam^+)

Figure 14. Southern Blot Analysis of T4 and T5 Generations: *gat4601* Probe and *Bgl* II Digest

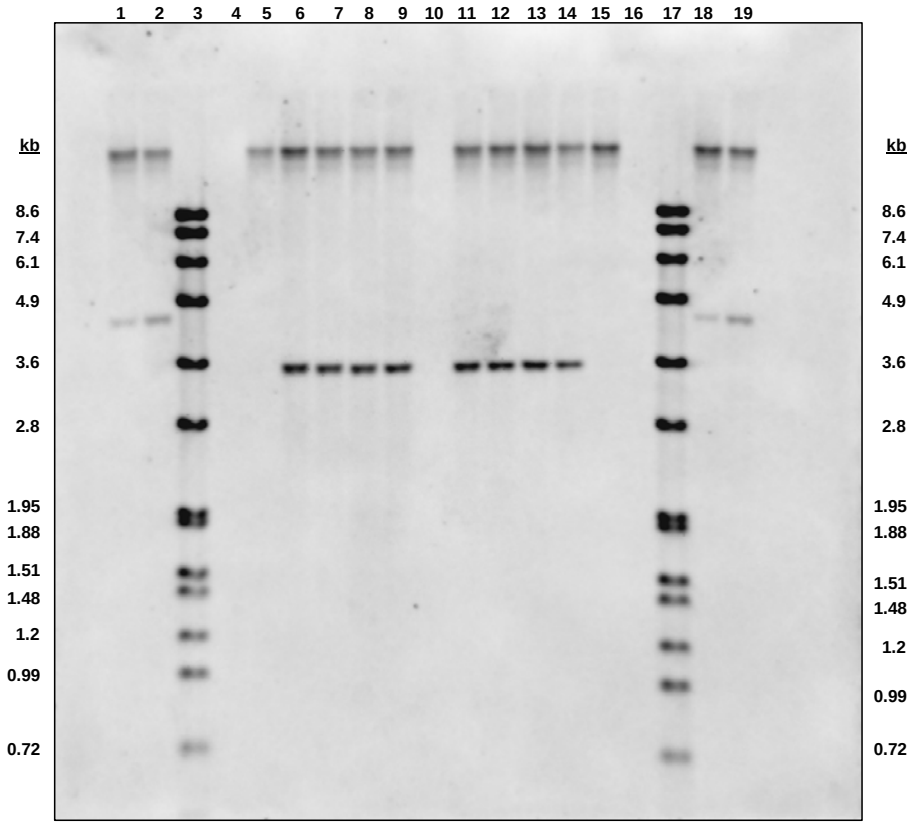


Genomic DNA isolated from leaf tissue of individual plants of 356043 soybean (T5 and T4 generation) and of control soybean (Jack) was digested with *Bgl* II and probed with the *gat4601* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 at the indicated approximate gene copy number and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Lane	Sample
1	Control (Jack) + 1 copy of PHP20163
2	Control (Jack) + 3 copies of PHP20163
3	DIG VII
4	blank
5	Control (Jack)
6	356043 soybean/ plant 1 (T5 generation)
7	356043 soybean/ plant 3 (T5 generation)
8	356043 soybean/ plant 5 (T5 generation)
9	356043 soybean/ plant 8 (T5 generation)
10	Blank

Lane	Sample
11	356043 soybean/ plant 16 (T4 generation)
12	356043 soybean/ plant 17 (T4 generation)
13	356043 soybean/ plant 18 (T4 generation)
14	356043 soybean/ plant 19 (T4 generation)
15	Control (Jack)
16	Blank
17	DIG VII
18	Control (Jack) + 1 copy of PHP20163
19	Control (Jack) + 3 copies of PHP20163

Figure 15. Southern Blot Analysis of T4 and T5 Generations: *gm-hra* Probe and *Bgl* II Digest

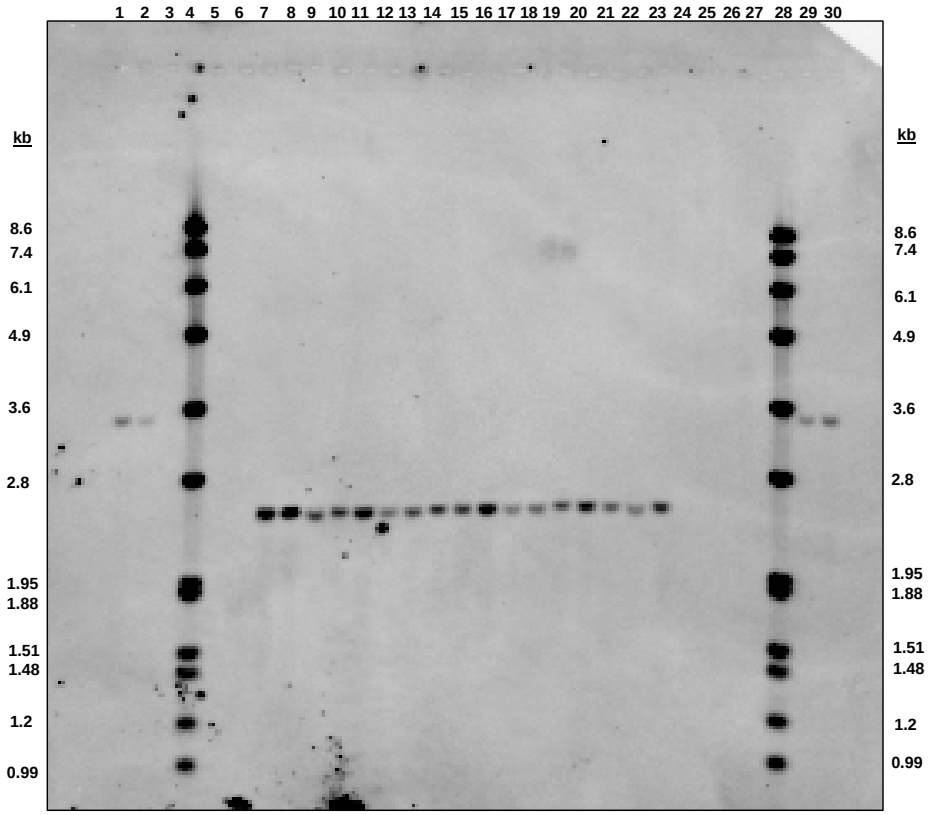


Genomic DNA isolated from leaf tissue of individual plants of 356043 soybean (T5 and T4 generation) and of control soybean (Jack) was digested with *Bgl* II and probed with the *gm-hra* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 at the indicated approximate gene copy number and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).
Note: Hybridizing bands due to endogenous sequences are visible in both 356043 soybean and control soybean lanes. Differing from those reported in Table 4, three endogenous bands are visible in all genomic samples at >8.6 kb due to slight differences in gel running, blotting and blot hybridization conditions. Certain endogenous bands may be difficult to discern on a printed copy but are visible on the original film and electronic version.

Lane	Sample
1	Control (Jack) + 1 copy of PHP20163
2	Control (Jack) + 3 copies of PHP20163
3	DIG VII
4	blank
5	Control (Jack)
6	356043 soybean/ plant 1 (T5 generation)
7	356043 soybean/ plant 3 (T5 generation)
8	356043 soybean/ plant 5 (T5 generation)
9	356043 soybean/ plant 8 (T5 generation)
10	Blank

Lane	Sample
11	356043 soybean/ plant 16 (T4 generation)
12	356043 soybean/ plant 17 (T4 generation)
13	356043 soybean/ plant 18 (T4 generation)
14	356043 soybean/ plant 19 (T4 generation)
15	Control (Jack)
16	Blank
17	DIG VII
18	Control (Jack) + 1 copy of PHP20163
19	Control (Jack) + 3 copies of PHP20163

Figure 16. Southern Blot Analysis of F3 Generation: *gat4601* Probe and *Bgl* II Digest

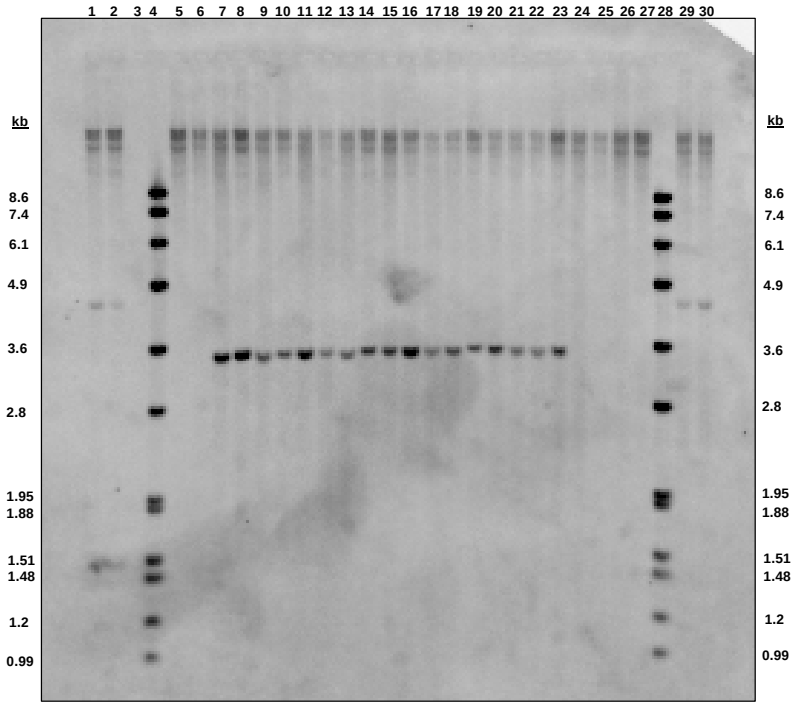


Genomic DNA isolated from leaf tissue of individual plants of 356043 soybean (F3 generation) and of control soybean (Jack, Elite 1) was digested with *Bgl* II and probed with the *gat4601* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 at the indicated approximate gene copy number and 2 µg of control soybean DNA (Jack or Elite 1). Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Lane	Sample
1	Control (Jack) + 3 copies PHP20163
2	Control (Jack) + 1 copy PHP20163
3	
4	DIG VII
5	Control (Jack)
6	Control (Elite 1)
7	356043 soybean / plant 24
8	356043 soybean / plant 25
9	356043 soybean / plant 26
10	356043 soybean / plant 27
11	356043 soybean / plant 28
12	356043 soybean / plant 30
13	356043 soybean / plant 31
14	356043 soybean / plant 33
15	356043 soybean / plant 34

Lane	Sample
16	356043 soybean / plant 35
17	356043 soybean / plant 36
18	356043 soybean / plant 37
19	356043 soybean / plant 38
20	356043 soybean / plant 39
21	356043 soybean / plant 40
22	356043 soybean / plant 41
23	356043 soybean / plant 18
24	356043 soybean / plant 29 (null)
25	356043 soybean / plant 32 (null)
26	Control (Elite 1)
27	Control (Jack)
28	DIG VII
29	Control (Elite 1) + 1 copy PHP20163
30	Control (Elite 1) + 3 copies PHP20163

Figure 17. Southern Blot Analysis of F3 Generation: *gm-hra* Probe and *Bgl* II Digest



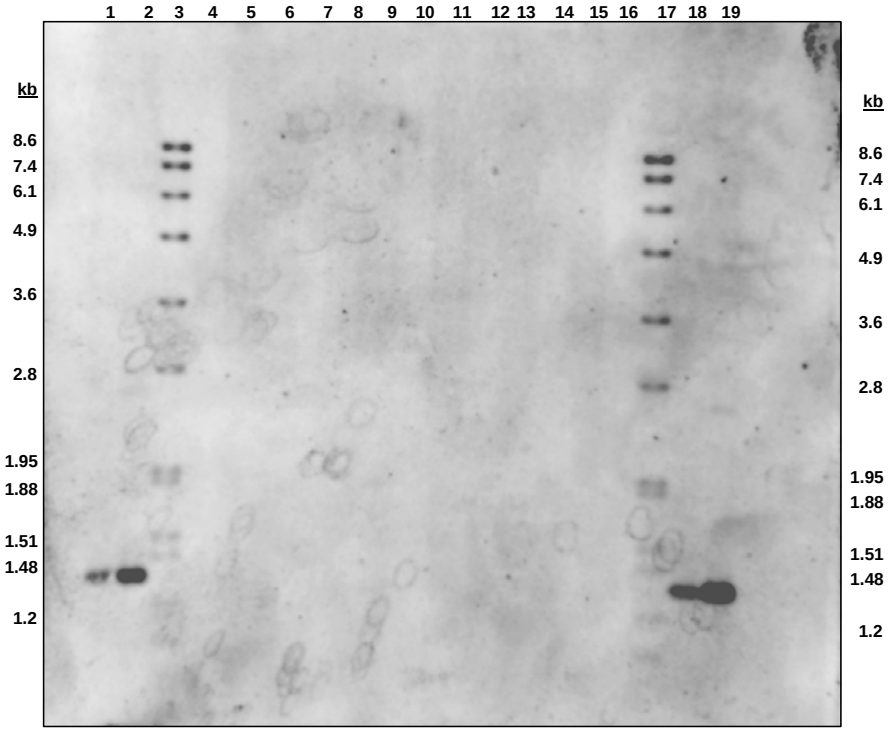
Genomic DNA isolated from soybean leaf tissue of individual plants of 356043 soybean (F3 generation) and of control soybean (Jack, Elite 1) was digested with *Bgl II* and probed with the *gm-hra* gene probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 at the indicated approximate gene copy number and 2 µg of control soybean DNA (Jack or Elite 1). Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Note: Hybridizing bands due to endogenous sequences are visible in both 356043 soybean and control soybean lanes. Differing from those reported in Table 3, four endogenous bands are visible in all samples at >8.6 kb due to slight differences in gel running, blotting and blot hybridization conditions. Certain endogenous bands may be difficult to discern on a printed copy but are visible on the original film and electronic version.

Lane	Sample
1	Control (Jack) + 3 copies PHP20163
2	Control (Jack) + 1 copy PHP20163
3	
4	DIG VII
5	Control (Jack)
6	Control (Elite 1)
7	356043 soybean / plant 24
8	356043 soybean / plant 25
9	356043 soybean / plant 26
10	356043 soybean / plant 27
11	356043 soybean / plant 28
12	356043 soybean / plant 30
13	356043 soybean / plant 31
14	356043 soybean / plant 33
15	356043 soybean / plant 34

Lane	Sample
16	356043 soybean / plant 35
17	356043 soybean / plant 36
18	356043 soybean / plant 37
19	356043 soybean / plant 38
20	356043 soybean / plant 39
21	356043 soybean / plant 40
22	356043 soybean / plant 41
23	356043 soybean / plant 18
24	356043 soybean / plant 29 (null)
25	356043 soybean / plant 32 (null)
26	Control (Elite 1)
27	Control (Jack)
28	DIG VII
29	Control (Elite 1) + 1 copy PHP20163
30	Control (Elite 1) + 3 copies PHP20163

Figure 18. Southern Blot Analysis: *backbone 20163* Probe and *Xba* I Digest

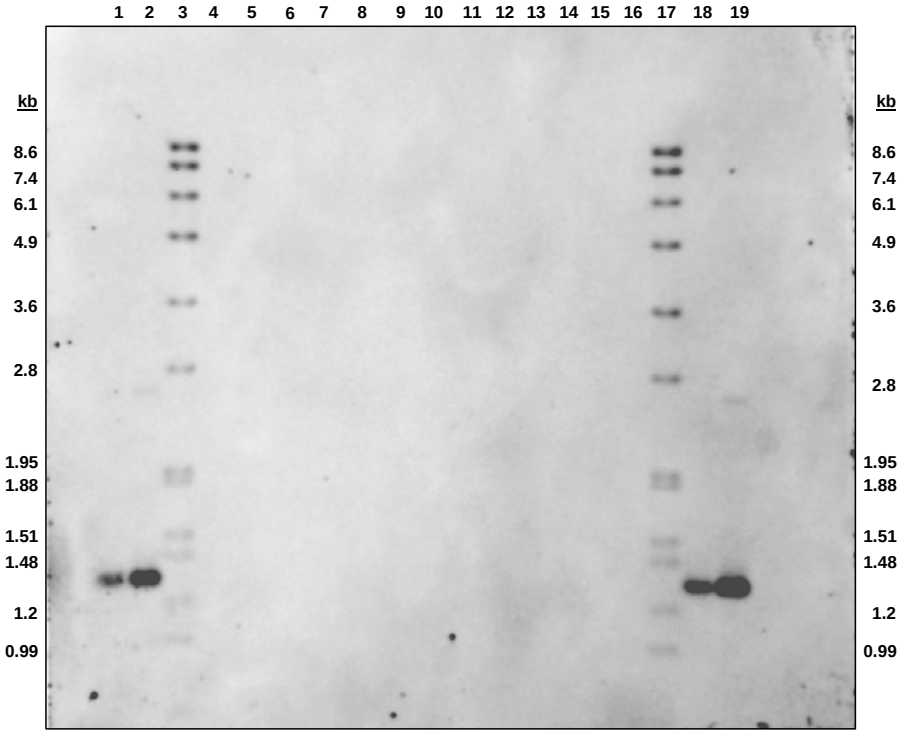


Genomic DNA isolated from leaf tissue of individual plants of 356043 soybean (T5 and T4 generation) and of control soybean (Jack) was digested with *Xba* I and probed with the *backbone 20163* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 at the indicated approximate gene copy number and 2 µg of control (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb). Note: Circulate spots visible on blot image are background hybridization due to the detection process.

Lane	Sample
1	Control (Jack) + 1 copy of PHP20163
2	Control (Jack) + 3 copies of PHP20163
3	DIG VII
4	blank
5	Control (Jack)
6	356043 soybean/ plant 1 (T5 generation)
7	356043 soybean/ plant 3 (T5 generation)
8	356043 soybean/ plant 5 (T5 generation)
9	356043 soybean/ plant 8 (T5 generation)
10	Blank

Lane	Sample
11	356043 soybean/ plant 16 (T4 generation)
12	356043 soybean/ plant 17 (T4 generation)
13	356043 soybean/ plant 18 (T4 generation)
14	356043 soybean/ plant 19 (T4 generation)
15	Control (Jack)
16	Blank
17	DIG VII
18	Control (Jack) + 1 copy of PHP20163
19	Control (Jack) + 3 copies of PHP20163

Figure 19. Southern Blot Analysis: *hyg 20163* Probe and *Xba* I Digest



Genomic DNA isolated from leaf tissue from individual plants of 356043 soybean (T5 and T4 generation) and of control soybean (Jack) was digested with *Xba I* and probed with the *hyg 20163* probe. Approximately 2 µg of genomic DNA was digested and loaded per lane. The positive controls included plasmid PHP20163 at the indicated approximate gene copy number and 2 µg of control soybean (Jack) DNA. Sizes of the DIG VII molecular weight markers are indicated adjacent to the blot image in kilobase pairs (kb).

Lane	Sample
1	Control (Jack) + 1 copy of PHP20163
2	Control (Jack) + 3 copies of PHP20163
3	DIG VII
4	blank
5	Control (Jack)
6	356043 soybean/ plant 1 (T5 generation)
7	356043 soybean/ plant 3 (T5 generation)
8	356043 soybean/ plant 5 (T5 generation)
9	356043 soybean/ plant 8 (T5 generation)
10	Blank

Lane	Sample
11	356043 soybean/ plant 16 (T4 generation)
12	356043 soybean/ plant 17 (T4 generation)
13	356043 soybean/ plant 18 (T4 generation)
14	356043 soybean/ plant 19 (T4 generation)
15	Control (Jack)
16	Blank
17	DIG VII
18	Control (Jack) + 1 copy of PHP20163
19	Control (Jack) + 3 copies of PHP20163

V-E. Inheritance of the Traits in 356043 Soybean

Chi square analysis of trait inheritance data from five different generations (T1, F2, F3, BC1F2, C2F2) was performed to determine the Mendelian heritability and stability of the *gat4601* and *gm-hra* genes in various generations of 356043 soybean. The breeding history of the five generations evaluated for Mendelian inheritance is described in Figure 2. For each of the generations tested, the plants were expected to segregate 1:2:1 (homozygous positive plants:hemizygous positive plants:homozygous negative [null] plants). Various approaches, as outlined below, were used to confirm this segregation ratio.

3:1 Positive (Homozygous + Hemizygous):Negative Segregation Ratio (Table 5)

In some studies, homozygous positive plants were not differentiated from hemizygous positive plants, resulting in a 3:1 positive:negative segregation pattern (Table 5).

For the three generations listed in Table 5, one of three methods was used to score the plants as positive or negative:

- qualitative PCR analysis to identify the plants containing the *gat4601* gene (T1 generation); or
- western analysis to score plants for GAT4601 protein expression followed by confirmation of those same plants by Southern analysis with both *gat4601* and *gm-hra* probes (F3 generation), or
- an ALS seed soak assay (BC1F2 generation). In this assay, seeds are soaked in an ALS-inhibiting herbicide containing the active ingredient chlorsulfuron. Only seed expressing the *gm-hra* gene will emerge after planting.

1:2 Homozygous Positive:Hemizygous Positive Segregation Ratio (Table 6)

In certain studies, plants that did not contain either the *gat4601* or *gm-hra* genes were removed from the study prior to the conduct of segregation analysis. Remaining plants were then scored as homozygous positive or hemizygous positive, resulting in a 1:2 homozygous positive:hemizygous positive segregation pattern (Table 6).

For the F2 generation listed in Table 6, two methods were used to remove the negative plants and score the remaining plants as homozygous positive or hemizygous:

- A glyphosate spray was applied after the plants had emerged, removing all of the homozygous negative (null) plants. This was followed by an ALS-inhibiting herbicide "ragdoll" assay for the *gm-hra* gene, where progeny seed from the F2 plants were screened to determine the F2 parent plant genotype. In the ragdoll assay, paper towels were wetted with an ALS-inhibiting herbicide containing the active ingredient chlorsulfuron. Ten progeny seeds from a single F2 parent plant were rolled into the wetted towel and allowed to germinate. An F2 parent plant was scored as homozygous positive for the *gm-hra* gene if all ten progeny seeds germinated and grew normally. A hemizygous positive F2 parent genotype was characterized by one or more of the germinated progeny seeds producing a seeding with unifoliate leaves exhibiting a wrinkled appearance; or
- An ALS seed soak assay of the seeds prior to planting removed all of the homozygous negative (null) plants, followed by a quantitative PCR (qPCR) assay to distinguish plants that were homozygous positive or hemizygous positive for the *gat4601* gene.

1:2:1 Homozygous Positive:Hemizygous:Homozygous Negative Segregation Ratio (Table 7)

In some studies, all plants were identified as homozygous positive, hemizygous positive, or homozygous negative to confirm a 1:2:1 expected segregation ratio (Table 7). Segregation analysis was conducted for the C2F2 generation using qPCR assays for both the *gat* and *gm-hra* genes.

Summary of Inheritance of Traits

Because the *gat4601* and *gm-hra* genes were physically linked in the DNA fragment used for transformation and are expected to have identical segregation ratios in the progeny of 356043 soybean, the *gat4601* results are applicable to the inheritance of *gm-hra* and vice versa. In generations where both traits were analyzed in the same plants (F3 and C2F2), identical segregation data would experimentally confirm co-segregation of the *gat4601* and *gm-hra* genes

Results from the Mendelian analysis are summarized in Tables 5, 6 and 7. Details of the statistical methodology can be found in Appendix 2. All P-values were greater than 0.05, indicating no statistically significant differences between the observed and expected frequencies of the *gat4601* and/or *gm-hra* genes in five generations of 356043 soybean.

The results of this analysis are consistent with the finding of a single locus of insertion of the *gat* and *gm-hra* genes that segregates in 356043 soybean progeny according to Mendel's laws of genetics. The stability of the insert has been demonstrated in five generations of self- and cross-pollinations.

Table 5. Comparison of Observed and Expected 3:1 Segregation Ratios for 356043 Soybean

Generation	Method	Observed		Expected		Chi-Square Test
		Positives +/+ or +/-	Negatives -/-	Positives +/+ or +/-	Negatives -/-	P-value
T1	*gat* PCR	59	23	61.5	20.5	0.610
F3 Elite 1 background	GAT westerns followed by Southern analyses with *gat4601* and *gm-hra* probes	75	15	67.5	22.5	0.088
BC1F2	ALS seed soak					
Elite 7 background		700	222	691.5	230.5	0.543
Elite 8 background		761	273	775.5	258.5	0.315
Elite 9 background		160	54	160.5	53.5	1.000
Elite 10 background		205	79	213	71	0.304

Table 6. Comparison of Observed and Expected 1:2 Segregation Ratios for 356043 Soybean

Generation	Method	Observed		Expected		Chi-Square Test
		Homo-zygous +/+	Hemi-zygous +/-	Homo-zygous +/+	Hemi-zygous +/-	P-value
F2						
Elite 1 background	Glyphosate spray to remove nulls and followed by ALS inhibitor ragdoll test	16	24	13.3	26.7	0.467
Elite 2 background		32	53	28.3	56.7	0.466
F2						
Elite 3 background	ALS seed soak to remove nulls followed by qPCR for *gat*	110	182	97.3	194.7	0.131
Elite 4 background		124	284	136	272	0.227
Elite 5 background		27	61	29.3	58.7	0.678
Elite 6 background		22	29	17	34	0.181

Table 7. Comparison of Observed and Expected 1:2:1 Segregation Ratios for 356043 Soybean

Generation	Method	Observed			Expected			Chi-Square Test
		Homo-zygous +/+	Hemi-zygous +/-	Homo-zygous -/-	Homo-zygous +/+	Hemi-zygous +/-	Homo-zygous -/-	P-value
C2F2								
Elite 44 background	*gat* and *gm-hra* qPCR	41	76	43	40	80	40	0.799
Elite 45 background		160	294	142	149	298	149	0.550

V-F. Summary and Conclusions

Southern blot analysis was conducted to characterize the transgene insertion in 356043 soybean. The analysis confirmed that a single, intact PHP20163A fragment was inserted into the soybean genome to produce 356043 soybean. A single copy of each of the genetic elements of the *gat4601* and the *gm-hra* expression cassettes was present and the integrity of the PHP20163A fragment was maintained. The analysis confirmed the stability of the insertion in 356043 soybean across the T4, T5, and F3 generations, thus confirming stability of inheritance during traditional breeding procedures. Southern blot analysis verified the absence of plasmid backbone sequences in 356043 soybean.

Inheritance studies confirmed that the insert segregated in normal Mendelian fashion. None of the P-values obtained in the studies indicated a statistically significant difference between observed and expected segregation ratios for the *gat4601* and/or *gm-hra* genes over five different plant generations. The results are consistent with the molecular characterization data, which indicates stable integration of the *gat4601* and *gm-hra* transgenes at a single site in the plant genome.

VI. Characterization of the Introduced GAT4601 and GM-HRA Proteins

VI-A. The GAT4601 Protein

A1. Identity of the GAT4601 Protein

GAT4601 is a protein based on N-acetyltransferase protein sequences from *Bacillus licheniformis*, a gram positive saprophytic bacterium that is widespread in nature and thought to contribute substantially to nutrient cycling due to the diversity of enzymes produced by members of its species. GAT4601 is 146 amino acids in length and has an approximate molecular weight of 17 kDa (Figure 20).

Figure 20. Deduced Amino Acid Sequence of the GAT4601 Protein

```
  1    MIEVKPINAE DTYELRHRIL RPNQPIEACM FESDLLRGAF HLGGFYRGKL
 51    ISIASFHQAE HSELQGQKQY QLRGMATLEG YREQKAGSTL VKHAEEILRK
101    RGADMLWCNA RTSASGYYKK LGFSEQGEIF DTPPVGPHIL MYKRIT
```

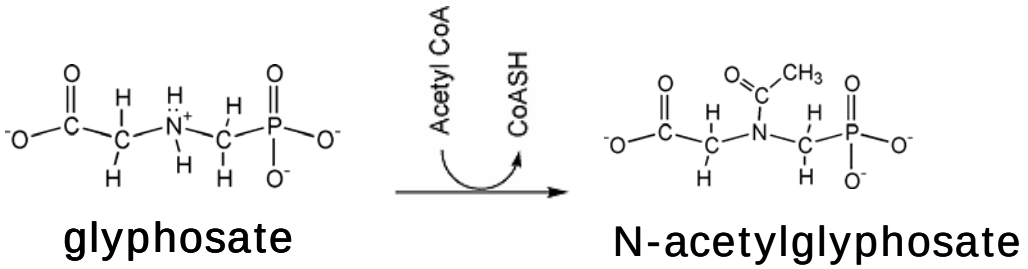
A2. Characteristics of GAT Proteins

Glyphosate acetyltransferase (GAT) proteins are members of the GCN 5-related family of N-acetyltransferases (also known as the GNAT superfamily). The GNAT superfamily is one of the largest enzyme superfamilies recognized to date with over 10,000 representatives from plants, animals and microbes. Members of the GNAT superfamily all contain a highly conserved GNAT motif but have high sequence diversity (Vetting *et al.,* 2005). GNAT proteins are known to have a number of metabolic functions including detoxification (Dyda *et al*., 2000). In particular, and as described below, GAT proteins can confer tolerance to the broad spectrum herbicide glyphosate.

Glyphosate has proven to be a very popular herbicide because of its effectiveness and safety. According to USDA, in 2005 glyphosate tolerant soybeans accounted for 87% of all soybeans planted in the United States (http://www.ers.usda.gov/Data/BiotechCrops/ExtentofAdoptionTable3.htm). Post emergence application of glyphosate effectively controls grasses and annual broadleaf weeds. Glyphosate inhibits the enzyme enolpyruvylshikimate-3-phosphate synthase (EPSPS) in the plant chloroplast-localized pathway that leads to the biosynthesis of aromatic amino acids. Some microbial EPSPS enzyme variants are insensitive to glyphosate inhibition. When expressed in chloroplasts of transgenic plants, the insensitive enzymes confer tolerance to glyphosate. This is the basis of current commercial glyphosate tolerant crops.

GAT proteins provide an alternative mechanism of resistance to glyphosate (Figure 21). They detoxify glyphosate to the non-phytotoxic form N-acetylglyphosate. This detoxification mechanism is similar to that of the phosphinothricin acetyltransferase (PAT or BAR) enzymes from *Streptomyces*, which detoxify glufosinate ammonium herbicides by adding an acetyl group (De Block *et al.,* 1987) and have been deregulated by USDA (Van Wert, 1994). As shown in Figure 21, GAT proteins acetylate the secondary amine of glyphosate using acetyl coenzyme A as an acetyl donor (Castle *et al.,* 2004). Although GAT4601, PAT and BAR enzymes are N-acetyltransferase proteins, expression of the *gat4601* gene in 356043 does not confer tolerance to glufosinate ammonium herbicides (Pioneer internal data). Transgenic expression of GAT proteins was shown to confer glyphosate tolerance in several plant species (Castle *et al.,* 2004). This strategy was used to develop glyphosate tolerant 356043 soybean.

Figure 21. Enzymatic Activity of GAT Proteins



A3. Generation of GAT4601

In order to develop a GAT protein that would confer commercial levels of tolerance to the herbicide glyphosate when expressed in plants, an in-house collection of several hundred bacterial isolates was screened. Because *Bacillus* species produce a wide variety of enzymes involved in secondary metabolism, we focused our search on bacilli isolated from non-extreme environments. A mass spectrometry method was developed to detect low levels of N-acetylglyphosate, the non-phytotoxic end product of glyphosate acetylation. The *Bacillus* isolates were grown to stationary phase, permeabilized, and incubated with glyphosate and acetyl coenzyme A, and the supernatants were screened for the presence of N-acetylglyphosate by spectrometry. Several strains of *Bacillus licheniformis* exhibited GAT activity and had the greatest reproducible accumulation of N-acetylglyphosate (http://www.isb.vt.edu/articles/sep0403.htm and Castle *et al*., 2004).

To isolate the gene encoding GAT, recombinant *E. coli* expressing genomic DNA fragments from *B. licheniformis* were assayed by the mass spectrometry method. DNA sequences of multiple genomic fragments specifying GAT activity from *B. licheniformis* strain B6 and *B. licheniformis* strain 401 (purchased from ATCC—catalog number 14580) were determined, and the corresponding genes were cloned. A polymerase chain reaction (PCR) survey of *B. licheniformis* strains revealed a third gene variant in isolate DS3. However, each of the GAT proteins identified demonstrated weak activity on glyphosate, and *gat* expression in transgenic tobacco and *Arabidopsis* did not confer glyphosate tolerance. Therefore the B6, 401 and DS3 *gat* genes were used as parents for fragmentation-based multigene shuffling to create enzymes with improved activity on the synthetic substrate glyphosate.

DNA shuffling is a process that recombines genetic diversity from parental genes to create libraries of gene variants that are then screened to identify those progeny with improved properties (Stemmer, 1994; Crameri *et al*., 1998). This process of fragmentation and recombination followed by selection can be repeated using those progeny with improved properties as parents for the next round of shuffling. In the case of the *gat4601* gene, this process was repeated seven times using a combination of multi-gene shuffling and the introduction of genetic diversity via PCR.

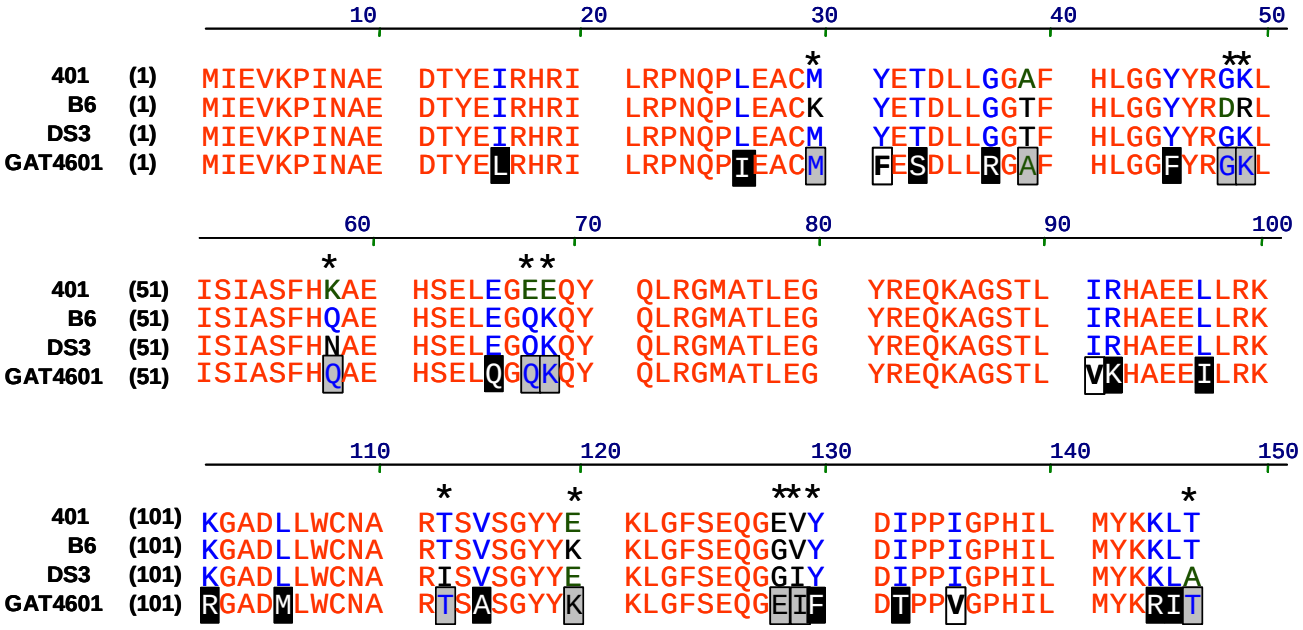
To initiate the first round of gene shuffling, the three native *B. licheniformis gat* genes were used as parental templates. The initial diversity represented in the three native GAT protein sequences was 12 out of 146 total amino acids (Figure 22). Libraries of shuffled gene variants were created, expressed in *E. coli*, and screened. Shuffled variants that resulted in the accumulation of more N-acetylglyphosate than the parental controls were selected for further rounds of shuffling. In each round of DNA shuffling, approximately 5,000 gene variants were screened and 24-48 purified enzymes were analyzed to determine their kinetic properties. Typically, three to twelve improved variants exhibiting a high k_{cat}, a low K_M, or a high k_{cat}/K_M ratio were chosen to be the parents for the next round.

In enzyme kinetics, k_{cat} is a measure of the turnover rate or speed of the reaction. The higher the k_{cat}, the faster the enzyme reaction. K_M is the affinity of the enzyme for a substrate or tightness of binding of the substrate to the enzyme. The lower the K_M, the greater the affinity of the enzyme for the substrate. A k_{cat}/K_M ratio is the common way to express the catalytic efficiency of the enzyme. The greater the k_{cat}/K_M ratio, the greater the catalytic efficiency of the enzyme for a given substrate.

Several GAT enzymes from the third round of gene shuffling had rate constants (k_{cat}/K_M ratios) about 100-fold improved over the original enzymes, but the gene variants were unable to confer glyphosate tolerance to transgenic plants. At the fifth round of shuffling, two advances were made, 1) the additional introduction of diversity by PCR incorporation of oligonucleotides based on related DNA sequences from *Bacillus cereus* and *Bacillus subtilis* during the fragment reassembly step that allowed for substitutions at 21 amino acid residues, and 2) a functional pre-screen based on resistance of GAT-expressing *E. coli* to glyphosate (Castle *et al*., 2004). At the sixth round of shuffling, a spectrophotometric assay was used that allowed the variants to be screened for relative k_{cat} and K_M values without protein purification.

At the end of the seventh round of gene shuffling, the rate constant of the best round 7 variant analyzed was approximately 2400-fold improved over the native enzymes (Siehl *et al*., 2005). The average k_{cat}/K_M of the parental enzymes was 4.2 min^{-1} mM^{-1}, and the average of the best round 7 variant was approximately 10,340 min^{-1} mM^{-1} (Figure 23). Genes from this round provided robust glyphosate tolerance when expressed in several plant species. It was from this round that the *gat4601* gene encoding the GAT4601 protein was identified. The GAT4601 protein is 84% homologous at the amino acid level to each of the three native GAT enzymes from which it was derived, compared to 94% homology of each of the native enzymes to each other (Table 8). A sequence pedigree showing the accumulation of changes attributed to either the three native *gat* genes from *Bacillus licheniformis* or the incorporation of diversity introduced by PCR based on related DNA sequences from *B. cereus* and *B. subtilis* in the GAT4601 protein is provided in Figure 22.

Figure 22. Comparison of Amino Acid Sequence Between GAT4601 and Parental GAT Proteins



GAT4601 pedigree showing diversity introduced from the native *gat* genes or by PCR using sequence information from *B. subtilis* and *B. cereus.*

■ Amino acids introduced from hypothetical proteins produc ed by relatedDNA sequences from *B.subtilis* or *B.cereus* in 5th round of shuffling.

X Amino acids introduced from random mutations during shuf fling (F31 6th round, V91 6th round, V135 2nd round).

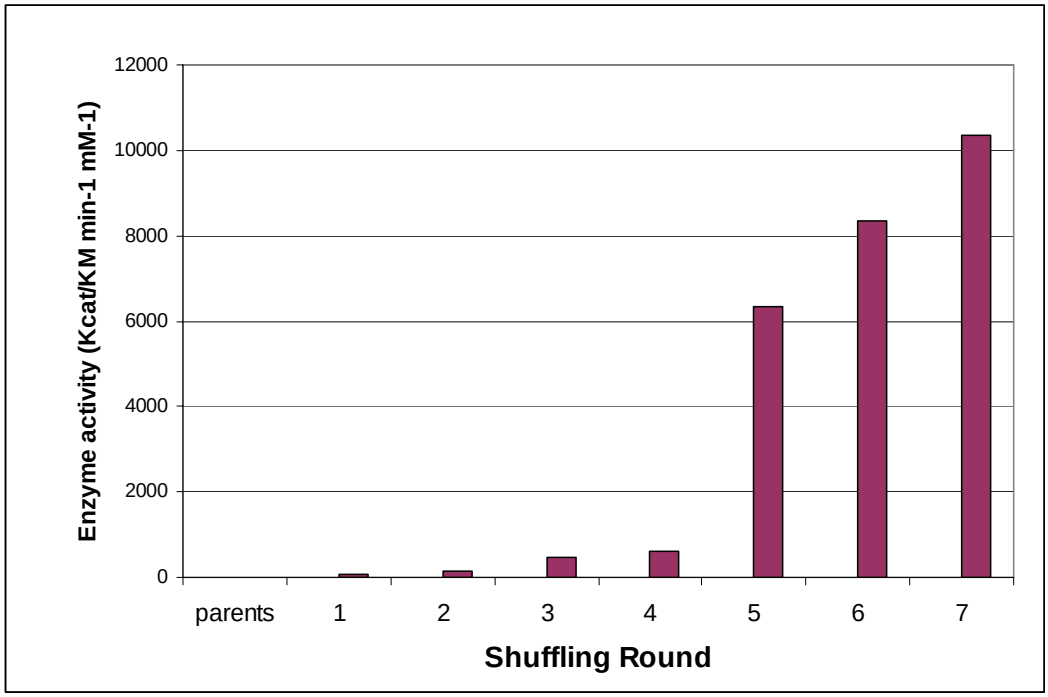
▒ Diversity of GAT 4601 from at least one of original pa rents

* Shows the initial diversity represented in the three na tive GAT proteins

Table 8. Comparison of Homology Between Parental GAT Proteins and GAT4601

	GAT from strain 401	GAT from strain B6	GAT from strain DS3	GAT4601
GAT from strain 401	100%	94%	94%	84%
GAT from strain B6		100%	94%	84%
GAT from strain DS3			100%	84%
GAT4601				100%

Figure 23. Kinetic Improvement of GAT Enzymes Through Shuffling



Taken from Siehl *et al*., 2005.

VI-B. The GM-HRA Protein

The GM-HRA protein is a modified version of the native acetolactate synthase (ALS) protein from soybean. When expressed in plants, GM-HRA confers tolerance to ALS-inhibiting herbicides. The herbicide tolerant *gm-hra* gene was made by isolating the herbicide sensitive soybean *gm-als* gene and introducing two specific amino acid changes in the mature protein.

B1. Characteristics of ALS Proteins

The native ALS, also known as acetohydroxyacid synthase (AHAS)[2], is a key enzyme that catalyzes the first common step in the biosynthesis of the essential branched-chain amino acids isoleucine, leucine, and valine (LaRossa and Schloss, 1984; LaRossa and Falco, 1984; Duggleby and Pang, 2000; Coruzzi and Last, 2000). 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 (Figure 24).

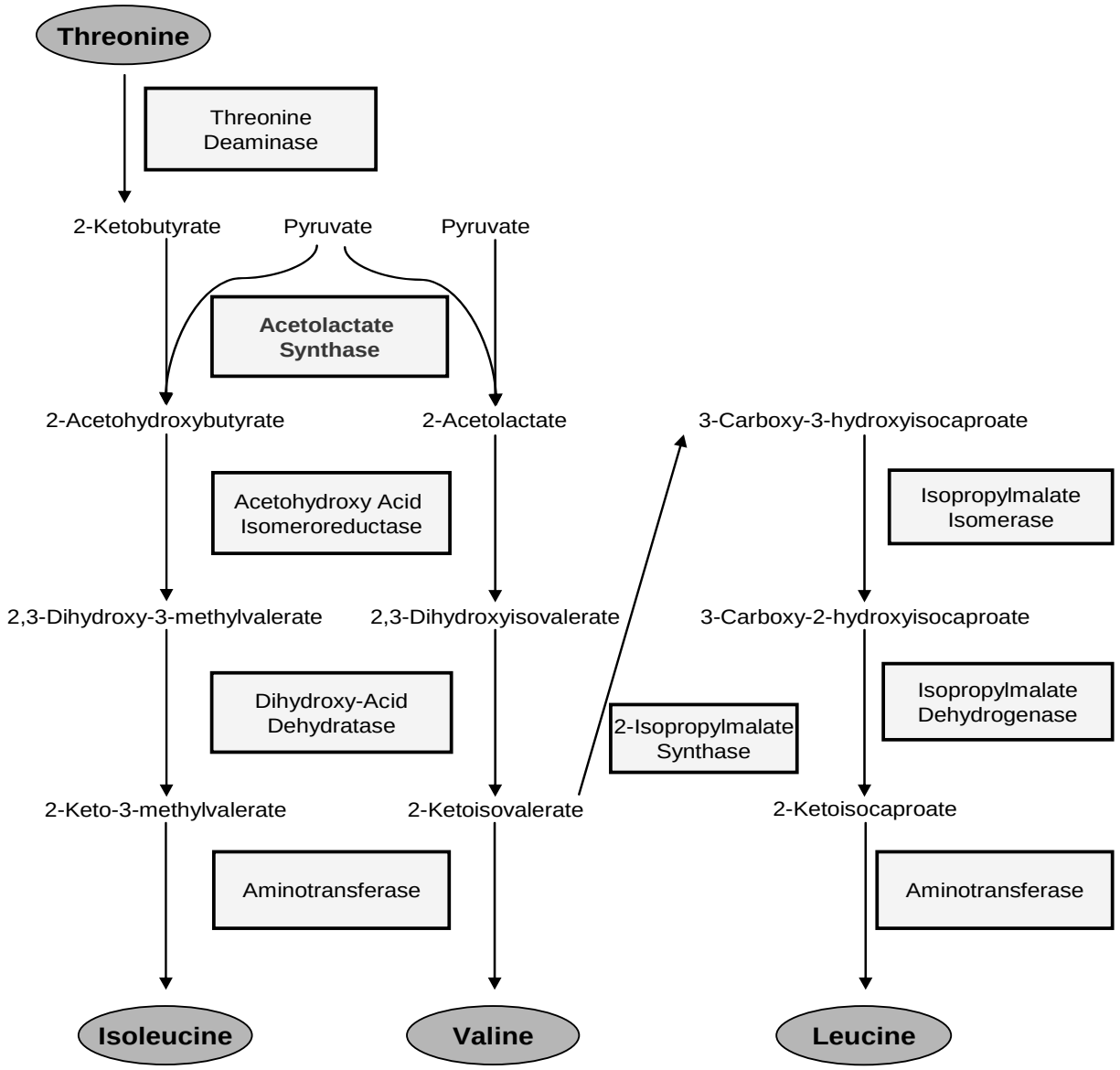
ALS enzymes are widely distributed in nature; ALS genes have been isolated from bacteria, fungi, algae and plants (Friden *et al.,* 1985; Falco *et al.,* 1985; Reith and Munholland, 1995; Mazur *et al.,* 1987). ALS enzymes generally consist of a larger catalytic subunit that is active alone and a smaller regulatory subunit. The large catalytic subunit of ALS in soybean was used to derive GM-HRA.

The large catalytic subunit of ALS typically functions as a homotetramer or homodimer with the active site located at a dimer interface. Recent studies on the crystal structures of ALS from yeast and *Arabidopsis thaliana* in the presence of ALS-inhibiting herbicides have provided evidence that the inhibitors are located at the entrance to and extending into a funnel-type channel leading to the active site, blocking access to the substrate (McCourt *et al*., 2006; Pang *et al*., 2002, 2003, 2004). These compounds inhibit plant growth by inactivating ALS, a key enzyme in the essential amino acid biosynthetic pathway; the negative effects on plant growth may be attributed to the physiological consequences of the enzyme inhibition (Duggleby and Pang, 2000). In addition to sulfonylureas (Levitt, 1978) and imidazolinones, there are three other chemical classes of herbicides identified to date that inhibit the activity of ALS: triazolopyrimidines, pyrimidinylthio (or oxy)-benzoates, and sulfonylamino-carbonyltriazolinones (Tan *et al.,* 2006).

In plants, ALS genes are nuclear encoded and the enzymes contain a chloroplast directed N-terminal transit peptide, estimated to be between 80 to 100 amino acids in length. The transit peptide serves to import the protein into the chloroplast, where the majority of branched chain amino acid biosynthesis occurs. Natural herbicide tolerance mutations are most commonly found at four amino acid locations in ALS: A122, P197, W574 and S653 (Duggleby and Pang, 2000). Several commercialized crops (soybean, maize, wheat, rice, canola and sunflower) that are tolerant to ALS inhibitors contain mutations derived through mutagenesis or selection (Tan *et al*. 2006).

[2] Although both 'ALS' and 'AHAS' are acceptable nomenclature for acetohydroxyacid synthase, recent literature more commonly refers to the enzyme in plants as 'AHAS' (EC 2.2.1.6, formerly EC 4.1.3.18). However, 'ALS' will be used throughout this submission for consistency with the early literature.

Figure 24. Branched Chain Amino Acid Biosynthesis in Plants



Adapted from Coruzzi and Last, 2000.

B2. Herbicide Tolerant ALS Enzymes in Different Species and Generation of GM-HRA Soybean

Following identification of ALS as the target site for sulfonylurea herbicides (Levitt, 1978), the first enzymes with herbicide tolerant activity were identified in bacteria (LaRossa and Schloss, 1984), yeast (*Sacchromyces cerevisiae*) (Falco and Dumas, 1985), and plants (Chaleff and Mauvais, 1984). The respective genes were then isolated from various species and amino acid sequence changes accountable for the tolerant phenotype were identified. Several reviews are available on amino acid substitutions that result in tolerance to ALS inhibitors (Hartnett *et al*., 1990, 1991; Falco *et al*., 1989; Duggleby and Pang, 2000). For example, a resistant tobacco (*Nicotiana tabacum*) line was isolated through two successive rounds of tissue culture selection in the presence of sulfonylureas herbicides (Creason and Chaleff, 1988). The gene responsible for the altered ALS was identified, and the sequence of the gene identified two amino acid substitutions that contributed to the herbicide tolerant activity: mutations P196A and W573L (Lee *et al*., 1988). These two mutations are equivalent to the locations of the commonly found natural tolerance mutations P197 and W574 (Duggleby and Pang, 2000). Both individual mutations conferred tolerance to herbicides, but the two mutations combined together typically resulted in a higher level of tolerance (Mazur and Falco, 1989; Hartnett *et al*. 1990; Creason and Chaleff, 1988).

In the case of 356043 soybean, mutations analogous to those described above in the double mutant tobacco were introduced into the sensitive version of the *gm-als I* gene (P183A and W560L) by site-directed mutagenesis in order to produce the *gm-hra* gene encoding the GM-HRA herbicide tolerant enzyme. The protein sequence of GM-HRA is compared to the GM-ALS I protein sequence in Figure 25.

B-3. The GM-HRA Protein

The GM-HRA protein is derived from the soybean acetolactate synthase protein (GM-ALS I). GM-HRA is 656 amino acids in length with a predicted molecular weight of 71 kDa for the full-length protein (Figure 25). ALS proteins contain N-terminal transit peptides and the mature protein is formed following transport into the chloroplast and subsequent cleavage of the transit peptide. The mature proteins start at residues S53 and S48 (double underlined in Figure 25), respectively, for GM-HRA and GM-ALS I, resulting in proteins of 604 amino acids with a predicted molecular weight of 65 kDa. The first five amino acids of GM-HRA in bold are derived from the translation of 15 nucleotides from the 5' untranslated region of the soybean *gm-als* gene. The two boxed residues in Figure 25, A183 and L560, in the GM-HRA sequence identify the differences from GM-ALS I.

Figure 25. Deduced Amino Acid Sequence of GM-HRA and its Alignment with GM-ALS 1

```
GM-HRA      1  MPHNTMAATA SRTTRFSSSS SHPTFPKRIT RSTLPLSHQT LTKPNHALKI
GM-ALS I    1  -----MAATA SRTTRFSSSS SHPTFPKRIT RSTLPLSHQT LTKPNHALKI

GM-HRA     51  KCSISKPPTA APFTKEAPTT EPFVSRFASG EPRKGADILV EALERQGVTT
GM-ALS I   46  KCSISKPPTA APFTKEAPTT EPFVSRFASG EPRKGADILV EALERQGVTT

GM-HRA    101  VFAYPGGASM EIHQALTRSA AIRNVLPRHE QGGVFAAEGY ARSSGLPGVC
GM-ALS I   96  VFAYPGGASM EIHQALTRSA AIRNVLPRHE QGGVFAAEGY ARSSGLPGVC

GM-HRA    151  IATSGPGATN LVSGLADALM DSVPVVAITG QVARRMIGTD AFQETPIVEV
GM-ALS I  146  IATSGPGATN LVSGLADALM DSVPVVAITG QVPRRMIGTD AFQETPIVEV

GM-HRA    201  SRSITKHNYL ILDVDDIPRV VAEAFFVATS GRPGPVLIDI PKDVQQQLAV
GM-ALS I  196  SRSITKHNYL ILDVDDIPRV VAEAFFVATS GRPGPVLIDI PKDVQQQLAV

GM-HRA    251  PNWDEPVNLP GYLARLPRPP AEAQLEHIVR LIMEAQKPVL YVGGGSLNSS
GM-ALS I  246  PNWDEPVNLP GYLARLPRPP AEAQLEHIVR LIMEAQKPVL YVGGGSLNSS

GM-HRA    301  AELRRFVELT GIPVASTLMG LGTFPIGDEY SLQMLGMHGT VYANYAVDNS
GM-ALS I  296  AELRRFVELT GIPVASTLMG LGTFPIGDEY SLQMLGMHGT VYANYAVDNS

GM-HRA    351  DLLLAFGVRF DDRVTGKLEA FASRAKIVHI DIDSAEIGKN KQAHVSVCAD
GM-ALS I  346  DLLLAFGVRF DDRVTGKLEA FASRAKIVHI DIDSAEIGKN KQAHVSVCAD

GM-HRA    401  LKLALKGINM ILEEKGVEGK FDLGGWREEI NVQKHKFPLG YKTFQDAISP
GM-ALS I  396  LKLALKGINM ILEEKGVEGK FDLGGWREEI NVQKHKFPLG YKTFQDAISP

GM-HRA    451  QHAIEVLDEL TNGDAIVSTG VGQHQMWAAQ FYKYKRPRQW LTSGGLGAMG
GM-ALS I  446  QHAIEVLDEL TNGDAIVSTG VGQHQMWAAQ FYKYKRPRQW LTSGGLGAMG

GM-HRA    501  FGLPAAIGAA VANPGAVVVD IDGDGSFIMN VQELATIRVE NLPVKILLLN
GM-ALS I  496  FGLPAAIGAA VANPGAVVVD IDGDGSFIMN VQELATIRVE NLPVKILLLN

GM-HRA    551  NQHLGMVVQL EDRFYKSNRA HTYLGDPSSE SEIFPNMLKF ADACGIPAAR
GM-ALS I  546  NQHLGMVVQW EDRFYKSNRA HTYLGDPSSE SEIFPNMLKF ADACGIPAAR

GM-HRA    601  VTKKEELRAA IQRMLDTPGP YLLDVIVPHQ EHVLPMIPSN GSFKDVITEG
GM-ALS I  596  VTKKEELRAA IQRMLDTPGP YLLDVIVPHQ EHVLPMIPSN GSFKDVITEG

GM-HRA    651  DGRTRY
GM-ALS I  646  DGRTRY
```

Double underlining (S53 in the GM-HRA sequence and S48 in the GM-ALS I sequence) indicates the start of the mature protein.

Boxed residues (A183 and L560 in the GM-HRA sequence) identify the differences from GM-ALS I.

VI-C. Levels of Transgenic Proteins in 356043 Soybean

356043 soybean contains the *gat4601* and *gm-hra* genes encoding the GAT4601 and GM-HRA proteins, respectively. The transcription of the *gat4601* gene is under the control of the SCP1 promoter, which directs GAT4601 expression constitutively. The transcription of the *gm-hra* gene is under the control of the S-adenosyl-L-methionine synthetase (SAMS) promoter from soybean, which directs GM-HRA expression constitutively.

Levels of the GAT4601 and GM-HRA proteins in 356043 and control (Jack) soybean were measured in replicated tissue samples collected from plants grown at six field locations in North America in 2005. The T5 generation of 356043 soybean was planted (Figure 2). Forage and root tissues were collected at the R3 stage of development, and grain was collected at the R8 stage of development (Gaska, 2006). Three replicated samples per tissue per location were collected for 356043 soybean, and one sample per tissue per location was collected for control soybean.

Forage tissue was analyzed because soybean plants are occasionally used as hay when alfalfa or clover are in short supply due to winter-killing or drought conditions. Forage tissue also allows measurement of transgenic protein levels in the above-ground portion of the plant. Root tissue was analyzed at the same growing stage as forage to obtain expression data on the below-ground portion of the plant. Protein expression in grain tissue was measured because soybean grain is the most commonly consumed portion of the plant for both humans and animals. Tissue samples were processed and analyzed using enzyme linked immunosorbent assay (ELISA) methods described in Appendix 3.

Results of the expression analyses are presented in Table 9 for GAT4601 and Table 10 for GM-HRA. The mean GAT4601 protein levels across six locations in 356043 soybean forage, root and grain were 1.6, 1.6 and 0.24 ng/mg of tissue (dry weight), respectively. The mean GM-HRA protein levels across six locations in 356043 soybean forage, root and grain were 27, 3.2 and 0.91 ng/mg of tissue (dry weight), respectively. Neither GAT4601 nor GM-HRA protein was detected in nontransgenic control (Jack) soybean tissues across the six locations.

These results confirm that protein expression of GAT4601 and GM-HRA in 356043 soybean is constitutive.

Table 9. Levels of GAT4601 Protein Measured in 356043 Soybean

Growth Stage/ Tissue	ng/mg Tissue Dry Weight[2]		Standard Deviation
	Mean	Range[1]	
356043 soybean			
R3/Forage	1.6	1.1 – 2.3	0.32
R3/Root	1.6	1.1 – 2.2	0.39
R8/Grain	0.24	0.14 – 0.39	0.072
Control			
R3/Forage	0	0	0
R3/Root	0	0	0
R8/Grain	0	0	0

[1]Range denotes the lowest and highest individual value across locations.
[2]For the soybean forage and root samples with results below the sample lower limit of quantitation (LLOQ), a value of zero was assigned for calculation purposes. Due to low expression, seven of eighteen grain results between the limit of detection (LOD) and the LLOQ were reported to avoid bias from assigning a zero value. Grain results below the sample LOD were assigned a value of 0.

Table 10. Levels of GM-HRA Protein Measured in 356043 Soybean

Growth Stage/ Tissue	ng/mg Tissue Dry Weight[2]		Standard Deviation
	Mean	Range[1]	
356043 soybean			
R3/Forage	27	13 – 42	8.0
R3/Root	3.2	0.32 – 7.6	2.2
R8/Grain	0.91	0.64 – 1.2	0.17
Control			
R3/Forage	0	0	0
R3/Root	0	0	0
R8/Grain	0	0	0

[1]Range denotes the lowest and highest individual value across locations.
[2]For results below the sample LLOQ, a value of zero was assigned for calculation purposes.

VI-D. Characterization of the GAT4601 and GM-HRA Proteins

Small amounts of GAT4601 and GM-HRA proteins were purified from 356043 leaf soybean tissue by immunoaffinity chromatography. Larger quantities of the two proteins were necessary to conduct safety assessment studies. Due to the low expression levels and complications involved with extracting and purifying sufficient quantities of the transgenic proteins from plant tissues, GAT4601 and GM-HRA were produced in and purified from a heterologous *E. coli* protein expression system. For GM-HRA, the mature form of the protein (excluding the chloroplast transit peptide sequence) was produced.

Prior to the use of purified microbially expressed proteins for toxicological and biochemical studies, characterization tests were done to confirm the identity and equivalency of the microbially expressed GAT4601 and GM-HRA proteins to the GAT4601 and GM-HRA proteins expressed *in planta* in 356043 soybean. A detailed description of methods used in these protein characterization studies and the resulting data are included in Appendix 4.

Characterization of the physicochemical properties of the microbially expressed and plant-derived GAT4601 and GM-HRA proteins was accomplished through the following analyses:

1) Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to determine the purity and confirm molecular weight;
2) Western blot analysis to confirm the molecular weight and immunoreactivity of the proteins;
3) Glycoprotein staining to determine post-translational modification (glycosylation);
4) Mass determination of tryptic peptides by matrix assisted laser desorption ionization mass spectrometry (MALDI-MS) to confirm the identity of the proteins;
5) N-terminal amino acid sequence analysis to further confirm the identity of the proteins.

Utilizing the above characterization methods, the identity of both the microbially expressed and the 356043 soybean plant-derived GAT4601 and GM-HRA proteins were verified. Also, the above characterization tests confirmed the equivalency of GAT4601 and GM-HRA proteins expressed in *E. coli* to the GAT4601 and GM-HRA proteins expressed *in planta* in 356043 soybean. Microbial GAT4601 and GM-HRA were used subsequently for *in vitro* and *in vivo* safety assessment studies summarized in Sections VI-E and VI-F.

VI-E. Summary of the Food and Feed Safety Assessment for the GAT4601 Protein

A detailed assessment of animal and human safety of the GAT4601 protein was submitted to FDA on June 16, 2006 as part of an Early Food Safety Evaluation. Further information on 356043 soybean will be submitted to FDA as part of the consultation process for bioengineered foods. The conclusions of the safety assessment are summarized below. The allergenic potential of the GAT4601 protein was assessed using a step-wise, weight-of-the evidence approach (Codex Alimentarius Commission, 2003) through the assessment of the *gat4601* gene source and history of use or exposure, bioinformatic comparison of the amino acid sequence of the GAT4601 protein with known protein allergen sequences, evaluation of the stability of the GAT4601 protein using *in vitro* gastric and intestinal digestion models, determination of the GAT4601 protein glycosylation status and heat lability of the GAT4601protein.

1) The donor organism, *B. licheniformis*, is a common soil bacterium widely distributed in the environment. Due to its ubiquitous presence as spores in soil and dust, *B. licheniformis* is widely known as a contaminant of food but is not associated with any adverse effects.
2) *B. licheniformis* has a history of safe use for the production of food enzymes in the United States, Canada, and Europe (*e.g.*, alpha-amylase, cyclodextrin glycosyltransferase, hemicellulase, proteases, pullulanase), biocontrol agents (EU Commission, 2000; FDA, 2001) and as a probiotic (Kritas *et al*., 2006; Alexopoulos *et al.,* 2004a and b). *B. licheniformis* was determined by EPA to present low risk of adverse effects to human health and the environment and was subsequently granted an exemption for use in certain industrial fermentation processes (Federal Register, 1997).
3) Bioinformatic analyses revealed no biologically relevant amino acid sequence identities between known or putative protein allergens and the GAT4601 protein sequence. Furthermore, no short (eight or more contiguous amino acids) polypeptide matches were shared between the GAT4601 protein and protein allergens. These data indicate the lack of both amino acid identity and immunologically relevant similarities between the GAT4601 protein and known protein allergens.
4) The GAT4601 protein is rapidly hydrolyzed in both simulated gastric and intestinal fluids (less than 30 seconds in simulated gastric fluid containing pepsin at pH 1.2 as demonstrated by SDS-PAGE analysis, and less than 2 minutes in simulated intestinal fluid containing pancreatin at pH 7.5 as demonstrated by SDS-PAGE analysis).
5) The GAT4601 protein is not glycosylated as demonstrated by glycoprotein staining.
6) The GAT4601 protein is heat labile (*i.e.,* heating at 56° C for 15 minutes completely inhibits the enzyme activity of the GAT4601 protein).
7) Bioinformatic analyses revealed GAT4601 to be similar to other N-acetyltransferase proteins. The GAT4601 protein retains the characteristics found in other N-acetyltransferases that are ubiquitous in plants and microorganisms (Neuwald and Landsman, 1997). GAT4601 contains the definitive motif for the GNAT family of N-acetyltransferases (Marchler-Bauer *et al.,* 2005). This superfamily of enzymes is present in all organisms, including plants, mammals, fungi, algae, and bacteria (Dyda *et al*., 2000).
8) No biologically relevant amino acid sequence identities were observed between known protein toxins and the GAT4601 protein sequence.
9) There was no evidence of acute toxicity in mice at a dose of 2000 mg purified protein preparation per kg of body weight (equivalent to approximately 1680 mg of the full-length GAT4601 protein per kg of body weight). Based on expression levels of GAT4601 protein in 356043 soybean grain, a child weighing 10 kg would have to consume 64,615 kg/day of 356043 soybean grain to match the dose used in the mouse acute toxicity test. An adult weighing 60 kg would have to consume 387,692 kg/day of 356043 soybean grain. Based on these simplistic worse-case calculations,

it is clear there is a wide margin of safety for the GAT4601 protein, especially if other factors such as market share or soybean grain processing are taken into account.

Results of the safety assessment indicate that the GAT4601 protein is unlikely to cause an allergic reaction in humans or be a toxin in humans or animals, and therefore is safe for animal and human consumption.

VI-F. Summary of the Food and Feed Safety Assessment for the GM-HRA protein

A detailed assessment of human and animal safety of the GM-HRA protein will be provided to FDA as part of the consultation on food and feed safety for 356043 soybean. The conclusions of the safety assessment are summarized below. As described above for the GAT4601 protein, the allergenic potential of the GM-HRA protein was assessed using a step-wise, weight-of-the evidence approach utilizing guidance provided by the Codex Alimentarius Commission (2003).

1) The donor organism is *Gycine max*. The GM-HRA protein is only two amino acids different from the endogenous soybean ALS protein from which it was derived, with an additional five amino acids on the N-terminus derived from the translation of 15 nucleotides from the 5' untranslated region of the soybean *gm-als* gene.
2) Bioinformatic analyses revealed GM-HRA to be similar to other acetolactate synthase proteins found in bacteria, fungi, algae and plants (Friden *et al.,* 1985; Falco *et al.,* 1985; Reith and Munholland, 1995; Mazur *et al.,* 1987).
3) Bioinformatic analyses revealed no biologically relevant amino acid sequence identities between known or putative protein allergens and the GM-HRA protein sequence. Furthermore, no short (eight or more contiguous amino acids) polypeptide matches were shared between the GM-HRA protein and protein allergens. These data indicate the lack of both amino acid identity and immunologically relevant similarities between the GM-HRA protein and known protein allergens.
4) The GM-HRA protein is rapidly hydrolyzed in both simulated gastric and intestinal fluids (less than 30 seconds in simulated gastric fluid containing pepsin at pH 1.2 as demonstrated by SDS-PAGE analysis, and less than 30 seconds in simulated intestinal fluid containing pancreatin at pH 7.5 as demonstrated by SDS-PAGE analysis or less than 1 minute as demonstrated by western blot analysis).
5) The GM-HRA protein is not glycosylated as demonstrated by glycoprotein staining.
6) The GM-HRA protein is heat labile (*i.e.,* heating at 50° C for 15 min completely inhibits the enzyme activity of the GM-HRA protein).
7) No biologically relevant amino acid sequence identities were observed between known protein toxins and the GM-HRA protein sequence.
8) There was no evidence of acute toxicity in mice at a dose of 2000 mg purified protein preparation per kg of body weight (equivalent to approximately 582 mg of full-length GM-HRA protein per kg of body weight). Based on expression levels of the GM-HRA protein in 356043 soybean grain, a child weighing 10 kg would have to consume 6,847 kg/day of 356043 soybean grain to match the dose used in the mouse acute toxicity test. An adult weighing 60 kg would have to consume 41,082 kg/day of 356043 soybean grain. Based on these simplistic worse-case calculations, it is clear there is a wide margin of safety for the GM-HRA protein, especially if other factors such as market share or soybean grain processing are taken into account.

Results of the safety assessment indicate that the GM-HRA protein is unlikely to cause an allergic reaction in humans or be a toxin in humans or animals, and therefore is safe for animal and human consumption.

VII. Agronomic Performance and Ecological Observations

The United States is the world's leading soybean producer and exporter. Farm value of U.S. soybean production in 2003/04 was $18.0 billion, the second-highest value among U.S.-produced crops, trailing only corn (USDA/ERS, 2006a).

An agronomic and phenotypic evaluation was conducted to assess the agronomic comparability of 356043 soybean to conventional soybean. This evaluation also included observed responses to insect and disease stressors. These evaluations form the basis for our conclusion that 356043 soybean is comparable to conventional soybeans and is therefore no more likely to pose a plant pest risk than conventional soybeans.

The agronomic and phenotypic evaluations were based on both laboratory experiments and replicated, multi-site field trials conducted by agronomists and scientists who are considered experts in the production and evaluation of soybean. In each of these assessments, 356043 soybean was compared to either a Jack control or a null segregant line that did not carry any transgenic insert DNA.

To evaluate the agronomic and phenotypic characteristics of 356043 soybean, data were collected to address specific characteristics that influence reproduction and survival characteristics of the crop.

VII-A. Germination and Dormancy Evaluations

In order to test germination and potential dormancy, seeds from the T7 generation (see Figure 2) of 356043 soybean were tested for germination under both cold and warm growing conditions using standard laboratory tests. The control was the nontransgenic Jack soybean variety.

The cold germination test consisted of four replicates containing 100 seeds each of 356043 soybean and Jack. The seeds were sown into a saturated soil/sand mixture and placed in a 10°C chamber for four days. After four days, the seed was moved to a 25°C chamber for an additional three days. At the end of the seven days, the number of germinated seeds was recorded.

The warm germination test consisted of eight replicates containing 50 seeds each of 356043 soybean and Jack. The seeds were placed between sheets of germination towels and placed in a 25°C chamber at approximately 90% relative humidity for five days. After five days, the number of germinated seeds was recorded.

Results of the seed germination testing are presented in Table 11. Both 356043 and control soybean showed a high rate of germination (99% or greater) in both the warm and cold germination tests. Out of the 800 total seeds tested from each of 356043 and control soybean, only one of the control seeds remained hard after the germination tests. Seeds that remain hard (*i.e.,* not swollen) can be indicative of potential dormancy. The hard seed was from the nontransgenic Jack line, so no further analysis was done.

These results indicate that the germination and dormancy characteristics of 356043 soybean were not altered when compared to nontransgenic soybean. Based on this information, 356043 soybean is unlikely to present an increased potential for weediness as compared to conventional soybeans.

Table 11. Summary of Seed Germination Results for 356043 Soybean

	Germination Test % germination (number of seeds germinated / number of seeds tested)		**Number of Hard Seeds**	
	Warm germination test	Cold germination test	Warm germination test	Cold germination test
356043 soybean	99.5% (398/400)	99.5% (398/400)	0	0
Control (Jack)	99.8% (399/400)	99.0% (396/400)	0	1

VII-B. Field Trial Evaluations

356043 soybean has been field tested in the United States since 2003 as authorized by USDA permits listed in Appendix 5. The list compiles a number of test sites in diverse regions of the U.S. including the major soybean growing areas of the Midwest and a winter nursery in Hawaii. Agronomic and phenotypic data were collected to assess agronomic comparability as it relates to plant pest potential. Certain agronomic data, for example seed shattering, can also be used for an assessment of enhanced weed potential for 356043 soybean.

Throughout the development process, additional qualitative phenotypic traits such as flower color, pubescence color, pod color, and hila color were also monitored. Flower color in soybean can be purple (dominant) or white (recessive). Pubescence color is based on the short hairs on soybean plant stems and pods at maturity. Pubescence color in soybean can be tawny (dominant) or grey (recessive). Pods on mature soybeans can be brown (dominant) or tan (recessive). The hilum is a scar visible on the soybean seed's coat; colors include yellow, grey, black, imperfect black, buff or brown.

The recipient cultivar for 356043 soybean, Jack, has white flowers, grey pubescence, brown pods at maturity and yellow hila. Throughout the breeding and development process for 356043 soybean, scientists skilled in the art of plant breeding and agricultural science monitored 356043 soybean for these expected phenotypic traits. Visual observations confirmed these qualitative phenotypic traits for 356043 soybean were unchanged from Jack and were stable through the breeding process over a three-year time period and at least six generations, when in a Jack genetic background.

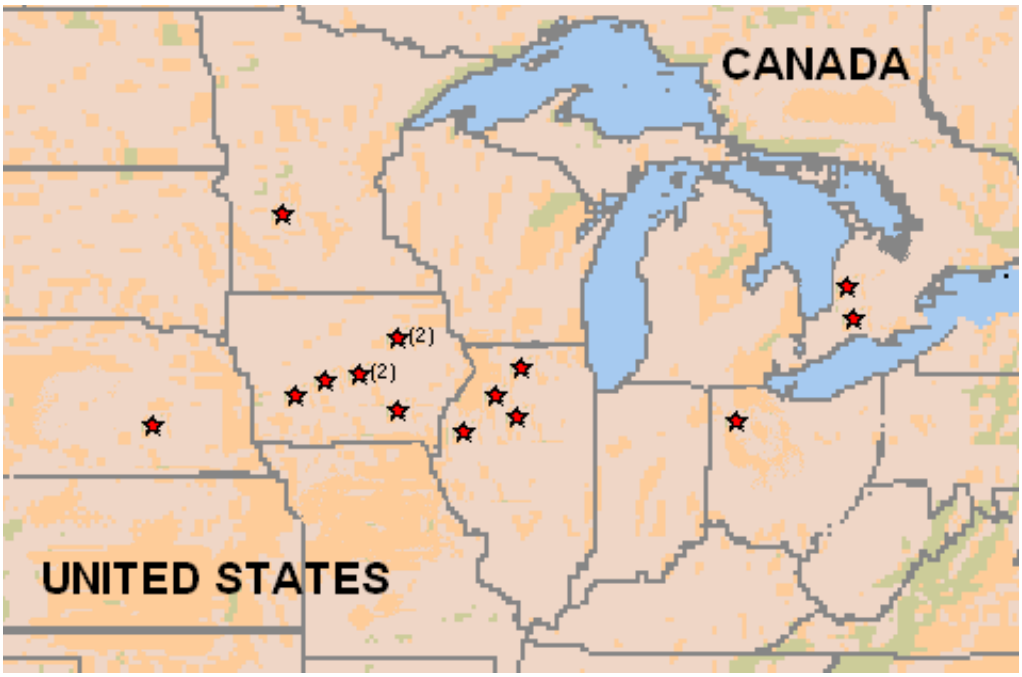
Also throughout the development process, 356043 soybean was observed for unexpected differences in responses to abiotic stress (*e.g*, drought, excess moisture, temperature extremes, *etc*.). These observations were qualitative and opportunistic, but 356043 soybean and controls (non-transgenic near isoline Jack, null segregants, and/or conventional soybean lines with similar genetics) were similar with respect to their response to abiotic stress.

Agronomic data were collected from four experiments (denoted A, B, C and D) conducted at 16 total field locations over two consecutive years (Figure 26). Categorical and quantitative data were collected from 356043 and control soybeans from three field locations in 2004 and an additional 13 field locations in 2005. The trial locations provided a range of environmental and agronomic conditions representative of the major soybean growing regions where commercial production of 356043 soybean is expected. Agronomic practices used to prepare and maintain each field site were characteristic of each respective region.

Table 12 outlines the quantitative phenotypic characteristics that were collected during the 2004 and 2005 growing seasons.

These findings provide evidence that 356043 soybean is comparable to its nontransgenic near isoline, Jack, and does not exhibit any unexpected qualitative phenotypic traits or unexpected responses to abiotic stress.

Figure 26. Map of 2004 and 2005 Locations for Agronomic Data Collection for 356043 Soybean



Experiment A, 2004 Purpose: Agronomic/yield evaluation 1. Dallas Center, IA 2. Johnston, IA 3. Cedar Falls, IA	Experiment B, 2005 Purpose: Agronomic/yield evaluation 1. Stuart, IA 2. Johnston, IA 3. Cedar Falls, IA 4. Princeton, IL 5. Rochelle, IL 6. Monmouth, IL 7. Napoleon, OH
Experiment C, 2005 Purpose: Nutrient composition, protein expression and agronomic evaluation 1. Wyoming, IL 2. Richland, IA 3. Paynesville, MN 4. York, NE 5. Thorndale, Ontario 6. Branchton, Ontario	Experiment D, 2005 Purpose: Yield evaluation 1. Wyoming, IL 2. Paynesville, MN 3. York, NE 4. Thorndale, Ontario 5. Branchton, Ontario Note: Stars for Experiment D are not noted on map above since the locations were also used for Experiment C during the 2005 growing season.

Table 12. Field Agronomic Characteristics Measured

General Characteristic	Characteristic Measured	Evaluation timing*	Data Description	Scale
Germination / Emergence	Emergence Score	VC-V2	Visual estimate of plant emergence	From 1 to 9, where 1=0-10% plants emerged, and 9= 90-100% plants emerged
	Early Population	VC-V2	Number of plants emerged per plot	Actual count per plot
	Seedling Vigor	VC-V2	Visual estimate of average vigor of emerged plants per plot	From 1 to 9, where 1=short plants with small leaves, and 9=tall plants with large leaves
Vegetative Parameters	Plant Height	Approximately R6	Height in cm from the soil surface to the tip of the highest leaf when extended by hand	Height in cm
	Lodging	Approximately R8	Visual estimate of lodging severity	From 1 to 9, where 1=plants laying flat, and 9=plants standing upright
	Final Population	Approximately R8	The number of plants remaining per plot	Actual count per plot
Reproductive Parameters	Days to Maturity	R8	Difference between maturity date and planting date	Number of days
	Shattering Score	Approximately R8	Visual estimate of grain shattering at maturity	From 1 to 9, where 1=high shattering and 9=no shattering
	Seed weight	R8	Weight of 100 random seeds	Grams
	Yield	Approximately R8	Harvest weight per area adjusted to 13% moisture content	Bushels per acre
Ecological Interactions	Disease Incidence	Approximately R6	Visual estimate of foliar disease incidence	From 1 to 9, where 1=poor disease resistance or high infection, and 9=best disease resistance or low infection
	Insect Damage	Approximately R6	Visual estimate of insect damage	From 1 to 9, where 1=poor insect resistance or high damage, and 9=best insect resistance or low damage

*Refer to Gaska, 2006 for a description of soybean growth stages.

Experiment A was planted at three locations in the major soybean growing region of the Midwest during the 2004 growing season. The purpose of Experiment A was to evaluate the agronomic characteristics of 356043 soybean relatively early in the breeding process (Figure 26). Seed of the T3 generation of 356043 soybean was used (Figure 2). The control plants were null segregants of the T3 generation.

The following characteristics were measured: emergence score, shattering score, yield, days to maturity, and seed weight. Descriptions of the characteristics and their measurement are found in Table 12. Seed was planted in rows 12 feet long and 30 inches apart, with 110 seeds per row. Plants were not thinned. Normal agronomic practices were employed throughout the growing season. There were three replicates of 356043 soybean and two replicates of null segregant control soybean at each location.

Results of Experiment A are summarized in Table 13. Categorical data collected during this study (*i.e*., emergence score and shattering score) were not statistically analyzed. Other agronomic data (yield, days to maturity and seed weight) were analyzed using Statistical Analysis Software (SAS). Refer to Appendix 2 for the statistical model used.

For all characteristics measured, 356043 soybean from the T3 generation was similar to T3 null segregants that did not contain the *gat4601* and *gm-hra* genes. No statistically significant differences were observed for mean yield, days to maturity and seed weight between the 356043 and control soybean when data were evaluated across locations. Results of Experiment A indicate 356043 soybean is agronomically comparable to null segregants that do not contain the *gat4601* and *gm-hra* genes.

Table 13. Summary of Experiment A: 2004 Agronomic Performance of 356043 Soybean Across Three Locations

Characteristic	T3 null segregant (control)	356043 soybean	Number of Locations	P-value
Emergence score	9	9	3	Not applicable
Shattering score	9	9	3	Not applicable
Yield (bushels/acre)	44.2	44.3	3	0.947
Days to Maturity	115.5	115.0	1	0.591
Seed weight (grams)	15.9	16.2	2	0.369

Experiment B was planted at seven locations in the major soybean growing region of the Midwest during the 2005 growing season. The purpose of Experiment B was to evaluate the agronomic characteristics of 356043 soybean (Figure 26). Seed from the T4 generation was used (Figure 2). The control plants were null segregants of the T4 generation.

The following characteristics were measured: emergence score, shattering score, yield, days to maturity, and seed weight. Descriptions of the characteristics and their measurement are found in Table 12. Similar to Experiment A, seed was planted in rows 12 feet long and 30 inches apart, with 110 seeds per row. Plants were not thinned. Normal agronomic practices were employed throughout the growing season. There were three replicates of 356043 soybean and two replicates of null segregant control soybean at each location.

Results of Experiment B are summarized in Table 14. Categorical data collected during this study (*i.e*., emergence score and shattering score) were not statistically analyzed. Other agronomic data (yield, days to maturity and seed weight) were analyzed using Statistical Analysis Software (SAS). Refer to Appendix 2 for the statistical model used.

For all characteristics measured, 356043 soybean from the T4 generation was similar to T4 null segregants that did not contain the *gat4601* and *gm-hra* genes. No statistically significant differences were observed for mean yield, days to maturity and seed weight between the 356043 and control soybean when data were evaluated across locations. Results of Experiment B indicate 356043 soybean is agronomically comparable to null segregants that do not contain the *gat4601* and *gm-hra* genes.

Table 14. Summary of Experiment B: 2005 Agronomic Performance of 356043 Soybean Across Seven Locations

Characteristic	T4 null segregant (control)	356043 soybean	Number of Locations	P-value
Emergence score	9	9	7	Not applicable
Shattering score	9	9	7	Not applicable
Yield (bushels/acre)	52.1	53.9	7	0.517
Days to Maturity (days)	123.5	124.5	5	0.104
Seed weight (grams)	14.1	14.2	6	0.585

Experiment C was planted at six locations in commercial soybean growing regions of North America during the 2005 growing season. The purpose of Experiment C was to evaluate the agronomic characteristics of 356043 soybean and collect tissue samples for transgenic protein levels and nutrient composition analyses (Figure 26). Seed from the T5 generation was used (Figure 2). The control was Jack, the non-transgenic near isoline.

The following characteristics were measured: early population, final population, seedling vigor, lodging, shattering score, disease incidence, insect damage, plant height and days to maturity. Descriptions of the characteristics and their measurement are found in Table 12. Seed was planted in rows 25 feet long and 30 inches apart, with 112 seeds per row. Plants were not thinned. Normal agronomic practices were employed throughout the growing season. There were three replicates of 356043 soybean and control soybean at each location, planted in a randomized complete block design of two-row plots. Each two-row plot was bordered on each side by one row of non-transgenic, commercial soybeans of a similar relative maturity.

The statistical analysis of agronomic data, including categorical data, was conducted to test for differences in the mean values between the 356043 soybean and the near isoline control Jack.

Because a greater number of characteristics were recorded in Experiment C, the statistical analysis of data was refined relative to that described for Experiments A and B in order to adjust for making multiple comparisons. When numerous comparisons are being made, it is important to control the rate of false positive results. Since the introduction of the false discovery rate (FDR) approach in the mid-1990's, it 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). A false positive result occurs when two means are deemed significantly different when, in fact, they are not. If one uses a 5% type I error rate for each agronomic characteristic measured, then the number of false positives increases as the number of characteristics increase. In order to help manage the false positive rate FDR method of Benjamini and Hochberg (1995) was applied to account for making multiple comparisons. P-values were adjusted accordingly. This resulted in the false positive rate being held to 5%. Both adjusted and unadjusted P-values are provided for Experiment C (Table 15). For a more detailed description of the statistical model used to analyze these data, refer to Appendix 2.

Results of Experiment C are summarized in Table 15. For all characteristics measured (early population, final population, seedling vigor, lodging, shattering score, disease incidence, insect damage, plant height and days to maturity), no statistical differences in mean values were seen between 356043 soybean and control Jack soybean across locations after applying the FDR adjustment. The results from Experiment C indicate 356043 soybean is agronomically comparable to control Jack soybean.

Table 15. Summary of Experiment C: 2005 Agronomic Performance of 356043 Soybean Across Six Locations

Characteristic		Control (Jack)	356043 soybean
Early Population (number of plants)	Mean[1]	140	145
	Range[2]	60.0 – 173	60.0 – 176
	Adjusted P-value [3]		0.459
	P-value[4]		0.204
Final Population (number of plants)	Mean	132	134
	Range	60.0 – 171	60.0 – 178
	Adjusted P-value		0.779
	P-value		0.779
Seedling Vigor (1-9 score)	Mean	8	7
	Range	7 – 9	5 – 9
	Adjusted P-value		0.2
	P-value		0.03
Lodging (1-9 score)	Mean	7	7
	Range	2 – 9	4 – 8
	Adjusted P-value		0.8
	P-value		0.7
Shattering Score (1-9 score)	Mean	9	9
	Range	7 – 9	8 – 9
	Adjusted P-value		0.7
	P-value		0.5
Disease Incidence (1-9 score)	Mean	7	7
	Range	5 – 8	5 – 9
	Adjusted P-value		0.8
	P-value		0.7
Insect Damage (1-9 score)	Mean	8	8
	Range	7 – 9	6 – 9
	Adjusted P-value		0.5
	P-value		0.2
Plant Height (cm)	Mean	101	97.3
	Range	65.0 – 135	63.0 – 135
	Adjusted P-value		0.391
	P-value		0.0868
Days to Maturity (days)	Mean	121	122
	Range	103 – 138	103 – 138
	Adjusted P-value		0.722
	P-value		0.481

[1] Least Square Mean
[2] Range denotes the lowest and highest individual values across locations.
[3] False Discovery Rate (FDR) adjusted P-value
[4] Non-adjusted P-value

Experiment D was planted at five locations in commercial soybean growing regions of North America during the 2005 growing season. The purpose of Experiment D was to compare the yield of 356043 soybean to a near-isogenic control line Jack (Figure 26). Seed from the T4 generation 356043 soybean was used (Figure 2). The control was Jack, the non-transgenic isoline. A description of the yield characteristic and its measurement is found in Table 12.

Seed from both lines was planted in rows 25 feet long and 30 inches apart with 112 seeds per row. There were three blocks of 356043 soybean and control soybean at each location, planted in a randomized complete block design of two-row plots. Each two-row plot was bordered on each side by one row of non-transgenic, commercial soybeans of a similar relative maturity. Plants were not thinned. Normal agronomic practices were employed throughout the growing season.

Results of Experiment D are summarized in Table 16. Refer to Appendix 2 for the statistical model used. No statistically significant difference was observed in the mean yield between the 356043 and control soybean when data were evaluated across locations. The results of Experiment D indicate that 356043 soybean is comparable in yield to control Jack soybean.

Table 16. Summary of Experiment D: 2005 Yield Results Across Five Locations

Statistic	Yield Data (bushels/acre)	
	Control (Jack)	356043 soybean
Mean[1]	37.3	37.2
Range[2]	20.9 – 45.7	23.4 – 46.0
P-value		0.849

[1] Least Square Mean
[2] Range denotes the lowest and highest individual values across locations.

VII-C. Ecological Observations

Ecological observations (plant interactions with insect pests and diseases) were recorded for all USDA-APHIS permitted field trials of 356043 soybean during the 2003, 2004 and 2005 growing seasons. Plant breeders, plant pathologists and entomologists, and others skilled in the art of agricultural science observed 356043 soybean and control lines at least every four weeks for insect and disease pressure and recorded the severity of any stressor seen. Any unexpected qualitative differences in response between 356043 soybean and various control lines (non-transgenic near isoline Jack, null segregants, and/or conventional soybean lines) were recorded.

A summary of the ecological observations is presented in Table 17 for insect stressors and in Table 18 for disease stressors. In every case, the severity of insect or disease stress on 356043 soybean was not qualitatively different from various control lines growing at the same location. These results support the conclusion that the ecological interactions for 356043 soybean were similar to control soybean lines with similar genetics or to conventional soybean lines.

Table 17. Insect Stressor Comparison of 356043 Soybean to Control Soybean Lines Across All Locations in 2003, 2004 and 2005

Insect Stressor	State (# of counties)	Range of severity in 356043 soybean	Differences with control?
Bean leaf beetle (*Certoma trifurcata*)	IA (5)	mild to moderate	no
	IL (2)	mild	no
	NE (1)	very mild	no
	OH (1)	mild	no
Grasshopper (Orthoptera)	IL (2)	mild	no
	MN (1)	mild	no
	NE (1)	very mild to mild	no
Japanese beetle (*Popillia japonica*)	IL (1)	mild	no
	OH (1)	mild	no
Rose beetle (*Adoretus sinicus*)	HI (1)	mild	no
Soybean aphid (*Aphis glycines*)	IA (4)	mild to moderate	no
	IL (3)	mild to moderate	no
	MN (1)	mild	no
	OH (1)	moderate	no
Soybean leaf miner (*Odontota horni*)	HI (1)	mild to moderate	no
Stink bug (Pentatomidae)	AR (1)	mild to moderate	no
	IL (1)	mild	no
Whitefly (*Bemisia sp.*)	HI (1)	mild to moderate	no

Note: In USDA final field trial reports, event DP-356Ø43-5 is called EAFS 3560.4.3.

Table 18. Disease Stressor Incidence of 356043 Soybean to Control Soybean Lines Across All Locations in 2003, 2004 and 2005

Disease Stressor	State (# of counties)	Range of severity in 356043 soybean	Differences with control?
Bacterial blight (*Pseudomonas savastanoi* pv. *glycinea*)	IA (3)	mild	no
Brown spot (*Septoria glycines*)	IA (4)	mild to moderate	no
	NE (1)	very mild to mild	no
	OH (1)	mild	no
Brown stem rot (*Phialophora gregata*)	OH (1)	mild	no
Downey mildew (*Peronospora manshurica*)	HI (1)	very mild to moderate	no
	MN (1)	mild	no
	OH (1)	mild	no
Phytophthora rot (*Phytophthora sojae*)	OH (1)	mild	no
Powdery mildew (*Microsphaera diffusa*)	HI (1)	mild	no
	IA (2)	mild	no
	MN (1)	mild	no

Note: In USDA final field trial reports, event DP-356Ø43-5 is called EAFS 3560.4.3.

VII-D. Conclusions on Agronomic Performance and Ecological Observations

356043 soybean was observed in laboratory experiments and at 16 field locations to measure agronomic and phenotypic data. Data generated from these studies represent observations that are typically recorded by plant breeders and agronomists to evaluate the qualities of soybean over a broad range of environmental conditions that 356043 soybean would encounter. The measured characteristics provide crop biology data useful in establishing a basis to assess phenotypic comparability and familiarity of 356043 soybean compared to conventional soybean in the context of ecological risk assessment.

The agronomic and phenotypic data showed no biologically meaningful differences between 356043 soybean and control soybean (the near isogenic control line Jack, null segregants of 356043 soybean, and conventional soybean lines) with respect to phenotype, germination, vegetative growth, reproductive parameters and response to biotic stressors. These data support the conclusion that 356043 soybean is comparable in agronomic characteristics to conventional soybean. The data also support the conclusion that 356043 soybean is unaltered with respect to interaction with beneficial organisms and plant pest risk. Likewise, assessment of the ecological data detected no biologically significant differences between 356043 soybean and control soybean lines indicative of a selective advantage that would result in increased weed potential for 356043 soybean.

VIII. Compositional Assessment

Compositional analysis of 356043 soybean was used to evaluate any changes in the levels of key nutrients, isoflavones, or antinutrients compared to the nontransgenic near isoline Jack. Along with agronomic data, compositional analysis is a general indicator that 356043 soybean will not exhibit unexpected or unintended effects with respect to plant pest risk. The U.S. FDA will review the details of the compositional analysis as a component of the safety assessment of 356043 soybean.

Comprehensive compositional analyses were performed on grain and forage tissues collected from the T5 generation of 356043 soybean and the nontransgenic near isoline Jack grown in 2005 at six field locations in soybean-growing areas of North America (Experiment C, Figure 26).

In a separate experiment, grain and forage tissues were also collected from four unique conventional commercial soybean varieties ("reference varieties") grown in 2005 at six field locations in soybean-growing areas of North America (Bagley, IA, York, NE, Glen Allen, VA, Germansville, PA, Larned, KS and Branchton, ON, Canada). The reference varieties were planted, harvested, processed, and analyzed using the same methods as were used for the Jack control and 356043 soybean. Composition analysis of the reference varieties was used to help determine the normal variation for the measured analytes.

In both experiments, seed was planted in rows 25 feet long and 30 inches apart, with 112 seeds per row. Normal agronomic practices were employed throughout the growing season. There were three replicates at each location, planted in a randomized complete block design of two-row plots. Each two-row plot was bordered on each side by one row of non-transgenic, commercial soybeans of a similar relative maturity.

The compositional assessment was conducted in accordance with the OECD consensus document on compositional considerations for new varieties of soybean (OECD, 2001). Compositional analysis of forage samples included proximates (protein, fat and ash), acid detergent fiber (ADF), and neutral detergent fiber (NDF). Compositional analysis of grain samples included proximates (protein, fat and ash), ADF, NDF, fatty acids, amino acids, isoflavones, and key antinutrients (stachyose, raffinose, lectins, phytic acid, and trypsin inhibitor).

Using the data obtained from the reference varieties, a tolerance interval was calculated to contain, with 95% confidence, 99% of the values contained in the population of commercial soybean. This statistical tolerance interval and the combined range of values for each analyte available from the published literature (OECD, 2001; ILSI 2004; Taylor *et al*., 1999) were used as the primary determinant for interpreting the composition results for 356043 soybean. In interpreting the compositional data, emphasis was placed on the analyte means (NRC/IOM, 2004). Analyte means that fell within the tolerance interval and/or combined literature range for that analyte were considered to be within the normal variability of commercial soybean varieties.

Statistical analysis of nutrient composition data was also conducted to test for differences in the analyte mean values between the 356043 soybean and the non-transgenic, near isoline control Jack (for details, refer to Appendix 2). In cases where mean values fell outside of the tolerance intervals, a statistical comparison of the means values between 356043 soybean and the control provided further context for data interpretation.

When numerous analytes are being evaluated on the same samples, controlling false positive outcomes is important. A false positive outcome occurs when the mean of the transgenic line is deemed significantly different from the mean of the control line analytical data, when in fact the two means are not different. If one uses a 5% type I error rate for each analyte, then the number of false positives increases as the number of analytes increase. In order to help manage the false positive rate, the false discovery rate (FDR) method of Benjamini and Hochberg (1995) was applied to account for making multiple comparisons. This results in the false positive rate being held to 5%. P-values were adjusted accordingly (refer to Appendix 2 for more information about the use of the false discovery rate adjustment). Both adjusted and non-adjusted P-values are provided in this petition. A significant difference between the mean of 356043 soybean and that of the near isoline was established with a FDR-adjusted P-value <0.05.

VIII-A. Key Nutrients in Soybean Forage and Grain

A1. Proximates and Fiber in Soybean Forage

Composition of forage was analyzed, since soybeans are occasionally used as an alternative forage source when alfalfa or clover are in short supply due to winter-killing or drought conditions. For both 356043 and control Jack soybean, the mean values for protein, fat, acid detergent fiber (ADF), neutral detergent fiber (NDF) and ash were found to be within the statistical tolerance intervals for commercial soybean varieties and/or the combined literature range (Table 19). For each analyte measured, there was no statistical difference between 356043 soybean and the control Jack (adjusted P-value >0.05).

In conclusion, proximate and fiber analysis of soybean forage demonstrated that 356043 soybean is comparable to near isogenic and reference soybean lines. No unexpected differences in the nutrient composition of 356043 soybean forage were seen.

Table 19. Proximates and Fiber in Soybean Forage

Analyte (% Dry Weight)		Control (Jack)	356043 soybean	Tolerance Interval[1]	Combined Literature Ranges[2]
Protein	Mean[3]	23.0	23.4	14.2 – 34.1	11.2 – 24.7
	Range[4]	19.6 – 26.3	20.3 – 26.5		
	Adjusted P-value[5]		0.727		
	P-value[6]		0.470		
Fat	Mean	4.53	4.59	2.53 – 6.80	1.30 – 5.13
	Range	3.40 – 5.41	3.58 – 5.28		
	Adjusted P-value		0.863		
	P-value		0.714		
ADF	Mean	28.4	27.4	14.3 – 40.8	32-38
	Range	22.7 – 36.8	23.3 – 30.8		
	Adjusted P-value		0.588		
	P-value		0.327		
NDF	Mean	42.3	40.4	23.8 – 61.5	34-40
	Range	32.6 – 48.7	34.7 – 49.4		
	Adjusted P-value		0.259		
	P-value		0.0832		
Ash	Mean	9.46	9.28	0 – 19.5	6.72 – 10.8
	Range	7.68 – 10.9	7.69 – 10.8		
	Adjusted P-value		0.668		
	P-value		0.388		

[1]Negative tolerance limits have been set to zero.
[2]Combined ranges are taken from published literature for soybeans (OECD, 2001; ILSI 2004; Taylor *et al*., 1999).
[3]Least Square Mean
[4]Range denotes the lowest and highest individual values across locations.
[5]False Discovery Rate (FDR) adjusted P-value
[6]Non-adjusted P-value

A2. Proximates and Fiber in Soybean Grain

Proximates and fiber were analyzed in soybean grain. For both 356043 and control Jack soybean, the mean values of protein, fat, acid detergent fiber (ADF), neutral detergent fiber (NDF) and ash were found to be within the statistical tolerance intervals established for commercial soybean varieties and/or the combined literature range for soybean grain (Table 20). No statistically significant differences were observed between the 356043 and control soybean variety mean values of protein, fat, ADF, NDF and ash (adjusted P-value >0.05).

In conclusion, proximate and fiber analysis of soybean grain demonstrate that 356043 soybean is comparable to near isogenic and reference soybean lines. No unexpected differences in the proximate and fiber composition of 356043 soybean grain were seen.

Table 20. Proximates and Fiber in Soybean Grain

Analyte (%Dry Weight)		Control (Jack)	356043 soybean	Tolerance Interval	Combined Literature Ranges[1]
Protein	Mean[2]	40.0	40.2	29.9 – 48.7	33.2 – 47.4
	Range[3]	38.0 – 41.9	38.7 – 42.1		
	Adjusted P-value[4]		0.844		
	P-value[5]		0.625		
Fat	Mean	16.0	15.3	7.01 – 24.2	8.10 – 23.6
	Range	12.7 – 18.8	11.6 – 18.2		
	Adjusted P-value		0.0779		
	P-value		0.0125		
ADF	Mean	17.6	18.6	8.51 – 22.1	7.81 – 18.6
	Range	7.65 – 28.4	8.95 – 24.2		
	Adjusted P-value		0.672		
	P-value		0.423		
NDF	Mean	16.1	18.2	8.07 – 21.9	8.53 – 21.3
	Range	10.6 – 21.4	12.0 – 24.8		
	Adjusted P-value		0.166		
	P-value		0.0417		
Ash	Mean	5.12	5.08	3.19 – 7.67	3.89 – 6.54
	Range	4.68 – 5.54	4.61 – 5.58		
	Adjusted P-value		0.672		
	P-value		0.419		

[1]Combined ranges are taken from published literature for soybeans (OECD, 2001; ILSI 2004; Taylor *et al*., 1999).
[2]Least Square Mean
[3]Range denotes the lowest and highest individual values across locations.
[4]False Discovery Rate (FDR) adjusted P-value
[5]Non-adjusted P-value

A3. Fatty Acids in Soybean Grain

Soybean has many uses in the food and industrial sectors and represents one of the major sources of edible vegetable oil. Triglycerides make up 99% of soybean oil. Soybean oil is noted for its high content of linoleic (C18:2) and linolenic (C18:3) polyunsaturated fatty acids. It also contains sizeable amounts of another unsaturated fatty acid, oleic (C18:1) and moderate amounts of the saturated fatty acids palmitic (C16:0) and stearic (C18:0) (OECD, 2001).

Levels of 24 fatty acids were measured in 356043 and control soybean grain. Levels of ten fatty acids were below the limit of quantitation for the assay: caprylic acid (C8:0), capric acid (C10:0), lauric acid (C12:0), myristoleic acid (C14:1), pentadecanoic acid (C15:0), pentadecenoic acid (C15:1), γ-linolenic acid (C18:3), eicosatrienoic acid (C20:3), arachidonic acid (C20:4), and erucic acid (C22:1). Therefore, no statistical analyses were conducted on these fatty acids and data are not shown. In addition, levels of myristic acid (C14:0), palmitoleic acid (C16:1) and eicosadienoic acid (C20:2) were less than 0.1% in both 356043 and control soybean; therefore, data are not shown.

Results of the fatty acid analysis are shown in Table 21. The mean values for palmitic acid (C16:0), stearic (C18:0), oleic acid (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), arachidic acid (C20:0), eicosenoic acid (C20:1), behenic acid (C22:0) and lignoceric acid (C24:0) for 356043 soybean were within the statistical tolerance intervals and/or the combined literature ranges for these fatty acids. The means values of minor fatty acids, heptadecanoic acid (C17:0) and heptadecenoic acid (C17:1), were above the upper range of the statistical tolerance interval and literature range for nontransgenic soybean varieties. Mean values for C17:0 and C17:1 were statistically significantly different from those of control soybean. However, levels of heptadecanoic and heptadecenoic acid are in general still very low; each represent less than 0.4% of the total fatty acid content in 356043 soybean.

An increase in two minor fatty acids, heptadecanoic acid (C17:0) and heptadecenoic acid (C17:1), was detected, but this is not unexpected as expression of the GM-HRA protein likely results in a slight shift in availability of the GM-HRA enzyme substrates, pyruvate and 2-ketobutyrate. These two compounds are also substrates for the enzyme complex that initiates oil biosynthesis. Refer to Appendix 6 for information regarding biosynthesis of 17-carbon fatty acids in soybean.

Both C17:0 and C17:1 are found in vegetable oils (USDA, 2005 and Pioneer data). C17:0 is found in corn, soybean, sunflower, peanut and olive oils, as well as butter. C17:1 is found in olive oil.

Both C17:0 and C17:1 fatty acids are also found in many different kinds of commonly consumed foods. The USDA (2005) nutrition database can be searched using the tools available at www.nutriondata.com to identify the foods highest in various components. C17:0 is commonly found in meat (lamb, beef, pork) and butter at levels comparable to those seen in the 356043 samples. The highest concentrations of C17:0 in foods are found in lamb, tofu, and butter. C17:1 is also found in a wide variety of foods such as tofu, beef, cheese, and various baked products. The highest concentrations of C17:1 in foods are found in tofu and ground beef. There is no evidence to indicate that exposure to either C17:0 or C17:1 from these sources is associated with adverse effects in humans. A poultry study confirmed the nutritional wholesomeness of feed made from 356043 soybeans to conventional diets (see Section VIII-D).

In conclusion, fatty acid analysis of soybean grain demonstrates that 356043 soybean is comparable to near isogenic and reference soybean lines. An increase in two minor fatty acids, C17:0 and C17:1, was detected in 356043 soybean, but this was not unexpected. This observation and the likely shift in substrate availability is further discussed in Appendix 6.

Table 21. Major Fatty Acids in Soybean Grain

% Total Fatty Acids		Control (Jack)	356043 soybean	Tolerance Interval[1]	Combined Ranges[2]
Palmitic acid (C16:0)	Mean[3]	9.98	9.30	2.93 – 19.6	7.00 – 12.7
	Range[4]	9.59 – 10.2	8.86 – 9.65		
	Adjusted P-value[5]		0.00270[7]		
	P-value[6]		0.000100		
Heptadecanoic acid (C17:0)	Mean	0.110	0.332	0.0722 – 0.131	0.0850 – 0.138
	Range	0.0896 – 0.130	0.280 – 0.391		
	Adjusted P-value		0.00270[7]		
	P-value		0.000100		
Heptadecenoic acid (C17:1)	Mean	0.0667	0.191	0.0351 – 0.0732	0.0730 – 0.0870
	Range	0.0571 – 0.0772	0.153 – 0.243		
	Adjusted P-value		0.00270[7]		
	P-value		0.000100		
Stearic acid (C18:0)	Mean	4.44	4.53	0.852 – 8.34	2.00 – 5.71
	Range	3.77 – 4.97	4.00 – 5.21		
	Adjusted P-value		0.381		
	P-value		0.174		
Oleic acid (C18:1)	Mean	21.1	22.2	11.3 – 32.6	14.3 – 34.0
	Range	18.4 – 23.6	19.5 – 24.9		
	Adjusted P-value		0.0135[7]		
	P-value		0.00100		
Linoleic acid (C18:2)	Mean	54.7	53.7	41.7 – 64.3	48.0 – 60.0
	Range	53.1 – 56.1	52.0 – 55.7		
	Adjusted P-value		0.0135[7]		
	P-value		0.000700		
Linolenic acid (C18:3)	Mean	8.35	8.54	1.15 – 14.7	2.00 – 12.5
	Range	6.85 – 10.2	7.01 – 10.4		
	Adjusted P-value		0.0415[7]		
	P-value		0.00410		
Arachidic acid (C20:0)	Mean	0.341	0.342	0.103 – 0.619	0 – 1.00
	Range	0.305 – 0.390	0.305 – 0.400		
	Adjusted P-value		0.959		
	P-value		0.913		
Eicosenoic acid (C20:1)	Mean	0.154	0.151	0.0549 – 0.319	0.140 – 0.316
	Range	0.122 – 0.191	0.115 – 0.186		
	Adjusted P-value		0.799		
	P-value		0.582		
Behenic acid (C22:0)	Mean	0.346	0.343	0.188 – 0.458	0.277 – 0.571
	Range	0.322 – 0.383	0.316 – 0.388		
	Adjusted P-value		0.784		
	P-value		0.561		
Lignoceric acid (C24:0)	Mean	0.172	0.173	0 – 0.310	Not Reported
	Range	0.103 – 0.244	0.0952 – 0.236		
	Adjusted P-value		0.959		
	P-value		0.924		

[1]Negative tolerance limits have been set to zero.
[2]Combined ranges are taken from published literature for soybeans (OECD, 2001; ILSI 2004; Taylor *et al*., 1999)
[3]Least Square Mean
[4]Range denotes the lowest and highest individual values across locations.
[5]False Discovery Rate (FDR) adjusted P-value
[6]Non-adjusted P-value
[7]Statistically significant difference; adjusted P-value< 0.05

A4. Amino Acids in Soybean Grain

Soybeans are a source of complete protein in human diets. A complete protein is one that contains significant amounts of all the essential amino acids that must be provided to the human body because of the body's inability to synthesize them. Soybean meal is also fed to animals primarily as a source of protein. Soybeans contain relatively high levels of certain essential amino acids that are deficient in many other common feedstuffs.

Total Amino Acids

Total levels of 18 amino acids were measured in soybean grain (Table 22). The mean values for methionine, cystine, tryptophan, threonine, isoleucine, histidine, valine, leucine, arginine, phenylalanine, glycine, alanine, aspartic acid, glutamic acid, proline, serine and tyrosine were within the statistical tolerance interval and/or the combined literature range. The mean value for lysine in 356043 soybean fell just outside the statistical tolerance interval (3.13% versus 3.10%). No statistically significant differences were observed between the 356043 and control soybean line for any of the amino acid mean values (adjusted P-value <0.05).

In conclusion, total amino acid analysis of soybean grain demonstrates that 356043 soybean is comparable to near isogenic and reference soybean lines. No unexpected differences in the amino acid composition of 356043 soybean grain were seen.

Table 22. Total Amino Acids in Soybean Grain

% Dry Weight Amino Acids		Control (Jack)	356043 Soybean	Tolerance Interval	Combined Ranges[1]
Methionine	Mean[2]	0.709	0.732	0.488 – 0.852	0.431 – 0.681
	Range[3]	0.664 – 0.745	0.620 – 0.978		
	Adjusted P-value[4]		0.487		
	P-value[5]		0.228		
Cystine	Mean	0.640	0.644	0.378 – 0.869	0.370 – 0.808
	Range	0.574 – 0.699	0.531 – 0.795		
	Adjusted P-value		0.909		
	P-value		0.797		
Lysine	Mean	3.07	3.13	1.98 – 3.10	2.29 – 2.86
	Range	2.79 – 3.38	2.88 – 3.55		
	Adjusted P-value		0.668		
	P-value		0.396		
Tryptophan	Mean	0.497	0.492	0.359 – 0.632	0.356 – 0.540
	Range	0.440 – 0.562	0.416 – 0.546		
	Adjusted P-value		0.863		
	P-value		0.708		
Threonine	Mean	1.91	1.94	1.57 – 2.21	1.25 – 1.89
	Range	1.69 – 2.09	1.76 – 2.09		
	Adjusted P-value		0.580		
	P-value		0.315		
Isoleucine	Mean	1.86	1.88	1.56 – 2.09	1.46 – 2.12
	Range	1.67 – 2.01	1.79 – 1.97		
	Adjusted P-value		0.780		
	P-value		0.539		
Histidine	Mean	1.28	1.33	0.897 – 1.41	0.878 – 1.22
	Range	1.14 – 1.43	1.13 – 1.42		
	Adjusted P-value		0.202		
	P-value		0.0622		
Valine	Mean	1.95	1.99	1.58 – 2.18	1.50 – 2.44
	Range	1.76 – 2.11	1.91 – 2.08		
	Adjusted P-value		0.356		
	P-value		0.150		
Leucine	Mean	3.12	3.16	2.53 – 3.52	2.20 – 4.00
	Range	2.87 – 3.38	3.05 – 3.35		
	Adjusted P-value		0.566		
	P-value		0.285		
Arginine	Mean	2.76	2.83	2.01 – 3.60	2.29 – 3.36
	Range	2.36 – 3.04	2.64 – 3.12		
	Adjusted P-value		0.176		
	P-value		0.0499		
Phenylalanine	Mean	2.10	2.11	1.74 – 2.43	1.60 – 2.24
	Range	1.86 – 2.30	1.98 – 2.29		
	Adjusted P-value		0.953		
	P-value		0.871		
Glycine	Mean	1.94	1.95	1.54 – 2.18	1.46 – 2.02
	Range	1.74 – 2.13	1.73 – 2.14		
	Adjusted P-value		0.784		
	P-value		0.555		

Table 22, continued. Total Amino Acids in Soybean Grain

% Dry Weight Amino Acids		Control (Jack)	356043 Soybean	Tolerance Interval	Combined Ranges[1]
Alanine	Mean	1.67	1.73	1.35 – 2.07	1.51 – 1.87
	Range	1.50 – 1.84	1.61 – 1.96		
	Adjusted P-value		0.152		
	P-value		0.0356		
Aspartic Acid	Mean	5.23	5.36	3.67 – 6.33	3.81 – 5.12
	Range	4.57 – 6.09	4.99 – 5.88		
	Adjusted P-value		0.565		
	P-value		0.272		
Glutamic Acid	Mean	7.92	8.00	6.04 – 9.54	5.84 – 8.72
	Range	7.21 – 8.73	7.54 – 8.59		
	Adjusted P-value		0.751		
	P-value		0.510		
Proline	Mean	2.55	2.55	1.85 – 2.70	1.69 – 2.61
	Range	2.26 – 2.75	2.42 – 2.71		
	Adjusted P-value		0.986		
	P-value		0.962		
Serine	Mean	2.24	2.29	1.85 – 2.71	1.63 – 2.48
	Range	2.07 – 2.47	2.14 – 2.43		
	Adjusted P-value		0.262		
	P-value		0.0873		
Tyrosine	Mean	1.49	1.50	0.908 – 1.69	1.02 – 1.62
	Range	1.30 – 1.66	1.36 – 1.69		
	Adjusted P-value		0.751		
	P-value		0.509		

[1]Combined ranges are taken from published literature for soybeans (OECD, 2001; ILSI 2004; Taylor *et al*., 1999)
[2]Least Square Mean
[3]Range denotes the lowest and highest individual value across locations
[4]False Discovery Rate (FDR) adjusted P-value
[5]Non-adjusted P-value

Acetylated Amino Acids

Because a similar GAT enzyme has been shown to acetylate some amino acids with low efficiency under certain *in vitro* conditions (Siehl *et al.,* 2005), the levels of the acetylated amino acids N-acetylglutamate (NAG) and N-acetylaspartate (NAA) were measured in 356043 and control soybean grain. Refer to Appendix 7 for more information on the potential for certain amino acids to act as substrates for GAT4601.

Results of the acetylated amino acid analyses are shown in Table 23. The mean values for NAA and NAG in 356043 soybean were above the statistical tolerance intervals for commercial soybeans. Mean values for NAA and NAG were also statistically different from those of control soybean. Literature values for NAA and NAG are not available.

NAA and NAG are normal components of human diets. NAA and NAG are present in a variety of foods including eggs, chicken, turkey, beef, chicken vegan chicken bouillon, mushrooms, hydrolyzed yeast, and soy sauce (Pioneer analysis of foods purchased in local grocery stores or through mail order suppliers; data not shown). There is no evidence to indicate that exposure to either NAA or NAG from these sources is associated with adverse effects in humans. A poultry study confirmed the nutritional wholesomeness and comparability of feed made from 356043 soybeans to conventional diets (see Section VIII-D). In addition, biological acetylation of amino acids is not a novel phenomenon. N-acetylglutamate is found in a number of organisms including plants. It is the first intermediate in the biosynthesis of arginine in prokaryotes, lower eukaryotes and plants (Caldovic and Tuckman, 2003). In eukaryotes, the most common post-translation modification to a protein is N-terminal methionine cleavage and subsequent N-terminal acetylation (Polevoda and Sherman, 2003). Acetylation of proteins is also commonly employed in the food industry to alter the solubility, water absorption capacity and emulsifying properties of protein concentrates (*e.g*., El-Adawy, 2000; Ramos and Bora, 2004).

In conclusion, analysis of N-acetylaspartate and N-acetylglutamate in soybean grain demonstrates that these two acetylated amino acids are elevated in 356043 soybean grain when compared to near isogenic control soybean grain. This observation is not unexpected, given the ability of GAT enzymes to acetylate glutamate and aspartate with low catalytic efficiency.

Table 23. N-Acetylaspartate and N-Acetylglutamate Levels in Soybean Grain

Analyte (mg/g dry weight)		Control (Jack)	356043 soybean	Tolerance Interval[1]
NAA	Mean[2]	0.00252	0.580	0 – 0.00227
	Range[3]	0.00106 – 0.0126	0.434 – 0.958	
	Adjusted P-value[4]		0.0001[6]	
	P-value[5]		0.0001	
NAG	Mean	0.00153	0.0116	0 – 0.00317
	Range	0.000876 - 0.00235	0.00484 – 0.0212	
	Adjusted P-value		0.0001[6]	
	P-value		0.0001	

[1]Negative tolerance limits have been set to zero
[2]Least Square Mean
[3]Range denotes the lowest and highest individual value across locations.
[4]False Discovery Rate (FDR) adjusted P-value
[5]Non-adjusted P-value
[6]Statistically significant difference, FDR adjusted P-value<0.05

Free Amino Acids

Amino acids in soybean seed are distributed between those which are incorporated into proteins (typically >99%) and those in the free amino acid pool (typically <1%) (Takahashi *et al*., 2003). Free amino acids are found in cells in amounts that vary according to the tissue and to the amino acid. The amino acids that play key roles in the incorporation and transfer of ammonia, such as glutamic acid, aspartic acid, and their amides, are often present in relatively high amounts, but the concentrations of the other free amino acids are very low.

In order to assess whether low levels of acetylation of aspartate and glutamate affected the overall amount or composition of the free amino acid pool, individual free amino acid levels were measured in 356043 and control soybean grain. Taurine, hydroxyl-L-proline, cysteine and L-cystine were analyzed and found to be nondetectable in both 356043 and control soybean grain, so these data are not shown. Results of the analysis for other free amino acids are presented in Table 24.

The means of each of the 23 individual free amino acids (L-aspartic acid, L-threonine, L-serine, L-asparagine, L-glutamate, L-glutamine, L-proline, L-glycine, L-alanine, L-valine, L-methionine, L-isoleucine, L-leucine, L-tyrosine, L-phenylalanine, γ-amino-n-butyric acid, ethanolamine, ammonia, L-ornithine, L-tryptophan), as well as the amount of total free amino acids, were within the statistical tolerance intervals for commercial soybeans. No statistically significant differences were observed between the 356043 and control soybean mean free amino acid values.

In conclusion, free amino acid analysis of soybean grain demonstrates that 356043 soybean is comparable to near isogenic and reference soybean lines. In particular, no differences were observed in the free amino acid levels of aspartate and glutamate.

Table 24. Free Amino Acids in Soybean Grain

Analyte (mg/g dry weight)		Control	356043 soybean	Tolerance Interval[1]
L-Aspartic Acid	Mean[2]	0.390	0.358	0 – 1.07
	Range[3]	0.158 – 0.780	0.149 – 0.635	
	Adjusted P-value[4]		0.567	
	P-value[5]		0.156	
L-Threonine	Mean	0.0307	0.0316	0 – 0.0553
	Range	0.0217 – 0.0495	0.0238 – 0.0431	
	Adjusted P-value		0.886	
	P-value		0.690	
L-Serine	Mean	0.0306	0.0333	0 – 0.0808
	Range	0.0215 – 0.0461	0.0246 – 0.0456	
	Adjusted P-value		0.712	
	P-value		0.334	
L-Asparagine	Mean	0.180	0.202	0 – 3.11
	Range	0.0837 – 0.322	0.0831 – 0.467	
	Adjusted P-value		0.752	
	P-value		0.383	
L-Glutamate	Mean	0.200	0.216	0 – 0.486
	Range	0.136 – 0.336	0.143 – 0.414	
	Adjusted P-value		0.567	
	P-value		0.133	
L-Glutamine	Mean	0.00940	0.0118	0 – 0.0724
	Range	0.000 – 0.0364	0.000 – 0.0448	
	Adjusted P-value		0.841	
	P-value		0.490	
L-Proline	Mean	0.0330	0.0407	0.00268 – 0.0762
	Range	0.0225 – 0.0486	0.0277 – 0.0662	
	Adjusted P-value		0.0765	
	P-value		0.00280[7]	
L-Glycine	Mean	0.0249	0.0272	0 – 0.0546
	Range	0.0212 – 0.0331	0.0203 – 0.0382	
	Adjusted P-value		0.567	
	P-value		0.151	
L-Alanine	Mean	0.0816	0.0819	0 – 0.301
	Range	0.0579 – 0.152	0.0515 – 0.155	
	Adjusted P-value		0.967	
	P-value		0.967	
L-Valine	Mean	0.169	0.190	0.0919 – 0.275
	Range	0.133 – 0.213	0.159 – 0.226	
	Adjusted P-value		0.112	
	P-value		0.00660[7]	
L-Methionine	Mean	0.0174	0.0168	0 – 0.0891
	Range	0.00887 – 0.0299	0.00964 – 0.0249	
	Adjusted P-value		0.886	
	P-value		0.698	
L-Isoleucine	Mean	0.0249	0.0271	0 – 0.108
	Range	0.0156 – 0.0393	0.0166 – 0.0521	
	Adjusted P-value		0.841	
	P-value		0.538	
L-Leucine	Mean	0.0222	0.0289	0 – 0.117
	Range	0.0106 – 0.0487	0.0120 – 0.0779	
	Adjusted P-value		0.525	
	P-value		0.0742	

Table 24, continued. Free Amino Acids in Soybean Grain

Analyte (mg/g dry weight)		Control	356043 soybean	Tolerance Interval[1]
L-Tyrosine	Mean	0.0296	0.0285	0 – 0.0526
	Range	0.0187 – 0.0472	0.0142 – 0.0484	
	Adjusted P-value		0.941	
	P-value		0.842	
L-Phenylalanine	Mean	0.0416	0.0413	0.0236 – 0.0900
	Range	0.0263 – 0.0747	0.0206 – 0.0799	
	Adjusted P-value		0.967	
	P-value		0.951	
γ-Amino-n-Butyric Acid	Mean	0.0880	0.0904	0 – 0.243
	Range	0.0497 – 0.148	0.0455 – 0.158	
	Adjusted P-value		0.959	
	P-value		0.884	
Ethanolamine	Mean	0.0880	0.0732	0 – 0.267
	Range	0.0188 – 0.206	0.0127 – 0.196	
	Standard Error	0.0317	0.0317	
	Adjusted P-value		0.525	
	P-value		0.0918	
Ammonia	Mean	0.0200	0.0202	0.0000726 – 0.0343
	Range	0.0147 – 0.0284	0.0147 – 0.0286	
	Adjusted P-value		0.967	
	P-value		0.913	
L-Ornithine	Mean	0.00210	0.00329	NA[6]
	Range	0.000916 – 0.00533	0.00159 – 0.00701	
	Adjusted P-value		0.525	
	P-value		0.0620	
L-Tryptophan	Mean	0.225	0.201	0 – 1.80
	Range	0.0744 – 0.454	0.0770 – 0.318	
	Adjusted P-value		0.791	
	P-value		0.430	
L-Lysine	Mean	0.0398	0.0427	0.00838 – 0.0654
	Range	0.0301 – 0.0498	0.0303 – 0.0520	
	Adjusted P-value		0.683	
	P-value		0.295	
L-Histidine	Mean	0.0532	0.0698	0 – 0.447
	Range	0.0209 – 0.108	0.0256 – 0.161	
	Adjusted P-value		0.432	
	P-value		0.0339[7]	
L-Arginine	Mean	0.537	0.646	0 – 2.19
	Range	0.297 – 1.24	0.280 – 1.59	
	Adjusted P-value		0.525	
	P-value		0.104	
Total Free Amino Acids	Mean	2.34	2.48	0 – 8.75
	Range	1.76 – 3.74	1.80 – 4.02	
	Adjusted P-value		0.712	
	P-value		0.335	

[1]Negative tolerance limits have been set to zero.
[2]Least Square Mean
[3]Range denotes the lowest and highest individual values across locations.
[4]False Discovery Rate (FDR) adjusted P-value
[5]Non-adjusted P-value
[6]Statistical analysis was not available (NA), due to no variation in the sample replicates for this analyte.
[7]Statistically significant difference, non-adjusted P-value<0.05

Amino Acid Conclusions

The distribution of total amino acids into three categories (amino acids that are incorporated into proteins, free amino acids, and acetylated amino acids) was calculated for 356043 and control soybean (Table 25). As expected, >99% of amino acids are incorporated into proteins and <1% of amino acids are in the free amino acid pool for both 356043 and control soybean. Together N-acetylaspartate and N-acetylglutamate make up less than 0.15% of the total amino acids in 356043 soybean. In addition, the levels of protein in forage and grain for 356043 soybean are comparable to control soybean (Tables 19 and 20, respectively). Therefore, the low levels of acetylation of aspartate and glutamate in 356043 soybean (Table 23) are not affecting amino acid incorporation into proteins or the level or composition of the free amino acid pool (Table 24), and 356043 soybean is comparable to control soybean with respect to amino acids.

Table 25. Distribution of Amino Acids in 356043 and Control Soybean Grain

	mg/g dry weight	**Control (Jack)**	**356043 soybean**
Total amino acids[1]	Mean (range) % of total amino acids	429.36 (384.68 – 474.16) (100%)	436.18 (404.97 – 475.89) (100%)
Free amino acids	Mean (range) % of total amino acids	2.340 (1.760 – 3.740) (0.54%)	2.480 (1.800 – 4.020) (0.57%)
Acetylated amino acids NAA + NAG	Mean (range) % of total amino acids	0.00405 (0.00194 – 0.01495) (0.0009%)	0.592 (0.439 – 0.979) (0.14%)
Incorporated amino acids (by subtraction[2])	Mean % of total amino acids	427.02 (99.45%)	433.11 (99.30%)

[1] Individual amino acids from Table 22 (% dry weight of tissue) were totaled and converted to mg/g to obtain total amino acid weight.
[2]The amount of incorporated amino acids was calculated by subtracting total free amino acids and acetylated amino acids from the total amino acid amount.

VIII-B. Isoflavones in Soybean Grain

Soybeans naturally contain a number of isoflavone compounds reported to possess biochemical activity, including estrogenic, anti-estrogenic, and hypocholesterolemic effects in mammalian species. Isoflavones (including phytoestrogens) have in the past been regarded as antinutrients; however, this is not universally accepted as they have also been reported to have beneficial anti-carcinogenic effects (Messina and Barnes, 1991).

Because phytoestrogens are key analytes in soybean, isoflavones were measured in 356043 and control soybean grain. Levels of acetylgenistin, acetyldaidzin, glycitein and acetylglycitin were below the limit of quantitation for the assay used in this analysis. Therefore, no statistical analyses were conducted on these analytes and data are not shown.

For all isoflavones analyzed (genistin, malonylgenistin, genistein, daidzin, malonyldaidzin, daidzein, glycitin and malonyglycitin) the mean values were within the statistical tolerance interval and/or combined literature range for commercial soybeans (Table 26).

In conclusion, isoflavone analysis of soybean grain demonstrates that 356043 soybean is comparable to near isogenic control soybean and reference soybean lines. No unexpected differences in the isoflavone composition of 356043 soybean grain were seen.

Table 26. Isoflavones in Soybean Grain

Analyte (mg/kg Dry Weight)		Control (Jack)	356043 soybean	Tolerance Interval[1]	Combined Literature Range
Genistin	Mean[2]	139	144	0 – 402	11.7 – 143
	Range[3]	71.0 – 223	<4.00 – 240		
	Adjusted P-value[4]		0.873		
	P-value[5]		0.733		
Malonylgenistin	Mean	1070	1120	0 – 2810	6.0 – 603
	Range	499 – 1560	652 – 1650		
	Adjusted P-value		0.331		
	P-value		0.128		
Genistein	Mean	10.4	12.3	0 – 32.3	0.1 – 22.6
	Range	<4.00 – 20.8	<4.00 – 31.2		
	Adjusted P-value		0.566		
	P-value		0.295		
Daidzin	Mean	58.6	67.3	0 – 343	0.7 – 83.6
	Range	34.5 – 86.3	51.4 – 92.3		
	Adjusted P-value		0.0779		
	P-value		0.0121		
Malonyldaidzin	Mean	703	790	0 – 2880	0.9 – 558
	Range	349 – 977	511 – 1190		
	Adjusted P-value		0.0437[6]		
	P-value		0.00540		
Daidzein	Mean	9.32	9.94	0 – 47.1	0.1 – 21.2
	Range	<4.00 – 18.3	<4.00 – 24.2		
	Adjusted P-value		0.863		
	P-value		0.711		
Glycitin	Mean	65.7	87.6	0 – 115	0.6 – 33.5
	Range	28.5 – 132	34.9 – 137		
	Adjusted P-value		0.141		
	P-value		0.0278		
Malonylglycitin	Mean	189	233	0 – 295	0.3 – 71.2
	Range	87.1 – 332	121 – 338		
	Adjusted P-value		0.0604		
	P-value		0.00820		

[1] Negative tolerance limits have been set to zero
[2] Least Square Mean
[3] Range denotes the lowest and highest individual values across locations.
[4] False Discovery Rate (FDR) adjusted P-value
[5] Non-adjusted P-value
[6] Statistically significant difference; adjusted P-value< 0.05

VIII-C. Key Antinutrients in Soybean Grain

Soybean grain contains several key antinutrients, such as oligosaccharides, lectins, phytic acid and protease inhibitors (OECD, 2001). Low molecular weight carbohydrates stachyose and raffinose are non-digestible oligosaccharides and are considered antinutrients. Lectins are proteins that bind to carbohydrate-containing molecules. Lectins in raw soybeans can inhibit animal growth and cause death. Phytic acid binds most of the phosphorus in soybeans which results in reduced bioavailability of phosphorus for nonruminant animals. In addition, phytic acid chelates mineral nutrients including calcium, magnesium, potassium, iron and zinc, rendering them unavailable to monogastric animals consuming beans. Protease inhibitors such as trypsin inhibitor interfere with the digestion of proteins, resulting in decreased animal growth.

Levels of key antinutrients were measured in 356043 soybean grain (Table 27). The mean values for all antinutrients measured (stachyose, raffinose, lectins, phytic acid and trypsin inhibitor) fell within the statistical tolerance interval for commercial soybean and/or the combined literature range.

In conclusion, antinutrient analysis of soybean grain demonstrates that 356043 soybean is comparable to near isogenic control soybean and reference soybean lines. No unexpected differences in the antinutrient composition of 356043 soybean grain were seen.

Table 27. Antinutrients in Soybean Grain

Analyte		Control (Jack)	356043 Soybeans	Tolerance Interval[1]	Combined Ranges[2]
Stachyose (% dry weight)	Mean[3]	3.14	3.12	2.65 – 4.78	1.21 – 3.50
	Range[4]	2.23 – 3.96	2.27 – 3.89		
	Adjusted P-value[5]		0.935		
	P-value[6]		0.842		
Raffinose (% dry weight)	Mean	0.619	0.637	0 – 1.99	0.212 – 0.661
	Range	0.344 – 0.986	0.351 – 1.11		
	Adjusted P-value		0.727		
	P-value		0.476		
Lectins (hemagglutinating units/mg)	Mean	5.80	5.97	0 – 11.4	0.105 – 9.04
	Range	1.95 – 12.4	0.615 – 13.7		
	Adjusted P-value		0.958		
	P-value		0.899		
Phytic acid (% dry weight)	Mean	1.22	1.20	0.459 – 1.78	0.634 – 1.96
	Range	0.924 – 1.80	0.830 – 1.57		
	Adjusted P-value		0.863		
	P-value		0.695		
Trypsin Inhibitor (trypsin inhibitor units/mg)	Mean	48.8	43.2	8.71 – 80.4	19.6 – 119
	Range	41.1 – 65.9	31.0 – 65.5		
	Adjusted P-value		0.0135[7]		
	P-value		0.000900		

[1]Negative tolerance limits have been set to zero
[2]Combined ranges are taken from published literature for soybeans (OECD, 2001; ILSI 2004; Taylor *et al*., 1999)
[3]Least Square Mean
[4]Range denotes the lowest and highest individual value across locations
[5]False Discovery Rate (FDR) adjusted P-value
[6]Non-adjusted P-value
[7]Statistically significant difference; adjusted P-value< 0.05

VIII-D. Poultry Feeding Study

Most soybean meal (97%) is used in animal feed, with the largest percentage of that (46%) going to poultry (OECD, 2001). To assess the wholesomeness of 356043 soybean grain when used as animal feed, a 42-day growth study was conducted in rapidly growing broiler chickens. The broiler chicken is a recognized model for assessing the wholesomeness of feedstuffs (ILSI, 2003).

The nutritional comparability of 356043 soybean was evaluated by comparing nutritional performance and carcass yield variables of broiler chickens fed diets containing processed fractions (meal, hulls, and oil) from the T7 generation of 356043 soybean (see Figure 2) with those fed diets produced with processed fractions from control non-transgenic near isoline (Jack) soybean. Diets produced from control soybean, 356043 soybean, and three non-transgenic commercial variety soybean fractions were fed to Ross x Cobb broilers (n = 120/group, 50% male and 50% female) for a period of 42 days. Diets were fed in three phases in accordance with standard commercial poultry farming practice: Starter (days 0-21), grower (days 22-35), and finisher (days 36-42). Starter diets contained 30% soybean meal, grower diets 26% soybean meal, and finisher diets 21.5% soybean meal; soybean hulls and oil were added at 1.0% and 0.5%, respectively, across all diets in each phase. Body weights were collected and feed intakes calculated every seven days during the growing period; weight gain, feed intake, and corrected feed:gain ratio (feed efficiency) were calculated at weekly intervals for days 0 to 42. Standard carcass and organ yield data were determined at the end of the feeding trial. Nutritional performance and carcass trait tolerance intervals were constructed using data from reference soybean groups. Data from control and 356043 soybean treatment groups were compared to each other and evaluated to determine if values were within the statistical tolerance intervals from this study. Statistical differences between broilers fed 356043 soybean and control soybean were also evaluated.

No statistically significant differences were observed in mortality, weight gain, feed efficiency (corrected for mortalities), and carcass yield variables between broilers consuming diets produced with 356043 soybean fractions and those consuming diets produced with near isoline control soybean fractions. Additionally, all response variables evaluated in control and 356043 soybean groups fell within the statistical tolerance intervals of the values observed in broilers fed diets produced with the reference soybean fractions. Based on the results from this study, 356043 soybean was nutritionally comparable to control soybean and commercial soybean varieties.

VIII-E. Conclusions on Compositional Assessment of 356043 Soybean

Extensive nutritional compositional analysis of forage and grain were conducted to evaluate the composition of 356043 soybean compared to a non-transgenic near isoline and four conventional soybean varieties. In total, data from 77 different analytical components (72 in grain, 5 in forage) were presented. Compositional analysis of 356043 soybean was used to evaluate any changes in the levels of key nutrients, isoflavones, or antinutrients.

Compositional analyses of forage included protein, fat, acid detergent fiber (ADF), neutral detergent fiber (NDF), and ash. Compositional analyses of grain included protein, fat, acid detergent fiber (ADF), neutral detergent fiber (NDF), ash, fatty acids, amino acids, two acetylated amino acids, free amino acids, isoflavones, and key antinutrients. Based on the compositional evaluation, the grain and forage of 356043 soybean are considered to be comparable to conventional soybean. An increase in two minor fatty acids, heptadecanoic acid (C17:0) and heptadecenoic acid (C17:1), was detected, but this is not unexpected as expression of the GM-HRA protein likely results in a slight shift in availability of the GM-HRA enzyme substrates, pyruvate and 2-ketobutyrate. These two compounds are also substrates for the enzyme complex that initiates oil biosynthesis.

Both C17:0 and C17:1 are found in vegetable oils. C17:0 is found in corn, soybean, sunflower, peanut and olive oils. C17:1 is found in olive oil. Both C17:0 and C17:1 fatty acids are also found in many different kinds of commonly consumed foods. C17:0 is commonly found in meat (lamb, beef, pork) and butter at levels comparable to those seen in the 356043 samples. The highest concentrations of C17:0 in foods are found in lamb, tofu, and butter. C17:1 is found in a wide variety of foods such as tofu, beef, cheese, and various baked products. There is no evidence to indicate that exposure to either C17:0 or C17:1 from these sources is associated with adverse effects in humans.

Levels of two acetylated amino acids, aspartate and glutamate, were also elevated. Again, this was not unexpected because GAT proteins are known to acetylate certain amino acids under specific *in vitro* conditions. Levels of N-acetylaspartate and N-acetylglutamate were very low in 356043 soybean (together less than 0.15% of the total amino acids), and these two analytes are components of commonly consumed foods including eggs, chicken, turkey, beef, chicken vegetable bouillon, mushrooms, yeast, and soy sauce.

In conclusion, 356043 soybean is comparable to commercial soybean varieties with respect to nutrient composition. A 42-day poultry broiler study confirmed the wholesomeness and nutritional comparability of 356043 soybean to conventional soybeans.

IX. Environmental Assessment and Impact on Agronomic Practices

IX-A. Environmental Assessment of the GAT4601 and GM-HRA Proteins

The GAT4601 protein sequence is derived from N-acetyltransferase protein sequences from *Bacillus licheniformis*, a gram positive saprophytic bacterium that is ubiquitous in soil and has a history of safe use for the production of food enzymes in the United States, Canada, and Europe (*e.g.,* alpha-amylase, cyclodextrin glycosyltransferase, hemicellulase, protease, pullulanase), biocontrol agents (EU Commission, 2000; FDA, 2001) and as a probiotic (Kritas *et al*., 2006; Alexopoulos *et al.,* 2004a and b). This indicates a previous history of exposure to and safe use of the source organism for GAT4601.

The GAT4601 protein in 356043 soybean is a member of the GNAT acetyltransferase superfamily that contains more than 10,000 representatives from plants, animals, bacteria, and fungi. Members of the GNAT superfamily all contain a highly conserved GNAT motif and high sequence diversity (Vetting *et al.,* 2005). GAT4601 is 84% identical and 94-95% similar at the amino acid level to the translated protein sequences of each of the three original *gat* genes from *Bacillus licheniformis* from which the *gat4601* gene was derived. GAT4601 retains the acetyltransferase enzyme function of the native proteins. The GAT4601 protein sequence does not have any homology to proteins that are toxic to humans or animals. Therefore, the GAT4601 protein is highly unlikely to pose a safety risk to beneficial organisms or the environment.

The GM-HRA protein in 356043 soybean encodes a modified version of the soybean acetolactate synthase (ALS) enzyme. ALS proteins are ubiquitous in bacteria, fungi, algae and plants (Friden *et al.,* 1985; Falco *et al*., 1985; Reith and Munholland, 1995; Mazur *et al*., 1987). Moreover, naturally occurring mutations in plant ALS proteins that confer herbicide tolerance have also been identified (for a review, see Duggleby and Pang, 2000). Several crops with herbicide tolerant ALS genes have been commercialized without any unexpected environmental consequences (for example, Clearfield®[3] wheat, Clearfield® sunflower, Clearfield® lentils, and STS®[4] soybean). None of these ALS enzymes or herbicide tolerant crops is known to pose an environmental safety concern. This indicates a previous history of exposure to and safe use of proteins similar to GM-HRA.

The GM-HRA protein introduced into 356043 soybean has minimal modifications compared to the endogenous soybean ALS enzyme from which it was derived. GM-HRA has two amino acid differences from the corresponding wild-type soybean ALS protein, plus an additional five amino acids derived from the translation of 15 nucleotides from the 5' untranslated region of the soybean *gm-als* gene. In addition to being derived from a soybean protein and having a high degree of similarity to other ALS plant enzymes, the GM-HRA protein sequence does not have any homology to proteins that are toxic to humans or animals. Therefore, the GM-HRA protein is highly unlikely to pose a safety risk to beneficial organisms or the environment.

In conclusion, there are likely to be negligible environmental effects due to the presence of the GAT4601 and GM-HRA proteins introduced in 356043 soybean.

[3] Clearfield® is a registered trademark of BASF.
[4] STS® is a registered trademark of DuPont or its affiliates.

IX-B. Weediness Potential of 356043 Soybean

Commercial soybean varieties in the United States are not considered weeds and are not effective in invading established ecosystems. Soybean has been grown throughout the world without any report that it is a serious weed. Cultivated soybean is unlikely to become a weed. Soybean seed rarely displays any dormancy characteristics and only under certain environmental conditions grows as a volunteer in the year following cultivation. If this should occur, volunteers do not compete well with the succeeding crop and can easily be controlled mechanically or chemically with herbicides other than glyphosate and ALS inhibitors. The soybean plant has no weedy tendencies and is non-invasive in natural habitats in the United States. It does not grow in unmanaged habitats.

There is little probability that 356043 soybean could become a problem weed. Various characteristics that might impart weediness potential were evaluated for 356043 and control soybean in comparative studies (Section VII). No differences were seen in characteristics such as seed germination, emergence, seedling vigor, seed shattering, yield, and disease/insect susceptibility. Assessment of these data detected no biologically significant differences between 356043 and control soybean indicative of a selective advantage that would result in increased weediness potential. Furthermore, post-harvest monitoring of field trial plots containing 356043 soybean has shown no differences in survivability or persistence of 356043 soybean as compared to control or conventional soybeans.

IX-C. Gene Flow Assessment

Due to the reproductive morphology (papilionaceous flower) of soybean, this crop exhibits a high percentage of self-fertilization. Natural outcrossing levels in soybean range from less than 0.5% to about 1% (Carlson and Lersten, 1987). As a reflection of this low potential for cross-pollination, Certified Seed Regulations allow foundation seed to be grown adjacent to other soybean cultivars as long as the distance is adequate to prevent mechanical mixing of the harvested seeds (see 7CFR 201.76, http://www.access.gpo.gov/nara/cfr/waisidx_01/7cfr201_01.html). Hence, the probability of gene transfer from 356043 soybean to other commercial soybean cultivars is very low.

Cultivated soybean can cross only with other members of its subgenus, *Soja* (reviewed in OECD, 2000). However, the potential for such gene flow to wild soybean relatives is limited by geographic isolation. Wild soybean species are native to China, Korea, Japan, Taiwan and the former USSR, and do not exist naturally in the United States. These species are not naturalized in North America, and although they could occasionally be grown in research plots, there are no reports of their escape from such plots to unmanaged habitats. There is therefore no potential for gene flow from cultivated 356043 soybean plants to wild soybean relatives in the United States.

There is no evidence of horizontal transfer of genetic material from soybean to other organisms through sexual mechanisms. The occurrence of potential horizontal gene transfer in the environment (bacteria, pathogens, *etc*.) has been studied using soil, water, and mammalian digestive tracts systems. These studies conclude that the risk of a possible transfer is irrelevant in an environmental risk assessment of transgenic soybean (Bogosian and Kane, 1991; Prins and Zadoks, 1994; Schluter *et al*., 1995; Jonas *et al*., 2001).

IX-D. Current Agronomic Practices for U.S. Soybean

D1. Soybean Production

The United States is the world's leading soybean producer and exporter. Farm value of U.S. soybean production in 2003/04 was $18.0 billion, the second-highest value among U.S.-produced crops, trailing only corn (USDA/ERS, 2006a). In 2004, soybean was planted on 75.1 million acres in 31 states, mostly in the Upper Midwest, Delta, and Southeast regions of the United States (USDA-NASS, 2006; USDA/ERS, 2006b). The average annual yields by state varied from 31 to 50 bushels/acre due to differences in rainfall, climate and soil productivity. However, yields in individual fields can be as high as 80 to 90 bushels/acre.

Soybeans are typically grown in the United States as row crops. Planting usually begins in late April or early-to-mid May, and harvesting generally occurs in late October to early November. Clean tillage has been the traditional method of field preparation, but no tillage and reduced tillage systems have become increasingly common. More than 60% of soybean acres are now grown under some form of conservation tillage (no-till, ridge-till or mulch-till). Conservation tillage practices provide the advantages of decreased soil compaction and fuel costs through reduction in use of heavy machinery, reduced soil erosion and better soil moisture conservation. Irrigation is not usually practiced (Van Doren and Reicosky, 1987); in 2002 only 7.5% of planted soybean acres were irrigated (USDA/ERS, http://www.ers.usda.gov/).

D2. Agricultural Chemical Use in Soybeans

Soybean is quite intensively managed, as evidenced by the chemical usage data from the 2004 USDA-NASS Agricultural Chemical Usage Report (http://www.nass.usda.gov/). Eleven states (81% of the total US soybean acreage) were included in this report: Arkansas, Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, Ohio and South Dakota. Phosphate and potash were the most commonly applied fertilizers in soybean; they were used on average in 26% and 23%, respectively, of the acreage in the states listed in the report. However, the acreage treated with fertilizer varies greatly by state, with some treating up to 63% of acres, and others only 5%. Because phosphate and potash are stable in soils, many growers don't apply them every year. Phosphate and potash are often applied on the corn crop in a corn-soybean rotation, using the same application for both crops. Herbicides were applied on 97% of the soybean acres. Glyphosate dominated herbicide usage, with application on 87% of planted acres. A total of 0.73 pounds of glyphosate were applied per acre per application (57.7 million total pounds). The next most frequently used herbicides were chlorimuron-ethyl (7% of acres treated), sulfentrazone (6%), trifluralin (5%), and pendimethalin (4%). Insecticides and fungicides are not widely used, with 4% of acres treated with insecticides and 1% of acres treated with fungicides.

D3. Weeds in Soybean

Growers must control weeds that compete with their crops for water, nutrients and sunlight. Depending on the crop and specific production region, weeds can decrease crop yields from 35% - 100% as well as interfere with harvest. More than 200 weed species, including 32 confirmed herbicide resistant species, threaten soybean yields in the United States each year (Weed Science Society of America, www.weedscience.org).

Soybeans are very sensitive to weed competition in the first few weeks after emergence. Weed scientists recommend that growers eliminate weed competition within three to five weeks after emergence. This helps insure that soybeans do not suffer irreversible yield

loss due to early competitive pressures from weeds (DeFelice and Butzen, 1997). Soybeans are very competitive with weeds once a canopy has developed, but early weeds can cause significant yield loss. Once weeds begin to impact soybean yield, each additional day they are allowed to compete can result in yield losses up to 1% per day (Hartzler, 2003). Narrow row planting has been helpful in increasing the competitive advantage over weeds, as the plants “shade out” later emerging weeds. However, because soybeans are relatively short, tall weeds can still shade the soybean plants, leading to yield losses.

Common weed problems in soybean fields include annuals (summer annual grass and broadleaf weeds such as foxtails or common ragweed), winter annuals and biennials such as marestail and biennial wormwood, and perennials such as quackgrass and Johnsongrass. Winter annuals, biennials and perennials are typically problems in no-till soybean fields. As growers shift to more no-till soybean production, weed specialists are seeing an increased frequency of these weeds. The perennials are particularly competitive and difficult to control, as these weeds re-grow every year from rhizomes or root systems. Dense populations of winter annuals can physically interfere with soybean planting and in some cases even reduce soil drying to such an extent that it delays soybean planting. Table 28 lists some of the most problematic weeds for soybean growers in 2004.

Table 28. Troublesome Weeds in Soybean (2004)

Weed Species	Total Area Treated[1] (Acres)	Infested Acres (%)
Annuals – Broadleaves		
Cocklebur	22,567,319	17.4
Velvetleaf	19,173,389	14.7
Pigweed, Redroot	17,483,951	13.4
Waterhemp, Common	16,350,256	12.6
Lambsquarters	16,257,487	12.5
Ragweed	10,870,854	8.4
Morningglory	8,197,887	6.3
Ragweed, Giant	7,902,766	6.1
Sunflower	6,110,363	4.7
Kochia	2,898,567	2.2
Smartweed	2,638,799	2.0
Nightshade, Black	2,315,295	1.8
Grasses		
Foxtail	39,420,858	30.3
Crabgrass	4,216,510	3.2
Barnyardgrass	3,547,755	2.7
Shattercane	2,117,204	1.6
Perennials		
Johnsongrass	8,763,039	6.7
Thistle, Canada	3,775,062	2.9
Thistle	2,270,897	1.7
Quackgrass	3,310,052	2.5

[1] The total soybean acreage in 2004 was 72 million acres. However, the total soybean herbicide-treated acreage was 130 million acres, due to multiple sprayings (USDA/ERS, 2004). Data from Doane Market Research, 2004.

D4. Weed Management in Soybean

Before the introduction of herbicides, weed control in soybean was through mechanical cultivation. Soil-incorporated and preemergence herbicides began to replace tillage and cultivation practices for soybeans in the 1960's (Carpenter and Gianessi, 1999). In the 1980's, postemergence herbicides became available to soybean growers, and their use has been steadily increasing. In 1988, 44% of soybean acres were treated with postemergence herbicides. By 2002, 91.8% of soybean acres were treated with postemergence herbicides (Carpenter and Gianessi, 1999; USDA/ERS, 2002).

The availability of postemergence herbicides together with herbicide tolerant crops has greatly facilitated the adoption of conservation tillage practices. Growers gained additional tools to control weeds; they could use a burndown treatment for early weed control and then apply one or two postemergence herbicides to control in-season weeds without tillage. This has ecological advantages in reducing soil erosion, lowering fuel and equipment costs, and allowing growers to plant larger acreages (since spraying is faster than tilling).

Narrow row spacing, which increases yields due to more efficient use of space, has also been made possible by postemergence herbicides. Mechanical cultivation is not possible with very narrow (~7.5 in) row spacing, but postemergence herbicides eliminate the need for mechanical cultivation. Narrow rows also improve weed control because the canopy closes more quickly and shades out later-emerging weeds.

Decisions about weed management may be the most complex ones that growers make, because each weed control option has trade-offs and affects the feasibility of using other options. Generally, growers must manage a wide array of broadleaf and grass weeds simultaneously. In selecting a weed management strategy, growers choose the most economical means to control weeds that does not decrease the quality or quantity of the crop. Therefore, growers will often use a combination of weed management techniques, including application of different herbicides, to effectively control weeds in their fields. The combination that a grower chooses to use depends upon factors such as: weed spectrum, level of infestation, soil type, cropping system, weather, and time and labor available for the treatment option. Data from the 2002 USDA-NASS Agricultural Chemical Usage Report indicate that, after herbicide usage, scouting was the most prevalent form of pest management practice for soybeans, with 92% of those surveyed scouting through general observation or deliberate scouting activities. Of those who completed scouting activities, 91% monitored for weeds. To avoid pest pressures (weeds, insects and diseases), 79% of growers rotated their crops. Of those surveyed, 55% used no-till/minimum till practices to manage pest pressures.

D5. Crop Rotation Practices

Most soybeans (67% in 2002, USDA/ERS, (2002)) are grown in rotation following corn, with another 18% in rotation with other row crops and small grains. About 14% are grown continuously. In areas with a longer-growing season such as the southern United States, soybeans are also double cropped, with soybeans planted after winter wheat is harvested (6% of planted acres in 2004, CTIC (2006)). Crop rotation aids in the management of diseases, insects and weeds and increases organic matter and soil fertility. In addition, crop rotation allows growers to diversify farm production to minimize market risks.

IX-E. Impact of the Introduction of 356043 Soybean on Agronomic Practices

E1. Impact on Cultivation and Management Practices

No negative impact is expected from the introduction of 356043 soybean on current cultivation and management practices for soybean. 356043 soybean has been shown to be comparable to conventional soybean in phenotypic, ecological and compositional characteristics (Sections VII and VIII). 356043 soybean is expected to be similar in its agronomic characteristics and have the same levels of resistance to insects and diseases as other commercial soybeans.

E2. Impact on Weed Control Practices

The commercialization of herbicide tolerant 356043 soybean is expected to have a beneficial impact on weed control practices, as growers will have another tool available to address their regional weed problems. 356043 soybean will enable growers to choose an optimal combination of glyphosate, ALS herbicide, and other complementary herbicides to best manage their individual weed populations. Growers value the glyphosate-resistant crop trait and the utility of glyphosate. The availability of 356043 soybean will enable growers to proactively manage weed populations while avoiding adverse population shifts of troublesome weeds or the development of resistance.

Glyphosate has proven to be the single most popular herbicide available for soybeans, and combinations of glyphosate with ALS-inhibiting herbicides or other herbicides with different modes of action and inherent crop tolerance will enable effective management of weed populations. Glyphosate and the ALS-inhibiting herbicide family are complementary in spectrum. Glyphosate is used to control weeds after emergence, whereas ALS inhibitors can be applied pre- or post-emergence. Both the glyphosate molecule and the ALS-inhibiting herbicide family are broad-spectrum herbicides. Both have excellent environmental profiles and low mammalian toxicity. Glyphosate has no soil residual, whereas the ALS inhibitors have a wide range of residual soil activity, which can be useful in controlling late-emerging weeds. Finally, ALS inhibitor herbicides are typically applied at very low use rates, in the ounces/acre or less range.

Approximately 87% of the U.S. soybean crop is already planted with glyphosate tolerant varieties. Therefore, no increase in the use of glyphosate over soybeans due to the commercialization of 356043 soybean is expected. The total per-season labeled use rate will not be higher than those used on other glyphosate tolerant soybeans.

Several ALS-inhibiting herbicides are currently registered for use in soybean and are used to control troublesome weeds, including those that are not effectively controlled with glyphosate. However, yields may suffer because soybean is not highly tolerant to ALS inhibitors. With 356043 soybean, crop tolerance will no longer be a problem. In addition, ALS-inhibiting herbicides are also used on sulfonylurea tolerant soybeans (STS®), which accounted for 3.8 million acres (~5% of soybeans planted) in 2006. Because current combined uses of ALS inhibitors on soybeans are relatively minor, we would expect the number of acres treated with ALS-inhibiting herbicides to increase with the introduction of 356043 soybean. Use rates of ALS inhibitors with 356043 soybean will be comparable to those currently labeled for use on soybean and other crops.

Growers will still select the best weed management for their specific situation, but with the availability of 356043 soybean, they will have more options available.

E3. Impact on Insect Control Practices

356043 soybean has been shown to be no different than conventional soybean in agronomic characteristics, and ecological observations have shown no changes in susceptibility to insect damage (Section VII). Therefore, there are no expected impacts on insect control practices for 356043 soybean.

E4. Impact on Crop Rotation Practices

We do not expect changes in crop rotation practices. There would be no expected increase in soybean-after-soybean plantings, and no projected increase in overall demand for glyphosate tolerant seed as the result of the introduction of an alternate tolerance gene. ALS-inhibiting herbicides can have recrop restrictions. However, as the ALS-inhibiting herbicides for 356043 soybean will be the same as the ones currently labeled for soybean, we would not expect any changes in crop rotation practices.

E5. Impact on Volunteer Management

We do not expect changes in volunteer management practices. 356043 soybean does not display any dormancy characteristics, and soybean in general grows as a volunteer in the year following cultivation only under certain environmental conditions. When this occurs, volunteers do not compete well with the succeeding crop and can easily be controlled mechanically or chemically with herbicides other than glyphosate or ALS inhibitors. Agronomic and phenotypic data confirm that, similar to conventional soybean, 356043 soybean does not have any weedy characteristics.

IX-F. Weed Resistance Management

F1. Evolution of resistant weeds

Crop pests respond to the repeated use of any mechanism that attempts to control them by evolving biological tactics to escape control. The widespread use of herbicides can lead to weed populations that are no longer susceptible. The first documented case of a weed evolving resistance in response to repeated use of an herbicide occurred in the mid 1960's (Ryan, 1970). During the 1970's, growers in the U.S. and Europe began to realize that one class of herbicides (triazines) that had successfully controlled many different weeds was no longer effective against certain populations of as many as 30 different weed species (LeBaron and McFarland, 1990; Bandeen *et al*., 1982). By 1990, weed scientists had evidence that at least 81 weed species contained individuals (biotypes) that had evolved resistance to one or more herbicides; 15 different classes of herbicides were no longer effective against at least one weed species (Holt and LeBaron, 1990). Currently more than 305 biotypes of herbicide resistant weeds occur around the world. Data from the international survey of herbicide resistant weeds can be found at http://www.weedscience.org/in.asp.

F2. Herbicide Resistance and Integrated Weed Management

The concerns associated with the use of herbicides and herbicide resistant crops include the evolution of resistant weeds, misapplication of the herbicide, herbicide drift, crop injury, carryover between growing seasons, costs, and the need for timely application. Herbicides are important tools for growers, and they should be used properly to preserve their effectiveness. Any weed management option that reduces herbicide-imposed selection pressure will reduce the rate of resistance development to the herbicide. By adopting practices such as mixing herbicides with different modes of action and crop rotation, selection pressure for resistant weeds can be reduced and the usefulness of herbicides preserved. These approaches are often part of Integrated Weed Management (IWM) programs.

Integrated Weed Management (IWM) is modeled after the more familiar Integrated Pest Management (IPM) used to control insects and plant pathogens. Both IWM and IPM are based on ecological and evolutionary principles.

IWM utilizes a range of weed control methods, including the following:

a. Avoid using the same herbicide or herbicides with the same mode of action multiple times per year or year after year.
b. Use tank-mixtures consisting of different herbicide types with overlapping weed spectra.
c. Use crop rotations because different crops allow different cultural and tillage options that compete much differently with weeds.
d. When using herbicides, use full label rates and tank mix partners.
e. Use clean seed and clean equipment to minimize spread of weed seed.
f. Monitor fields after herbicide applications for appearance of resistant weeds.
g. Control weeds before they form seed.
h. Where practical, use cover crops, set-aside programs, and other methods to reduce weed seed in soil.

IWM relies on using a variety of control measures to slow the evolution of resistance to a single control measure; therefore IWM is maximized when growers have access to the widest possible array of weed control tools. Pioneer and DuPont through its Crop Protection Chemicals business actively promote IWM techniques through communication, research, education and participation in industry coalitions such as the Herbicide Resistance Action Committee (HRAC, see Appendix 8, Section 8.4 on stewardship of 356043 soybean). With the introduction of 356043 soybean, new weed control options will be available to growers that are compatible with integrated weed management practices. This technology advances agricultural sustainability by helping growers achieve higher profits per acre while providing improved environmentally sound options for integrated weed management.

IX-G. Overall Environmental and Agronomic Practices Conclusions

A thorough characterization of 356043 soybean was performed, including molecular analysis, GAT4601 and GM-HRA protein level analysis, phenotypic and ecological evaluation, and nutrient composition evaluation. Assessment of the data generated supports the conclusion of no increased plant pest potential, phenotypic comparability, and familiarity as they relate to ecological risk assessment.

Due to the previous history of exposure to and safe use of organisms containing proteins similar to GM-HRA and GAT4601, as well as the safety assessments on the GAT4601

and GM-HRA proteins, no environmental effects due to the presence of the GAT4601 and GM-HRA proteins introduced in 356043 soybean are expected.

356043 soybean has been shown to be agronomically and ecologically similar to conventional soybeans, which have no weedy tendencies and are non-invasive in natural habitats. No differences were seen in characteristics such as seed germination, dormancy, emergency, seedling vigor, seed shattering, yield, and disease/insect susceptibility. Assessment of these data detected no biologically significant differences between 356043 and control soybean indicative of a selective advantage that would result in increased weediness or outcrossing potential. On the basis of these data, it is concluded that there is no increased plant pest potential of 356043 soyeans.

With the introduction of herbicide tolerant 365043 soybean, we do not expect a significant change in agronomic practices, with the exception of current weed control practices. We anticipate no increase in the usage of glyphosate, but we do expect an increase in the use of ALS-inhibiting herbicides, as this family of herbicides currently is not as widely used for control of weeds in soybean. With the introduction of 356043 soybean, new weed control options will be available to growers that are compatible with integrated weed management practices. This technology advances agricultural sustainability by helping growers achieve higher profits per acre while providing improved environmentally sound options for integrated weed management.

X. Adverse Consequences of Introduction

Pioneer Hi-Bred International, Inc. is unaware of any information indicating that 356043 soybean may pose a greater plant pest risk than conventional soybean. There are no adverse environmental consequences anticipated with its introduction. Thus we make the statement “unfavorable information: NONE,” and on the basis of the substantial benefits that this product offers for weed control options, Pioneer requests that 356043 soybean be granted nonregulated status under 7 CFR Part 340.

XI. References

Agresti, A. (2002) Categorical Data Analysis. A.J. Wiley and Sons, Second Edition, pp 14-23.

Alexopoulos, C., Georgoulakis, I.E., Tzivara, A., Kritas, S.K., Stochu, A. and Kyriakis, S.C. (2004a) Field evaluation of the efficacy of a probiotic containing *Bacillus licheniformis* and *Bacillus subtilis* spores, on the health status and performance of sows and their litters. Journal of Animal Physiology and Animal Nutrition 88: 381-392.

Alexopoulos, C., Georgoulakis, I.E., Tzivara, A., Kyriakis, C.S., Govaris, A. and Kyriakis, S.C. (2004b) Field evaluation of the effect of a probiotic-containing *Bacillus licheniformis* and *Bacillus subtilis* spores on the health status, performance, and carcass quality of grower and finisher pigs. Journal of Veterinary Medicine Series A Physiology Pathology Clinical Medicine 51: 306-312.

An, G., Mitra, A., Choi, H.K., Costa, M.A., An, K., Thornburg, R.W. and Ryan, C.A. (1989) Functional analysis of the 3' control region of the potato wound-inducible proteinase inhibitor II gene. Plant Cell 1: 115-122.

Bandeen, J.D., Stephenson, G.R. and Cowett, E.R. (1982) Discovery and distribution of herbicide-resistant weeds in North America. In LeBaron H.M. and Gressel J., eds, Herbicide Resistance in Plants. John Wiley & Sons, Inc., New York, pp 9-30.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57(1): 289-300.

Bernasconi, P., Woodworth, A.R., Rosen, B.A., Subramanian, M.V. and Siehl, D.L. (1995) A naturally occurring point mutation confers broad range of tolerance to herbicides that target acetolactate synthase. Journal of Biological Chemistry 270: 17381-17385.

Bogosian, G. and Kane, J.F. (1991) Fate of recombinant *Escherichia coli* K-12 strains in the environment. Advances in Applied Microbiology 36: 87–131.

Bowen, B.A., Bruce, W.B., Lu, G., Sims, L.E. and Tagliani, L.A. (2000) Synthetic promoters. United States Patent No. 6072050.

Bowen, B.A., Bruce, W.B., Lu, G., Sims, L.E. and Tagliani, L.A. (2003) Synthetic promoters. United States Patent No. 6555673 B1.

Bradshaw, L. D., Padgette, S.R., Kimbal, S.L. and Wells, B.H. (1997) Perspectives on glyphosate resistance. Weed Technology 11: 189-198.

Braerson, S. R., Rodriquez, D. J., Tran, M., Feng, Y., Biest, N. A. and Dill, G. M. (2002) Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. Plant Physiology 129: 1265-1275.

Caldovic, L. and Tuchman, M. (2003) N-Acetylglutamate and its changing role through evolution. Biochemical Journal 372: 279-290.

Carlson, J. B. and Lersten, N. R. (1987) Reproductive morphology. In Wilcox J.R., ed, Soybeans: Improvement, Production, and Uses, Second Edition. American Society of Agronomy, Madison, Wisconsin, pp 95-134.

Carpenter, J. and Gianessi, L. (1999) Herbicide tolerant soybeans: why growers are adopting Roundup Ready varieties. AgBioForum 2(2): 65-72.

Castle, L.A., Siehl, D.L., Gorton, R., Patten, P.A., Chen, Y.H., Bertain, S., Cho, H.-J., Duck, N.B., Wong, J., Liu, D. and Lassner, M.W. (2004) Discovery and directed evolution of a glyphosate tolerance gene. Science 304: 1151-1154.
Chaleff, R.S. and Mauvais, C.J. (1984) Acetolactate synthase is the site of action of two sulfonylurea herbicides in higher plants. Science 224: 1443-1445.

Chen, J. and Dellaporta, S. (1994) Urea-based plant DNA miniprep. In Freeling M. and Walbot V., eds, The Maize Handbook. Springer-Verlag, New York, pp 526-527.

Codex Alimentarius Commission (2003) Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, Italy, June 30 – July 5, 2003. Report of the third session of the Codex *ad hoc* intergovernmental task force on foods derived from biotechnology. Appendix III, Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants, and Appendix IV, Annex of the assessment of possible allergenicity, pp 47-60.

Coruzzi, G. and Last, R. (2000) Amino acids. In Buchanan, B., Gruissem, W. and Jones, R., eds, Biochemistry and Molecular Biology of Plants, Chapter 8. American Society of Plant Physiologists, Rockville, MD, pp 358-410.

Crameri, A., Raillard, S. A., Bermudez, E. and Stemmer, W.P.C. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391: 288-291.

Creason, G.L. and Chaleff, R.S. (1988) A second mutation enhances resistance of a tobacco mutant to sulfonylurea herbicides. Theoretical and Applied Genetics 76: 177-182.

Cuin, T.A., Miller, A.J., Laurie, S.A. and Leigh, R.A. (2003) Potassium activities in cell compartments of salt-grown barley leaves. Journal of Experimental Botany 54: 657-661.

CTIC (2006) Conservation Technology Information Center, National Summary Survey data, year 2004, http://www.conservationinformation.org/index.asp?site=1&action=crm_results.

Culpepper, A.S. (2004) Glyphosate-induced weed shifts. Abstracts of the Weed Science Society of America 44: 92.

De Block, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gosselé, V., Rao Movva, N., Thompson, C., Van Montagu, M. and Leemans, J. (1987) Engineering herbicide resistance in plants by expression of a detoxifying enzyme. EMBO Journal 6: 2513-2518.

DeFelice, M. and Butzen, S. (1997) Critical period of weed interference in soybeans and the use of postemergence herbicides. Crop Insights 7(1): 1-7.

Duggleby, R.G. and Pang, S.S. (2000) Acetohydoxyacid synthase. Journal of Biochemistry and Molecular Biology 33: 1-36.

Dyda, F., Klein, D.C. and Hickman, A.B. (2000) GCN5-related N-acetyltransferases: a structural overview. Annual Review of Biophysics and Biomolecular Structure 29: 81-103.

Dyer, W. E. (1994) Resistance to glyphosate. In Powles S.B. and Holtum J.A.M., eds, Herbicide Resistance in Plants. Lewis Publishers, Boca Raton, FL, pp 229-241.

El-Adawy, T.A. (2000) Functional properties and nutritional quality of acetylated and succinylated mung bean protein isolate. Food Chemistry 70(1): 83-91.

EU Commission (2000) Opinion of the Scientific Committee on Food on β-cyclodextrin produced using cycloglycosyltransferase from a recombinant *Bacillus licheniformis.* SCF/CS/ADD/AMI 52 Final, 13 July 2000, pp1-5, http://ec.europa.eu/food/fs/sc/scf/out58_en.pdf

Falco, S.C. and Dumas, K.D. (1985) Genetic analysis of mutants of *Saccharomyces cerevisiae* resistant to the herbicide sulfometuron methyl. Genetics 109: 21-35.

Falco, C.S. and Li, Z., (2003) S-adenosyl-L-methionine synthetase promoter and its use in expression of transgenic genes in plants. United States Patent Application No 2003/0226166

Falco, S.C., Dumas, K.D. and Livak, K.J. (1985) Nucleotide sequence of the yeast *ILV2* gene which encodes acetolactate synthase. Nucleic Acids Research 13: 4011-4027.

Falco, S.C., McDevitt, R.E., Chui, F.-C., Hartnett, M.E., Knowlton, S., Mauvais, C.J., Smith, J.K. and Mazur, B.J. (1989) Engineering herbicide-resistance acetolactate synthase. Developments in Industrial Microbiology 30: 187-194.

FDA (2001) Partial list of enzyme preparations that are used in foods. Center for Food Safety & Applied Nutrition, Office of Food Additive Safety. July 2001, pp 1-4, http://www.cfsan.fda.gov/~dms/opa-enzy.html.

Federal Register (1997) Environmental Protection Agency. 40 CFR Parts 700, 720, 721, 723, and 725, Microbial Products of Biotechnology, Final Regulations under the Toxic Substances Control Act; Final Rule, Vol 62, pp 17910-17958.

Feng, P. C. C., Tran, M., Chiu, T., Sammons, R., Heck, G. and Cajacob, C. (2004) Investigations into glyphosate-resistant horseweed (*Conyza canadensis*): retention, uptake, translocation and metabolism. Weed Science 52: 498-505.

Franz, J. E., Mao M.K. and Sikorski J.A. (1996) Introduction. In Franz, J. E., M. K. Mao and J. A. Sikorski, eds, Glyphosate: A Unique Global Pesticide. American Chemical Society, Washington, DC: pp 1-16.

Friden, P., Donegan, J., Mullen, J., Tsui, P., Freundlich, M., Eoyang, L., Weber, R. and Silverman, P.M. (1985) The *ilvB* locus of *Escherichia coli* K-12 is an operon encoding both subunits of acetohydroxyacid synthase I. Nucleic Acids Research 13: 3979-3993.

Gallie, D.R. and Walbot, V. (1992) Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. Nucleic Acids Research 20: 4631-4638.

Gaska, J. (2006) Soybean growth and development, http://soybean.uwex.edu/library/soybean/morphology/documents/04_soy_growth_development.pdf.

Hartnett, M.E., Chui, C.-F., Mauvais, C.J., McDevitt, R.E, Knowlton, S., Smith, J.K., Falco, S.C. and Mazur, B.J. (1990) Herbicide-resistant plants carrying mutated acetolactate synthase genes. In Managing Resistance to Agrochemicals. ACS Symposium Series 421: 459-473.

Hartnett, M.E., Chui, C.-F., Falco, S.C., Knowlton, S., Mauvais, C.J. and Mazur, B.J. (1991) Molecular analysis of sulfonylurea herbicide resistant ALS genes. In Caseley J.C., Cussans G.W. and Atkin R.K., eds, Herbicide Resistance in Weeds and Crops. Butterworth Heinemann, Oxford, pp 343-353.

Hartzler, B. (2003) Is your weed management program reducing your economic return? Iowa State University, Weed Science, http://www.weeds.iastate.edu/mgmt/2003/economics.shtml.

Heap, I. (2006) The International Survey of Herbicide Resistant Weeds, www.weedscience.com.

Heap, I. M. and LeBaron, H. (2001) Introduction and overview of resistance. In Powles S. B. and Shaner D. L., eds, Herbicide Resistance and World Grains. CRC Press, Boca Raton. pp 1-22.

Hilgenfield, K.L., Martin, A.R., Mortensen, D.A. and Mason, S.C. (2004) Weed management in glyphosate resistant soybean system: weed species shifts. Weed Technology 18: 284-291.

Holt, J.S. and LeBaron H.M. (1990) Significance and distribution of herbicide resistance. Weed Technology 4: 141-149.

Ibdah, M., BarIlan, A., Livnah, O., Schloss, J.V., Barak, Z. and Chipman, D.M. (1996) Homology modeling of the structure of bacterial acetohydroxy acid synthase and examination of the active site by site-directed mutagenesis. Biochemistry 35: 16282-16291.

ILSI (2003) Best practices for the conduct of animal studies to evaluate crops genetically modified for input traits. International Life Sciences Institute, Washington, DC, http://orig.ilsi.org/publications/pubslist.cfm?publicationid=575.

ILSI (2004) ILSI Crop Composition Database Version 3.0. International Life Sciences Institute, Washington, DC, http://www.cropcomposition.org/.

Jasieniuk, M. (1995) Constraints on the evolution of glyphosate resistance in weeds. Resistant Pest Management 7: 31-32.

Jonas, D.A., Elmadfa, I., Engel, K.H., Heller, K.J., Kozianowski, G., König, A., Müller, D., Narbonne, J.F., Wackernagel, W. and Kleiner, J. (2001) Safety considerations of DNA in food. Annals of Nutrition & Metabolism 45: 235-254.

Keil, M., Sanches-Serrano, J., Schell, J. and Willmitzer, L. (1986) Primary structure of a proteinase inhibitor II gene from potato. Nucleic Acids Research 14: 5641-5650.

Kishore, G. M. and Shah, D.M. (1988) Amino acid biosynthesis inhibitors as herbicides. Annual Review of Biochemistry 57: 627-663.

Klein, T.M., Wolf, E.D., Wu, R. and Sanford, J.C. (1987) High velocity microprojectiles for delivering nucleic acids into living cells. Nature 327: 70-73.

Kritas, S.K., Govaris, A., Christodoulopoulos, G. and Burriel, A.R. (2006) Effect of *Bacillus licheniformis* and *Bacillus subtilis* supplementation of ewe's feed on sheep milk production and young lamb mortality. Journal of Veterinary Medicine Series A Physiology Pathology Clinical Medicine 53: 170-173.

LaRossa, R.A. and Falco, S.C. (1984) Amino acid biosynthetic enzymes as targets of herbicide action. Trends in Biotechnology 2: 158-161.

LaRossa, R.A. and Schloss, J.V. (1984) The sulfonylurea herbicide sulfometuron methyl is an extremely potent and selective inhibitor of acetolactate synthase in *Salmonella typhimurium*. Journal of Biological Chemistry 259(4): 8753-8757.

LeBaron, H.M. and McFarland, J. (1990) Herbicide resistance in weeds and crops: overview and prognosis. American Chemical Society Symposium Series 421: 336-352.

Lee, K.Y., Townsend, J., Tepperman, J., Black, M., Chui, C.F., Mazur, B., Dunsmuir, P. and Bedbrook, J. (1988) The molecular basis of sulfonylurea herbicide resistance in tobacco. EMBO Journal 7(5): 1241-1248.

Levitt, G. (1978) Herbicidal sulfonamides. United States Patent No. 4127405.

Lorraine-Colwill, D.F., Powles, S.B., Hawkes, T.R., Hollinshead, P.H., Warner, S.A.J. and Preston, C. (2003) Investigations into the mechanism of glyphosate resistance in *Lolium rigidum*. Pesticide Biochemistry and Physiology 74: 62-72.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Marchler, G.H., Mullokandov, M., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Yamashita, R.A., Yin, J.J., Zhang, D. and Bryant, S.H. (2005) CDD: a Conserved Domain Database for protein classification. Nucleic Acids Research 33: D192-D196.

Mazur, B.J. and Falco, S.C. (1989) The development of herbicide resistant crops. Annual Review of Plant Physiology and Plant Molecular Biology 40: 441-470.

Mazur, B., Chui, C.-F. and Smith, J.E. (1987) Isolation and characterization of plant genes coding for acetolactate synthase, the target enzyme for two classes of herbicides. Plant Physiology 85: 1110-1117.

McCourt, J.A., Pang, S.S., King-Scott, J., Gudda, L.W. and Duggleby, R.G. (2006) Herbicide-binding sites revealed in the structure of plant acetohydroxyacid synthase. Proceedings of the National Academy of Sciences of the United States of America 103(3): 569-573.

Messina, M. and Barnes, S. (1991) The role of soy products in reducing risk of cancer. Journal of National Cancer Institute 83: 541-546.

Nandula, V. K., Reddy, K.N., Duke, S.O. and Poston, D.H. (2005) Glyphosate-resistant weeds: Current status and future outlook. Outlooks on Pest Management 12: 183-197.

Neuwald, A.F. and Landsman, D. (1997) GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. Trends in Biochemical Sciences 22(5):154-155.

Nickell, C.D., Noel, G.R., Thomas, D.J. and Waller, R. (1990) Registration of "Jack" soybean. Crop Science 30: 1365.

NRC/IOM (2004). Frameworks, findings, and recommendations. In Safety of Genetically Engineered Foods, Approaches to Assessing Unintended Health Effects, The National Academies Press, Washington, D.C., pp 175-190.

Odell, J.T., Nagy, F. and Chua, N.-H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313: 810-812.

OECD (2000) Consensus Document on the Biology of *Glycine max* (L.) Merr. (Soybean). Organisation for Economic Co-operation and Development, Paris, France. Series on harmonization of regulatory Oversight in Biotechnology No.15. ENV/JM/MONO(2000)9, http://www.olis.oecd.org/olis/2000doc.nsf/LinkTo/env-jm-mono(2000)9.

OECD (2001) Consensus document on compositional considerations for new varieties of soybean: key food and feed nutrients and anti-nutrients. Organisation for Economic Co-operation and Development, ENV/JM/MONO(2001)15, pp 1-30, http://www.olis.oecd.org/olis/2001doc.nsf/LinkTo/env-jm-mono(2001)15.

Owen, M. D. K. (1997) Risks and benefits of weed management technologies. In De Prado R., Jorrin J. and Garcia-Torres L., eds, Weed and Crop Resistance to Herbicides. Kluwer Academic Publishers, London, pp 291-297.

Owen, M.D.K. (2002) Glyphosate resistant waterhemp in Iowa. Proceedings of the North Central Weed Science Society 57: 210.

Owen, M.D.K. and Zelaya, I.A. (2005) Herbicide-resistant crops and weed resistance to herbicides. Pest Management Science 61: 301-311.

Padgette, S.R., Kolacx, K.H., Delannay, X., Re, D.B., Lavallee, B.J., Tinius, C.N., Rhodes, W.K., Otero, Y.I., Barry, G.F. and Eicholtz, D.A. (1995) Development, identification, and characterization of glyphosate-tolerant soybean line. Crop Science 35: 1451-1461.

Pang, S.S., Duggleby, R.G. and Guddat, L.W. (2002) Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors. Journal of Molecular Biology 317: 249-262.

Pang, S.S., Guddat, L.W. and Duggleby, R.G. (2003) Molecular basis of sulfonylurea herbicide inhibition of acetohydroxyacid synthase. Journal of Biological Chemistry 278(9): 7639-7644.

Pang, S.S., Guddat, L.W. and Duggleby, R.G. (2004) Crystallization of *Arabidopsis thaliana* acetohydroxyacid synthase in complex with the sulfonylurea herbicide chlorimuron ethyl. Acta Crystallographica D60: 153-155.

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

Polevoda B. and Sherman F. (2003) N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. Journal of Molecular Biology 325: 595-622.

Powles, S.B., Lorraine-Colwill, D.F., Dellow, J.J. and Preston, C. (1998) Evolved resistance to glyphosate in rigid ryegrass (*Lolium rigidum*) in Australia. Weed Science 46: 604-607.

Pratley, J.E., Urwin, N.A.R., Stanton, R.A., Baines, P.R., Broster, J.C., Cullis, K., Schafer, D.E., Bohn, J.A. and Krueger, R.W. (1999) Resistance to glyphosate in *Lolium rigidum*: I. Bioevaluation. Weed Science 47: 405–411.

Prins, T. and Zadoks, J. (1994) Horizontal gene transfer in plants, a biohazard? Outcome of a literature review. Euphytica 76(1/2): 133-138.

Ramos, C.M.P. and Bora, P.S. (2004) Functional characterization of acetylated Brazil nut *(Bertholletia excelsa* HBK) kernel globulin. Cienica e Tecnologia de Alimentos 24(1): 134-138.

Reith, M. and Munholland, J. (1993) Two amino-acid biosynthetic genes are encoded on the plastid genome of the red alga *Porphyra umbilicalis*. Current Genetics 23(1): 59-65.

Reith, M. and Munholland, J. (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Molecular Biology Reporter 7: 333-335.

Rey, M., Ramaiya, P., Nelson, B.A., Brody-Karpin, S.D., Zaretsky, E.J., Tang, M., de Leon, A.L, Xiang. H., Gusti, V., Clausen, I.G., Olsen, P.B., Rasmussen, M.D., Andersen, J.T., Jørgensen. P., Larsen, T.S., Sorokin, A., Bolotin. A., Lapidus, A., Galleron, N., Ehrlich, S.D. and Berka, R.M. (2004) Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. Genome Biology 5: R77.

Ryan, G.F. (1970) Resistance of common groundsel to simazine and atrazine. Weed Science 18: 614-616.

Santarem, E.R., Trick, H.N., Essig, J.S. and Finer, J.J. (1998) Sonication-assisted *Agrobacterium*-mediated transformation of soybean immature cotyledons: optimization of transient expression. Plant Cell Reports 17: 752-759.

Schluter, K., Futterer, J. and Potrykus, I. (1995) 'Horizontal' gene transfer from a transgenic potato line to a bacterial pathogen (*Erwinia chrysanthemi*) occurs - if at all - at an extremely low frequency. Bio/Technology 13(10): 1094-1098.

Shaner, D.L. (2000) The impact of glyphosate-tolerant crops on the use of other herbicides and on resistance management. Pest Management Science 56: 320-326.

Shaner, D.L. and Heap, I. (2002) Herbicide resistance in North America: the case for resistance to ALS Inhibitors in the United States. In: Agrochemical Resistance, ACS Symposium Series, 808, American Chemical Society, Washington, DC, pp 161-167.

Shaner, D.L. and O'Connor, S.L. (2000) The Imidazoline Herbicides. CRC Press, Boca Raton, FL, 290 pages.

Shaner, D. L., Feist, D.A. and Retzinger, E.J. (1997) SAMOA^{TM}: one company's approach to herbicide-resistant weed management. Pesticide Science 51: 367-370.

Siehl, D.L., Castle, L.A., Gorton, R., Chen, Y.H., Bertain, S., Cho, H.-J., Keenan, R., Liu, D. and Lassner, M.W. (2005) Evolution of a microbial acetyltransferase for modification of glyphosate: a novel tolerance strategy. Pest Management Science 61: 235-240.

Smeda, R.J. and Schuster, C.L. (2002) Differential sensitivity to glyphosate among biotypes of common waterhemp (*Amaranthus ridus* Sauer.). In Jacob, S., Dodd, H.J. and Moore, J.H., eds, 13th Australian Weeds Conference, Council of Australian Weed Science Societies, Perth, Australia, p 642.

Spelman, R.J. and Bovenhuis, H. (1998) Animal Genetics 29(2):77-84.

Stemmer, W.P.C. (1994) DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. Proc. Natl. Acad. Sci. USA 91: 10747-10751.

Stetter, J. (1994) Introduction. In Stetter J., ed, Herbicide Inhibiting Branched Chain Amino Acid Biosynthesis: Recent Developments. Springer-Verlag, New York, pp 1-2.

Stewart, C.N. Jr., Adang, M.J., All, J.N., Boerma, H.R., Cardineau, G., Tucker, D. and Parrott, W.A. (1996) Genetic transformation, recovery, and characterization of fertile soybean transgenic for a synthetic *Bacillus thuringiensis crylAc* gene. Plant Physiology 112: 121-129.

Takahashi, M., Uematsu, Y., Kashiwaba, K., Yagasaki, K., Hajika, M., Matsunaga, R., Komatsu., K., and Ishimoto, M. (2003) Accumulation of high levels of free amino acids in soybean seeds through integration of mutations conferring seed protein deficiency. Planta 217: 577-586.

Tan, S.Y., Evans, R.R., Dahmer, M.L., Singh, B.K. and Shaner D.L. (2005) Imidazolinone-tolerant crops: history, current status and future. Pest Management Science 61: 246-257.

Tan, S., Evans, R. and Singh, B. (2006) Herbicidal inhibitors of amino acid biosynthesis and herbicide-tolerant crops. Amino Acids 30: 195-204.

Taylor, N., Fuchs, R., MacDonald, J., Shariff, A. and Padgette, S. (1999) Compositional analysis of glyphosate-tolerant soybeans treated with glyphosate. Journal of Agricultural and Food Chemistry 47: 4469-4473.

Tranel, P.J. and Wright, T.R. (2002) Resistance of weeds to ALS-inhibiting herbicides: what have we learned? Weed Science 50: 700-712.

USDA/ERS (2002), http://www.ers.usda.gov/Data/ARMS/app/Crop.aspx.

USDA/ERS (2004), http://www.ers.usda.gov/Browse/FarmPracticesManagement/.

USDA (2005) USDA Nutrient Database for Standard Reference, Release 18. Nutrient Data Laboratory Home Page, http://www.ars.usda.gov/ba/bhnrc/ndl.

USDA/ERS (2006a), http://www.ers.usda.gov/Briefing/SoybeansOilCrops.

USDA/ERS (2006b), http://www.ers.usda.gov/Briefing/SoybeansOilcrops/background.htm.

USDA-NASS (2006), http://www.nass.usda.gov/index.asp.

Van Doren, D.M. and Reicosky, D.C. (1987) Tillage and irrigation. In Wilcox J.R., ed, Soybeans: Improvement, Production, and Uses, Second Edition. American Society of Agronomy, Madison, Wisconsin, pp 391-428.

VanGessel, M.J. (2001) Glyphosate-resistant horseweed from Delaware. Weed Science 49: 703-705.

Van Wert, S. (1994) Petition for the determination of nonregulated status for glufosinate resistant corn transformation events T14 and T25, http://www.aphis.usda.gov/brs/aphisdocs/94_35701p.pdf.

Vetting, M., de Carvalho, L.P.S., Yu, M., Hegde, S.S., Magnet, S, Roderick, S. and Blanchard, J. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Archives of Biochemistry and Biophysics 433: 212-226.

Yan, B., Srinivasa Reddy, M.S. and Collins, G.B. (2000) *Agrobacterium tumefaciens* – mediated transformation of soybean [*Glycine max* (L.) Merrill] using immature zygotic cotyledon explants. Plant Cell Reports 19: 1090-1097.

XII. Appendices 1-8

Appendix 1. Materials and Methods for Molecular Characterization of 356043 Soybean

To characterize the DNA insertion in 356043 soybean, Southern blot analysis was conducted. Individual plants of the T4 generation were analyzed to determine the number of each of the genetic elements of the expression cassettes inserted and to verify the integrity of the PHP20163A fragment was maintained upon integration. The integration pattern of the insertion in 356043 soybean was investigated with *Bgl* II and *Xba* I restriction enzymes. Southern blot analysis was conducted on individual plants of two generations, T4 and T5, to confirm insert stability across generations and to verify the absence of backbone sequences from plasmid PHP20163. The F3 generation was analyzed to verify insertion stability after crossing to an elite variety and to confirm Mendelian segregation of the insertion (discussed in Section V-E). All probes used for the analysis are indicated on the schematic maps of PHP20163A and PHP20163 (Figures 5 and 6) and outlined in Table 2.

1.1. 356043 Soybean Material

Seeds from the T4, T5, and F3 generations (Figure 2) of 356043 soybean were planted and leaf tissue harvested from individual plants was used for genomic DNA extraction.

1.2. Control Soybean Material

Seeds from the unmodified Jack soybean variety and the Elite 1 variety were planted and leaf tissue harvested from individual plants was used for genomic DNA extraction. Jack and Elite 1 control DNA was used as a negative control to help interpret hybridization results since the *gm-hra* cassette probes (SAMS, *gm-hra* and *gm-als* terminator) cross-hybridize with endogenous soybean sequences.

1.3. Reference Material

Plasmid DNA from PHP20163 was prepared from either an *E. coli* strain expressing dam methylase (Dam^+) or lacking dam methylase (Dam^-) (*E. coli* strains, Invitrogen, Carlsbad, CA) and was used as a positive control for Southern analysis to verify probe hybridization and to verify sizes of internal fragments. The plasmid stock was a copy of the plasmid used for microprojectile bombardment experiments to produce 356043 soybean and was digested with restriction enzymes to confirm the plasmid map. The probes used in this study (Table 2) were derived from plasmid PHP20163 or from a plasmid containing equivalent genetic elements.

DNA molecular weight markers for gel electrophoresis and Southern blot analysis were used to determine approximate molecular weights. For Southern analysis, DNA Molecular Weight Marker VII, digoxigenin (DIG) labeled (Roche, Indianapolis, IN), was used as a size standard for hybridizing fragments. ΦX174 RF DNA/Hae III Fragments (Invitrogen, Carlsbad, CA) was used as a molecular weight standard to determine sufficient migration and separation of the fragments on the gel.

1.4. Genomic DNA Extraction

Genomic DNA was extracted from leaf tissue harvested from individual plants as described above. The tissue was pulverized in tubes containing grinding beads using a Geno/Grinder™ (SPEX CertiPrep, Inc., Metuchen, NJ) instrument and the genomic DNA

isolated using a urea-based procedure (modification from Chen and Delalporta, 1994). Approximately 1 gram of ground tissue was extracted with 5 mL Urea Extraction Buffer (7 M Urea, 0.34 M NaCl, 0.05 M Tris-HCl, pH 8.0, 0.02 M EDTA, 1% N-Lauroylsarcosine) for 12-30 minutes at 37°C, followed by two extractions with phenol/chloroform/isoamyl alcohol (25:24:1) and one extraction with water saturated chloroform. The DNA was precipitated from the aqueous phase by the addition of 1/10 volume of 3 M NaOAc (pH 5.2) and 1 volume of isopropyl alcohol, followed by centrifugation to pellet the DNA. After washing the pellet twice with 70% ethanol, the DNA was dissolved in 0.5 mL TE buffer (10mM Tris, 1 mM EDTA, pH 7.5) and treated with 10 μg Ribonuclease A for 15 minutes at 37°C. The sample was extracted once with phenol:chloroform:isoamyl alcohol (25:24:1) and once with water saturated chloroform, followed by precipitation with isopropyl alcohol and washing with 70% ethanol. After drying, the DNA was re-dissolved with 0.5 mL TE buffer and stored at 4°C.

1.5. Quantitation of Genomic DNA

Following extraction, the DNA was quantified on a spectrofluorometer using PicoGreen® reagent (Molecular Probes, Inc., Eugene, OR) following a standard procedure. The DNA was also visualized on an agarose gel to confirm quantitation values from the PicoGreen® analysis and to determine DNA quality.

1.6. Phenotypic Identification of 356043 Soybean

Phenotypic analysis of 356043 soybean plants and control plants was carried out by western blot analysis using antibodies to the GAT4601 protein to confirm the absence or presence of the GAT4601 protein in material used for Southern blot analysis.

Total protein was extracted by grinding several leaf punches to homogeneity in 150 µl of protein extraction buffer (50mM Tris-HCl (pH 7.5), 0.1% SDS, and 10mM β-mercaptoethanol). An aliquot of each crude extract was mixed with LDS Sample Buffer and reducing agent (Invitrogen) and heated to approximately 95°C for 5 minutes. Proteins were separated by size under denaturing conditions through a NuPAGE™ Bis/Tris Gel system as described (Invitrogen). Selected molecular weight standards were used to determine sufficient migration in the gel and for molecular weight determination on the western blot (Invitrogen). The Bis/Tris gel was transferred to a nitrocellulose membrane using the method as described for the Novex™ XCell II™ Blot Module Western Transfer (Invitrogen). Alternatively, gels were stained with SimplyBlue™ SafeStain (Invitrogen) to visualize the proteins to verify equivalent sample loading.

The GAT4601 protein band was detected using the WesternBreeze® Chemiluminescent Western Blot Immunodetection Kit as described (Invitrogen). Primary monoclonal antibodies specific for the GAT4601 protein were used with the WesternBreeze® Kit. Bands were then visualized using a chemiluminescent substrate. Blots were exposed to X-ray film for one or more time points to detect protein bands. Purified GAT4601 protein was used as a positive control on the western blots. Plants were scored as positive for GAT4601 when a band of the appropriate size was present and scored as negative when the band was absent on the western blots.

A preliminary Southern blot analysis of DNA isolated from all 356043 soybean plants was used to verify the presence of both the *gat4601* and *gm-hra* genes. Methods for this preliminary characterization are described below. Final Southern blot analysis was carried out on a subset of 356043 soybean plants.

1.7. Digestion of DNA for Southern Blot Analyses

Genomic DNA samples extracted from selected 356043 soybean and control soybean plants were digested with restriction enzymes following a standard procedure. Approximately 2 μg of genomic DNA was digested in a volume of 100 μL using 50 units of enzyme according to manufacturer's recommendations. The digestions were carried out at 37°C for three hours, followed by ethanol precipitation with 1/10 volume of 3 M NaOAc (pH 5.2) and 2 volumes of 100% ethanol. After incubation at 4°C and centrifugation, the DNA was allowed to dry and re-dissolved in TE buffer. The reference plasmid, PHP20163, was spiked into a control plant DNA sample in an amount equivalent to approximately one or three gene copies per soybean genome and digested with the same enzyme to serve as a positive control for probe hybridization and to verify sizes of internal fragments on the Southern blot.

1.8. Electrophoretic Separation and Southern Transfer

Following restriction enzyme digestion, the DNA fragments produced were electrophoretically separated by size through an agarose gel and a molecular weight standard [ΦX174 RF DNA/Hae III Fragments (Invitrogen)] was used to determine sufficient migration and separation of the fragments on the gel. DIG labeled DNA Molecular Weight Marker VII (Roche), visible after DIG detection as described below, was used to determine hybridizing fragment size on the Southern blots.

Agarose gels containing the separated DNA fragments were depurinated, denatured, and neutralized in situ, and transferred to a nylon membrane in 20x SSC buffer (3M NaCl, 0.3 M Sodium Citrate) using the method as described for the TURBOBLOTTER™ Rapid Downward Transfer System (Schleicher & Schuell, Keene, NH). Following transfer to the membrane, the DNA was bound to the membrane by UV crosslinking (Stratalinker, Stratagene, La Jolla, CA).

1.9. DNA Probe Labeling for Southern Blot Hybridization

Probes for the SCP1 promoter, *gat4601*, *pin*II terminator, SAMS, *gm-hra*, and *gm-als* terminator were used to detect genes and elements within the insertion (Table 2). Backbone and hygromycin resistance gene cassette regions (*backbone 20163* and *hyg20163* probes) of the PHP20163 plasmid were used to verify absence of plasmid backbone DNA in 356043 soybean (Table 2). DNA fragments of the probe elements were generated by PCR from plasmid PHP20163 (Figure 6) or a plasmid with equivalent elements using specific primers. PCR fragments were electrophoretically separated on an agarose gel, excised and purified using a gel purification kit (Qiagen, Valencia, CA). DNA probes were generated from these fragments by PCR that incorporated a DIG labeled nucleotide, [DIG-11]-dUTP, into the fragment. PCR labeling of isolated fragments was carried out according to the procedures supplied in the PCR DIG Probe Synthesis Kit (Roche).

1.10. Probe Hybridization and Visualization

The DNA fragments bound to the nylon membrane were detected as discrete bands when hybridized to a labeled probe. Labeled probes were hybridized to the target DNA on the nylon membranes for detection of the specific fragments using the procedures essentially as described for DIG Easy Hyb solution (Roche). After stringent washes, the hybridized DIG-labeled probes and DIG-labeled DNA standards were visualized using CDP-Star Chemiluminescent Nucleic Acid Detection System with DIG Wash and Block Buffer Set (Roche). Blots were exposed to X-ray film for one or more time points to detect hybridizing fragments and to visualize molecular weight standards. Images were

then digitally captured by detection with the Luminescent Image Analyzer LAS-3000 (Fujifilm Medical Systems, Stamford, CT). Digital images were compared to original X-ray film exposures as verification for use in this report. The sizes of detected bands were documented for each digest and each probe.

1.11. Stripping of Probes and Subsequent Hybridizations

Following hybridization and detection, membranes were stripped of DIG-labeled probe to prepare the blot for subsequent re-hybridization to additional probes. Membranes were rinsed briefly in distilled, de-ionized water and then stripped in a solution of 0.2 M NaOH and 1.0% SDS at 40°C with constant shaking. The membranes were then rinsed in 2x SSC and either used directly for subsequent hybridizations or stored at 4°C or -20°C for later use. The alkali-based stripping procedure effectively removes probes labeled with the alkali-labile DIG.

Appendix 2. Description of Statistical Analyses

2.1. Trait Inheritance Data (Section V-E)

Based on Mendel's segregation law, the expected segregation ratios listed in Tables 5, 6 and 7 were tested by the statistic:

$$\chi^2 = \Sigma [(|o - e| - 0.5)^2 / e]$$

For a two-genotype case, the statistic can be expressed as:

$$\chi^2 = (| n_{obs(pos)} - n_{exp(pos)} | - 0.5)^2 / n_{exp(pos)} + (| n_{obs(neg)} - n_{exp(neg)} | - 0.5)^2 / n_{exp(neg)}$$

where o = observed frequency of the genotype; e= expected frequency of the genotype; and 0.5 = Yates correction for continuity (Yates, 1934). χ^2 follows a chi-squared distribution with one degree of freedom (df) for two-genotype cases (Agresti, 2002).

2.2. Agronomic Data from Experiments A and B (Section VII-B)

Data presented in Tables 13 and 14 were statistically analyzed using the following linear mixed model:

$y_{ijk} = \mu_i + r_k + \delta_{ik} + \varepsilon_{ijk}$

$r_k \sim iid\ N(0,\sigma^2_{loc})$, $\delta_{ik} \sim iid\ N(0,\sigma^2_{loc\times genotype})$, and $\varepsilon_{ik} \sim iid\ N(0, \sigma^2_{resid})$

Where μ_i denotes the mean of the ith genotype (pos or neg, fixed effect), r_k denotes the effect of the kth location (random effect), δ_{ik} denotes the interaction between the ith genotype and the kth location (random effect), and ε_{ijk} denotes the effect of the jth plot assigned the ith genotype in the kth location (random effect or residual). Notation ~iid $N(0, \sigma_a^2)$ denotes random variables that are identically independently distributed (iid) as normal with zero mean and variance σ_a^2.

2.3. Agronomic Data from Experiments C and D (Section VII-B)

Data were analyzed using a linear mixed model designed to account for the design effects of location and blocks within location. The linear mixed model assumed the entries were a fixed effect while the locations, blocks within locations and the entry by location interaction were random effects. A significant difference between the mean of 356043 soybean and the means of a set of comparable control lines was established with a P-value <0.05.

Early population, seedling vigor, disease incidence, insect damage, days to maturity, lodging, shattering score, final population (presented in Table 15), and yield data (presented in Table 16) were analyzed using the following linear mixed model:

$y_{ijk} = \mu_i + \ell_j + r_{k(j)} + (\mu\ell)_{ij} + \varepsilon_{ijk}$

$\ell_j \sim iid\ N(0,\sigma^2_{Loc})$, $r_{k(j)} \sim iid\ N(0,\sigma^2_{Rep})$, $(\mu\ell)_{ij} \sim iid\ N(0,\sigma^2_{Loc\times Ent})$, and $\varepsilon_{ijk} \sim iid\ N(0,\sigma^2_{plot})$

Where μ_i denotes the mean of the i^{th} entry (fixed effect), ℓj denotes the effect of the j^{th} location (random effect), $r_{k(j)}$ denotes the effect of the k^{th} block within the j^{th} location (random effect), $(\mu\ell)_{ij}$ denotes the interaction between the entries and locations (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} location (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 σ^2_a.

Plant height data (Table 15) were analyzed using the following linear mixed model:

$$y_{ijkl} = \mu_i + \ell_j + r_{k(j)} + (\mu\ell)_{ij} + \varepsilon_{ijk} + \delta_{ijkl}$$

$\ell_j \sim iid\ N(0,\sigma^2_{Loc})$, $r_{k(j)} \sim iid\ N(0,\sigma^2_{Rep})$, $(\mu\ell)_{ij} \sim iid\ N(0,\sigma^2_{Loc\times Ent})$, $\varepsilon_{ijk} \sim iid\ N(0,\sigma^2_{plot})$, and $\delta_{ijkl} \sim iid\ N(0,\sigma^2_{observations})$

Where μ_i denotes the mean of the i^{th} entry (fixed effect), ℓj denotes the effect of the j^{th} location (random effect), $r_{k(j)}$ denotes the effect of the k^{th} block within the j^{th} location (random effect), $(\mu\ell)_{ij}$ denotes the interaction between the i^{th} entries and j^{th} locations (random effect), ε_{ijk} denotes the effect of the plot assigned the i^{th} entry in the k^{th} block of the j^{th} location (random plot), and δ_{ijkl} denotes the effect of the plant assigned the l^{th} plant in the i^{th} entry in the k^{th} block of the j^{th} location (observational error).

2.4. Nutrient Composition Data from Experiment C (Section VIII-A, B and C)

Data were analyzed using a linear mixed model designed to account for the design effects of location and blocks within location. The linear mixed model assumes the entries are a fixed effect while the locations, blocks within locations and the entry by location interaction are random effects. A significant difference between the mean of 356043 soybean and the means of a set of comparable control lines was established with an FDR-adjusted P-value <0.05 for each analyte.

Composition data presented in Tables 19-24 and Tables 26-27 were analyzed using the following linear mixed model:

$$y_{ijk} = \mu_i + \ell_j + r_{k(j)} + (\mu\ell)_{ij} + \varepsilon_{ijk}$$

$\ell_j \sim iid\ N(0,\sigma^2_{Loc})$, $r_{k(j)} \sim iid\ N(0,\sigma^2_{Rep})$, $(\mu\ell)_{ij} \sim iid\ N(0,\sigma^2_{Loc\times Ent})$, and $\varepsilon_{ijk} \sim iid\ N(0,\sigma^2_{plot})$

Where μ_i denotes the mean of the i^{th} entry (fixed effect), ℓj denotes the effect of the j^{th} location (random effect), $r_{k(j)}$ denotes the effect of the k^{th} block within the j^{th} location (random effect), $(\mu\ell)_{ij}$ denotes the interaction between the entries and locations (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} location (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 σ^2_a.

Appendix 3. Materials and Methods for Determination of the GAT4601 and GM-HRA Protein Concentrations

Soybean plant tissue concentrations of the GAT4601 and GM-HRA proteins were determined using enzyme linked immunosorbent assays (ELISA) developed at Pioneer Hi-Bred International, Inc., Johnston, IA, USA.

3.1. Storage and Processing of ELISA samples

Upon receipt, all plant tissue samples were stored in temperature-monitored freezers at <-10°C.

Forage samples were coarsely homogenized on dry ice using a Stephan VCM 12 (Stephan Machinery Singapore Pte Ltd, Singapore) blender for approximately 2 minutes and sub-sampled.

All samples were lyophilized at <-12°C under vacuum. The lyophilization time varied between 18 to 72 hours depending on the sample size and tissue type.

Forage and grain tissues were finely ground for approximately 60 seconds using a GenoGrinder (BT&C/OPS Diagnostics, Metuchen, NJ, USA). Root samples were finely ground for approximately two minutes using a Harbil 5G High-Speed Mixer (IDEX Corporation, Northbrook, IL).

Between lyophilization and grinding, samples were stored frozen in temperature-monitored freezers at <-10°C.

3.2. Protein Extraction from Processed Soybean Tissues

Processed soybean tissues were weighed into 1.2 ml tubes at the following target weights: 10 mg for grain and 20 mg for root and forage tissues. Each sample was extracted with 600 microliters (µl) of chilled (2-8°C) GAT ELISA Buffer (GEB) which is comprised of 25 millimolar (mM) HEPES (Sigma-Aldrich, Inc., St. Louis, MO), 25 mM CAPS (Mallinckrodt Baker, Inc., Phillipsburg, NJ), 150 mM NaCl, 0.1% Tween-20 (ICI Americas, Inc., London, UK), 0.5% PVP-40 (Sigma-Aldrich, Inc.), 10% ethylene glycol, and 5% StabilZyme Select (SurModics, Inc., Eden Prairie, MN). Two 5/32" steel balls were added to the tube and the samples were homogenized with a single 30 second cycle at 1500 strokes per minute using a SPEX Certiprep GenoGrinder^3 (BT&C/OPS Diagnostics, Metuchen, NJ). Following centrifugation, supernatants were removed, diluted and analyzed for extractable GAT4601 and GM-HRA protein concentrations using a specific enzyme linked immunosorbent assay (ELISA) for each protein.

3.3. Determination of GAT4601 Protein Concentrations

The GAT4601 ELISA utilizes a sequential "sandwich" format for the quantitation of GAT4601 in soybean plant tissue extracts. Standards (analyzed in triplicate wells) and diluted tissue extracts (analyzed in duplicate wells) were incubated in a 96-well stabilized ELISA plate that was pre-coated with a GAT4601-specific antibody. Unbound substances were washed from the plate and a different GAT4601-specific antibody that was conjugated to the enzyme horseradish peroxidase (HRP) was added to each well. Bound GAT4601 protein was sandwiched between the antibody coated on the plate and the antibody-HRP conjugate. At the end of the incubation, unbound substances were washed from the plate. Detection of the bound GAT4601-antibody complex was

accomplished by the addition of a substrate that generated a colored product in the presence of HRP. The reaction was stopped with stop solution (hydrochloric acid) and the optical density of each well was determined using a Molecular Devices plate reader (Molecular Devices Corporation, Sunnyvale, CA) with a wavelength setting of 450 nm minus 650 nm. SoftMax Pro software (Molecular Devices Corporation) was used to perform the calculations that generated the quadratic fit of the standard curve and converted the sample optical density (OD) values to GAT4601 protein concentration values. The mean duplicate well values in ng/ml were used in the calculation of the reported GAT4601 concentration of each sample (ng/mg dry weight). A characterized sample extract (QCE) was included on each plate as a control for data acceptance.

The quantitative range for the assay was 0.4 ng/ml ± 10% to 8.0 ng/ml ± 10% which allows a full standard curve range of 0.36 to 8.8 ng/ml. The lower limit of quantitation (LLOQ) in ng/mg dry weight for each tissue was based on extraction volume (µl) to weight ratios, the limit of quantitation for the ELISA in ng/ml, and the dilutions used for analysis. In this study, the sample LLOQ on a ng/mg dry weight basis for GAT4601 was 0.22 ng/mg dry weight for grain and 0.11ng/mg dry weight for root and forage.

3.4. Determination of GM-HRA Protein Concentrations

The GM-HRA ELISA method utilized a sequential “sandwich” format for the determination of the presence of GM-HRA protein in soybean plant tissue extracts. In this assay, sample extracts were incubated for one hour in stabilized 96-well plates that were precoated with a GM-HRA-specific antibody. Unbound substances were washed from the plate, and a different GM-HRA- antibody that had been conjugated to the enzyme horseradish peroxidase (HRP) was added to the wells. Bound GM-HRA protein was sandwiched between the antibody coated on the plate and the antibody-HRP conjugate. At the end of the one hour incubation, unbound substances were washed from the plate. Detection of the bound GM-HRA protein-antibody complex was accomplished by the addition of a substrate solution, which generated a colored product in the presence of HRP. The reaction was stopped with stop solution (hydrochloric acid) and the optical density of each well was determined using a Molecular Devices (Molecular Devices Corporation) plate reader with a wavelength setting of 450 nm minus 650 nm. SoftMax Pro software (Molecular Devices Corporation) was used to perform the calculations that generated the quadratic fit for the standard curve and converted the sample OD values to GM-HRA protein concentration values. The mean concentration from the duplicate wells in ng/ml was used in the calculation of the reported GM-HRA concentration of each sample (ng/mg dry weight).

The quantitative range for the assay was 1.0 ng/ml ± 10% to 20 ng/ml ± 10% which allows a full range of 0.9 to 22 ng/ml. The lower limit of quantitation (LLOQ) in ng/mg dry weight for each tissue was based on extraction volume (µl) to weight ratios, the limit of quantitation for the ELISA in ng/ml, and the dilutions used for analysis. In this study, the sample LLOQ on a ng/mg dry weight basis for GM-HRA was 0.54 ng/mg dry weight for grain and 0.27 ng/mg dry weight for root and forage.

3.5. Calculations for Determining Concentrations of GAT4601 and GM-HRA Proteins

SoftMax Pro software (Molecular Devices Corporation) was used to perform the calculations required to convert the OD values obtained by the microtiter plate reader to concentration values.

A standard calibration curve was included on each ELISA plate. The equation for the standard curve was generated by the software, which used a quadratic fit to relate the mean OD values obtained for the standards to the respective standard concentrations (ng/ml).

The regression equation was applied as follows: $y = Cx^2 + Bx + A$

Where x = known standard concentration and y = respective mean absorbance value (OD).

Interpolation of the sample concentration (ng/ml) was done by solving for x using the values for A, B, and C that were determined for the standard curve.

$$\text{Sample concentration (ng/ml)} = \frac{-B + \sqrt{B^2 - 4C * (A - \text{sample OD})}}{2C}$$

i.e.: given curve parameters of A = 0.143, B = 0.00625, C= -0.00000399 and a sample OD = 0.249

$$\text{Concentration} = \frac{-0.00625 + \sqrt{0.00625^2 - 4\ (-0.00000399)\ (0.143 - 0.249)}}{2\ (-0.00000399)} = 17.1 \text{ ng/ml}$$

Sample concentration values obtained from the SoftMax Pro software were converted from ng/ml to ng/mg tissue dry weight as follows:

ng/mg tissue dry weight = ng/ml *extraction volume (ml)/mg tissue dry weight

i.e.: if the concentration = 17.1 ng/ml, extraction volume = 0.60 ml, and a tissue weight = 10.0 mg

ng/mg tissue dry weight = 17.1 ng/ml * 0.60 ml/10.0 mg = 1.03 ng/mg tissue dry weight

Appendix 4. Methods and Results for Characterization of the GAT4601 and GM-HRA Proteins

4.1. Protein Purification from 356043 Soybean Plant Tissue

4.1a. GAT4601

The GAT4601 protein was extracted from approximately 30 grams of 356043 soybean leaf tissue in 180 ml of homogenization buffer consisting of 50 mM HEPES, pH 7.8, 150 mM NaCl, 1 mM EDTA, 5% glycerol, and Complete Protease Inhibitor Cocktail (Roche Applied Science, Indianapolis, IN). Particulate material in the resulting slurry was removed by filtering through four layers of cheese cloth. The filtrate was further clarified by centrifugation for 20 minutes at 30,500*g*.

The GAT4601 protein was partially purified from the extract using an immunoaffinity column that had been prepared by coupling GAT-specific mouse monoclonal antibodies to AminoLink Plus Coupling Gel (Pierce Biotechnology Inc., Rockford, IL) according to the manufacturer's instructions. Collected fractions (2 ml each) were analyzed by SDS-PAGE and western blot. The fractions containing the GAT4601 protein were pooled and concentrated using a Nanosep 3K concentrator (Pall Corporation, East Hills, NY).

4.1b. GM-HRA

For final purification, the GM-HRA protein was extracted from approximately 30 grams of 356043 soybean tissue in 180 milliliters (ml) of homogenization buffer consisting of 50 mM Tris-HCl, pH 7.5, 5 mM sodium pyruvate, 10 µM flavin adenine dinucleotide, 1 mM EDTA, 5% glycerol, 5 mM magnesium chloride, 50 mM sodium chloride, 5% (w/w) polyvinylpolypyrrolidone, and EDTA-free Complete Protease Inhibitor Cocktail (Roche Applied Science, Indianapolis, IN, USA). Particulate material in the resulting slurry was removed by filtering through four layers of cheese cloth. The filtrate was further clarified by centrifugation for 20 minutes at 30,500*g*.

The GM-HRA protein was partially purified from the extract using an immunoaffinity column that had been prepared by coupling GM-HRA-specific mouse monoclonal antibodies to AminoLink Plus Coupling Gel (Pierce Biotechnology, Inc.) according to the manufacturer's instructions. Collected fractions (2 ml each) were analyzed by SDS-PAGE and western blot. The GM-HRA protein-containing fractions were pooled and concentrated using a Nanosep 10K concentrator (Pall Corporation).

4.2. Protein Expression in *E. coli* and Purification

4.2a. GAT4601

The GAT4601 protein was expressed in *E. coli* strain BL21 (DE3) as a soluble protein and was purified at Aldveron (Fargo, ND) using cation exchange chromatography, anion exchange chromatography, and hydrophobic interaction chromatography, followed by diafiltration into 0.1 M ammonium bicarbonate, pH 7.8. The purified protein solution was either used for testing (endotoxin content) or lyophilized, and the resulting powder was stored at -80°C. For all the characterization studies the protein was used in the following form: approximately 1 mg of the lyophilized powder was dissolved in 1 ml of 100 mM KCl, 10% methanol, and 25 mM HEPES, pH 7.2.

4.2b. GM-HRA

For GM-HRA, the mature form of the protein (excluding the chloroplast transit peptide sequence) was produced. The GM-HRA protein was expressed in *E coli* strain BL21 (DE3) RIPL as a fusion protein containing a His-T7 tag and was purified using a immobilized metal affinity column (Ni-NTA His Bind resin, Novagen, EMD Biosciences, Inc., San Diego, CA) followed by cleavage of the fusion tag with thrombin (Calbiochem, # 605195, EMD Biosciences, Inc.), which was used in-solution at 10 units per mg of protein, after it was eluted from the Ni-NTA column. The cleaved tag and the thrombin were removed from the purified protein by diafiltration. The thrombin cleavage of the tag resulted in one extra N-terminal glycine residue as the N-terminal amino acid, which is not found in the mature GM-HRA protein sequence. The material was then dialyzed into 0.1 M ammonium bicarbonate pH 7.5 and then lyophilized.

4.3. Method for Determination of Protein Concentration

Total protein concentration of GM-HRA and GAT4601 proteins was determined using Coomassie Plus Protein Assay Reagent (Pierce Biotechnology, Inc.). Bovine serum albumin (BSA) was used as the protein standard. After incubation for approximately five minutes, absorbance was measured at 595 nm.

4.4. SDS-PAGE and Western Blot Methods

4.4a. SDS-PAGE Method

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed by first mixing GAT4601 and GM-HRA protein samples from both sources with Laemmli sample buffer (Gradipore Limited, Australia) containing 100 mM dithiothreitol and heating the solution at 100°C for approximately five minutes. The prepared protein samples were loaded into a 10-20% gradient Ready Gel Tris-HCl gel (Bio-Rad Laboratories, Inc., Hercules, CA). PageRuler pre-stained Protein Ladder (Fermentas, Inc., Hanover, MD) molecular weight markers were loaded on the gel to provide a visual estimate of molecular weight. Electrophoresis was conducted using the Ready Gel Cell system (Bio-Rad Laboratories, Inc) with Tris-glycine running buffer (Gradipore Limited) and a constant 150 volts (V) for approximately 60 minutes or until the dye front neared the bottom of the gel. Upon completion of electrophoresis, the gels were either removed from the gel cassette and prepared for western blot analysis or were stained with Coomassie Blue.

4.4b. Western Blot Method

Following electrophoresis, the resulting gel was soaked in transfer buffer (48 mM Tris-HCl at pH 8.6, 39 mM glycine, 0.0375% sodium dodecyl sulfate, and 20% methanol) for approximately 10 to 20 minutes. The polyvinylidene difluoride (PVDF) membrane was placed in 100% methanol briefly, followed by immersion in transfer buffer for 10-15 minutes. The Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell system (Bio-Rad Laboratories, Inc.) was used to electrophoretically transfer the protein bands from the gel to the membrane at 10-11 V for approximately 30-45 minutes for microbial and plant-derived GAT4601 proteins and 60 minutes for microbial and plant-derived GM-HRA.

Following transfer, the membrane was washed and then blocked for approximately 60 minutes by incubating in phosphate-buffered saline solution with 0.5% Tween-20 containing 3% or 5% (w/v) non-fat dry milk. The blocked membrane was washed three to four times for approximately five minutes each in Tris-buffer and then incubated for approximately 60 minutes with the primary antibody (mouse monoclonal GM-HRA-specific antibody R6961 or mouse monoclonal GAT-specific antibody 18F9). The

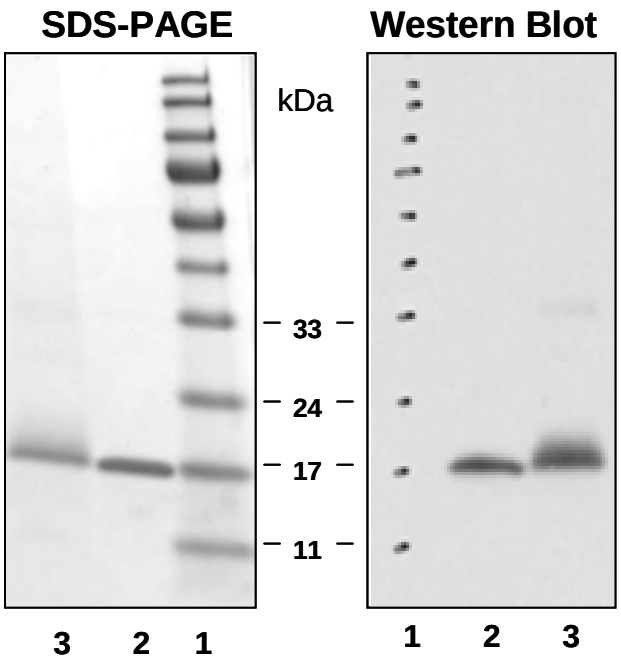
unbound antibody was washed from the membrane and then incubated for approximately 60 minutes with the secondary antibody (anti-mouse IgG-horseradish peroxidase (HRP) conjugate). The membrane was then washed for approximately five minutes in Tris-buffer. Prior to the detection, the blot was soaked in phosphate buffered saline solution for approximately five minutes each and developed with SuperSignal West Pico Chemiluminescent Substrate (Pierce Biotechnology, Inc.) according to the manufacturer's instructions. The molecular weight of the recognized protein band was estimated using PageRuler prestained protein ladder (Fermentas, Inc.).

4.5. Results of SDS-PAGE and Western Analyses

4.5a. GAT4601

SDS-PAGE analysis of the microbially derived and 356043 soybean plant-derived GAT4601 resulted in a protein band migrating as expected at approximately 17 kDa (Appendix 4, Figure 1). Because only one band was observed in the lane, purity of the microbial GAT4601 relative to total protein present was visually estimated as greater than 95%. Western blot analysis of the microbial and 356043 soybean plant-derived GAT4601 proteins resulted in the detection of an immunoreactive protein band migrating approximately at 17kDa (Appendix 4, Figure 1). Therefore, SDS-PAGE and western analysis confirmed that both microbially expressed and 356043 soybean plant-derived GAT4601 proteins had the expected molecular weight and immunoreactivity and were equivalent in these analyses.

Appendix 4, Figure 1. SDS-PAGE and Western Blot Analysis of Plant-Derived and Microbially Expressed GAT4601



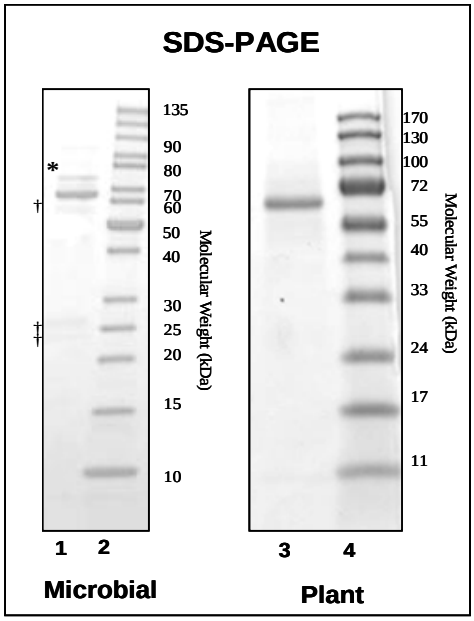
Lane #	Description
1	Fermentas Page Ruler Pre-Stained Protein Ladder (#SM0671)
2	GAT4601 protein derived from *E. coli* expression system (~1 µg load for SDS-PAGE and ~10 ng load for western blot)
3	GAT4601 protein derived from 356043 soybean (~1 µg load for SDS-PAGE and ~15 ng load for western blot)

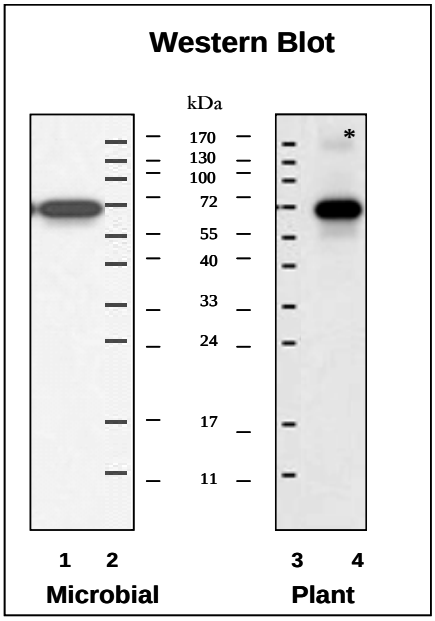
4.5b. GM-HRA

SDS-PAGE analysis of the microbial and 356043 soybean plant-derived GM-HRA proteins revealed both proteins migrating as expected at approximately 65 kDa (Appendix 4, Figure 2). Microbial GM-HRA protein has greater than 75% purity following densitometry analysis. Upon western blot analysis, both the microbial and 356043 soybean-derived GM-HRA proteins were recognized by GM-HRA-specific antibodies, and confirmed the presence of GM-HRA protein by detection of a single band recognized at approximately 65 kDa.

SDS-PAGE and western analysis confirmed that both microbially expressed and 356043 soybean-derived GM-HRA had the expected molecular weight and immunoreactivity and were equivalent to each other in these analyses.

Appendix 4, Figure 2. SDS-PAGE and Western Blot Analysis of Plant-Derived and Microbially Expressed GM-HRA





* Unidentified protein
† Identified as putative GM-HRA degradative peptide products

* Unidentified protein

Lane #	Description
1	Microbial GM-HRA Protein (~1 µg)
2	Fermentas Protein Molecular Marker (#SM0661)
3	Plant Derived GM-HRA Protein (~1 µg)
4	Fermentas Protein Molecular Marker (#SM0671)

Lane #	Description
1	Microbial GM-HRA Protein (~20 ng)
2 & 3	Fermentas Protein Molecular Marker (#SM0671)
4	356043 Soybean Plant Derived GM-HRA Protein (~20ng)

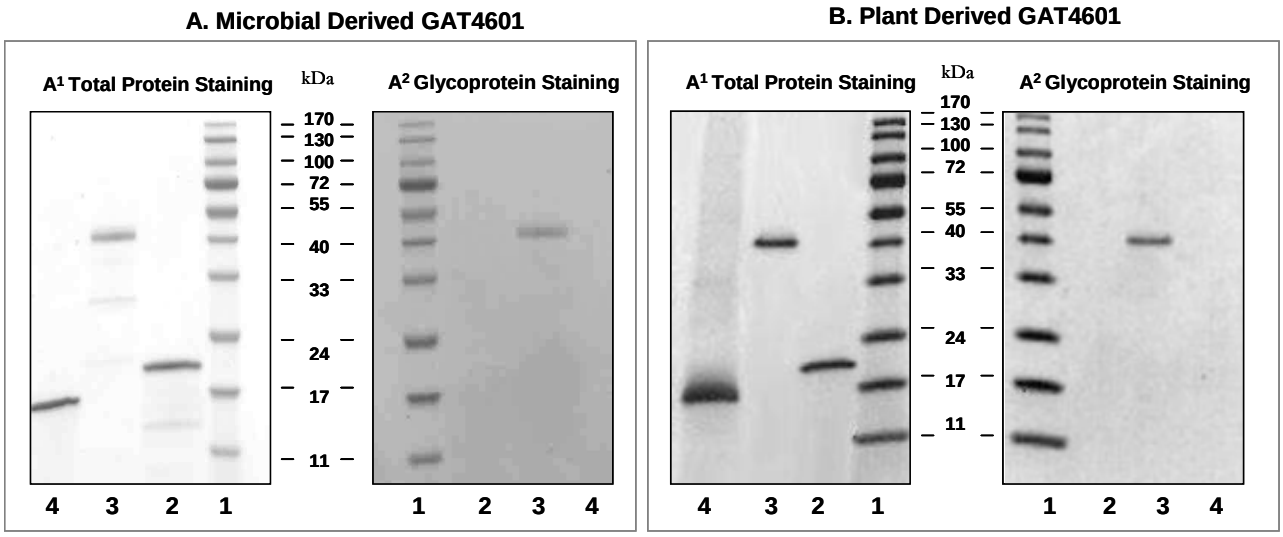
4.6. Method for Detection of Protein Glycosylation

A GelCode Glycoprotein staining kit (Pierce Biotechnology, Inc.) was used according to the manufacturer's instructions to determine whether the GAT4601 or GM-HRA proteins were glycosylated. GAT4601 and GM-HRA proteins and the control proteins were loaded at equivalent concentrations (~1μg/lane). SDS-PAGE was conducted as described previously. Following electrophoresis, the gel was fixed with 50% methanol, washed with 3% acetic acid, and then incubated with an oxidizing solution. The gel was incubated with GelCode glycoprotein staining reagent, treated with the reducing reagent, and washed. Glycoproteins were visualized as magenta colored bands on the gel. Following glycoprotein detection, the gel was scanned and the image captured electronically. The same gel was then stained with Coomassie Blue to visualize the total protein content.

4.7. Results of Protein Glycosylation Analysis

Glycosylation analysis demonstrated that there was no detectable glycosylation of either the microbial expressed GAT4601or GM-HRA proteins, or 356043 soybean plant-derived GAT4601 and GM-HRA proteins. The glycoprotein positive control (horseradish peroxidase) was stained and clearly visible as a magenta colored band and the negative control (soybean trypsin inhibitor) did not show any staining. This confirms that the microbially expressed and 356043 soybean plant-derived, GAT4601 and GM-HRA proteins are non-glycosylated and hence are equivalent with respect to protein glycosylation.

Appendix 4, Figure 3. Glycosylation Analysis of Microbially Expressed (Panel A) and Plant-Derived (Panel B) GAT4601

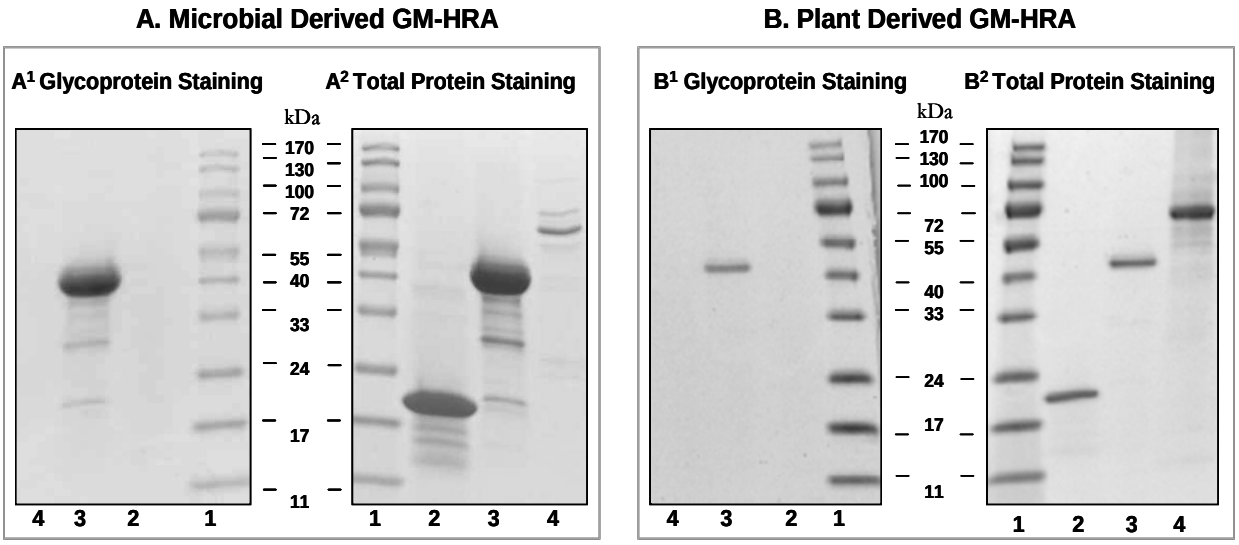


[1] Stained with the glycoprotein stain followed by staining with Coomassie Blue (GelCode Blue Total Protein Stain)
[2] Stained with the glycoprotein staining kit (GelCode Glycoprotein Stain)

Lane #	Description of the Sample
1	Fermentas PageRuler Prestained Protein Ladder (#SM0671)
2	Soybean Trypsin Inhibitor (~1 µg load)
3	Horseradish Peroxidase (~1 µg load)
4	Microbial GAT 4601 Protein (~1 µg load)

Lane #	Description of the Sample
1	Fermentas PageRuler Pre-stained Protein Ladder (#SM0671)
2	Soybean Trypsin Inhibitor (~1 µg load)
3	Horseradish Peroxidase (~1 µg load)
4	356043 Soybean-Derived GAT4601 Protein (~1.5 µg load)

Appendix 4, Figure 4. Glycosylation Analysis of Microbially Expressed (Panel A) and Plant-Derived (Panel B) GM-HRA



[1] Stained with the glycoprotein staining kit (GelCode Glycoprotein Stain)
[2] Stained with the glycoprotein stain followed by staining with Coomassie Blue (GelCode BlueTotal Protein Stain)

Lane #	Description of the Sample
1	Fermentas PageRuler Unstained Protein Ladder (#SM0661)
2	Soybean Trypsin Inhibitor (~20 µg load)
3	Horseradish Peroxidase (~20 µg load)
4	Microbial GM-HRA, Lot PCF-0008 (~1 µg load)

Lane #	Description of the Sample
1	Fermentas PageRuler Prestained Protein Ladder (#SM0671)
2	Soybean Trypsin Inhibitor (~1 µg load)
3	Horseradish Peroxidase (~1 µg load)
4	356043 soybean-derived GM-HRA Protein (~1.5 µg load)

4.8. Matrix Assisted Laser Desorption Ionization-Mass Spectrometry (MALDI-MS) Method

The GAT4601 and GM-HRA protein bands were excised from a SDS-PAGE gel and stained with Coomassie Blue as described above. The gel slice was placed in a labeled tube and shipped overnight on dry ice to the Keck Biotechnology Resource Laboratory (Yale University, New Haven, CT) for trypsin digestion and MALDI-MS analysis. MALDI-MS was performed on a Waters MALDI-L/R spectrometer (Waters Corporation, Milford, MA) in the reflecton mode of operation. Detected peptide peaks were considered a match if the observed experimental mass of peptides derived from *in silico* trypsin cleavage of the GM-HRA and GAT4601 proteins were within 100 parts per million (ppm) of the theoretical mass of the peptides derived from trypsin cleavage of the proteins. Allowances were made for the following potential modifications to the peptides: oxidation of methionine or tryptophan (15.995 Da greater than the theoretical value) and modification of cysteine residues by acrylamide free radicals during SDS-PAGE (71.037 Da greater than the theoretical value).

4.9. Results of MALDI-MS

MALDI-MS analysis of the trypsin digest of the 356043 soybean plant-derived GAT4601 protein identified 5 peptides that were within 100 ppm of theoretical peptide masses predicted from the *in silico* trypsin digestion of the GAT4601 protein (data not shown). Two additional peptide matches could be made by allowing for a modification of a cysteine residue by acrylamide (observed mass increases by 71.037 Da) and oxidation of a methionine residue (observed mass increases by 15.995 Da). The seven identified peptides accounted for 76 of the 146 amino acids, or 52% of the GAT4601 amino acid sequence. MALDI-MS analysis of the trypsin digest of the microbial GAT4601 protein identified six peptides that were within 100 ppm of theoretical peptide masses predicted from the *in silico* trypsin digestion of the GAT4601 protein. One additional peptide match could be made by allowing for a modification of a cysteine residue by acrylamide (observed mass increases by 71.037 Da). Similar to the 356043 soybean plant derived GAT4601, the seven peptides identified by MALDI-MS for microbial GAT4601 accounted for 76 of the 146 amino acids, or 52% of the GAT4601 amino acid sequence (data not shown).

MALDI-MS analysis of the trypsin digestion of the plant-derived GM-HRA protein identified 14 peptides that were within 100 ppm of theoretical peptide masses predicted from the *in silico* trypsin digestion of the GM-HRA protein (data not shown). The 14 matched peptides covered 26% or 157 of the 604 amino acids in the GM-HRA amino acid sequence. MALDI-MS analysis of the trypsin digest of the microbial GM-HRA protein identified twelve peptides that were within 100 ppm of theoretical peptide masses predicted from the *in silico* trypsin digest of the GM-HRA protein. One additional peptide match could be made by allowing for a modification of cysteine residues by acrylamide free radicals during SDS-PAGE (71.037 Da greater than the theoretical mass) and six additional peptide matches could be made by allowing for the oxidation of methionine or tryptophan (15.995 Da greater than the theoretical mass), which has been observed during this type of analysis. The 18 identified peptides account for 232 of the 605 amino acids, or 38%, of the microbial GM-HRA sequence (data not shown).

These data support the equivalency of the microbial GAT4601 and GM-HRA proteins to the 356043 soybean plant-derived GAT4601 and GM-HRA proteins, respectively.

4.10. N-Terminal Sequencing Method

Protein samples separated by SDS-PAGE and electrophoretically transferred to a PVDF membrane as described earlier were stained with Ponceau S solution (Sigma-Aldrich, St. Louis, MO) to visualize the protein bands. The observed bands were excised and shipped to the Keck Biotechnology Resource Laboratory (Yale University, New Haven, CT) for Edman N-terminal amino acid sequencing using the Procise 494 cLC analyzer (Applied Biosystems, Inc., Foster City, CA) equipped with an online high performance liquid performance chromatography (HPLC) system

4.11. Results from N-Terminal Sequencing

Results of the N-terminal amino acid sequence analysis of the first 13 amino acids of 356043 soybean plant-derived GAT4601 protein and microbial GAT4601 were consistent with the expected N-terminal sequence for GAT4601 (data not shown). Also, results of the N-terminal amino acid sequence analysis of the first 13 amino acids of 356043 soybean plant-derived GM-HRA protein were consistent with the expected N-terminal sequence for GM-HRA. The N-terminal sequence of the microbial GM-HRA protein was consistent with the expected N-terminal sequence after accounting for the additional N-terminal glycine (data not shown).

On the basis of MALDI-MS analysis and N-terminal sequencing, microbially derived GAT4601 and GM-HRA proteins were found to be identical to 356043 soybean plant-derived GAT4601 and GM-HRA proteins, respectively.

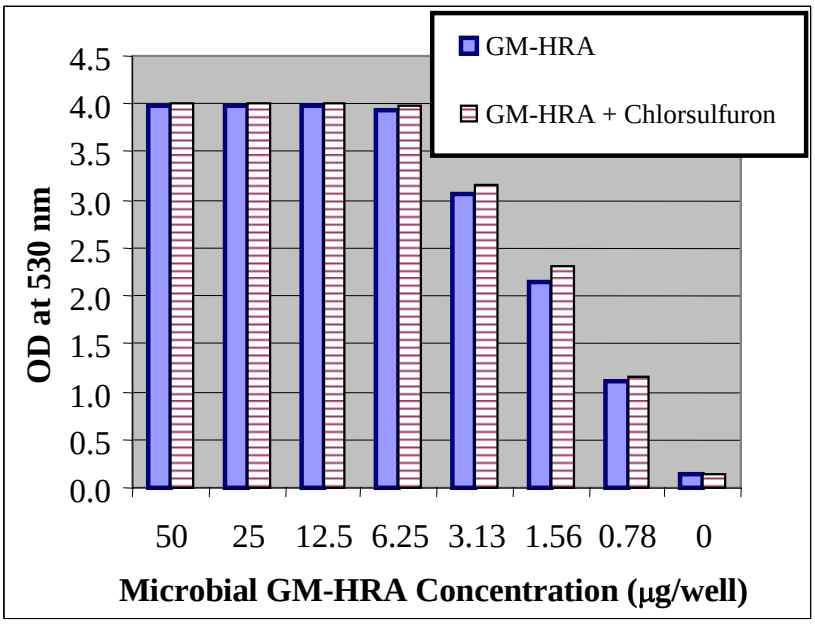
4.12. Method for ALS Activity Assay

A known amount (0 to 50 ug) of 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 $MgCl_2$ 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 $MgCl_2$). Acetone standard was diluted in 0.1 M phosphate buffer, pH 7.5 and triplicate 100 µl aliquots of protein and standard dilution were dispensed into a 96 well plate. Then 5 µl of 200 mM phosphate buffer or chlorsulfuron (5µg/ml) was added to each well. This was followed by the addition of 10 µl of a 1.1 M pyruvate solution to each well. The plates were incubated at 37° C for one hour, at which point 5 µl of 2 N H_2SO_4 was added to stop the reaction. The plates were incubated at 60° C for 15 minutes. Prior to the addition of 50 µl of creatine/naphtol solution per well, the plates were allowed to cool at room temperature for 15 minutes. The plates were incubated at 60° C for 15 minutes and immediately were read at 530 nm using a SpectraMax Model 190 spectrophotometer (Molecular Devices Corporation, Sunnyvale, CA).

4.13. Results of the ALS Activity Measurement of GM-HRA Protein

The activity of the microbial GM-HRA protein was determined by a manual assay for activity measurement of acetolactate synthesis. As expected, results demonstrated that the GM-HRA protein has equivalent ALS biochemical activity (Appedix 4, Figure 5) both in the presence and absence of 100ng/ml chlorosulfuron (an inhibitor of ALS activity).

Appendix 4, Figure 5. Acetolactate Synthase (ALS) Activity Assay on Microbial GM-HRA Protein in the Presence or Absence of Herbicide



4.14. Electrospray Mass Spectroscopy Method for Microbial GAT4601 and GM-HRA Proteins

The GAT4601 and GM-HRA microbially expressed proteins were sent to the Keck Biotechnology Resource Laboratory (Yale University, New Haven, CT) and subjected to reverse-phase desalting procedure on a C4 ZipTip (Millipore Corporation, Billerica, MA). The protein was eluted in 50% acetonitrile containing 0.2% formic acid and then analyzed using electrospray ionization mass spectroscopy (ESMS) with a Waters/Micromass Q-tof Micro instrument (Waters Corporation, Milford, MA). The ESMS spectra were recorded in the positive ion mode, and the raw data was transformed using the maximum entropy algorithm that calculates molecular size to charge.

4.15. Results of Electrospray Mass Spectroscopy Analysis of Microbial GAT4601 and GM-HRA Proteins

Analysis of microbial GAT4601 by electrospray mass spectroscopy identified a major peak at a molecular weight of 16716.00 Da. The average was determined to be 16714.72 Da from molecular size to charge peaks (15 total) observed in the spectrum. This is consistent with the theoretical value of 16713.9 Da for the GAT4601 protein (data not shown). Analysis of the microbial GM-HRA protein by electrospray mass spectroscopy identified a major peak at 65316 Da. This was consistent with the expected molecular mass of 65312 Da for the mature GM-HRA protein after accounting for the expected extra N-terminal glycine residue (data not shown).

4.16. Method for Amino Acid Composition Analysis of Microbial GAT4601 and GM-HRA Proteins

The microbial GAT4601 and GM-HRA proteins were sent to Keck Biotechnology Resource Laboratory (Yale University, New Haven, CT) for amino acid analysis. Both proteins were hydrolyzed at 115°C in 100 µl of 6 N HCl, and 0.2% phenol containing 2 nM norleucine as an internal standard. Amino acid composition analysis was completed in triplicate using a Beckman Model 7300 (Beckman Coulter, Inc., Fullerton, CA) ion-exchange instrument. The instrument was calibrated with a 2 nM mixture of amino acids and was operated according the manufacturer's instructions. Using this type of analysis method glutamine/glutamic acid and asparagine/aspartic acid were not individually quantified and the methionine value is predicted to be less than the theoretical value.

The results from the amino acid composition analysis were used to determine the concentration of the microbial GAT4601 and GM-HRA proteins in the sample. Three (for GAT4601) and/or four (for GM-HRA) amino acids were chosen for the quantification based on the closeness of the fit of the observed number of residues to the expected value.

4.17. Results of Amino Acid Composition Analysis of Microbial GAT4601 and GM-HRA Proteins

Based on the amino acid composition determination, the calculated concentration of GAT4601 and GM-HRA in the lyophilized protein powders were found to 84% (0.84mg in 1mg of lyophilized powder), and 29.1% (0.291 mg in 1 mg of lyophilized powder), respectively (data not shown).

Appendix 5. USDA Field Trials of 356043 Soybean

Year of Planting	Permit Name	Permit Valid Date	State	Number of Counties Where 356043 Soybean Was Planted
2003	03-022-04R	5/1/2003	HI	1
2004	03-022-04R	5/1/2003	HI	1
	04-020-01R	4/13/2004	IA	3
	04-020-02R	4/15/2004	HI	2
2005	04-020-01R	4/13/2004	IL	3
			IA	2
			OH	1
	05-024-01R	5/12/2005	IL	2
			IA	4
			MN	1
			NE	1
			AR	1
	05-024-02R	5/25/2005	HI	1
2006	06-019-01R*	5/1/2006	AR	2
			FL	1
			GA	1
			IL	8
			IN	2
			IA	7
			MN	5
			MS	1
			NE	3
			NJ	1
			ND	2
			OH	2
			VA	1
			WI	2
	06-019-02R*	5/22/2006	HI	1

Note: In USDA final reports, 356043 soybean (event DP-356Ø43-5) is called EAFS 3560.4.3.

*Final field test report not yet submitted to USDA.

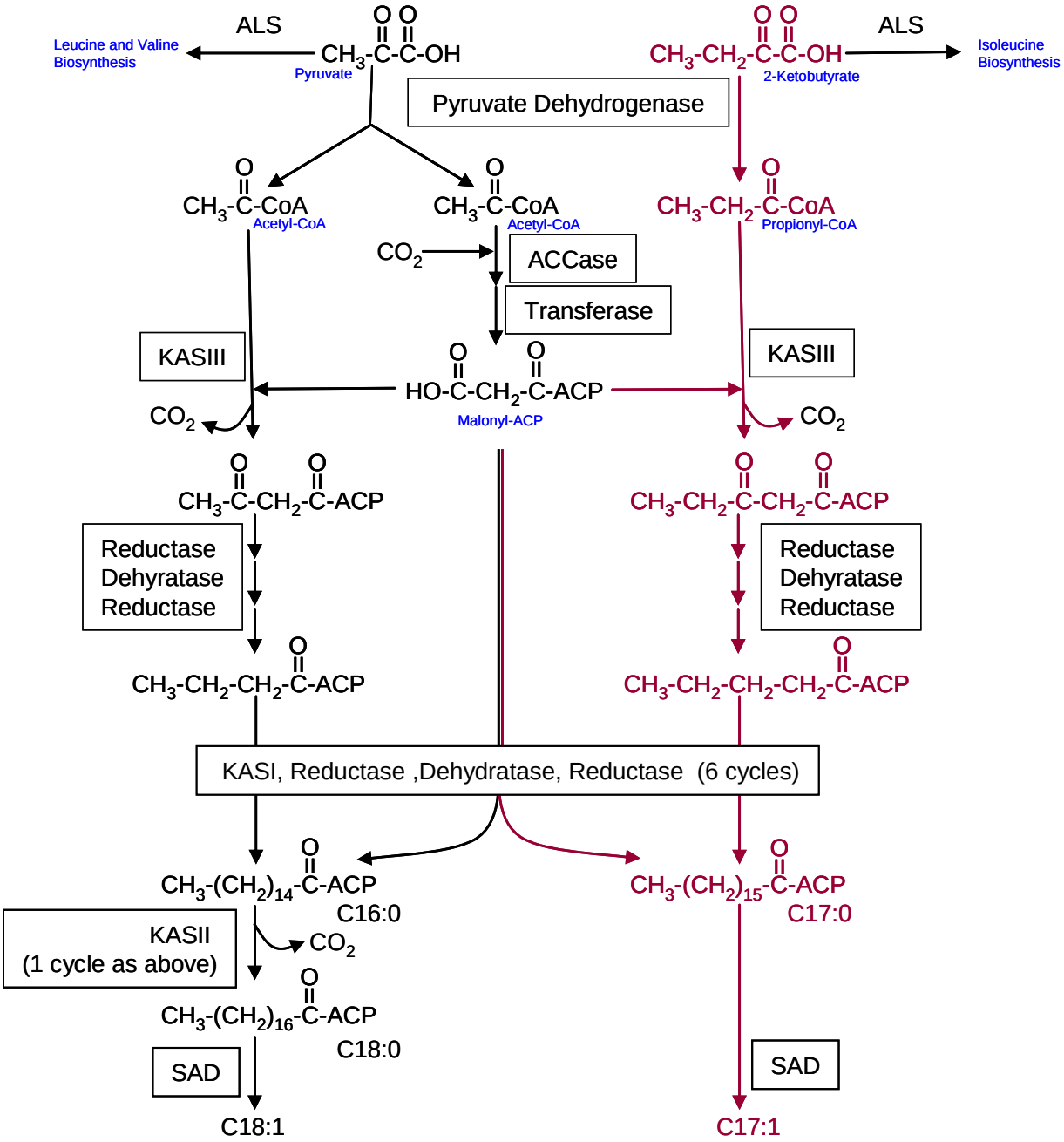
Appendix 6. Synthesis of 17-Carbon Fatty Acids in Soybean

The biosynthetic pathway for fatty acids in plants begins with the conversion of pyruvate to acetyl-CoA. One acetyl-CoA molecule is then converted to malonyl-CoA and joined with a second acetyl-CoA molecule forming a four carbon fatty acid. Subsequent additions of two carbon moieties from malonyl-CoA result in fatty acids with an even number of carbons, such as 16 carbon (C16:0) and 18 carbon (C18:0) fatty acids. Odd numbered fatty acids result from the conversion of 2-ketobutyrate to a three carbon compound, propionyl-CoA, followed by addition of a two carbon moiety from malonyl-CoA forming a five carbon fatty acid. Subsequent additions of two carbon moieties from malonyl-CoA result in fatty acids with an odd number of carbons such as C17:0.

The primary pathway for both even-chain and odd-chain fatty acids begins with the pyruvate dehydrogenase complex (Appendix 6, Figure 1). Pyruvate is converted to acetyl-CoA by the pyruvate dehydrogenase complex and then used in the production of fatty acids with even chain numbers. However it is known that in *E. coli,* for example, 2-ketobutyrate can serve as a substrate for pyruvate dehydrogenase with a K_M that is about ten times higher than that of pyruvate. K_M is the affinity of the enzyme for a substrate or tightness of binding of the substrate to the enzyme. The higher the K_M, the lower the affinity of the enzyme for the substrate. The product of the 2-ketobutyrate reaction with pyruvate dehydrogenase is the three carbon compound propionyl-CoA instead of the two carbon acetyl-CoA from pyruvate. Fatty acid biosynthesis that initiates with propionyl-CoA results in fatty acids with an odd number of carbons.

One of the specific herbicide resistance mutations (tryptophan 560 to leucine) introduced into GM-ALS to form the GM-HRA enzyme may increase the 2-ketobutyrate pool available for C17:0 fatty acid synthesis. In studying the structure of *E. coli* ALS, it was shown that a mutation at amino acid position 464 changing the native tryptophan to leucine resulted in a 50-fold decrease in substrate preference for 2-ketobutyrate versus pyruvate (Ibdah *et al.*, 1996). The mutation also confers a high level of resistance to inhibition by ALS inhibiting herbicides on the *E. coli* ALS enzyme. Tryptophan 464 in the *E. coli* ALS is equivalent to tryptophan 560 in the soybean ALS. The mutation in the GM-HRA allele of GM-ALS is from tryptophan to leucine and occurs at position 560, and thus could have a similar loss in substrate preference. With the loss of substrate preference for 2-ketobutyrate it is reasonable to expect that the concentration of 2-ketobutyrate relative to pyruvate is increased in soybeans expressing GM-HRA. The increased levels of 2-ketobutyrate would become more available for pyruvate dehydrogenase and subsequent formation of odd chain fatty acids. Although fatty acids containing 17 carbons (C17:0, heptadecanoic acid and C17:1, heptadecenoic acid) are commonly found in soybean oil, their levels could be increased by the mutation in GM-HRA.

Appendix 6, Figure 1. Fatty Acid Biosynthesis – Even and Odd Chain Pathway



Even chain fatty acid biosynthesis is depicted in black arrows (left side of diagram) and odd chain fatty acid biosynthesis is depicted in red arrows (right side of diagram).

KAS I = 3-ketoacyl-acyl carrier protein synthase I
KAS II = 3-ketoacyl-acyl carrier protein synthase II
KAS III = 3-ketoacyl-acyl carrier protein synthase III
SAD = stearoyl-acyl carrier protein desaturase
ACCase = acetyl coenzyme A carboxylase

Appendix 7. Survey of Potential Substrates for GAT4601

In enzyme kinetics, k_{cat} is a measure of the turnover rate or speed of the reaction. The higher the k_{cat}, the faster the enzyme reaction. K_M is the affinity of the enzyme for a substrate or tightness of binding of the substrate to the enzyme. The lower the K_M, the greater the affinity of the enzyme for the substrate. A k_{cat}/K_M ratio is the common way to express the catalytic efficiency of the enzyme. The greater the k_{cat}/K_M ratio, the greater the catalytic efficiency of the enzyme for a given substrate. Specificity of an enzyme can be judged by comparing the values of k_{cat}/K_M for various enzyme substrates.

A survey of enzyme substrates was undertaken to determine the kinetic properties and specificity of GAT enzymes, including a Round 7 GAT protein similar to the GAT4601 protein in 356043 soybeans (Siehl *et al*., 2005). The survey was done, in part, to determine the physiological substrate of the native enzymes from *Bacillus licheniformis*. A broad sampling of amino acids, antibiotics and nucleotides was surveyed with the native enzyme. Among antibiotics, no detectable activity was seen with fosmidomycin, D,L-phosphinothricin or kanamycin, erythromycin, carbenicillin, spectinomycin, streptomycin, chloramphenicol or ampicillin. Other biological amines that supported no activity were D-glucosamine, serotonin, anthranilate, ornithine, purine and pyrimidine bases, nucleosides, nucleotides, histone and tRNA. None of the amino acids, antibiotics or nucleotides surveyed exhibited high catalytic efficiency with the native enzyme, from which it was concluded that the physiological role of GAT in *B. licheniformis* is still unknown.

Native GAT enzymes from *Bacillus licheniformis* strains did exhibit weak activity with seven of the common amino acids: L-aspartate, L-serine, phospho-L-serine, L-threonine, L-glutamic acid, L-asparagine and L-cysteine (Siehl *et al.,* 2005). Activity with the other protein amino acids was either nil or less than 3% of that of glyphosate. Shuffling of the enzyme towards an improved k_{cat}/K_M for glyphosate resulted in increased activity toward aspartate, although overall catalytic efficiency was still very low (~3% of glyphosate for a Round 7 GAT protein). Activity with serine and phosphoserine almost completely vanished.

In order to examine the specific catalytic efficiency of the GAT4601 enzyme expressed in 356043 soybean, a substrate specificity study was done with the microbial GAT4601 protein that had previously been determined to be equivalent to plant-derived protein (see Section VI-D and Appendix 4). A dithiobis-2-nitrobenzoate enzymatic end-point assay was used to detect any enzyme activity in a survey of various substrates (Siehl *et al*., 2005) using assay buffer containing 100 mM KCl to represent physiological conditions in the plant cytosol and chloroplast (Coruzzi and Last, 2000; Cuin *et al*., 2003). Twenty agrochemicals, 21 amino acids and 11 antibiotics were tested as potential substrates for the GAT4601 protein. As expected from the previous survey, no significant activity was seen with the majority of the substrates.

Of the substrates surveyed for activity with GAT4601, only five amino acid substrates indicated low but measurable enzyme activity (sufficiently above the limit of quantitation of the assay): L-aspartate, L-glutamate, L-serine, L-glycine and L-threonine. The remaining amino acids and substrates surveyed (other than glyphosate) produced levels of end product near or below the limit of quantitation of the end-point assay.

GAT4601 activity on the five amino acids was further characterized using a continuous spectrophotometric assay (Siehl *et al*., 2005) for characterization of kinetic properties. (Appendix 7, Table 1). The affinity of the GAT4601 enzyme for L-serine, L-threonine and L-glycine was so low that a K_M value could not be estimated. Therefore, a k_{cat}/K_M ratio

could not be calculated. The catalytic efficiency of GAT4601 on aspartate and glutamate was about 3% of that of glyphosate.

Appendix 7, Table 1. k_{cat}/K_M for the GAT4601 Enzyme on Glyphosate, Aspartate, Glutamate, Serine and Threonine

Substrate	**k_{cat}/K_M (+/- standard deviation) min^{-1} mM^{-1}**	**% k_{cat}/K_M of glyphosate**
Glyphosate	267 (+/- 21.1)	100
Aspartate	7.81 (+/- 1.17)	2.93
Glutamate	7.72 (+/- 2.53)	2.89
Serine	ND[1]	ND[1]
Threonine	ND[1]	ND[1]
Glycine	ND[1]	ND[1]

[1] ND – Not able to determine because the K_M was too high to estimate.

The results of the GAT4601 substrate survey confirmed the results of the earlier published substrate survey for the GAT enzyme. The GAT4601 enzyme was able to use five amino acids as substrates (aspartate, glutamate, serine, glycine and threonine), although inefficiently (~0.2 - 3% of that of GAT4601 on glyphosate). Further kinetic characterization indicated that the K_M of GAT4601 for serine, glycine and threonine was too high to estimate (low affinity of the GAT4601 enzyme for serine, glycine and threonine as substrates). The level of catalytic efficiency of GAT4601 on aspartate and glutamate was about 3% of that of GAT4601 on glyphosate. The presence of N-acetylaspartate and N-acetylglutamate in 356043 and control soybean grain was examined in the nutrient composition studies due to the GAT4601 enzyme activity on aspartate and glutamate as substrates.

Appendix 8. Herbicide Resistant Weeds

8.1. Evolution of Herbicide Resistant Weeds

Weeds will eventually adapt and circumvent most control mechanisms. Herbicide resistant weeds are a well-established aspect of weed control for many herbicide classes including acetyl-CoA carboxylase and ALS inhibitors, dinitroanilines, triazines and other photosystem II inhibitors.

Herbicide resistance usually evolves in only one or two weed species in an area, even though a much larger number of weeds are exposed to the same herbicide selection intensity. Nonetheless, weed resistance to herbicides currently affects hundreds of thousands of fields and the most widely used herbicides (Heap, 2006). According to a recent survey, more than 305 types of herbicide resistant weeds are present in agricultural fields around the world (www.weedscience.org). Resistant weeds often increase the cost of crop production and limit the effectiveness of herbicides that can be used and the crops that can be grown. Despite these challenges, growers have used a variety of approaches to limit the impact of resistant weeds on crop productivity.

When growers say that their "weeds have become resistant," they really mean that the population of resistant weed biotypes, which formerly existed at low numbers, has increased. The spread of a resistance phenotype, which leads to a weed population that is not susceptible to the herbicide, depends primarily upon the exposure to the herbicide that the weed is able to tolerate. When an herbicide is applied, most of the susceptible weeds die, while the resistant weeds survive, mature, and produce seed. Even though they may still be few in number, repeated application of the same herbicide continues to increase the proportion of resistant weeds in the population.

Not all cases of weed shifts that are driven by herbicide use can be explained by herbicide selective pressure that favors weeds with a genetically based biochemical capacity to survive exposure. Weeds with delayed emergence and slower development are able to avoid exposure to the herbicide (Hilgenfield *et al*., 2004).

Herbicide resistance can become an ecological problem if the resistant weed biotype replaces the non-resistant biotype in the weed population. Even then, the shift to an herbicide resistant population of weeds has ecological consequence only if the resistant population cannot be controlled with other herbicides or other control practices. This is generally not the case. Many hundreds of cases of resistant weeds have been documented worldwide, but resistance is usually not a limiting factor for crop production. There are examples, such as some locations in Australia, where biotypes of rigid ryegrass (*Lolium rigidum*) are resistant to many herbicides in several different classes (Heap, 2006).

In spite of the evolution of herbicide resistance in weed populations, US soybean growers continue to have many herbicides and management options for weed control. Even so, growers must always be concerned about herbicide sustainability and the economic consequences of losing an herbicide due to the evolution of resistant weeds.

8.2. Characteristics of Glyphosate and ALS-Inhibiting Herbicides

8.2a. Glyphosate

Glyphosate is a broad spectrum herbicide that was introduced in the 1970s for management of annual, perennial and biennial herbaceous grasses, sedges, and broadleaves, as well as woody brush and trees (Franz *et al*., 1996).

Glyphosate controls plants by inhibiting the enzyme EPSPS (5-enolpyruvylshikimate-3-phosphate synthase). EPSPS is an essential enzyme in the shikimate pathway that ultimately leads to the production of aromatic amino acids (tryptophan, tyrosine, and phenylalanine). The shikimate pathway for synthesizing aromatic amino acids, and therefore the enzyme EPSPS, is found in plants, bacteria and fungi, but not animals.

The structure of glyphosate resembles the structure of the substrate for EPSPS, which is phosoenolpyruvate (PEP). Therefore, glyphosate competes with PEP for the enzyme's active site and prevents conversion of PEP to the precursor that is required in the synthesis of aromatic amino acids. Aromatic amino acids are essential for many plant processes such as protein synthesis, cell wall formation, pathogen defense and hormone production. At high rates, glyphosate is toxic to virtually all plants.

8.2b. ALS-Inhibiting Herbicides

ALS-inhibiting herbicides were discovered in 1975 (Stetter, 1994; Shaner and O'Connor, 2000; Tan *et al*., 2005). They inhibit a plant enzyme called acetolactate synthase (ALS) that is required for the production of essential branched-chain amino acids such as valine, leucine, and isoleucine. There are five different chemical classes of ALS herbicides that have been commercialized: sulfonylureas (SU), imidazolinones (IMI), triazolopyrimidines (TP), pyrimidinylthiobenzoates (PTB), and sulfonylamino-carbonyl-triazolinones (SCT).

ALS inhibitors control a wide spectrum of grass and broadleaf weeds at very low application rates (typically, fraction of ounces per acre). In addition, they generally have very low mammalian toxicity and possess a favorable environmental profile. Today, about 56 different ALS inhibitor active ingredients are marketed with registrations in all major crops. Significant changes in herbicide potency, crop selectivity, and weed control can be made with small chemical alterations within the ALS-inhibiting herbicide class.

8.3. Evolution of Resistance to Glyphosate and ALS-Inhibiting Herbicides

8.3a. Glyphosate

For more than two decades, the evolution of glyphosate resistance was not perceived as a problem (Bradshaw *et al*., 1997). This opinion was based on the difficulty of discovering a fully functional 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) that was insensitive to glyphosate, the inability of plant species to enzymatically deactivate glyphosate, the lack of soil activity of glyphosate (thus reducing the selection pressure due to exposure), and the empirical observation that no resistant weeds had appeared after at least 20 years of use.

The most likely theoretical ways for weeds to develop resistance did not seem likely for glyphosate (Jasieniuk, 1995). Over-production of the EPSPS target site did not increase resistance enough for plants to survive glyphosate amounts used in agricultural settings (Kishore and Shah, 1988). The EPSPS modifications that conferred glyphosate resistance in bacteria were inside the enzyme's active site, which reduced its catalytic

efficiency and thus would probably reduce plant fitness (Padgette *et al*., 1995). No higher plants could be found with even low levels of ability to metabolically inactivate glyphosate (Dyer, 1994).

However, the views about the capacity of weeds to develop glyphosate resistance changed in 1996 when a glyphosate-resistant weed, *Lolium rigidum*, was discovered in Australia (Powles *et al*., 1998; Pratley *et al.*, 1999). Since then, glyphosate-resistant biotypes in at least eight other weed species have been confirmed (Nandula *et al*., 2005), and more than a million hectares are now affected by resistant weeds. Some of these biotypes exhibit resistance at application rates four to 13 times higher than susceptible populations.

The most widespread glyphosate-resistant weed is marestail (*Conyza canadensis*). Glyphosate-resistant marestail was confirmed in 2000 (VanGessel, 2001). Marestail produces very large numbers of light, wind dispersed seed and can cross-pollinate, leading to rapid spreading of resistant weeds to no-tillage crop and non-crop land. As a result, five years after its first occurrence, it was found on a half-million hectares across the U.S. Midwest, South and Atlantic states (VanGessel, 2001; Heap, 2006). Dose response analysis showed these populations were eight to 13-fold more resistant than susceptible marestail populations.

In 2002 and 2003, growers, consultants and extension agronomists frequently observed common lambsquarters (*Chenopodium album*), a significant problem in row crops, as the only surviving weed in soybean fields after glyphosate usage. Growers and extension agents thought factors other than endogenous resistance, such as an inability of glyphosate to penetrate lambsquarters thick cuticle, might explain persistence after glyphosate exposure.

Significant populations of common and tall waterhemp (*Amarathus sp*.) that survived glyphosate application were first observed in fields in Iowa, Illinois, and Missouri (Owen, 2002; Smeda and Schuster, 2002). Studies indicate that plants survived glyphosate rates 2.6 times the label rate and some waterhemp plants have now been classified as glyphosate resistant (Owen and Zelaya, 2005).

The molecular basis for weed resistance to glyphosate is not understood in most cases. Initial studies of various weed species revealed EPSPS target site insensitivity (Braerson *et al*., 2002). Differences in translocation and transport to the chloroplast are also important in some weeds (Lorraine-Colwill *et al*., 2003; Feng *et al*., 2004). The mechanisms of glyphosate resistance in other weeds appear to be complex and polygenic.

The effectiveness, economic benefits, and ease of using glyphosate have led to repeated applications, year after year, in areas where glyphosate tolerant biotech crops are grown. This intensive use has resulted in a high selection pressure for weeds that inherently are difficult to control with glyphosate (Culpepper, 2004). Eventually, this selection pressure can lead the spectrum of weeds in the fields to a shift to those weeds that inherently can tolerate glyphosate. In soybeans, some growers have difficulty controlling waterhemp (*Amaranthus* spp.), copperleaf (*Acalypha* spp.), velvetleaf (*Abutilon theophrasti*), giant ragweed (*Ambrosia trifida*), marestail (*Conyza canadensis*), winter annuals, and lambsquarters (*Chenopodium spp*.) with glyphosate.

As predicted, spectrum shifts to weed populations with endogenous glyphosate resistance have occurred more rapidly than evolved resistance in response to glyphosate exposure (Shaner, 2000). For example, in Iowa, common waterhemp (*Amaranthus rudis*)

and velvetleaf (*Abutilon theophrasti*) became a concern in glyphosate-resistant soybean soon after crop commercialization (Owen, 1997).

8.3b. ALS-Inhibiting Herbicides

The five classes of ALS-inhibiting herbicides differ chemically, but all bind to the same target site on the ALS enzyme. The primary molecular basis of weeds resistant to ALS inhibitors is reduction of target site sensitivity, although there are biotypes that are resistant through more rapid detoxification of the herbicide to inactive metabolites.

To date, weed scientists have identified 93 weed species containing biotypes that are resistant to ALS-inhibiting herbicides (www.weedscience.org, 2006). The large number of resistant biotypes is due, in part, to the relatively large number of amino acid substitutions that can change the ALS enzyme from a sensitive to a resistant form. Five different mutations sites have been identified in naturally occurring resistant weed populations (Bernasconi *et al.,* 1995; Tranel and Wright, 2002). However, not all resistant weeds are resistant to all classes of ALS-inhibiting herbicides (Tranel and Wright, 2002). ALS inhibitor resistance generally falls into three categories:

1. Broad resistance to sulfonylureas (SU), imidazolinones (IMI), triazolopyrimidines (TP), and pyrimidinylthiobenzoates (PTB);
2. Resistance to IMI and PTB only; and
3. Resistance to SU and TP only.

The evolution of ALS inhibitor-resistant weeds has typically been after five to eight years of repeated, if not continuous, use of herbicides with that mode of action. Resistance has generally not been selected where ALS-inhibiting herbicides have been used as part of an integrated program (Shaner *et al*., 1997). Fortunately, most ALS inhibitor-resistant weed populations are localized. Exceptions include ALS-resistant kochia (*Kochia scoparia* (L.) Schrad.) and Russian thistle (*Salsola iberica* Sennen & Pau) that are now present in over 60% of the wheat fields in the northern United States (Heap and LeBaron, 2001), wild sunflowers (*Helianthus annuus* L.) in the upper Midwest, and waterhemp species in the Midwest. Now growers must mix herbicides with other modes of action to control these weeds. Fortunately, most resistant weed populations can still be controlled with commonly used, low cost mixtures and the value of this mode of action has largely been retained. Therefore, despite the evolution of resistance, ALS-inhibiting herbicides are still among the most efficacious and widely used weed control agents in the world (Shaner and Heap, 2002). Use of ALS-inhibiting herbicides has increased dramatically over the past 25 years, and is continuing to increase. These products account for about 17.5% of the total worldwide herbicide market. Within the past five years, eight new active ingredients from the ALS-inhibiting herbicide family have been introduced. ALS inhibitors are becoming an increasing good solution to the growing problem of glyphosate resistance.

8.4. Stewardship of Herbicide Tolerant 356043 Soybean

DuPont's Pioneer Hi-Bred International, Inc. and Crop Protection Chemicals businesses both have long histories of product stewardship. Because of the unique nature of the dual herbicide tolerance in 356043 soybean, stewardship efforts for this product will be a joint collaboration between the two businesses. Stewardship principles will be incorporated into marketing, positioning, promotional and communications strategies for herbicide tolerant 356043 soybean. Examples of these efforts are detailed below.

8.4a. Local Weed Management

In collaboration with university investigators, private consultants, other manufacturers, and growers, DuPont Crop Protection field development personnel conduct hundreds of field trials annually to refine existing recommendations and to investigate potential new active ingredients and herbicide combinations deployed in an Integrated Weed Management context to improve solutions for existing and emerging weed problems. Within Pioneer, field agronomists give presentations to local audiences about resistant weeds and best management practices.

8.4b. Product Labeling

For at least 20 years, all DuPont herbicide labels have carried voluntary statements regarding resistant weed management. An example of the language used on the labels follows. These labels actively promote an integrated weed management philosophy as seen in the second paragraph:

RESISTANCE

When herbicides that affect the same biological site of action are used repeatedly over several years to control the same weed species in the same field, naturally-occurring resistant biotypes may survive a correctly applied herbicide treatment, propagate, and become dominant in that field. Adequate control of these resistant weed biotypes cannot be expected. If weed control is unsatisfactory, it may be necessary to retreat the problem area using a product affecting a different site of action.

To better manage herbicide resistance through delaying the proliferation and possible dominance of herbicide resistant weed biotypes, it may be necessary to change cultural practices within and between crop seasons such as using a combination of tillage, retreatment, tank-mix partners and/or sequential herbicide applications that have a different site of action. Weed escapes that are allowed to go to seed will promote the spread of resistant biotypes.

It is advisable to keep accurate records of pesticides applied to individual fields to help obtain information on the spread and dispersal of resistant biotypes. Consult your agricultural dealer, consultant, applicator, and/or appropriate state agricultural extension service representative for specific alternative cultural practices or herbicide recommendations available in your area.

8.4c. Training and Education of Sales Representatives and Agronomists

DuPont Crop Protection: In addition to the formal academic training of technical, marketing, and sales professionals, DuPont Crop Protection has a mandatory on-line interactive training course with intensive and comprehensive coverage of product stewardship generally and herbicide resistance specifically. This training is administered globally to all employees involved in weed management recommendations and is also provided to other interested parties as a community education contribution. The technical, marketing, and sales professionals are required to be certified in this training by receiving a passing grade of 90% or higher in each of the eight modules. A person unfamiliar with this topic is expected to take about twelve hours to complete the internal version of the training. This is the first training of its kind to be awarded Certified Crop Adviser credits by the American Society of Agronomy. The CCA standard allows growers, employers, and other organizations to help manage risk by enabling them to have assurance that a person has the appropriate professional qualifications.

Pioneer: The Pioneer sales force, made up of agronomists, account managers and sales representatives, and customers receive ongoing mandatory and voluntary stewardship training throughout the sales season through a variety of tactics. These include stewardship training sessions administered in person or communicated electronically by Pioneer Stewardship personnel. The Pioneer sales force also utilizes a proprietary computer software system to access information about products, traits, crop management practices, sales transactions and required stewardship documentation. This formal training is in addition to the practical experience of the Pioneer sales force working with growers on the positioning and use of products containing biotech traits, including herbicide resistant traits.

8.4d. Technical Bulletins Provided to Seed Customers and the Public

Direct-Mail

Pioneer's *Growing Point* magazine is mailed to all Pioneer customers, prospects and employees. This magazine contains information about products, marketing programs, production and management practices and advice on agronomic application of traits and technologies. Information on 356043 soybean and weed resistance management practices will be included in future issues of *Growing Point*.

Information on Websites

- Pioneer's "Growing Point" site (https://www.pioneer.com/growingpoint/login/login.jsp) has an extensive agronomy section, with access to information about many different crops and crop management practices. Technical bulletins such as "Crop Insights" have been published for the past 15 years, and are a good example of how Pioneer makes the latest information available to growers.

- The DuPont Biotechnology website (http://www2.dupont.com/Biotechnology/en_US/) contains an in-depth "Science Knowledge" section addressing "Herbicide Resistant Crops and Weed Management: Scientific Summary and the DuPont Perspective", with a section devoted to Integrated Weed Management, as well as an FAQ section that addresses weed control and Integrated Weed Management. There is also a Scientific Summary on Integrated Weed Management and Herbicide Resistance: (http://www2.dupont.com/Biotechnology/en_US/science_knowledge/herbicide_resistance/moreinfo8.html).

8.4e. Involvement in Industry Groups

DuPont Crop Protection is a long-standing participant in the Herbicide Resistance Action Committee (HRAC), an industry-based group supported by Croplife International. Their stated mission is to "*Facilitate the effective management of herbicide resistance by fostering communication and co-operation between industry, government and farmers*." They work towards fostering responsible attitudes towards herbicide use, communicating herbicide resistance management strategies and support their implementation through practical guidelines, and they seek active collaboration with public and private researchers, especially in the areas of problem identification and devising and implementing management strategies.

8.4f. Customer Satisfaction and Weed Resistance Management Plan

Pioneer and DuPont Crop Protection are committed to active market presence wherever we sell our products. Consistent with our stewardship principles, all of our business teams are required to maintain an active contact with customers and awareness of end-user practices and a capability to respond rapidly to issues arising. DuPont Crop Protection maintains a customer satisfaction database to track and address any complaints, including for resistant weeds. Field employees are trained and provided tools and processes for responding to inquiries regarding product performance (or any potential impacts on human health or environment).

PIONEER®
A DUPONT COMPANY

Addendum to Petition for the Determination of Nonregulated Status for Herbicide Tolerant 356043 Soybean 06-271-01p

Submitting Company:

Pioneer Hi-Bred International, Inc.
7100 NW 62nd Avenue
PO Box 1014
Johnston, IA 50131-1014

Submitted by:

Tracy A. Rood, Registration Manager
Pioneer Hi-Bred International, Inc.
7250 NW 62nd Ave.
PO Box 552
Johnston, IA 50131
Telephone: 515-270-4036
Fax: 515-334-4478

Prepared by:

Bryan Delaney, Ray Layton and Tracy Rood

OECD Unique Identifier: DP-356Ø43-5

July 19, 2007

NO CBI

Analysis of Impact of Potential Dietary Increase in NAA from 356043 Soybean on Individuals with Canavan Disease

N-acetyl-L-aspartate (NAA) is an essential substance of the human brain. NAA is continuously produced and metabolized within the brain. It is present in relatively high concentrations and is the second most abundant free amino acid in the brain (Tallan *et al.*, 1956; Miyake *et al.*, 1981). NAA is normally converted to acetate and the free amino acid L-aspartate by the enzyme aspartoacylase. The free acetate resulting from the hydrolysis of NAA is required for proper myelination of the neurons of the brain, which is critical for development of a functional CNS (Birken and Oldendorf, 1989).

Canavan Disease (CD) is caused by heritable mutations in the aspartoacylase gene (Zeng *et al.,* 2002; Hershfield *et al.,* 2007). These mutations result either in the absence of, or expression of a non-functional, aspartoacylase enzyme (Matalon and Michals-Matalon, 1999). The lack of aspartoacylase activity causes an inability to hydrolyze NAA, resulting in a deficiency of acetate leading to hypomyelination in the brain of infants and children (Madhavarao *et al.,* 2005; Kirmani *et al.,* 2002; and Mehta and Namboodiri, 1995) and an accumulation of excess NAA. Concentrations of NAA in the brain increase approximately two-fold relative to healthy persons (Blüml, 1999; Leone *et al.,* 2000; Janson *et al.,* 2006), and persons with CD excrete large amounts of excess NAA (10-100-fold higher than normal) in the urine (Kvittingen *et al.*, 1986).

NAA is also a natural constituent of commonly consumed foodstuffs (*e.g.,* meat and eggs). Raw, unprocessed 356043 soybeans contain an increased concentration of NAA relative to control soybeans because of the aspartate acetylation activity of the GAT4601 enzyme. This elevated concentration of NAA translates to modest increases in estimated dietary exposure when factors such as soybean processing and soybean consumption patterns are taken into consideration (see Table 4 of the June 8, 2007 amendment to the 356043 soybean petition). The mean baseline dietary exposure of NAA to children ages 1-6 from existing foodstuffs is 0.0054 mg NAA/kg body weight/day. Predicted mean exposure after commercialization of 356043 soybean to children ages 1-6 is 0.0135 mg NAA/kg body weight/day.

Since the brain is the principal source of NAA, the levels of NAA being excreted in the urine of persons with CD provides us with a minimum estimate for the body's endogenously produced levels of NAA against which we can evaluate any contribution from dietary NAA. Estimates of total urinary NAA elimination in healthy and CD individuals can be calculated based on reported urinary concentrations of NAA and estimated daily urine output (Table 1). The concentration of NAA excreted by persons with CD (31-126 mg/kg body weight/day) relative to healthy persons (maximum of 2 mg/kg body weight/day) is indicative of the large amount of NAA naturally synthesized in the body.

To understand the potential impact of NAA dietary exposure on persons with CD, daily dietary NAA exposure estimates were compared to the amount of NAA excreted in the urine. Even using conservative assumptions, both the baseline and predicted daily dietary exposure to NAA are at least 2000-fold lower than quantities excreted in the urine on a daily basis by persons with CD. Therefore, the overwhelming majority of NAA in the urine of persons with CD is being produced endogenously, and the background dietary exposure to NAA is insignificant relative to the whole body pool of NAA in persons with CD. Further, the modest potential increase in dietary exposure to NAA from the commercialization of 356043 soybean is negligible when compared to the significant concentrations of NAA being synthesized in the body and to the whole body pool of NAA in persons with CD.

In conclusion, NAA synthesized by the brain is an essential substance that provides acetate for proper brain development in healthy individuals. NAA is present at relatively high levels in the brain. Although there could be modest increases in dietary exposure to NAA resulting from commercialization of 356043 soybeans, this increase is negligible compared to the amount of NAA being produced in the brain, as evidenced by the large amount of NAA being excreted in the urine of individuals with CD. Thus, we conclude that there will be no adverse impact to persons with CD from the potential increase in dietary NAA as a result of commercializing 356043 soybean.

Table 1. Comparison of NAA Excretion in Urine

	Amount of NAA in urine (mg NAA/1000 mg of creatinine) [1]	**Typical urine output in children (ml/day)**[2]	**Calculated amount of NAA excreted in urine (mg/kg body weight/day)**[3]
Healthy individuals	0 – 47	250-500	0 – 2
Canavan individuals	1,239 – 2,529	250-500	31 – 126

[1] From Janson *et al.*, 2006.

[2] Typical urine output in infants and children is in the range of 1-2 ml/kg/hour (WHO). Using a default body weight of 10 kg for infants and small children, they produce approximately 250-500 ml of urine over a 24 hour period.

[3] The mass of NAA excreted in urine was calculated using an assumption of 100 mg/dL of creatinine in the urine (Barr *et al.*, 2004) and the following sample equation:

(1,239 mg NAA/1000 mg creatinine) * (100 mg/dL) * (250 ml/day) * (dL/100 ml) = 310 mg NAA/day per 10 kg body weight of a child or 31 mg NAA/kg body weight/day

REFERENCES

Barr, D., Wilder, L., Caudill, S., Gonzalez, A., Needham, L. and Pirkle, J. (2005) Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. Environmental Health Perspectives 113: 192-200.

Birken, D.L. and Oldendorf, W.H. (1989) N-acetyl-L-aspartic acid: a literature review of a compound prominent in ^{1}H-NMR spectroscopic studies of brain. Neuroscience & Biobehavioral Review 13:23-31.

Blüml, S. (1999) In vivo quantitation of cerebral metabolite concentrations using natural abundance 13C MRS at 1.5 T. Journal of Magnetic Resonance 136:219-225.

Hershfield, J.R., Pattabiraman, N., Madhavarao, C.N. and Namboodiri, M.A.A. (2007) Mutational analysis of aspartoacylase: Implications for Canavan disease. Brain Research 1148:1-14.

Janson, C.G., McPhee, S.W.J., Francis, J., Shera, D., Assadi, M., Freese, A., Hurh, P., Haselgrove, J., Wang, D.J., Bilaniuk, L. and Leone, P. (2006) Natural history of Canavan disease revealed by proton magnetic resonance spectroscopy (^{1}H-MRS) and diffusion weighted MRI. Neuropediatrics 37:209-221.

Kirmani, B.F., Jacobowitz, D.M., Kallarakal, A.T. and Namboodiri, M.A. (2002) Aspartoacylase is restricted primarily to myelin synthesizing cells in the CNS: therapeutic implications for Canavan disease. Brain Res Mol Brain Res 107:176-82.

Kvittingen E., Guldal, G., Borsting, S., Skalpe, I., Stokke, O. and Jellum, E. (1986) N-acetylaspartic aciduria in a child with progressive cerebral atrophy. Clinica Chimica Acta 15, 158(3): 217-227.

Leone, P., Janson, C.G., Bilaniuk, L., Wang, Z., Sorgi, F., Huang, L., Matalon, R., Kaul, R., Zeng, Z., Freese, A., McPhee, S.W., Mee, E. and During, M.J. (2000) Aspartoacylase gene transfer to the mammalian central nervous system with therapeutic implications for Canavan disease. Annals of Neurology 48:27-38.

Madhavarao, C.N., Arun, P., Moffett, J.R., Szucs, S., Surendran, S., Metalon, R., Garbern, J., Hristova, D., Johnson, A., Jiang, W. and Namboodiri, M.A.A. (2005) Defective N-acetylaspartate catabolism reduces brain acetate levels and myelin lipid synthesis in Canavan's disease. Proceedings of the National Academies of Science USA 102:5221-5226.

Matalon, R. and Michals-Matalon, K. (1999) Recent advances in Canavan disease. Adv Pediatr 46:493-506.

Mehta, V. and Namboodiri, M.A. (1995) N-acetylaspartate as an acetyl source in the nervous system. Brain Res Mol Brain Res 31:151-7.

Miyake, M., Kakimoto, Y. and Sorimachi, M. (1981) A gas chromatographic method for the determination of N-acetyl-L-aspartic acid, N-acetyl-alpha-aspartylglutamic acid and beta-citryl-L-glutatmic acid and their distributions in the brain and other organs of various species of animals. Journal of Neurochemistry 36:804-810.

Tallan, H.H., Moore, S., Stein, W.H. (1956) N-acetyl-L-aspartic acid in brain. Journal of Biological Chemistry 219(1):257-264.

WHO. Essential surgical care manual. http://www.steinergraphics.com/surgical/002_03.2A.html

Zeng, B.J., Wang, Z.H., Ribeiro, L.A., Leone, P., De Gasperi, R., Kim, S.J., Raghavan, S., Ong, E., Pastores, G.M., Kolodny, E.H. (2002) Identification and characterization of novel mutations of the aspartoacylase gene in non-Jewish patients with Canavan disease. Journal of Inherited Metabolic Disease 25:557-70.

Addendum 2 to Petition for the Determination of Nonregulated Status for Herbicide Tolerant 356043 Soybean 06-271-01p

Submitting Company:

Pioneer Hi-Bred International, Inc.
7100 NW 62nd Avenue
PO Box 1014
Johnston, IA 50131-1014

Submitted by:

Tracy A. Rood, Registration Manager
Pioneer Hi-Bred International, Inc.
7250 NW 62nd Ave.
PO Box 552
Johnston, IA 50131
Telephone: 515-270-4036
Fax: 515-334-4478

OECD Unique Identifier: DP-356Ø43-5

June 8, 2007

NO CBI

Safety Assessment for Acetylated Amino Acids NAA and NAG

Because the levels of two acetylated amino acids, N-acetylaspartate (NAA) and N-acetylglutamate (NAG), in 356043 soybean grain are greater than the levels in control or reference soybean varieties (see Table 23 of petition 06-271-01p), potential health risks from dietary consumption of NAA and NAG from 356043 soybean were evaluated for both humans and livestock. These risks were evaluated by reviewing the metabolism of acetylated amino acids in humans and animals, analyzing levels of NAA and NAG in selected foods to understand the current sources and levels of exposure, analyzing levels of NAA and NAG in processed soybean fractions to understand the fate of these analytes during processing, and conducting an exposure assessment for humans to compare the dietary intake of NAA and NAG with and without exposure from 356043 soybean. The expected dietary exposures were then compared with levels of exposure in the broiler study where no adverse effects were seen.

1. Acetylated Amino Acid Metabolism in Humans and Animals

Acetylated amino acids are naturally occurring substances that have been identified in many biological systems including plants and animals. N-acetylglutamate is the first intermediate in the biosynthesis of arginine in prokaryotes, lower eukaryotes and plants (Caldovic and Tuckman, 2003).

Acetylation of N-terminal amino acids is the most commonly observed posttranslational modification of cytosolic proteins (Persson *et al.,* 1985; Polevoda and Sherman, 2002). It has been estimated that up to 80% of all cytosolic proteins in mammalian systems are α-N-acetylated (Brown and Roberts, 1976; Driessen *et al*., 1985). Enzymatic acetylation of amino acids using acetyl-CoA as the acetyl donor group can occur either cotranslationally or posttranslationally depending on the biological system (Polevoda and Sherman, 2000). Enzymes responsible for intracellular acetylation of amino acids (N^{α}-acetyltransferases) have been identified in rat, yeast, and other eukaryotic organisms (Lee *et al*., 1988 and 1989; Mullen *et al*., 1989; Yamada and Bradshaw, 1991a and 1991b). The biological role of acetylation of N-terminal amino acids of cytosolic proteins has been investigated and evidence indicates that this modification protects proteins from proteolysis by intracellular aminopeptidases (Jörnvall, 1975; Brown, 1979; Berger *et al*., 1981). A variety of additional roles for N-acetylation of amino acids in biological systems have been established (Polevoda and Sherman, 2002).

In light of the wide distribution and biological roles of acetylation, it is not surprising that a number of enzymes responsible for deacetylation of N-acetylated amino acids (*i.e*., acylases) have also been described. It has long been speculated that enzymatic deacetylation of amino acids is a "general" phenomenon in mammals because this reaction has been observed in numerous organs (Neuberger and Sanger, 1943). It was later hypothesized that this enzymatic reaction plays a role in the salvage of N-acetylated amino acids formed during the metabolic degradation of N-terminal acetylated proteins (Endo, 1980; Gade and Brown, 1981). To date, four classes of acylases (Types I – IV) that mediate deacetylation of N-acetylated amino acids have been described in mammalian systems that differ with regard to distribution and specificity.

2. Safety of Acetylated Amino Acids

Acetylation of proteins is commonly employed in the food industry to alter the solubility, water absorption capacity and emulsifying properties of protein concentrates (*e.g.*, El-Adawy, 2000; Ramos and Bora, 2004).

One well-characterized industrial application of acetylated amino acids is the use in livestock feed in cases when it is unsuitable to use free amino acids. For example, the quality of soy protein fractions can be limited by the concentrations of the essential amino acid L-methionine. This deficiency can be overcome by supplementation of diets with free L-methionine; however this can lead to development of objectionable odors and flavors from conversion of L-methionine to methional by Strecker degradation (Balance, 1961). Therefore, feed may be supplemented with N-acetyl-L-methionine.

In rats fed diets produced with soy protein isolates, growth and weight gains were similar regardless of whether they were supplemented with L-methionine or α-N-acetyl-L-methionine (Boggs *et al*., 1975; Amos *et al*., 1975).

In the case of α-N-acetyl-L-methionine, the basis of substitution for L-methionine is metabolic deacetylation. Metabolism studies have demonstrated that N-acetyl-^{14}C-L-methionine is readily metabolized to L-methionine in rats and in human infants (Boggs, 1978; Stegink *et al.,* 1980 and 1982). Similarly, metabolic deacetylation of α-N-acetyl-L-methionine has been reported in *in vitro* studies using rabbit intestinal epithelial cells (Brachet *et al.*, 1991). While these reports demonstrate that this enzymatic deacetylation occurs within the digestive system, there is also evidence that N-acetyl amino acids are deacetylated in other tissues (Yoshita and Lin, 1972).

Nutritional and metabolic studies with the α-N-acetyl forms of some other amino acids have been conducted in humans, rats, and pigs. In most cases, these studies have reported that the α-N-acetyl form of amino acids substitute for the constituent amino acid via metabolic deacetylation. Such results have been reported for glutamate (Magnusson *et al*., 1989; Neuhauser and Bassler, 1986; Arnaud *et al*., 2004), phenylalanine and tryptophan (Du Vigneaud *et al.,* 1934) and threonine (Boggs, 1978). Although specific information is not available for aspartate and glutamate, there is no reason to believe these amino acids would not also be biologically available when acetylated.

3. History of Consumption of Acetylated Amino Acids in Common Foods

Several different foods were analyzed for NAA and NAG content. These foods were selected because they had high concentrations of aspartic acid and glutamic acid (USDA, 2006) and it was assumed that this might be correlated with high levels of NAA and NAG, respectively. Foods were purchased either at local grocery stores or from on-line retailers. NAA was found to be present in a variety of foods including autolyzed yeast, chicken bouillon (vegan), eggs, ground turkey, ground chicken and ground beef (Table 1). NAG was found to be present in a variety of foods including autolyzed yeast, ground beef, ground turkey, and dried egg powder (Table 2). Since only a small number of selected foods were tested, it is likely NAA and NAG can be found in many other commonly consumed foods.

In conclusion, NAA and NAG are normal components of human diets, based on their presence in common foods. There is no evidence to indicate that exposure to either NAA or NAG from these sources is associated with adverse effects in humans.

Table 1. Levels of NAA in Selected Foods

Description	NAA (mg/kg fresh weight)
Autolyzed Yeast	12.57
Chicken Bouillon (vegan)	12.11
Dried Egg Powder	6.94
Ground Turkey	3.98
Ground Chicken	1.53
Whole Egg	1.38
Ground Beef	1.07
Egg White	0.55

Table 2. Levels of NAG in Selected Foods

Description	NAG (mg/kg fresh wt)
Autolyzed Yeast	159.75
Ground Beef	1.53
Ground Turkey	0.79
Dried Egg Powder	0.70
Chicken Bouillon (vegan)	0.36
Whole Egg	0.05

4. Quantitative Dietary Exposure Assessment for NAA and NAG

The dietary exposure of NAA and NAG was estimated using DEEM/FCID (Dietary Exposure Evaluation Model – Food Commodity Intake Database, Version 2.14, Exponent Inc., Washington DC). This model is commonly used by EPA Office of Prevention, Pesticides and Toxic Substances to estimate human dietary exposure. Annual average and 90th percentile exposures were calculated for the US population and several subpopulations.

Two scenarios each were modeled for NAA and NAG. The baseline scenarios used NAA and NAG concentrations in common foods (chicken, turkey, beef and eggs for NAA and turkey, beef and eggs for NAG) and control soybeans. The 356043 soybean scenarios used the same NAA and NAG concentrations in control soybeans and selected foods as above, with the addition of 356043 soybeans. In the 356043 soybean scenarios, 356043 soybean made up 45% of the consumed soybeans.

DEEM/FCID categorizes soy consumption as: soybean seed, soybean flour, soybean flour-babyfood, soymilk, soymilk-babyfood, soybean oil, and soybean oil-babyfood. Therefore, a processing study was conducted to determine the levels of NAA and NAG in various soybean fractions. Soybean raw grain values were used for soybean seed. In the dietary exposure assessment, processed categories NAA and NAG concentrations in soybean seed were adjusted using processing factors (Table 3). Processing of whole soybeans showed reductions in NAA and NAG levels in several fractions. The results of the assessments for the US population and selected population subgroups are summarized for NAA in Table 4 and for NAG in Table 5.

Table 3. Levels of NAA and NAG in Processed 356043 Soybean Fractions[1]

Fraction	NAA mg/kg	NAA Processing Factor	NAG mg/kg	NAG Processing Factor
Whole Soybeans (Raw Agricultural Commodity)	636.05		19.97	
Hull Material	1766.54	2.78	34.66	1.74
Defatted Raw Flakes	595.87	0.94	22.67	1.14
Defatted Toasted Meal	550.05	0.87	24.84	1.24
Defatted Flour	479.96	0.76	21.68	1.09
Refined, Bleached and Deodorized Oil[2]	<4	0.003	<4	0.10
Protein Isolate[2]	<4	0.003	<4	0.10
Protein Concentrate[2]	23.39	0.037	<4	0.10
Soy Milk[3]	30.71	NC	1.58	NC

[1]The 356043 soybean grain used in the processed fraction study was from the 2005/06 growing season in South America. The grain was the T7 generation (harvested from T6 plants).
[2]The lower limit of quantitation for NAA and NAG is 4 mg/kg. For the purposes of the exposure assessment, these values were set to 2.0, which is half of the LLOQ.
[3]Units for soy milk are in mg/L. Processing factors for soy milk were not calculated (NC).

Table 4. DEEM Dietary Exposure Analysis for NAA

Population Subgroup	NAA Exposure mg/kg body weight/day			
	Baseline, Mean	Baseline, 90th Percentile	356043 Soybean, Mean	356043 Soybean, 90th percentile
U.S. Population	0.0026	0.0058	0.0066	0.0128
Hispanics	0.0031	0.0070	0.0061	0.0134
Non-hispanic whites	0.0024	0.0053	0.0060	0.0119
Non-hispanic blacks	0.0032	0.0072	0.0062	0.0130
Non-hisp/non-white/non-black	0.0030	0.0065	0.0175	0.0313
All infants	0.0021	0.0077	0.0040	0.0103
Children 1-6 yrs	0.0054	0.0120	0.0135	0.0244
Children 7-12 yrs	0.0034	0.0077	0.0085	0.0160
Youth 13-19 yrs	0.0025	0.0056	0.0055	0.0118
Adults 20-49 yrs	0.0023	0.0049	0.0060	0.0111
Adults 50+ yrs	0.0020	0.0042	0.0048	0.0086

The estimated annual mean baseline exposure to NAA for the US population was 0.0026 mg/kg body weight/day. With the addition of 356043 soybeans, the estimated exposure increased to 0.0066 mg/kg body weight/day. The estimated 90th percentile baseline exposure was 0.0058 mg/kg body weight/day which increased to 0.0128 mg/kg body weight with the addition of 356043 soybeans.

These estimated exposures can be compared to the poultry study where broiler chickens were exposed to an average of 24.4 mg/kg body weight/day throughout a 42-day test (Part VII, Section 5a). The predicted human exposures are less than 0.1% of the NAA exposure in the poultry study where no adverse effects were observed.

The population subgroup with the highest predicted exposure is the non-hispanic/non-white/non-black subgroup (*i.e.*, primarily Asian) who had a mean annual exposure 0.0175 mg/kg body weight/day and a 90th percentile exposure of 0.0313 mg/kg body weight/day. The higher rates of soybean consumption among people of Asian descent are likely associated with consumption of soyfood products such as tofu, edamame, natto, miso, soy sauce and tempeh that are derived from special identity preserved soybean varieties (commonly called “food grade” soybeans). These “food grade” soybeans are distinguished from their grain- or oil-type counterparts by a larger seed and a clear or transparent hilum. Soybean varieties grown in the US currently for the Asian food market are typically non-GMO. Therefore, the predicted exposure of acetylated amino acids from 356043 soybeans to the non-hispanic/non-white/non-black subgroup is likely a large overestimate. Assuming this overestimate, predicted exposure in this group will be less than 0.1% of the NAA exposure in the poultry study where no adverse effects were observed.

These results indicate that there is a large margin of safety between predicted dietary exposures and levels of exposure observed in the poultry study.

Table 5. DEEM Dietary Exposure Analysis for NAG

Population Subgroup	NAG Exposure mg/kg body weight/day			
	Baseline, Mean	Baseline, 90th percentile	356043 Soybean, Mean	356043 Soybean, 90th percentile
U.S. Population	0.0019	0.0043	0.0032	0.0064
Hispanics	0.0020	0.0046	0.0032	0.0068
Non-hispanic whites	0.0018	0.0040	0.0030	0.0060
Non-hispanic blacks	0.0022	0.0051	0.0034	0.0075
Non-hisp/non-white/non-black	0.0025	0.0056	0.0047	0.0089
All infants	0.0021	0.0050	0.0049	0.0111
Children 1-6 yrs	0.0036	0.0079	0.0063	0.0117
Children 7-12 yrs	0.0027	0.0057	0.0045	0.0085
Youth 13-19 yrs	0.0020	0.0041	0.0032	0.0060
Adults 20-49 yrs	0.0017	0.0038	0.0028	0.0054
Adults 50+ yrs	0.0013	0.0029	0.0021	0.0041

The estimated annual mean baseline exposure to NAG for the US population was 0.0019 mg/kg body weight/day. With the addition of 356043 soybeans the estimated exposure increased to 0.0032 mg/kg body weight/day. The estimated 90th percentile baseline exposure was 0.0043 mg/kg body weight/day which increased approximately 50% to 0.0064mg/kg body weight/day with the addition of 356043 soybeans. These estimated exposures can be compared to the poultry study where broiler chickens were exposed to an average of 1.34 mg/kg body weight/day throughout a 42-day test. The predicted human exposures are less than 1% of the NAG exposure in the poultry study where no adverse effects were observed.

The population subgroups with the highest predicted exposure were the childrens groups between the ages of 1-6 yrs (average mean exposures of approximately 0.0063 mg/kg body weight/day, 90th percentile 0.0117 mg/kg body weight/day, 356043 scenario). Predicted exposure for these groups is less than 1% of the NAG exposure in the poultry study where no adverse effects were observed.

These results indicate that there is a large margin of safety between predicted dietary exposures and levels of exposure observed in the poultry study.

5. Animal Exposure Assessment for NAA and NAG

NAA and NAG can persist through the toasting process and be detected in defatted, toasted meal (Table 3) that is fed to livestock. Because low levels of NAA and NAG are currently found in commercial soybeans, livestock animals are already being exposed to these compounds.

Animal exposure to NAA and NAG would likely increase with the commercialization of 356043 soybeans. Approximately 50% of soybean meal is consumed by poultry. The safety of 356043 soybeans to poultry is supported by the results of the 42-day broiler study where no adverse effects were found. NAA and NAG levels were measured in the diets, and dietary consumption and body weight data were used to calculate exposure. The average exposure to NAA during the 42-day broiler study was 24.4 mg/kg body weight/day. The average exposure to NAG during the 42-day broiler study was 1.3 mg/kg body weight/day. No adverse effects were seen during the study at these levels. In general practice, exposure levels would be expected to be lower due to mixing of 356043 soybeans in general commerce with non-356043 soybeans.

The next largest use for soybean meal is swine (26%). Soybean meal makes up approximately 30% of starter diets and 20% of grow-finish diets (Iowa State University, 2002). Consumption of diet for young pigs (4 to 15 kg) is approximately 583 g/day and for larger pigs (35 to 100 kg) is approximately 2500 g/day (National Research Council, 1998). Assuming an NAA concentration of 513 mg/kg (Table 3) in toasted meal, this would result in an estimated exposure of 10.0 mg/kg body weight/day in young pigs and 4.3 mg/kg body weight/day in older pigs. Assuming an NAG concentration of 24.8 mg/kg in toasted meal, this would result in an estimated exposure of 0.49 mg/kg body weight/day in young pigs and 0.21 mg/kg body weight/day in older pigs. These estimated exposures are much lower than the exposure levels seen in the poultry study where no effects were seen.

Beef and dairy cattle consume approximately 18% of soybean meal. Because microorganisms in the rumen make microbial protein using nitrogen from both protein and non-protein sources, NAA and NAG are expected to be readily metabolized in the rumen.

6. Conclusions

The estimated daily intake of N-acetylaspartate (NAA) and N-acetylglutamate (NAG) from 356043 soybean (assuming 45% of the commodity soybeans grown in the US were 356043 soybeans) was calculated for humans. On a body weight basis, the mean and 90th percentile intakes of NAA are determined to be 0.007 and 0.013 mg/kg body weight/day, respectively. The mean and 90th percentile intakes of NAG are determined to be 0.003 and 0.006 mg/kg body weight/day, respectively.

NAA and NAG are typical constituents of the human diet present in soybean, eggs and meat. The current mean and 90th percentile estimated daily intakes of naturally occurring NAA in the general US population are 0.003 and 0.006 mg/kg body weight/day, respectively. For NAG, the mean and 90th percentile estimated daily intakes of naturally occurring NAG are 0.002 and 0.004 mg/kg body weight/day.

Commercialization of 356043 soybean may increase dietary exposure to NAA and NAG above current levels of exposure. Because acetylated amino acids are metabolizable and have been safely consumed by humans and animals, there are no safety issues that would be expected to result from this increase in potential exposure.

Additional References
(not included with original petition)

Amos, H.E., Schelling, G.T., Digenis, G.A., Swintosky, J.V., Little, C.O., and Mitchell Jr., D.E. (1975) Methionine replacement value of N-acetylmethionine and homocysteinethiolactone hydrochloride for growing rats. Journal of Nutrition 105:577-580.

Arnaud, A., Ramirez, M., Baxter, J.H. and Angulo, A.J. (2004) Absorption of enterally administered N-acetyl-l-glutamine versus glutamine in pigs. Clinical Nutrition 23:1303-1312.

Ballance P.E. (1961) Production of volatile compounds related to the flavor of foods from the Strecker degradation of DL-methionine. Journal of the Science of Food and Agriculture 12:532-536.

Berger, E.M., Cox, G., Weber, L. and Kenney, J.S. (1981) Actin acetylation in Drosophila tissue culture cells. Biochemical Genetics 19:321-331.

Boggs, R.W. (1978) Bioavailability of acetylated derivatives of methionine, threonine, and lysine. Advances in Experimental Medicine and Biology 105:571-586.

Boggs, R.W., Rotruck, J.T. and Damico, R.A. (1975) Acetylmethionine as a source of methionine for the rat. J Nutr 105:326-330.

Brachet, P., Gaertner, H., Tome, D., Dumontier, A.M., Guidoni, A. and Puigserver, A. (1991) Transport of N-α(or, -ε)-L-methionyl-L-lysine and acetylated derivatives across the rabbit intestinal epithelium. The Journal of Nutritional Biochemistry 2:387-394.

Brown, J.L. (1979) A comparison of the turnover of α-N-acetylated and nonacetylated mouse L-cell proteins. Journal of Biological Chemistry 254:1447-1449.

Brown, J.L. and Roberts, W.K. (1976) Evidence that approximately 80% of the soluble proteins from Ehrlich ascites cells are N-α-acetylated. Journal of Biological Chemistry 251:1009-1014.

Driessen, H.P.C., Dejong, W.W., Tesser, G.I. and Bloemendal, H. (1985) The mechanism of N-terminal acetylation of proteins. CRC Critical Reviews in Biochemistry 18:281-335.

Du Vigneaud, V., Loring, H.S. and Craft, H.A. (1934) The oxidation of the sulfur of the acetyl and formyl derivatives of *d*- and *l*-cystine in the animal body. Journal of Biological Chemistry 107:519-525.

Endo, Y. (1980) In vivo deacetylation of N-acetyl amino acids by kidney acylases in mice and rats. A possible role of acylase system in mammalian kidneys. Biochimica et Biophysica Acta 628: 13-18.

Gade, W. and Brown, J.L. (1981) Purification, characterization and possible function of α-N-acetylamino acid hydrolase from bovine liver. Biochimica et Biophysica Acta 662: 86-93.

Iowa State University (2002) Life Cycle Swine Nutrition.
http://www.ipic.iastate.edu/LCSN/LCSNutrition.pdf.

Jörnvall, H. (1975) Acetylation of protein N-terminal amino-groups. Structural observations on α-amino acid acetylated proteins. Journal of Theoretical Biology 55:1-12.

Lee, F.-J.S., Lin, L.-W. and Smith, J.A. (1988) Purification and characterization of an N alpha-acetyltransferase from *Saccharomyces cerevisiae*. Journal of Biological Chemistry 263: 14948-14955.

Lee, F.-J.S., Lin, L.-W. and Smith, J.A. (1989) Molecular cloning and sequencing of a cDNA encoding N alpha-acetyltransferase from *Saccharomyces cerevisiae*. J Biol Chem 264:12339-12343.

Magnusson, I., Kihlberg, R., Alvestrand, A., Wernerman, J., Ekman, L. and Wahren, J. (1989) Utilization of intravenously administered N-acetyl-L-glutamine in humans. Metabolism 38 (Suppl 1): 82-88.

Mullen, J.R., Kayne, P.S., Moerschell, R.P., Tsunasawa, S., Gribskov, M., Colavito-Shepanski, M., Grunstein, M., Sherman, F. and Sternglanz, R. (1989) Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. EMBO Journal 8:2067-2075.

National Research Council (1998) Nutrient Requirements of Swine. Tenth Revised Edition. Nutrient requirements of domestic animals. National Academy Press, Washington, DC.

Neuberger, A. and Sanger, F. (1943) The bioavailability of the acetyl derivatives of lysine for growth. The Biochemical Journal 37:515-518.

Neuahuser, M. and Bassler, K.H. (1986) Biological availability of glutamine from N-acetyl-L-glutamine in intravenous administration. Studies in the rat. Infusionstherapie und Klinische Ernährung 13(6): 292-296.

Persson, B., Flinta, C., von Heijne, G. and Jörnvall, H. (1985) Structures of N-terminally acetylated proteins. European Journal of Biochemistry 152:523-527.

Polevoda, B. and Sherman, F. (2000) N^{α}-terminal acetylation of eukaryotic proteins. Journal of Biological Chemistry 375:36479-36482.

Polevoda, B. and Sherman, F. (2002) The diversity of acetylated proteins. Genome Biology 3:1-6.

Stegink, L.D., Filer Jr., L.J. and Baker, G.L. (1980) Plasma methionine levels in normal adult subjects after oral loading with L-methionine and N-acetyl-L-methionine. Journal of Nutrition 110:42-49.

Stegink, L.D., Filer Jr., L.J. and Baker ,G.L. (1982) Plasma and urinary methionine levels in one-year-old infants after oral loading with L-methionine and N-acetyl-L-methionine. J Nutr 112:597-603.

USDA (2006) USDA Nutrient Database for Standard Reference, Release 19. Nutrient Data Laboratory Home Page, http://www.ars.usda.gov/ba/bhnrc/ndl.

Yamada, R. and Bradshaw, R.A. (1991a) Rat liver polysome N^{α}-acetyltransferase: Isolation and characterization. Biochemistry 30:1010-1016.

Yamada, R. and Bradshaw, R.A. (1991b) Rat liver polysome N^{α}-acetyltransferase: Substrate specificity. Biochemistry 30:1017-1021.

Yoshida, A. and Lin, M. (1972) NH_2-terminal formylmethionine- and NH_2-terminal methionine cleaving enzymes in rabbits. Journal of Biological Chemistry 247: 952-957.