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

Supporting document 1

Safety assessment – Application A1064

Food derived from Herbicide-tolerant Soybean Line CV127

SUMMARY AND CONCLUSIONS

Background

A genetically modified (GM) soybean line BPS-CV127-9, hereafter referred to as CV127, has been developed that is tolerant to the imidazolinone class of herbicides. This is achieved through expression of an imidazolinone-tolerant acetohydroxyacid synthase (AHAS) catalytic subunit encoded by the *csr1-2* gene derived from the plant *Arabidopsis thaliana*.

In conducting a safety assessment of food derived from soybean line CV127, a number of criteria have been addressed including: a characterisation of the transferred gene, its origin, function and stability in the soybean genome; the changes at the level of DNA, protein and in the whole food; compositional analyses; evaluation of intended and unintended changes; and the potential for the newly expressed protein to be either allergenic or toxic in humans.

This safety assessment report addresses only food safety and nutritional issues. It therefore does not address:

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

History of Use

Soybean (*Glycine max*), the host organism is grown as a commercial crop in over 35 countries worldwide. Soybean-derived products have a range of food and feed as well as industrial uses and have a long history of safe use for both humans and livestock. Oil, in one form or another, accounts for the major food use of soybean and is incorporated in salad and cooking oil, bakery shortening, and frying fat as well as processed products such as margarine.

Molecular Characterisation

Comprehensive molecular analyses of soybean line CV127 indicate that a single copy of the *csr1-2* gene expression cassette has been inserted at a single locus in the soybean genome. The introduced genetic elements are stably inherited from one generation to the next.

No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred during the transformation event.

The *csr1-2* gene expression cassette in CV127 is identical in sequence to the transforming plasmid DNA except for three point mutations, one of which occurred in the AHAS coding sequence, resulting in a conservative amino acid change. This mutation does not affect the function or activity of the AHAS enzyme. The remaining two point mutations did not occur in either a coding or regulatory region of the expression cassette and therefore do not have any impact on the expression of the inserted DNA.

The transformation event also resulted in a partial duplication of the *csr1-2* coding sequence directly before the 3' integration point, generating a 501 bp open reading frame (ORF) that extends into the 3' flanking sequence of the inserted DNA. There is no detectable transcription of this ORF in CV127. The inserted DNA also contains the majority of the *A. thaliana SEC61γ* (*AtSEC61γ*) subunit gene, which was inadvertently included in the DNA fragment used for the transformation. This gene is weakly transcribed in CV127. This transcription does not lead to the expression of any detectable protein (see below) and does not raise any safety concerns.

Characterisation of Novel Protein

Soybean line CV127 expresses the AHAS catalytic subunit from *A. thaliana*. This protein is immunologically indistinguishable from the endogenous imidazolinone-sensitive soybean AHAS, therefore protein expression levels were measured as total AHAS (endogenous soybean AHAS plus *A. thaliana* AHAS). The highest AHAS levels were found in young leaves and plants but typically at levels that were too low to be quantified. The levels in soybean seed were also too low to be quantified and no AHAS protein was able to be detected in any processed soybean fraction.

Soybean line CV127 also contains the *SEC61γ* subunit gene from *A. thaliana* which was shown in the molecular characterisation to be weakly transcribed. No AtSEC61γ subunit protein was able to be detected in CV127 therefore, if it is expressed, it is below the level of detection (< 15 ppb in seed).

Several studies were done to confirm the identity and physicochemical and functional properties of AHAS expressed in CV127. These studies demonstrated that the AHAS protein expressed in CV127 is as expected in terms of its physicochemical and functional properties but the mature form of protein is slightly larger (by 34 amino acids) than anticipated due to the N-terminal signal peptide being cleaved at a different site to what had been previously predicted. The AHAS protein expressed in CV127 is not glycosylated and exhibits the expected enzymatic activity.

An assessment was done to determine the potential toxicity and allergenicity of the AHAS protein as well as the AtSEC61γ subunit protein (should it be expressed). Both proteins are highly homologous to proteins that have been safely consumed in food. Bioinformatic analyses confirmed the lack of any significant amino acid sequence similarity of either protein to known protein toxins or allergens and digestibility studies demonstrated that both proteins would be rapidly digested in the gastrointestinal tract. The AHAS protein was also shown to be rapidly inactivated at temperatures > 60˚C and is not detectable in processed products such as meal, protein isolate, protein concentrate and oil. Taken together, the evidence indicates that both proteins are unlikely to be toxic or allergenic to humans.

Herbicide Metabolites

Herbicide tolerance in soybean line CV127 is achieved by the introduction of a gene encoding a herbicide-insensitive form of the AHAS enzyme. Studies have shown that the expression of such an enzyme has no impact on the uptake, translocation and metabolism of imidazolinone herbicides by the plant. No novel metabolites would therefore be expected as a result of the genetic modification.

Compositional Analyses

Detailed compositional analyses were done to establish the nutritional adequacy of seed from soybean line CV127 sprayed with imidazolinone herbicides. Analyses were done of proximate (moisture, crude protein, fat, ash, fibre), amino acid, fatty acid, vitamin, mineral, phytic acid, trypsin inhibitor, lectin, isoflavone, stachyose, raffinose and phospholipid content. The levels were compared to levels in the seeds of a control line grown alongside the GM line.

These analyses did not indicate any differences of biological significance between the seed from CV127 soybean and the control. Significant differences were noted in a number of constituents. However the differences were typically small and almost all mean values were within the range reported for conventional soybean varieties. Any observed differences therefore represent the natural variability that exists within soybean. The spraying of CV127 soybean with imidazolinone herbicides did not have a significant effect on seed composition.

In addition, no significant differences were identified in endogenous allergen content of seed from CV127 soybean compared to the non-GM counterpart.

Conclusion

No potential public health and safety concerns have been identified in the assessment of soybean line CV127. On the basis of the data provided in the present Application, and other available information, food derived from soybean line CV127 is considered to be as safe for human consumption as food derived from conventional soybean cultivars.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADF	acid detergent fibre
AHAS	acetohydroxyacid synthase
BLAST	Basic Local Alignment Search Tool
bp	base pairs
CDS	coding sequence
CTP	chloroplast transit peptide
DNA	deoxyribonucleic acid
dw	dry weight
ELISA	enzyme linked immunosorbent assay
EMBRAPA	Brazilian Agricultural Research Corporation
EMS	ethyl methane sulphonate
ER	endoplasmic reticulum
FAD	flavin adenine dinucleotide
FASTA	Fast Alignment Search Tool - All
FSANZ	Food Standards Australia New Zealand
fw	fresh weight
GM	genetically modified
ILSI	International Life Sciences Institute
kb	kilobases
kDa	kilo Dalton
LCMS	liquid chromatography mass spectrometry
LOD	limit of detection
LOQ	limit of quantitation
mg	milligram
NDF	neutral detergent fibre
ng	nanogram
OECD	Organisation for Economic Co-operation and Development
ORF	open reading frame
PCR	polymerase chain reaction
ppb	parts per billion
PVDF	polyvinylidene fluoride
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
ThDP	thiamine diphosphate
µg	microgram
US	United States of America
UTR	untranslated region

1. Introduction

A genetically modified (GM) soybean line BPS-CV127-9, hereafter referred to as CV127, has been developed that is tolerant to the imidazolinone class of herbicides.

Tolerance to imidazolinone herbicides in soybean line CV127 is achieved through expression of an imidazolinone-tolerant acetohydroxyacid synthase (AHAS)[1] catalytic subunit encoded by the *csr1-2* gene derived from the plant *Arabidopsis thaliana*. AHAS, which catalyses the first step in the biosynthesis of the branched-chain amino acids (valine, leucine, and isoleucine), is the target of several classes of structurally unrelated herbicides, including the imidazolinones, the sulfonylcarboxamides, the sulfonylureas and the triazolopyrimidines.

The AHAS catalytic subunit combines with a smaller regulatory subunit to form the AHAS enzyme complex. The regulatory subunit is necessary for full enzymatic activity and also for end-product feedback inhibition by the branched chain amino acids. The *A. thaliana* AHAS catalytic subunit expressed in soybean line CV127 has altered herbicide binding properties such that imidazolinone herbicides are unable to bind, and therefore inhibit, its activity. It is able to combine with the endogenous soybean regulatory subunit to form an imidazolinone-tolerant AHAS enzyme which is able to function in the presence of imidazolinone herbicides.

2. History of use

2.1 Host organism

The host organism for the transferred genes is soybean (*Glycine max* (L.) Merr.), belonging to the family Leguminosae. The soybean variety Conquista was used as the parent for the genetic modification described in this application. Conquista is a high-yielding non-GM commercial cultivar in Maturity Group VIII developed by EMBRAPA[2] for cultivation mainly in Brazil.

Soybean is grown as a commercial food and feed crop in over 35 countries worldwide (OECD 2001) and has a long history of safe use for both humans and livestock. The major producers of soybeans, accounting for 90% of world production, are the United States of America (US), Argentina, Brazil and China. Australia and New Zealand are net importers of soybean, however Australia grows crops extending from the tropics to temperate regions, mainly in the eastern states and as a rotational crop (James and Rose 2004).The seed is used mainly to produce meal for use in animal feed (Grey 2006).

Soybean food products are derived either from whole or cracked soybeans. Whole soybeans are used to produce soy sprouts, baked soybeans, roasted soybeans and traditional soy foods such as miso, tofu, soy milk and soy sauce. Cracked soybeans have the hull (seed coat) removed and are then rolled into flakes which undergo solvent extraction to remove the oil. This crude oil is further refined to produce cooking oil, shortening and lecithins as well as being incorporated into a variety of edible and technical/industrial products. The flakes are dried and undergo further processing to form products such as meal (for use in e.g. livestock, pet and poultry food), protein concentrate and isolate (for use in both edible and technical/industrial products), and textured flour (for edible uses).

[1]Also known as acetolactate synthase (ALS)
[2]The Brazilian Agricultural Research Corporation

The hulls are used in mill feed. Unprocessed (raw) soybeans are not suitable for food use, and have only limited feed uses, as they contain toxicants and anti-nutritional factors, such as lectins and trypsin inhibitors (OECD 2001). Appropriate heat processing inactivates these compounds.

Soybean oil constitutes approximately 30% of global consumption of edible fats and oils (American Soybean Association, 2011), and is currently the second largest source of vegetable oil worldwide (USDA, 2009). Oil, in one form or another, accounts for the major food use of soybean (Shurtleff and Aoyagi, 2007) and is incorporated in salad and cooking oil, bakery shortening, and frying fat as well as processed products such as margarine.

2.2 Donor organism

The donor organism for the AHAS catalytic subunit gene *csr1-2,* and associated regulatory elements, is *A. thaliana* (common names: thale cress, mouse ear cress). *A. thaliana* is a small flowering plant belonging to the mustard (Brassicaceae) family, which includes cultivated species such as cabbage and radish. *A. thaliana* is widely used as a model organism in plant biology and genetics and its genome was the first plant genome to be fully sequenced. The 125 megabase genome of *A. thaliana* contains 25,498 predicted genes, organised into five chromosomes, encoding proteins from 11,000 families (Arabidopsis Genome Initiative 2000). Although a member of the mustard family, *A. thaliana* is not commonly cultivated or harvested for food due to its small size; it therefore does not have a history of significant human consumption. There are however no reports of *A. thaliana* being allergenic or a source of toxins.

3. Molecular characterisation

Molecular characterisation is necessary to provide an understanding of the genetic material introduced into the host genome and helps to frame the subsequent parts of the safety assessment. The molecular characterisation addresses three main aspects: the transformation method together with a detailed description of the DNA sequences introduced to the host genome; a characterisation of the inserted DNA including any rearrangements that may have occurred as a consequence of the transformation; and the genetic stability of the inserted DNA and any accompanying expressed traits.

Studies submitted:

BASF Report No BPS-001-06 (2007) Molecular Characterization of Cultivance Soybean Event 127.

BASF Report No BPS-015-O8 (2008) Bioinformatics Analysis of Deduced Amino Acid Sequences of Open Reading Frames Contained in the Transgene Insert of Herbicide-Tolerant Soybean BPS-CV127-9.

BASF Report No BPS-006-09 (2009) Bioinformatics Analysis of Deduced Amino Acid Sequences of Open Reading frames Created by the Junctions of the Insert with Genomic DNA in Herbicide-Tolerant Soybean BPS-CV127-9.

BASF Report No BPS-005-09 (2009) Bioinformatics Analysis of the Genomic Area Surrounding the Transgene in Herbicide-Tolerant BPS-CV127-9 Soybean.

BASF Report No BPS-014-07 (2007) Molecular Bridging Study for herbicide Tolerant Soybean Herbicide-Tolerant Soybean BPS-CV127-9.

BASF Report No BPS-002-06 (2007) Determination of the 5' End of the *Arabidopsis thalianaSec61γ* Subunit Transcript in Cultivance Soybean Event 127.

3.1 Method used in the genetic modification

Embryonic tissue from a single seed from the soybean cultivar Conquista was transformed with a purified linear 6.2 kb *PvuII* fragment from plasmid pAC321 containing the *csr1-2* gene cassette (Figure 1). The transformation method used was microprojectile bombardment (Aragão et al 1996). Following bombardment, the tissue was transferred to selective media containing imazapyr; only those cells transformed with the *csr1-2* gene cassette continued to grow.

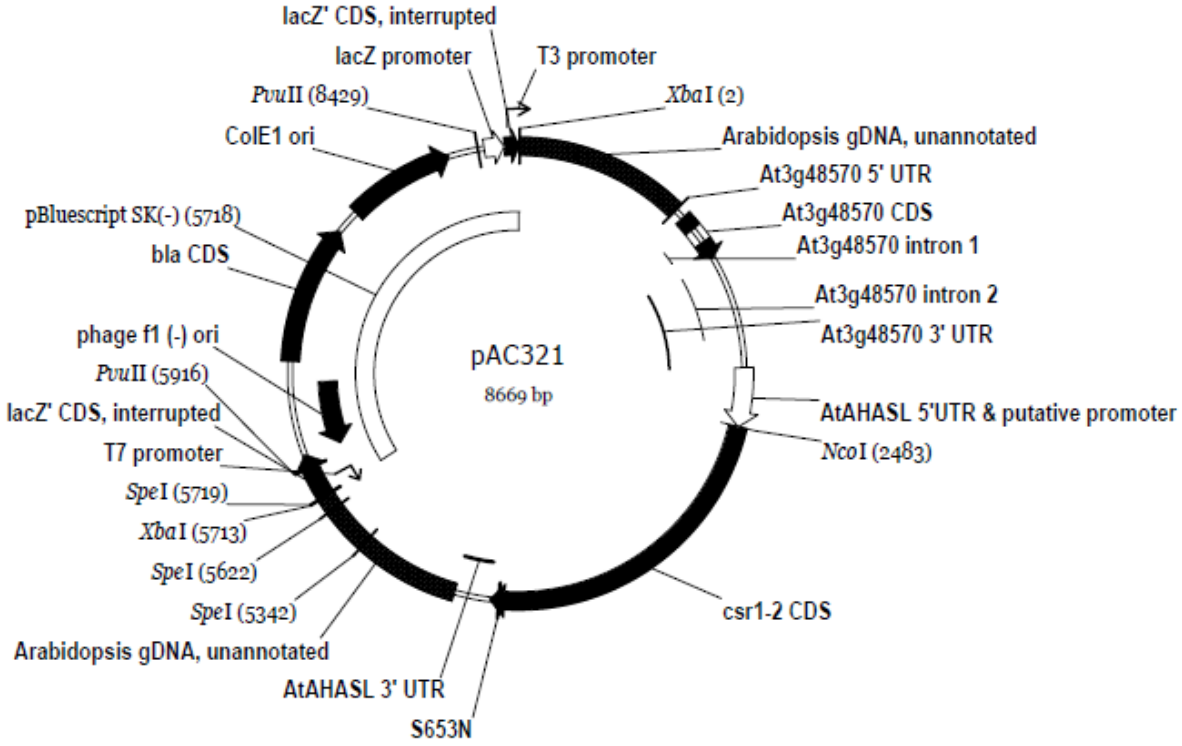

Figure 1: Map of plasmid pAC321. See Table 1 for a full description of all the genetic elements contained within the 6.2 kb PvuII fragment.

3.2 Description of the introduced genes

A full description of the genetic elements contained within the *PvuII* fragment from pAC321 is provided in Table 1.

3.2.1 csr1-2 gene

The *csr1-2* gene, encoding the AHAS catalytic subunit, was isolated from an imidazolinone-tolerant mutant of *A. thaliana* that had been obtained by direct selection of seedlings from an ethyl methane sulphonate (EMS) mutagenised population (Haughn and Somerville 1990). The coding region of the gene is 2013 nucleotides in length, encoding 670 amino acids with no identifiable introns (Mazur et al 1987). The gene is nuclear encoded. Imidazolinone tolerance is produced by a point mutation (G for A) resulting in a single amino acid substitution – serine to asparagine – at amino acid position 653 (Sathasivan et al 1990). The *csr1-2* gene that was introduced into soybean is under the regulatory control of the native *csr1-2* gene promoter and termination of transcription is also by the native *csr1-2* transcription termination signal.

3.2.2 Other sequences

In addition to the *csr1-2* gene, the *PvuII* fragment introduced into soybean also includes the majority of the *A. thaliana SEC61γ* (At*SEC61γ*) subunit gene[3], including the entire 207 bp At*SEC61γ* coding sequence. The AtSEC61γsubunit, together with the α and β subunits, is part of a heterotrimeric complex which forms a protein translocation channel associated with the endoplasmic reticulum (Osborne et al 2005).

In addition, two regions (approximately 1 kb each in length) of *A. thaliana* genomic DNA were also included in the *PvuII* fragment. No genes are currently annotated to these regions of the *A. thaliana* genome.

The inclusion of the At*SEC61γ* subunit gene and two regions of unannotated genomic DNA was unintentional. A 5717 bp *XbaI* fragment from *A. thaliana* that contains the *csr1-2* gene was cloned to create plasmid pAC321. In order that complete and functional *A. thaliana* promoter and terminator regions flanking the *csr1-2* coding sequence were contained within the genomic fragment, approximately 2.5 kb of unannotated DNA 5' to the start of the *csr1-2* coding sequence and 5' untranslated region (UTR) and approximately 1 kb of unannotated DNA 3' to the end of the *csr1-2* coding sequence was included. At the time of transformation a precise description of the *csr1-2* gene promoter region was not available as the sequencing of the *A. thaliana* genome had not been completed. The presence of the At*SEC61γ* subunit gene, as part of the unannotated DNA flanking the 5' end of the *csr1-2* coding sequence, only became apparent once additional *A. thaliana* genome sequence information became available. The 6.2 kb *PvuII* transformation fragment contained about 1 kb of unannotated DNA 5' to the At*SEC61γ* subunit 5'UTR, most of which is not present in CV127 soybeans. The At*SEC61γ* 5'UTR begins 18 nucleotides downstream from the 5' transgene integration site. As such, the Applicant considers it unlikely the insert contains the complete native promoter for the At*SEC61γ* gene.

3.3 Breeding to obtain CV127

An initial imidazolinone-tolerant transformation event (T_0) was advanced to the T_4 generation via self pollination then backcrossed to the original parent Conquista. The resulting progeny (F_1 generation) were then advanced to the F_8 generation using successive rounds of self pollination and selection. A single imidazolinone-tolerant line was selected for further breeding and was designated CV127 line 603. This line was used as a parent for crosses in a breeding programme to further the development of varieties expressing imidazolinone tolerance. One such cross was to the commercial non-GM soybean line Conquista x BRI98-641 to produce commercial lines suitable for cultivation in Brazil (Maturity Group VIII-IX). From this cross, individual plants from within the same F_3 family were evaluated by PCR to identify those that were homozygous for the *csr1-2* gene or were null segregants (i.e., did not contain either *csr1-2* alleles). From these selected plants, two F_4 seed bulks were created, one bulk for progeny of F_3 plants homozygous for *csr1-2* (line CV127) and one bulk for progeny of null F_3 plants (used as the control line for the field studies). Each F_4 bulk was multiplied in a greenhouse to obtain the F_5 seed of CV127 line 127 and the null segregant control line used in field studies. Another cross was made with the soybean line DX x BRS184[2], resulting in progeny adapted to the higher latitudes of Argentina (Maturity Group IV). The breeding steps are presented in Figure 2.

[3]Arabidopsis Genome Initiative locus code At3g48570

Table 1: Genetic elements present in the PvuII fragment from pAC321

Genetic Element	Range (bp)	Function
pBluescript SK(-) phagemid	8429-8669	*Escherichia coli* cloning vector; Stratagene Corporation (Short et al 1988).
PvuII restriction site	8429	
lacZ' CDS, interrupted	8590-8669	*E. coli* β-galactosidase alpha fragment coding sequence, interrupted by *A. thaliana* genomic DNA in pAC321; allows blue-white screening for DNA inserts in pBluescript SK(-) multiple cloning site by alpha-complementation.
T3 promoter	8632	Bacteriophage T3 promoter transcription initiation site; allows *in vitro* synthesis of RNA from DNA cloned in phagemid by T3 polymerase.
A. thaliana gDNA, unannotated	1-1051	*A. thaliana* genomic DNA; no genes currently annotated in this region.
A. thaliana locus At3g48570	1052-2119	Protein translocation complex SEC61 gamma chain-like protein from *A. thaliana*.
At3g48570 5' UTR	1052-1113	5' untranslated region for putative *A. thaliana* SEC61 gamma chain.
At3g48570 CDS	1114-1207, 1307-1422	Putative *A. thaliana* SEC61 gamma chain coding sequence.
At3g48570 intron 1	1208-1306	Putative *A. thaliana* SEC61 gamma chain intron 1, interrupts CDS.
At3g48570 3' UTR	1423-1442, 1916-2119	3' untranslated region for putative *A. thaliana* SEC61 gamma chain.
At3g48570 intron 2	1443-1915	Putative *A. thaliana* SEC61 gamma chain intron 2.
At*AHASL* 5' UTR and putative promoter	2120-2483	Putative promoter and 5' untranslated region for *A. thaliana* acetohydroxyacid synthase catalytic subunit.
csr1-2 CDS	2484-4496	Coding sequence for *A. thaliana* acetohydroxyacid synthase catalytic subunit with (S653N) point mutation which confers tolerance to imidazolinones (Sathasivan et al 1990).
At*AHASL* 3' UTR	4497-4714	3' untranslated region for *A. thaliana* acetohydroxyacid synthase catalytic subunit.
A. thaliana gDNA, unannotated	4715-5717	*A. thaliana* genomic DNA; no genes currently annotated in this region.
pBluescript SK(-) phagemid	5718-5916	*Escherichia coli* cloning vector; Stratagene Corporation (Short et al 1988).
lacZ' CDS, interrupted	5718-5916	*E. coli* β-galactosidase alpha fragment coding sequence, interrupted by *A. thaliana* genomic DNA in pAC321; allows blue-white screening for DNA inserts in pBluescript SK(-) multiple cloning site by alpha-complementation.
PvuII restriction site	5916	

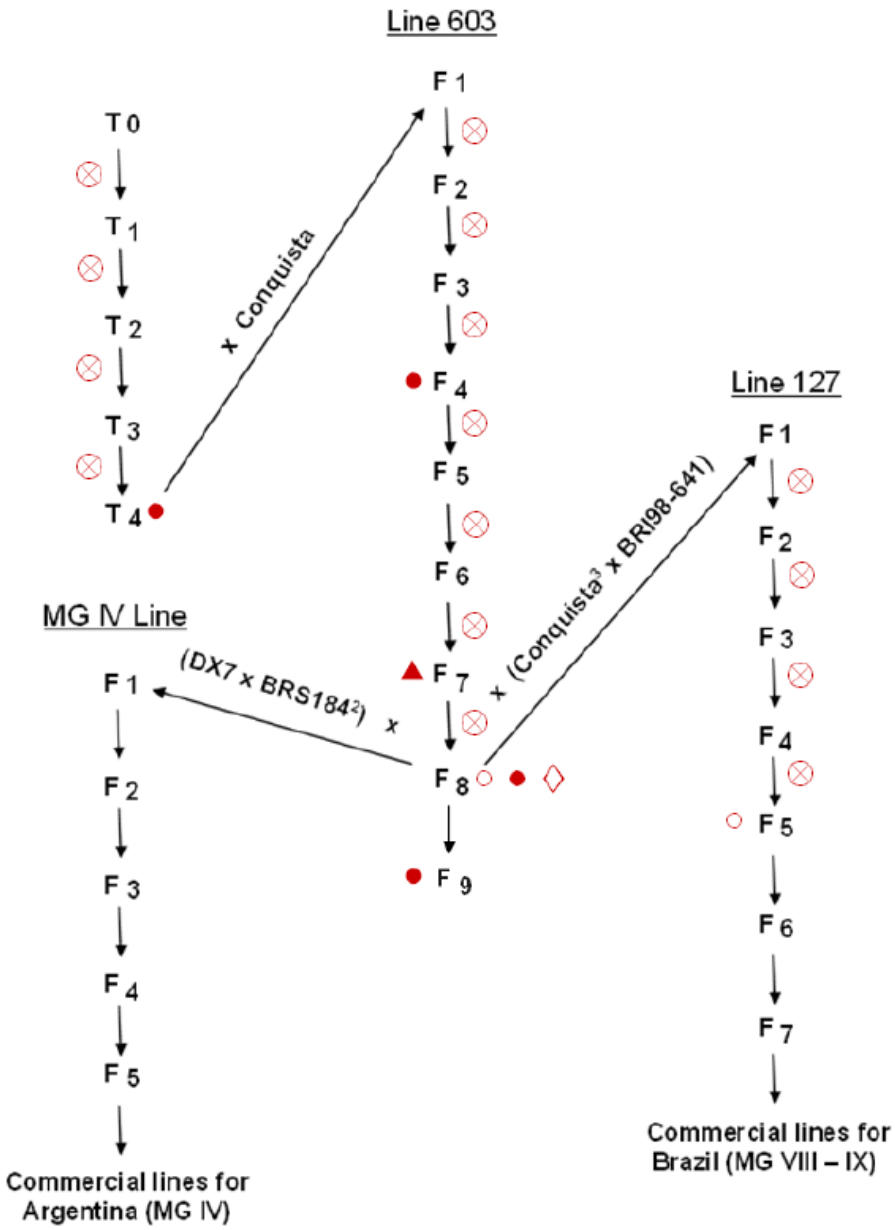

Figure 2: Breeding history of CV127. [Key: ⊗ = self pollination, ● = intergenerational stability, ○= Southern blot analysis, ◊ = Southern blot analysis for molecular characterisation, ▲= DNA sequence analysis]

3.4 Characterisation of the genes in the plant

A range of analyses was undertaken to characterise the genetic modification in soybean line CV127. These included: determination of insert copy number and integrity; DNA sequence and ORF analysis of inserted DNA as well as flanking and junction regions; and transcription analysis of newly created ORFs.

3.4.1 Insert copy number and integrity

Genomic DNA was isolated from young leaf tissue taken from the F_8 generation of soybean line CV127 (see Figure 2) and from the parental line Conquista, which served as the negative control. Two additional Conquista genomic DNA samples were spiked with one- and two-genome copy equivalents of pAC321 plasmid DNA and used as positive controls.

Southern blot analysis was used to determine the number of copies of inserted DNA, integrity of the *csr1-2* expression cassette as well as to confirm the absence of plasmid backbone DNA. Isolated genomic DNA from CV127 and Conquista (negative and spiked positive control) was digested with the restriction enzymes *NcoI*, *SpeI*, and *XbaI* (see Figure 1 for restriction site location) before being separated on agarose gels for Southern blotting (Sambrook et al 1989). Blots were hybridised separately with three different probes (the At*AHASL* 5' UTR; the *csr1-2* coding sequence; and the At*AHASL* 3' UTR) that spanned the entire *PvuII* fragment used for transformation.

No hybridising bands were detected when genomic DNA from Conquista was probed with each of the three probes, indicating that the endogenous soybean *ahas1* and *sec61γ* subunit genes were not detected under the stringency conditions used for the Southern blot. Bands of the expected size were readily detected with the pAC321-spiked genomic DNA from Conquista, indicating the Southern blot was sensitive enough to detect a single copy of inserted DNA.

Genomic DNA from CV127 probed with each of the three probes produced single hybridising bands, except for the *SpeI* digest probed with the coding sequence probe where a second hybridising band of about 800 bp was observed. The Applicant states the presence of this small hybridising band is consistent with the duplication of a small fragment of the *csr1-2* coding sequence at the 3' flanking sequence junction in CV127.

The results of the Southern blotting are consistent with the presence of a single copy of the *csr1-2* expression cassette, containing a functional copy of the *csr1-2* gene, along with the coding sequence of the gene for the AtSEC61γ subunit protein at the 5' end of the *csr1-2* gene and a small fragment of the *csr1-2* gene at the 3' end of the insert.

These same set of blots were also used to confirm the absence of plasmid backbone DNA, using two probes that spanned the entire pAC321 plasmid backbone. No hybridising bands were detected in genomic DNA from Conquista or CV127. Hybridising bands of the expected size were detected in pAC321-spiked Conquista genomic DNA.

3.4.2 Analysis of inserted DNA

Sequence analysis

DNA sequence analysis was done to confirm the organisation of the inserted DNA in CV127. Six overlapping PCR fragments were generated that spanned the entire inserted DNA in CV127 as well as the junction regions with adjacent soybean genomic DNA.

DNA sequence analysis revealed the total insert size in CV127 to be 4758 bp. Other than the insertion of a 376 bp fragment from the *csr1-2*coding sequence at the 3' integration point, the inserted DNA is identical to the sequence in the transformation vector, except for three point mutations. One of the mutations is a G to A in the *csr1-2* coding sequence, resulting in a change from arginine to lysine at amino acid position number 272. This is regarded as a conservative amino acid change and the Applicant states there was no impact on herbicide tolerance or enzymatic activity of the encoded AHAS protein (see Section 4.8.3). The two other mutations (G to A and G to C) were both downstream of the 3' UTR of the *csr1-2* gene, and had no impact on the function of the inserted DNA as they did not occur in either a coding or regulatory region. DNA sequence analysis also revealed that certain parts of the *PvuII* transforming fragment are not contained within the inserted DNA in CV127. Deletions of unannotated *A. thaliana* genomic DNA occurred at both the 5' end (1274 bp) and the 3' end (500 bp) upon the insertion of the *PvuII* fragment into the soybean genome. An alignment of the *PvuII* transformation fragment with the DNA insert in CV127 is provided in Figure 3.

Open reading frame analysis

The inserted DNA in CV127 was analysed for additional ORFs greater than 29 amino acids in length using Vector NTI, version 10.3.0 (Invitrogen). Identified ORFs were then subject to bioinformatic analyses where their deduced amino acid sequences were compared to the amino acid sequences of known protein toxins and allergens to determine if there was any significant similarity. These analyses are completely theoretical as there is no reason to expect that any of the identified ORFs would actually be expressed in CV127. Such analyses can provide assurance however that any of the putative encoded polypeptides are not significantly similar to proteins that may raise safety concerns.

An ORF was defined as any contiguous nucleic acid sequence which began at a start codon (i.e. ATG) and continued until the first termination codon (i.e. TAA, TAG or TGA). Any ORF from any of the six potential reading frames with a deduced amino acid sequence of 29 amino acids or greater which was located either entirely or partly within the inserted DNA was chosen for further analysis. The minimum ORF size of 29 amino acids was selected by the Applicant because this is the minimum polypeptide size that has been calculated to be capable of eliciting the clinical symptoms of an allergic reaction (Bannon et al 2002).

A total of 27 ORFs were identified including the two ORFs corresponding to the duplicated portion of the *csr1-2*coding sequence and the At*SEC61*γ coding sequence. Bioinformatic analysis of the full *csr1-2* coding sequence was not undertaken as part of this study (see Sections 4.9.2 and 4.10.2 for that analysis). The deduced amino acid sequences of the 27 ORFs were compared to entries in the Food Allergy Research and Resource Program (FARRP) Allergen Protein Database (Version 8.00)[4] and the GenBank® non-redundant peptide sequence database (downloaded October 26, 2008).The FARRP database comprises known or putative food, respiratory, venom/salivary or contact allergens. The database also includes a number of wheat glutenins and gliadins (proteins involved with coeliac disease) but only if there is evidence of IgE binding. The version of the database used by the Applicant contained 1313 entries categorised into 483 protein taxonomic groups representing 229 species. GenBank® is the NIH genetic sequence database which is available through the National Center for Biotechnology Information (NCBI) website. The protein sequence database contains all non-redundant GenBank coding sequence translations, protein sequences from NCBI's Reference Sequence Project, sequences derived from the Protein Data Bank (PDB), the last major release of the SWISS-PROT protein sequence database, the Protein Information Resource (PIR) Protein Sequence Database, and the Protein Research Foundation (PRF) Protein Sequence Database.

[4] www.allergenonline.com

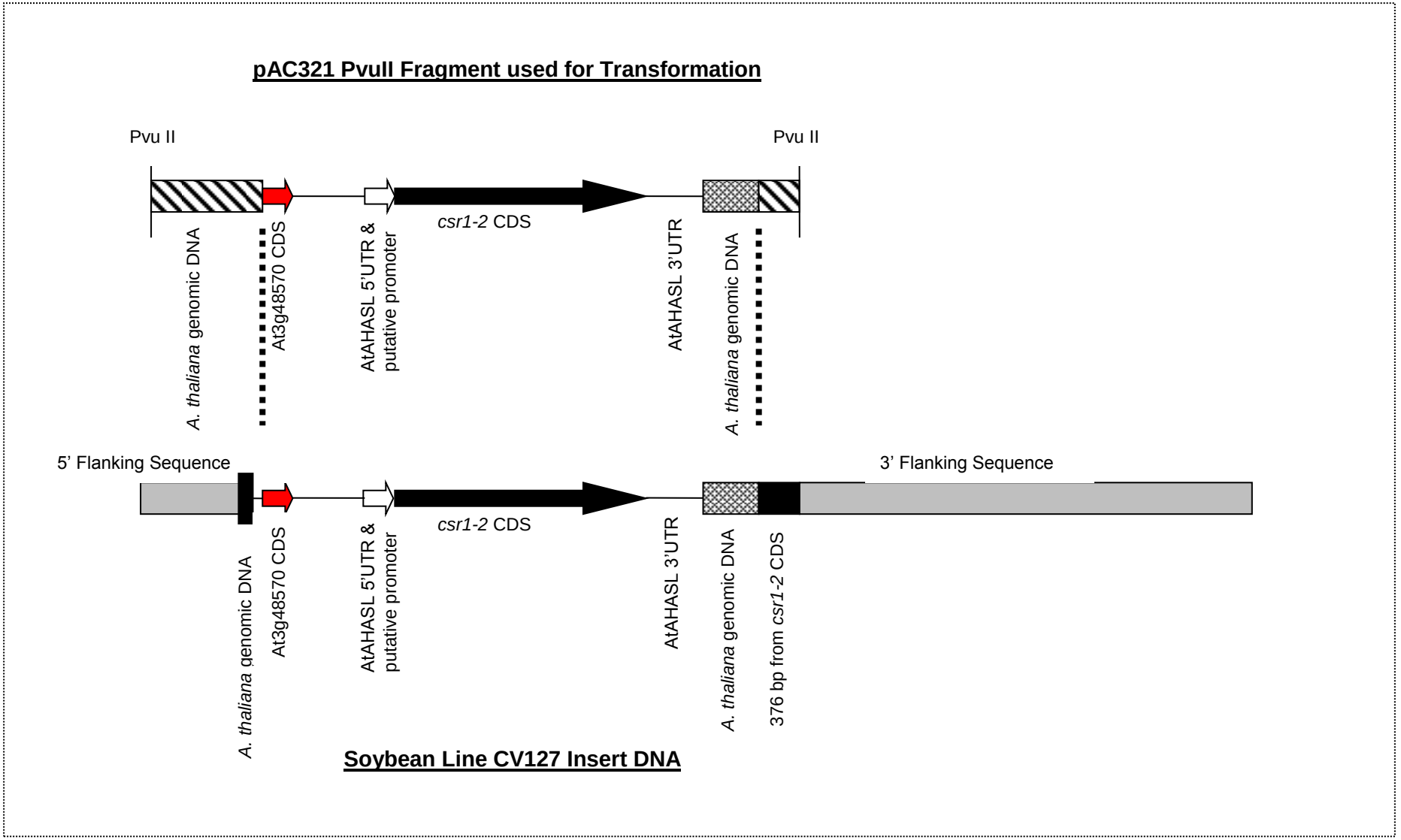

Figure 3: Alignment of pAC321 PvuII transformation fragment with CV127 insert DNA.

To determine sequence similarity to known or putative allergens, two different bioinformatic comparisons were done. The first involved an 8-amino acid short sequence identity search, the purpose of which is to identify potential linear IgE or possibly also T-cell epitopes. The second involved a search for 35% sequence identity or greater using a sliding window of 80 amino acids, which aims to also detect potential conformational IgE epitopes. ORFs with a deduced amino acid sequence greater than 80 amino acids were subdivided into all possible overlapping 80 amino acid segments. ORFs of 80 amino acids or less were analysed as a single intact sequence.

To determine sequence similarity to known protein toxins, as well as general similarity to known proteins, the deduced amino acid sequence of each ORF was submitted to a protein-protein Basic Local Alignment Search Tool (BLASTP)(NCBI Version 2.2.18) analysis. The BLASTP analysis identifies regions of local similarity between the putative polypeptide and proteins in the GenBank® non-redundant peptide sequence database. The top 250 alignments or all alignments with an expect (E) value less than 1, whichever is fewer, were manually compared to known toxins which act on humans, as listed in the United States Code of Federal Regulations (40 CFR Part 725.421).

None of the identified ORFs shared a sequence of 8 or more consecutive amino acids or 35% or greater identity over 80 amino acids with any known or putative allergens in the FARRP database. The identified ORFs also did not show significant similarity to any of the protein toxins listed in the US CFR. Only the ORFs encoding the At*SEC61*γ coding sequence and a portion of the *csr1-2* coding sequence showed any significant similarity to known proteins. The top BLAST result for the ORF encoding the At*SEC61*γ coding region was protein transport protein SEC61 gamma subunit, putative, *A. thaliana* (accession number NP_566909) and the top BLAST result for the ORF encoding the duplicated portion of the *csr1-2* coding sequence was acetolactate synthase catalytic subunit, *A. thaliana* (accession number ABJ80681). The remaining 25 ORFs did not show any significant similarity to any of the proteins in the GenBank® database.

3.4.3 Analysis of flanking sequences

Inverse-PCR, and subsequent DNA sequencing of the amplified fragments, was used to obtain the sequence of soybean DNA flanking the inserted *csr1-2* expression cassette. Approximately 1.3 kb of 5' flanking sequence and 4.6 kb of 3' flanking sequence was obtained.

BLAST analysis of the 5' flanking sequence queried against available public DNA databases (GenBank + EMBL[5] + DDBJ[6] + PDB sequences) and BASF Plant Sciences proprietary DNA databases identified a region of sequence identity with a proprietary expressed sequence tag (EST), confirming that the flanking sequence is native soybean DNA. The flanking sequence was further analysed for the presence of ORFs. This revealed the presence of a 315 bp ORF upstream of the 5' end of the insertion. The inserted DNA integrated 60 bp downstream from the stop codon of this predicted ORF.

Analysis of the 4.6 kb 3' flanking sequence showed that, upstream of the 3' integration point, there is a 376 bp sequence which corresponds to a portion of the *csr1-2* coding sequence, differing by only a single nucleotide. The duplication of this 376 bp portion of the *csr1-2* sequence at the 3' junction region created an ORF of 501 bp that extends from the inserted DNA into the 3' flanking sequence (see also 3.4.2 above).

---

[5]European Molecular Biology Laboratory database
[6]DNA Data Base of Japan

BLAST analysis of the 3' flanking sequence identified a region with sequence similarity to a soybean catalase gene (accession number Z12021). The inserted DNA integrated about 500 bp upstream of the putative gene, with the putative coding sequence being about 2.4 kb downstream from the integration point. The Applicant states that even if the putative catalase gene is active, it is unlikely to be affected by the insertion. BLAST analysis also identified a region of the distal 3' flanking sequence that is shared with a proprietary soybean EST.

Attempts by the Applicant to PCR amplify the integration site in the non-GM parental line Conquista using primers specific to the flanking regions in CV127 were unsuccessful. The Applicant believes this indicates the flanking DNA present in CV127 does not exist in the same context in the genome of Conquista, suggesting that a rearrangement at the insertion site occurred during transformation.

This rearrangement was confirmed following bioinformatic analysis of the 5' and 3' flanking regions of the inserted DNA in CV127. This showed that the *csr1-2* expression cassette has most likely inserted into soybean chromosome 2, with 141 bp in the 5' junction having sequence identity to soybean chromosome 18.

3.4.4 Analysis of junction regions

The junction regions between the inserted DNA and soybean genomic DNA were analysed to determine if additional ORFs had been created following insertion of the *csr1-2* expression cassette. Identified ORFs were subjected to bioinformatics analysis as above (section 3.4.2) to determine if any of the deduced amino acid sequences showed sequence similarity to known protein toxins or allergens. As stated previously, such analyses are completely theoretical.

The junction regions in CV127 were analysed for additional ORFs using Vector NTI, version 10.3.0 (Invitrogen). This analysis identified 24 ORFs at the 5' junction and six ORFs at the 3' junction. The six ORFs at the 3' junction region included the duplicated portion of the *csr1-2* coding region, which created an ORF of 501 bp extending into the 3' flanking region.

None of the 30 ORFs were found to share a sequence of 8 or more consecutive amino acids or 35% or greater identity to any known or putative allergens in the FARRP database. The identified ORFs also did not show any significant similarity to known protein toxins. As expected, the ORF corresponding to the duplicated portion of the *csr1-2* coding sequence showed significant similarity with *A. thaliana* acetolactate synthase. The remaining 29 ORFs did not show significant similarity with any of the proteins in the GenBank® database.

3.4.5 Reverse-transcription PCR analysis

Reverse-transcription (RT) PCR was done to determine if either the 376 bp duplication of the *csr1-2* coding sequence (generating an ORF of 501 bp extending into the 3' flanking region) or the At*SEC61*γ subunit gene coding sequence in CV127 was transcribed. Total RNA was extracted from young leaf tissue obtained from the F_7 (At*SEC61*γ subunit gene assay) or F_8 (*csr1-2* coding sequence assay) generation of soybean line CV127 (see Figure 2) and used as template for the RT-PCR.

AtSEC61γ *subunit gene coding sequence*

The endogenous soybean genes *Sec61γ* and *Iota* were used as positive controls and total RNA isolated from *A. thaliana* leaf and root tissues was also used as a positive control. Reactions without template RNA were used as negative controls.

Results showed the presence of a band of 393 bp which was amplified from total RNA obtained from CV127. This band was identical in size to that amplified from total RNA obtained from *A. thaliana* root and leaf tissue. The band amplified in CV127 was very faint compared to that amplified from *A. thaliana*, suggesting the At*SEC61γ* subunit gene is only weakly transcribed in CV127. The Applicant stated that DNA sequencing of the amplified band in CV127 confirmed its identity as At*SEC61γ* coding sequence (data not provided).

It was not obvious from either the RT-PCR or DNA sequencing which promoter was driving the transcription of the At*SEC61γ* subunit gene as only 18 bp of the native promoter had been included in the original genomic fragment from *A. thaliana*. The Applicant therefore undertook additional studies to further characterise the At*SEC61γ* subunit gene transcript produced in CV127, and in particular to determine its 5' end using RNA-ligase mediated rapid amplification of 5' complementary DNA (cDNA) ends (RLM-5'-RACE). This technique involves using reverse transcription of isolated messenger RNA to obtain a cDNA copy of the transcript. In this process, the unknown 5' end portion of the transcript is copied using a gene specific primer that recognises a known sequence from the middle of the transcript. Following cDNA synthesis, a homopolymeric tail is added to the 3' end of the synthesised cDNA. A second gene specific primer used in combination with a universal primer that recognises the homopolymeric tail can then be used to PCR amplify the cDNA. The PCR fragments can then be cloned and sequenced.

A total of 12 PCR fragments of various sizes were cloned and sequenced following the RLM-5'-RACE protocol. The two longest fragments were both 419 bp in length and turned out to be identical, with both being a chimera of the flanking soybean sequence and the At*SEC61γ* subunit gene. The At*SEC61γ* transcript is thus a transcriptional fusion with a short segment of adjacent soybean flanking sequence. The longest RLM-5'-RACE cDNA clones contained 89 nucleotides derived from CV127 5' flanking sequence, followed by 18 nucleotides of unannotated *A. thaliana* genomic sequence (which is presumably a small segment of the At*SEC61γ* subunit gene promoter), followed by the *A. thaliana* At*SEC61γ* coding sequence through to the binding site of the gene specific promoter used in the RLM-5'-RACE. The Applicant states it would be expected that the actual transcript in CV127 would continue through the native *A. thaliana* UTR of the At*SEC61γ* subunit gene, but this was not addressed in this particular study.

There were no nucleotide differences between the DNA sequence of the longest RLM-5'-RACE cDNA clones and the corresponding flanking soybean DNA and At*SEC61γ* subunit gene coding sequence of CV127, with allowances being made for intron splicing in the At*SEC61γ* subunit gene. The Applicant states the fact that introns 1 and 2 are properly spliced out of the transcript, in agreement with the full genomic sequence for *A. thaliana*, confirms that the RLM-5'-RACE cDNA clones are derived from mRNA transcript, rather than inadvertently cloned genomic DNA fragments.

csr1-2 coding sequence duplication

Primers specific for the soybean *Iota* gene were used in positive control reactions. Genomic DNA from the F_8 generation of soybean line CV127 was also used as a positive control. Reactions without template RNA were used as negative controls.

Results show the amplification of a 435 bp fragment from CV127 genomic DNA, with no detectable PCR product observed using total RNA as the template. These results suggest that the new 501 bp ORF generated by the duplication of a portion of the *csr1-2* coding sequence is not transcribed.

3.5 Stability of the genetic changes

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

In the case of soybean line CV127, the stability of the genetic changes across nine breeding generations was demonstrated by molecular methods and by progeny segregation analysis using traditional breeding methods.

3.5.1 Genetic stability

Genetic stability of the inserted DNA in CV127 was established using Southern blot analysis of genomic DNA from four different generations (T_4, F_4, F_8 and F_9 – see Figure 2).

Genomic DNA was digested with *NcoI* and *SpeI* and probed with the same three probes used in the Southern blot analyses to map the inserted DNA in CV127 (Section 3.4.1). Conquista genomic DNA was used as a negative control and Conquista genomic DNA spiked with one and two-genome copy equivalents of pAC321 was used as a positive control.

Multiple hybridising bands in genomic DNA from the T_4 generation were detected with all three probes, indicating that the T_4 generation contains multiple copies of the *csr1-2* expression cassette. In contrast, genomic DNA from the F_4, F_8 and F_9 generations produced the same Southern pattern (i.e. single hybridising bands) as previously observed for the insert and copy number analyses (Section 3.4.1). The Applicant states this result clearly indicates there were multiple copies of the insert in the T_4 generation and these segregated in the progeny of the cross between T_4 and Conquista (see Figure 2) with only a single copy being retained in the segregant that was selected for further breeding. The results of the Southern blot analyses indicate this single copy is stably inherited in subsequent generations.

Additional Southern blot analyses were also provided by the Applicant to demonstrate the stability of the insert when introduced into commercial soybean varieties.

3.5.2 Phenotypic stability

Segregation studies were used to establish the phenotypic stability of the insert and to confirm the molecular characterisation data.

The inheritance pattern of the *csr1-2* gene was determined by crossing the F_8 generation of CV127 (homozygous for the *csr1-2* gene) with a commercial non-GM line (Conquista[3] x BRI98-641). The heterozygous F_1 progeny was self fertilised to generate an F_2 generation of segregating individuals. The seeds from each individual F_2 parent plant were used to plant separate rows in the greenhouse. Each row of F_3 plants was considered to be a discrete family and six to eight plants from each family were sprayed with imazapyr to assign the families to three genetic groups: homozygous null; heterozygous for the *csr1-2* gene; or homozygous for the *csr1-2* gene. Two F_3 families that were found to be segregating for the *csr1-2* gene were selected for further studies.

Individuals were analysed by PCR to determine the number of null, heterozygous, and homozygous *csr1-2* plants. Chi-square analysis demonstrated a 1:2:1 segregation pattern which supports the inheritance of a single gene.

In further analyses, the F_3 plants of the selected families were self fertilised and seed was planted to generate an F_4 population of each family. The F_4 plants were sprayed with imazapyr and plants were rated as resistant or sensitive to the herbicide. As all the families were derived from plants heterozygous for the *csr1-2* gene, a ratio of three tolerant to one sensitive plant was expected within the F_4 progeny. The observed ratio was confirmed by the chi-square test and supports Mendelian inheritance of a single dominant gene.

3.6 Antibiotic resistance marker genes

No antibiotic resistance marker genes are present in soybean line CV127. Plasmid backbone analysis shows that no plasmid backbone has been integrated into the soybean genome during transformation, i.e. the β-lactamase gene, which was used as a bacterial selectable marker gene during laboratory manipulations, is not present in soybean CV127.

3.7 Conclusion

Comprehensive molecular analyses of soybean line CV127 indicate that a single copy of the *csr1-2* gene expression cassette has been inserted at a single locus in the soybean genome. No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. The *csr1-2* gene expression cassette in CV127 is identical in sequence to the transforming DNA except for three point mutations, one of which occurs in the AHAS coding sequence and results in a conservative amino acid change. This mutation does not affect the function or activity of the AHAS enzyme. The remaining two point mutations did not occur in either a coding or regulatory region of the expression cassette and do not affect the function of the introduced gene. The insertion of the *csr1-2* gene expression cassette in the soybean genome also resulted in a partial duplication of the *csr1-2* coding sequence directly before the 3' integration point, generating a 501 bp ORF that extends into the 3' flanking sequence of the inserted DNA. There is no detectable transcription of this ORF in CV127. The inserted DNA also contains the majority of the *A. thaliana SEC61γ* subunit gene, which was inadvertently included in the genomic fragment used for the transformation. This gene is weakly transcribed in CV127. The introduced genetic elements are stably inherited from one generation to the next.

4 Characterisation of novel proteins

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

4.2 Function and phenotypic effects of the novel proteins

There are potentially two novel proteins expressed in soybean line CV127. The first is the AHAS catalytic subunit encoded by the *csr1-2* gene. The second is the AtSEC61γ subunit, the gene for which was unintentionally included as part of the 5' UTR of the *csr1-2* gene. Studies undertaken as part of the molecular characterisation indicate this gene is weakly transcribed which raises the possibility that the encoded protein may also be expressed in CV127.

4.1.1 AHAS catalytic subunit

The AHAS enzyme (EC 2.2.1.6) is ubiquitous in plants and microorganisms and catalyses the first common step in the metabolic pathway for the biosynthesis of the branched chain amino acids valine, leucine and isoleucine (Stidham and Singh 1991, Duggleby and Pang 2000).The activity of AHAS is feedback-inhibited by the end product amino acids.

Biosynthesis involves two parallel reactions starting with the decarboxylation of pyruvate and its condensation with either a second molecule of pyruvate to yield 2-acetolactate or 2-ketobutyrate to yield 2-aceto-2-hydroxybutyrate. Additional enzymatic reactions convert 2-acetolactate to valine and leucine and 2-aceto-2-hydroxybutyrate to isoleucine (Figure 4).

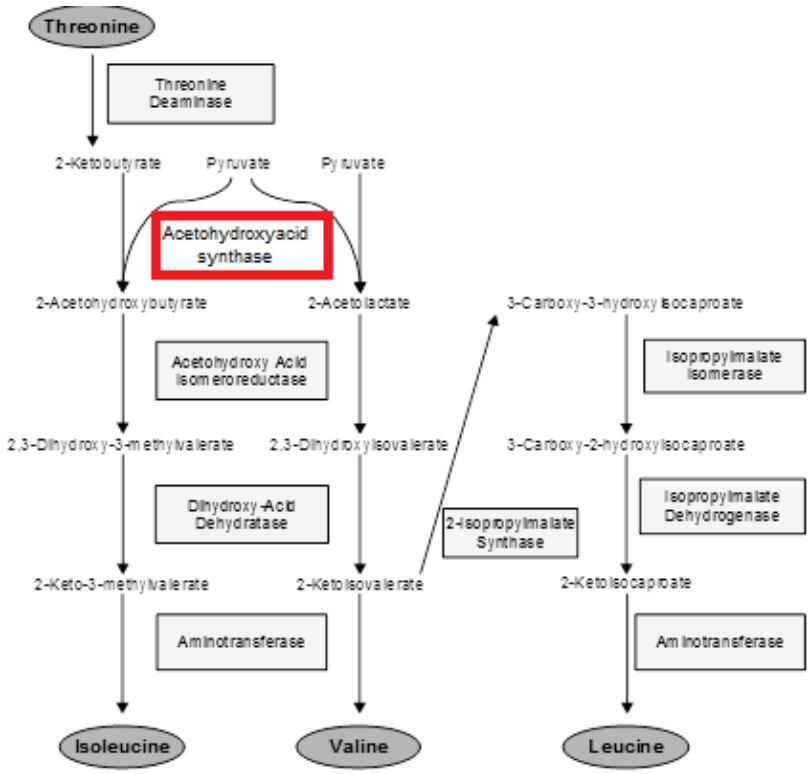

Figure 4: Biosynthetic pathway for valine, leucine and isoleucine in plants (Coruzzi and Last 2000).

Biochemical studies have shown that AHAS requires a cofactor FAD (flavin adenine dinucleotide), thiamine diphosphate (ThDP) and a bivalent metal ion, Mg^{2+} or Mn^{2+}, for activity (Davies 1963, Singh et al 1988). The enzyme uses ThDP as the coenzyme in the condensation reactions and Mg^{2+} is presumed to be required for binding of ThDP to the active site.

The structure of the AHAS enzyme was recently reviewed by Duggleby et al (2008). The enzyme is composed of two subunits, each with a specific function. The catalytic subunit, which contains ThDP and is active alone, has a molecular mass in the 59 – 66 kDa range, although in eukaryotes it is synthesised as a larger precursor protein. An N-terminal signal peptide, which is subsequently removed, is required to direct the enzyme to mitochondria in fungi and to chloroplasts in plants. The regulatory subunit possesses no catalytic activity but greatly stimulates the activity of the catalytic subunit and is also necessary for conferring end-product feedback inhibition by the branched chain amino acids. The molecular mass of the regulatory subunit varies greatly across species. In bacteria it is usually quite small and generally ranges between 10 and 20 kDa. In eukaryotes it is larger; 34 kDa in yeast and over 50 kDa in plants. Like the catalytic subunit, it is also synthesised as a larger precursor protein with an N-terminal signal peptide for targeting to a specific organelle.

The *A. thaliana* AHAS catalytic subunit expressed in soybean line CV127 has a deduced molecular mass of 72.6 kDa, which includes the native chloroplast transit peptide (CTP) at the N-terminus (Mazur et al 1987). During transport into the chloroplast, the CTP is removed to produce the mature protein. Based on a sequence comparison of the deduced amino acid sequence of plant and microbial AHAS proteins, the CTP of the *A. thaliana* AHAS catalytic subunit is predicted to be 85 amino acids in length and once cleaved produces a mature catalytic subunit of 585 amino acids and a molecular mass of about 64 kDa (Mazur et al 1987).

The *A. thaliana* AHAS catalytic subunit encoded by the *csr1-2* gene is identical to the native *A. thaliana* AHAS catalytic subunit except for its tolerance to the imidazolinone herbicides due to a single point mutation that results in the substitution of a serine with an asparagine at residue 653 (S653N). This amino acid substitution is known to interfere with the binding of imidazolinone herbicides (McCourt et al 2006, Duggleby et al 2008), resulting in imidazolinone tolerance, with no effect on either feedback inhibition by the branched-chain amino acids or normal catalytic function of the enzyme (Newhouse et al 1992).

The *A. thaliana* AHAS catalytic subunit shares 81.8% identity with the endogenous soybean AHAS catalytic subunit, which increases to 90.6% identity if conservative amino acid replacements are considered. Similar levels of amino acid identity also exist between the *A. thaliana* AHAS catalytic subunit and those present in corn and wheat.

4.1.2 AtSEC61γ subunit protein

The 69 amino acid AtSEC61γ subunit protein, together with the α and β subunits, is part of a heterotrimeric complex which forms a channel across the endoplasmic reticulum (ER) through which newly synthesised secretory proteins are translocated and via which membrane proteins are integrated (Rapoport et al 1996). This complex is highly conserved among eukaryotes and is also structurally related to the Sec proteins in bacteria which translocate proteins across the cytoplasmic membrane (Hartmann et al 1994).

Extensive characterisation of the mammalian Sec61 complex has been undertaken. Multiple copies of the heterotrimer surround a central aqueous pore which spans the ER membrane (Beckmann et al 1997, Hanein et al 1996). The α subunit is the largest of the subunits in the complex and spans the membrane ten times, whereas the smaller β and γ subunits span the membrane only once (Rapoport et al 1996).

The SEC61γ subunit protein from *A. thaliana* is 86% similar to the endogenous soybean SEC61γ subunit protein, differing only in 10 amino acids, with five of the differences being conservative.

4.3 Protein expression analysis

Studies submitted:

BASF Report No RF-1247-07 (2007) Analysis of Expression Levels of Arabidopsis Acetohydroxyacid Synthase (AHAS) Protein, by ELISA, in the Cultivance Soybean Event 127, Plants Grown in Brazilian Field Trials during the Summer 2006/2007 Season.

BASF Report No RF-1383-07 (2007) Analysis of Expression Levels of Arabidopsis Acetohydroxyacid Synthase (AHAS) Protein, by ELISA, in the BPS-CV127-9 Soybean, Plants Grown in Brazilian Field Trials during the 2007 Season.

BASF Report No BPS-013-08 (2008) Compositional Analysis of Fractions produced during Processing of Grain from Imidazolinone-Tolerant Soybean BPS-CV127-9 produced in Brazil in 2006/2007 and Fate of AtAHAS in these Fractions.

BASF Report No BPS-010-07A (2007) AtSEC61γ Subunit Protein Expression in Cultivance Soybean Event 127.

4.2.1 AHAS

Expression levels of the AHAS enzyme in different tissues of soybean line CV127 were determined using enzyme-linked immunosorbent assay (ELISA). Tissues were harvested from plants grown in field trials conducted in geographically distinct locations that were representative of commercial soybean production areas in Brazil. Seven field trials were conducted in the 2006/2007 growing season (Season 1) and six in the 2007 short growing season (Season 2).

Each field trial location had two replications (two separate plots) of the imidazolinone-tolerant CV127 soybean and the imidazolinone-sensitive control soybean[7]. All plots of the CV127 soybean were sprayed with an imidazolinone herbicide, with all plots of the imidazolinone-sensitive control being sprayed with a mixture of Bentazone + Acifluorfen-sodium as necessary to control weeds. In each of the two field trial seasons, leaf and seed samples were collected at all field trial sites at the V2 (plants 15 - 20 cm tall with three nodes and two unfolded leaflets) and R8 (full maturity) growth stages, respectively. In addition, six whole plants per plot, including roots, were collected at two of the trial sites at three different developmental stages, including the V2, R2 (plants in full bloom), and R8 stages. Three whole plants were maintained as such, and the remaining three whole plants were dissected into plant parts, including leaves, stems, roots, flowers, and pods, depending on the stage of plant development.

In addition, seed from soybean line CV127 and the control were processed using standard methods to produce oil, meal and protein isolate and concentrate fractions, with each processed fraction being analysed by ELISA for AHAS expression levels.

Expression levels of the AHAS protein in different tissues of CV127 soybean plants were determined by ELISA using immunoaffinity-purified polyclonal rabbit anti-AHAS peptide 2 antibody and Protein G-purified goat antibodies specific for AHAS. Due to the high amino acid homology between the *A. thaliana* AHAS enzyme encoded by the *csr1-2* gene and the endogenous soybean AHAS enzyme, the antibody used in the ELISA assay was not capable of distinguishing between the two enzymes but instead measures total AHAS protein in the samples.

[7] F5 null segregant for the 2006/2007 season; F6 null segregant for the 2007 season

The Applicant stated that it was expected tissues from CV127 soybean expressing both the *A. thaliana* AHAS catalytic subunit and the endogenous soybean AHAS catalytic subunit would have higher levels of total AHAS protein compared to the control expressing only the endogenous soybean AHAS protein. The difference between the AHAS amount in CV127 plants and the near-isogenic control can therefore be attributed to the expression of the *csr1-2* gene.

The lower limit of quantification (LOQ) and limit of detection (LOD) for the AHAS ELISA assay were determined experimentally for each tissue type in each experiment.

The results of the protein expression analyses are provided in Tables 2 and 3 below and indicate that the expression levels of the *A. thaliana* AHAS protein in CV127 soybean are extremely low. The highest total AHAS levels were found in young leaves and plants (V2 growth stage), with the levels declining as the plants matured (Table 2). For the vast majority of samples, however, AHAS was either non-detectable (level below the LOD) or was detectable but at a level below the LOQ. In Season 1, total AHAS levels in young leaves ranged from 53-128 ng/g fresh weight (335-714 ng/g dry weight) and in Season 2 ranged from 18-59 ng/g fresh weight (<LOQ-254 ng/g dry weight) (Table 3). In all seed samples, AHAS protein levels were either at or below the LOQ in both CV127 soybean and the near-isogenic control (Table 2). Not surprisingly, the AHAS protein was not detectable in any of the processed fractions produced from the seed of CV127 or the control (Table 4).

In the vast majority of cases where quantifiable levels of AHAS could be measured, and as anticipated, CV127 soybean contained higher levels of total AHAS compared to the control line. These higher levels can be attributed to the expression of the *csr1-2* gene.

Table 2: Total AHAS protein levels in different tissues and growth stages of CV127 soybean and the control line

Growth Stage	Soybean Line	Tissue	Mean ng AHAS/g Fresh Weight				Mean ng AHAS/g Dry Weight			
			Season 1		Season 2		Season 1		Season 2	
			Londrina	Santo Antônio de Posse	Brasília	Santo Antônio de Goiás	Londrina	Santo Antônio de Posse	Brasília	Santo Antônio de Goiás
V2	CV127	Whole Plant	61	40	<LOQ	<LOQ	314	214	<LOQ	<LOQ
		Leaves	103	128	27	18	511	714	133	113
		Roots	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		1st Trifoliate	55	60	21	21	278	300	111	126
	Control	Whole Plant	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		Leaves	15	14	<LOQ	<LOQ	78	73	<LOQ	<LOQ
		Roots	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		1st Trifoliate	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
R2	CV127	Whole Plant	ND[1]	34	<LOQ	<LOQ	ND	160	<LOQ	<LOQ
		Leaves	<LOQ	24	<LOQ	<LOQ	<LOQ	106	<LOQ	<LOQ
		Roots	<LOQ	17	<LOQ	<LOQ	<LOQ	50	<LOQ	<LOQ
		Flowers	<LOQ	22	<LOQ	-[2]	<LOQ	125	<LOQ	-
	Control	Whole Plant	ND	<LOQ	<LOQ	<LOQ	ND	<LOQ	<LOQ	<LOQ
		Leaves	ND	<LOQ	<LOQ	<LOQ	ND	<LOQ	<LOQ	<LOQ
		Roots	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		Flowers	<LOQ	<LOQ	<LOQ	-	<LOQ	<LOQ	<LOQ	-
R8	CV127	Whole Plant	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		Leaves	24	26	<LOQ	<LOQ	26	30	<LOQ	<LOQ
		Roots	<LOQ	<LOQ	15	17	<LOQ	<LOQ	42	48
		Seed	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
	Control	Whole Plant	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		Leaves	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
		Roots	ND	ND	15	<LOQ	ND	ND	41	<LOQ
		Seed	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ

[1] Non-detectable – value was below the LOD
[2] Not tested due to insufficient flower tissue at this site

Table 3: Total AHAS protein levels in young leaves of CV127 soybean and the control line

Growth Stage/Tissue	Location	Soybean Line	Mean ng AHAS/g fw		Mean ng AHAS/g dw	
			Season 1	Season 2	Season 1	Season2
V2/Leaves	Uberaba	CV127	80	59	427	254
		Control	16	<LOQ	86	<LOQ
	SeteLagoas	CV127	53	59	335	220
		Control	<LOQ	<LOQ	<LOQ	<LOQ
	Santo Antônio de Goiás	CV127	61	18	337	113
		Control	<LOQ	<LOQ	<LOQ	<LOQ
	Brasília	CV127	61	27	363	133
		Control	<LOQ	<LOQ	<70	<LOQ
	Ponta Grossa	CV127	92	-[1]	478	-
		Control	<LOQ	-	<LOQ	-
	Teresina	CV127	-	14	-	80
		Control	-	<LOQ	-	<LOQ
	Vilhena	CV127	-	39	-	207
		Control	-	14	-	83
	Londrina	CV127	103	-	511	-
		Control	15	-	78	-
	Santo Antônio de Posse	CV127	128	-	714	-
		Control	14	-	73	-

[1] Sample not tested in this location

Table 4: Total AHAS levels in processed soybean fractions from soybean line CV127 and the control line

Fraction	AHAS (ng/g dw)	
	CV127	Control
Defatted toasted meal	ND[1]	ND
Defatted untoasted meal	ND	ND - <13
Protein isolate	ND	ND
Protein concentrate	ND	ND
Oil	ND	ND
Seed	<14	<14

[1] Non-detectable – value was below the LOD

4.2.2 AtSEC61γ subunit

Western blot analysis of microsomal membrane protein fractions prepared from leaf and seed samples from glasshouse and field grown CV127 was done to determine if any *A. thaliana* SEC61γ subunit can be detected in CV127 soybean tissues. Microsomal membranes include the ER and were selected to enrich the AtSEC61γ subunit in the samples to be analysed. The leaf tissue was obtained from CV127 lines 603 and 127 and the seed was obtained from CV127 line 603 (see Figure 2). Control material for the study consisted of microsomal membrane preparations from leaf and seed tissue from the null segregant of CV127 line 603 (field grown) as well as leaf and seed tissue from the parental variety Conquista (glasshouse grown). The reference material for the study was a microsomal membrane preparation from leaves of *A. thaliana*. The *A. thaliana* SEC61γ subunit protein used as a positive control in the study was produced in an *E. coli* expression system.

Aliquots of microsomal membrane preparation from each of the samples were separated on an SDS polyacrylamide gradient gel then transferred to a polyvinylidene fluoride (PVDF) membrane which was subject to Western blotting using rabbit anti-AtSEC61γ polyclonal antibody and a chemiluminescent substrate. The LOD for this method was found to be about 2.5 ng AtSEC61γ subunit protein.

The *E. coli*-produced AtSEC61γ subunit protein was readily detected by Western blot analysis and the endogenous SEC61γ subunit protein in the microsomal membrane preparation from *A. thaliana* seed (but not leaves) was also able to be detected. In contrast, Western blotting was unable to detect any *A. thaliana* SEC61γ subunit protein in either seed or leaves from CV127 soybean and nor was it able to detect the endogenous soybean SEC61γ subunit protein in either seed or leaves from either the parental line (Conquista) or the CV127 null segregant. This was unexpected given the high level of amino acid sequence similarity (86%) between the two proteins and the fact that the antibody readily detected the *A. thaliana* SEC61γ subunit protein in the microsomal membrane preparation from *A. thaliana* seed. On the basis of these results, the Applicant estimates that if the *A. thaliana* SEC61γ subunit protein is expressed in CV127 the level will be no greater than 5 ppb in leaf or 15 ppb in seed.

4.3 Protein characterisation studies

A range of biochemical studies were done to confirm the identity as well as the physicochemical and functional properties of the *A. thaliana* catalytic subunit expressed in soybean line CV127. The plant extracts used in these studies were obtained from young leaves from the F_6 generation of CV127 line 127 (see Figure 2). In some studies, immunoaffinity purified AHAS was used and in others an ammonium sulphate AHAS-enriched extract was used. An ammonium sulphate precipitated AHAS-enriched extract from the F_6 null segregant (expressing only endogenous soybean AHAS) was used as a control.

A. thaliana AHAS catalytic subunit purified from an *E. coli* over-expression system was used as a reference substance for these studies. The AHAS protein produced in *E. coli* lacked the N-terminal 85 amino acid CTP but included an additional 38 N-terminal amino acids comprising a histidine tag to assist purification, the bacteriophage T7 gene 10 leader sequence to enhance expression and an Xpress™ tag for detection. The remainder of the protein is identical to that encoded by the *csr1-2* gene in CV127, including the S653N and R272K amino acid substitutions. The AHAS protein was purified by affinity chromatography, precipitated with ammonium sulphate then dialysed to remove salts. The resulting preparation was designated test substance AtAHAS-0107.

As the expression levels of the AHAS protein in soybean CV127 were extremely low, making purification of large quantities technically difficult, test substance AtAHAS-0107 was also used as a surrogate for plant-produced AHAS in a number of studies (see Section 4.9 and 4.10). The biochemical and immunological equivalence of the AtAHAS-0107 test substance to the CV127-produced AHAS protein was therefore also determined as part of the protein characterisation studies.

Studies submitted:

BASF Report No BPS-013-07 (2007) Characterization of AtAHAS Protein Produced in Imidazolinone-Tolerant Soybean BPS-CV127-9 and Comparison with AtAHAS Protein expressed in Recombinant *Escherichia coli*.

BASF Report No BPS-011-07 (2007) Characterization of Test Substance Arabidopsis Acetohydroxyacid Synthase (Lot#ATAHAS-01 07).

4.3.1 Western blot analysis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) followed by Western blot analysis was used to confirm that the *A. thaliana* AHAS catalytic subunit expressed in soybean line CV127 had the predicted molecular mass. Immunoreactive bands of about 64 kDa were detected in leaf extracts from both CV127 and the null segregant. Since endogenous soybean AHAS and *A. thaliana* AHAS have a number of epitopes in common, the Western blot revealed only a single band in CV127 leaf material. A smaller immunoreactive band of about 47 kDa was also detected in both the CV127 and control plant extracts, which is most likely a degradation product of the mature AHAS protein.

The microbially-produced AHAS protein is 623 amino acids in length consisting of the predicted mature AHAS protein of 585 amino acids plus an additional 38 N-terminal amino acids (see Section 4.7.1). The apparent molecular mass of the microbially-produced AHAS protein was therefore expected to appear slightly greater on Western blot than that of the plant-produced AHAS proteins. Both the microbially and plant-produced proteins however exhibited similar electrophoretic mobility suggesting that the mature plant-produced AHAS protein is larger than the predicted 585 amino acids. The Applicant states these additional amino acids are most likely derived from the CTP component at the N-terminal end of the protein which suggests a different CTP cleavage site to that predicted. This conclusion is supported by the results of amino acid sequencing of the plant-produced AHAS protein (Section 4.8.2).

4.3.2 Amino acid sequence determination

Attempts to obtain N-terminal amino acid sequence using Edman degradation methods were unsuccessful. Alternative methods, involving a combination of SDS-PAGE and liquid chromatography coupled with tandem mass spectroscopy (LC/MS/MS) were therefore used to obtain amino acid sequence data. Immunoaffinity-purified AHAS from the young leaves of soybean CV127 was subjected to SDS-PAGE and stained with Coomassie blue. The bands were excised then digested with trypsin to generate peptide fragments which were then subjected to MS analysis. The amino acid sequences obtained were then compared to the deduced amino acid sequence of the *A. thaliana* AHAS catalytic subunit.

The amino acid sequence obtained represented about 23% of the entire *A. thaliana* AHAS amino acid sequence (Figure 5). All the amino acid sequences obtained were identical to the deduced amino acid sequence of the *A. thaliana* AHAS catalytic subunit produced in *E. coli*. This confirms that the amino sequence obtained was for protein encoded by the introduced *csr1-2* gene, not from the endogenous soybean AHAS. The only anomaly was the production of a peptide fragment corresponding to amino acid residues 52 to 73. These amino acid residues were predicted to be part of the CTP and therefore were not expected to remain in the mature protein. The Applicant states the occurrence of this peptide fragment suggests the CTP cleavage site is located on the C-terminal side of residue 51, not residue 85 as had previously been predicted.

```
1     MAAATTTTTT   SSSISFSTKP   SPSSSKSPLP   ISRFSLPFSL   NPNKSSSSSR
51    RRGIKSSSPS   SISAVLNTTT   NVTTTPSPTK   PTKPETFISR   FAPDQPRKGA
101   DILVEALERQ   GVETVFAYPG   GASMEIHQAL   TRSSSIRNVL   PRHEQGGVFA
151   AEGYARSSGK   PGICIATSGP   GATNLVSGLA   DALLDSVPLV   AITGQVPRRM
201   IGTDAFQETP   IVEVTRSITK   HNYLVMDVED   IPRIIEEAFF   LATSGRPGPV
251   LVDVPKDIQQ   QLAIPNWEQA   MKLPGYMSRM   PKPPEDSHLE   QIVRLISESK
301   KPVLYVGGGC   LNSSDELGKF   VELTGIPVAS   TLMGLGSYPC   DDELSLHMLG
```

```
351   MHGTVYANYA   VEHSDLLLAF   GVRFDDRVTG   KLEAFASRAK   IVHIDIDSAE
401   IGKNKTPHVS   VCGDVKLALQ   GMNKVLENRA   EELKLDFGVW   RNELNVQKQK
451   FPLSFKTFGE   AIPPQYAIKV   LDELTDGKAI   ISTGVGQHQM   WAAQFYNYKK
501   PRQWLSSGGL   GAMGFGLPAA   IGASVANPDA   IVVDIDGDGS   FIMNVQELAT
551   IRVENLPVKV   LLLNNQHLGM   VMQWEDRFYK   ANRAHTFLGD   PAQEDEIFRN
601   MLLFAAACGI   PAARVTKKAD   LREAIQTMLD   TPGPYLLDVI   QPHQEHVLPM
651   IPNGGTFNDV   ITEGDGRIKY
```

Figure 5: Predicted amino acid sequence of the A. thaliana AHAS catalytic subunit. The amino acids underlined represent the sequences obtained by LC/MS/MS. The two amino acid substitutions (S653N and R272K) are boxed. The glutamate (E) residue at the previously predicted cleavage site of the CTP is highlighted in black. The four potential N-glycosylation sites [NX(S,T)X] are indicated by shading.

As AHAS proteins are present at only very low levels in plant tissues, the amino acid sequence of the *A. thaliana* AHAS catalytic subunit had not previously been experimentally determined, and so the precise location of the cleavage site for the CTP is unknown. However, based on sequence comparisons between the deduced amino acid sequence of plant and microbial AHAS proteins, the CTP of the *A. thaliana* AHAS catalytic subunit was predicted to be cleaved at the C-terminal side of residue 85 (Mazur et al 1987). Based on this information, the *E. coli* expression system for the production of large quantities of the *A. thaliana* AHAS catalytic subunit was designed to produce a protein lacking the first 85 amino acids corresponding to the predicted CTP. The amino acid sequencing data however suggest that the mature *A. thaliana* AHAS protein expressed in CV127 is potentially at least 619 amino acids in length, which would explain why it appears similar in molecular mass to the 623 amino acid AHAS protein produced in *E. coli*. The *E. coli*-produced AHAS protein therefore lacks the first 34 amino acids of the mature form of the AHAS protein, as expressed *in planta*.

4.3.3 Glycosylation

Analyses were done to assess whether any post-translational glycosylation of the plant-produced AHAS protein had occurred. The *A. thaliana* AHAS catalytic subunit expressed in soybean line CV127 contains four potential glycosylation sites (Figure 5). At least one of these is present in the CTP and therefore would not be present in the mature protein. The Applicant states however that there are numerous *O* glycosylation sites (any serine or threonine). It was not anticipated that AHAS would be glycosylated as it does not enter the secretory pathway but rather is transported to the chloroplast where proteins typically are not glycosylated.

As glycosylation of proteins by *E. coli* is rare, the *E. coli*-produced AHAS was used as a negative control, along with creatinase, a non-glycosylated enzyme. Transferrin, which is known to contain 5% glycan moieties by weight, corresponding to about 25 glucose equivalents per molecule, was used as a positive control. Samples were separated on a polyacrylamide gel, transferred to a PVDF membrane, where any glycan moieties were oxidised using periodate, labelled with digoxigenin, and then detected using an anti-digoxigenin antibody coupled to alkaline phosphatase. The limit of detection for this method was 2-4 molecules of glucose equivalents per molecule of AHAS.

There was no evidence of glycosylation in either the plant or *E. coli*-produced AHAS.

4.3.4 Enzymatic activity

A number of studies were done which examined the enzymatic activity of the AHAS proteins present in soybean line CV127 (a mixture of endogenous (sensitive) soybean AHAS and imidazolinone-tolerant *A. thaliana* AHAS) and the control plant line (the null segregant) and compared this to the purified *E. coli*-produced AHAS protein. These studies determined the catalytic activity of the enzymes, their response to feedback inhibition by the end-product amino acids, valine and leucine, as well as their sensitivity to inhibition by an imidazolinone herbicide.

The AHAS reaction involves the condensation of two molecules of pyruvate to form acetolactate (an acetohydroxy acid). The activity of AHAS can be measured using an assay which indirectly detects the formation of acetolactate (Singh et al 1988). The assay involves conversion of acetolactate to acetoin which is detected colourimetrically (A_{530nm}) via the formation of a complex with creatine and naphthol. One unit of AHAS activity is defined as 1 µmole acetoin produced per minute.

The *E. coli*-produced AHAS and both plant-produced AHAS proteins exhibited the expected enzymatic activity with the substrate pyruvate, confirming they all have the same functional activity. However, the specific activity of the ammonium sulphate AHAS-enriched plant extracts (5.227 units/µg AHAS for the CV127 extract and 6.431 units/µg AHAS for the control plant extract) was much higher than that of the *E. coli*-produced AHAS preparation (0.005 units/µg AHAS). The Applicant states this is most likely due to a loss of enzyme activity as a result of the extensive purification of the *E. coli*-produced AHAS.

Both the CV127-produced and *E. coli*-produced AHAS proteins showed tolerance to inhibition by the imidazolinone herbicide imazethapyr, compared to the control plant AHAS protein. An intermediate level of inhibition of the CV127-produced AHAS proteins was observed which was expected due to this preparation being a mixture of endogenous (sensitive) soybean AHAS and imidazolinone-tolerant *A. thaliana* AHAS. The Applicant noted that the degree of herbicide inhibition observed for the plant-produced AHAS proteins (CV127 derived and control) was lower than anticipated but may reflect the tolerance that soybean naturally has to the particular imidazolinone herbicide (imazethapyr) used for the assay.

Feedback inhibition by the branched chain amino acids leucine and valine was monitored by measuring AHAS activity of each of the preparations in the presence of increasing concentrations of the amino acids. No feedback inhibition was observed with the *E. coli*-produced AHAS. This was expected due to the absence of the regulatory subunit, which is responsible for mediating feedback regulation. The level of feedback inhibition observed for both the CV127-produced AHAS and the control (endogenous) AHAS was very similar.

A study was also done to compare the enzymatic activity of wild type (imidazolinone sensitive) *A. thaliana* AHAS to that of the imidazolinone-tolerant mutant AHAS (bearing the S653N substitution) and the AHAS expressed in CV127 (bearing the S653N and R272K substitutions). The main purpose of this study was to determine whether the R272K substitution, which occurred during the plant transformation process, had any impact on the function or activity of the AHAS catalytic subunit.

The coding regions of each of the three AHAS proteins were cloned into a bacterial expression vector for *in vitro* expression and purification. Each protein was then assayed for enzyme activity in the absence and presence of imazethapyr. ELISAs were done in parallel to determine the relative quantities of expressed recombinant AHAS protein. These assays indicated that expression and quantity of recombinant AHAS protein was comparable in each of the samples tested.

The enzymatic activity of the wild type AHAS was significantly inhibited by imazethapyr. At the lowest concentration of imazethapyr tested (9.75 µM) there was less than 20% of activity remaining compared to the control (wild type AHAS in absence of imazethapyr). In contrast, the S653N AHAS and S653N R272K AHAS proteins were equally tolerant to imazethapyr. These results indicate that the R272K substitution has no affect on the function or activity of the AHAS catalytic subunit.

4.4 Potential toxicity

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

The assessment focuses on: whether the novel protein has a prior history of safe human consumption, or is sufficiently similar to proteins that have been safely consumed in food; amino acid sequence similarity with known protein toxins and anti-nutrients; structural properties of the novel protein including whether it is resistant to heat or processing and/or digestion. Appropriate oral toxicity studies in animals may also be considered, particularly where results from the biochemical, bioinformatic, digestibility or stability studies indicate a concern.

While the At*SEC61γ* gene is only weakly transcribed (Section 3.4.6) and no protein can be detected in CV127 (Section 4.2.2), there remains the possibility that trace amounts of the protein are present in soybean below the level of detection. The AtSEC61γ subunit was therefore included in the assessment of potential toxicity.

Studies submitted:

BASF Report No BPS-014-08 (2008) Bioinformatics Analysis of Deduced Amino Acid Sequences of *Arabidopsis thaliana ahasl* and *SEC61γ* from Herbicide-Tolerant Soybean BPS-CV127-9 for Allergenicity and Toxicity Potential.

BASF Report No BPS-018-07 (2008) Heat Stability of Arabidopsis Acetohydroxyacid Synthase Present in Test Substance AtAHAS-0107.

BASF Report No 99C0295/07011 (2008) Acute Oral Toxicity Study in CD®-1-mice.

4.4.1 History of human consumption

Both the AHAS catalytic subunit and the SEC61γ subunit are derived from *A. thaliana* which is not a commonly consumed plant. However both proteins share a high degree of amino acid sequence similarity with their soybean counterparts (81.8 % identity for the AHAS catalytic subunit and 86 % identity for the AtSEC61γ subunit). The *A. thaliana* AHAS catalytic subunit also has similar levels of homology with the AHAS catalytic subunits present in corn and wheat. Both proteins can thus be regarded as sufficiently similar to proteins that have been safely consumed in food.

4.4.2 Amino acid sequence similarity to known protein toxins

Bioinformatics analyses were done to determine the amino acid sequence similarity of the *A. thaliana* AHAS catalytic subunit (containing the S653N and R272K substitutions) and AtSEC61γ subunit to known protein toxins.

The deduced amino acid sequences of both proteins were compared to proteins in the GenBank non-redundant peptide sequence database using BLASTP analysis, as previously described in Section 3.4.2. The top 250 alignments or all alignments with an expect (E) value less than 1, whichever is fewer, were manually compared to known toxins which act on humans, as listed in the United States Code of Federal Regulations (40 CFR Part 725.421).

Neither of the protein sequences showed significant similarity to proteins known to be toxic to humans.

4.4.3 In vitro digestibility

See Section 4.5.3.

4.4.4 Stability to heat and processing

Stability to heat and processing was determined for the *A. thaliana* AHAS catalytic subunit using *E. coli*-produced test substance AtAHAS-0107. An *E. coli*-produced AHAS catalytic subunit lacking the S653N and R272K substitutions (i.e. wild type AHAS which is imidazolinone sensitive; AtAHAS-0207 test substance) was used as a reference substance. No studies were done using the AtSEC61γ subunit protein.

The enzymatic activity (as described in Section 4.3.4) of the AtAHAS-0107 test substance and reference substance was monitored during incubation of a solution of the protein preparation at 4, 37, 60, 75 and 100˚C for 2, 10, 30 and 60 minutes.

AHAS activity in both the test and reference substances was found to be stable at both 4 and 37˚C after 60 minutes of incubation (Figure 6). At 60˚C, activity had decreased to zero by 30 minutes and when incubated at 75 and 100˚C activity decreased to zero within 2 minutes. These results demonstrate that the AHAS catalytic subunit from *A. thaliana* is rapidly inactivated at temperatures above 60˚C. There was no difference in heat stability between the test and reference substances.

As soybean processing involves heat treatments greater than 60˚C, the Applicant states that the *A. thaliana* AHAS catalytic subunit in soybean line CV127 will be functionally inactive in any processed food products. It has already been established that AHAS is not detectable in any of the processed fractions from soybean (Section 4.2.1, Table 4).

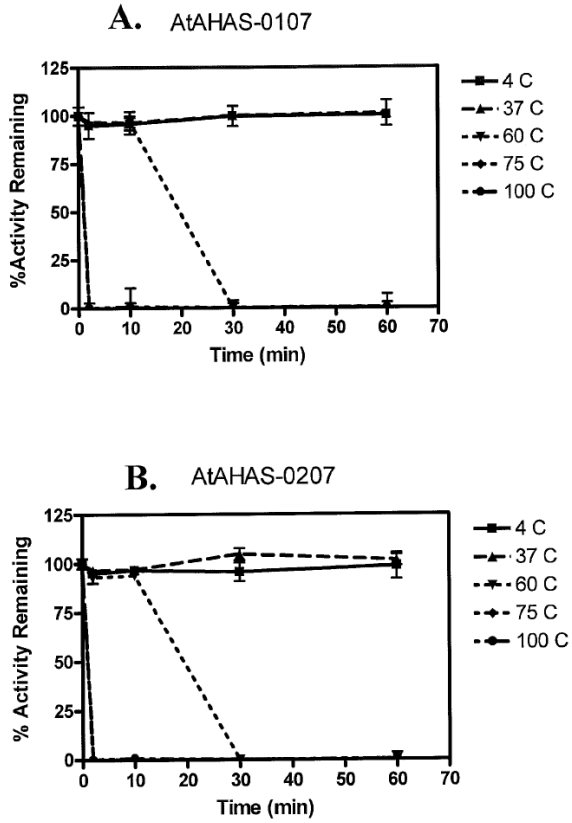

Figure 6: Per cent AHAS activity remaining as a function of time at different incubation temperatures. Results are shown for the imidazolinone-tolerant (Panel A) and wild type (Panel B) A. thaliana AHAS catalytic subunits.

4.4.5 Acute oral toxicity study

An acute oral toxicity study in mice using *E. coli*-produced test substance AtAHAS-0107 was submitted by the Applicant but is not included in this safety assessment. Such studies are only deemed necessary if the results of the biochemical, bioinformatic, digestibility or stability studies indicate further investigation of potential toxicity is warranted. In this case, no safety concerns were identified in any of the other studies.

4.5 Potential allergenicity

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

- the source of the novel protein;
- any significant amino acid sequence similarity between the novel protein and known allergens;
- the structural properties of the novel protein, including susceptibility to digestion, heat stability and/or enzymatic treatment; and
- specific serum screening if the novel protein is derived from a source known to be allergenic or has amino acid sequence similarity to a known allergen, additional *in vitro* and *in vivo* immunological testing may be warranted.

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

The AtSEC61γ subunit was also included in the assessment of potential allergenicity.

Studies submitted:

BASF Report No BPS-014-08 (2008) Bioinformatics Analysis of Deduced Amino Acid Sequences of *Arabidopsis thaliana ahasl* and *SEC61γ* from Herbicide-Tolerant Soybean BPS-CV127-9 for Allergenicity and Toxicity Potential.

BASF Report No BPS-012-07 (2007) Digestive Fate of Test Substance Arabidopsis Acetohydroxyacid Synthase (Lot #AtAHAS-0107) and AtAHAS Produced in Imidazolinone Herbicide-Tolerant Soybean.

BASF Report No BPS-002-08 (2008) Digestive Fate of Arabidopsis SEC61γ Subunit Protein.

4.5.1 Source of the protein

Both the AHAS catalytic subunit and the SEC61γ subunit proteins were derived from *A. thaliana*, a member of the Brassicaceae. *A. thaliana* is not commonly used for food therefore it has no history of widespread human consumption. However it has a long history of being used in research without any reports of it being associated with allergenic reactions in humans.

4.5.2 Amino acid sequence similarity to known allergens

Bioinformatics analyses were done to determine the amino acid sequence similarity of the *A. thaliana* AHAS catalytic subunit (containing the S653N and R272K substitutions) and AtSEC61γ subunit to known allergens. The deduced amino acid sequences of both proteins were compared to entries in the FARRP Allergen Protein Database (Version 8.00), as previously described in Section 3.4.2.

Neither protein shared a sequence of eight or more consecutive identical amino acids or 35% or greater identity over 80 amino acids with any potential allergens in the FARRP Allergen Protein Database.

4.5.3 In vitro digestibility

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

A pepsin digestibility assay (Thomas *et al*., 2004) was conducted to determine the digestive stability of the AHAS catalytic subunit as well as the AtSEC61γ subunit. In addition to the pepsin protocol using simulated gastric fluid (SGF), a second assay was done using simulated intestinal fluid (SIF) containing pancreatin, which is a mixture of enzymes including amylase, trypsin, lipase, ribonuclease and protease.

The relevance of the SIF study however, is only meaningful for proteins that are resistant to pepsin digestion because ordinarily an ingested protein would first need to survive passage through the stomach before being subject to further digestion in the small intestine.

AHAS catalytic subunit

Studies were done to determine the digestibility of the *E. coli*-produced AtAHAS-0107 test substance as well as AHAS protein in leaf and seed extracts from soybean line CV127. The plant extracts are a mixture of the endogenously encoded soybean AHAS protein and the *A. thaliana* AHAS catalytic subunit encoded by the introduced *csr1-2* gene. Due to the high amino acid sequence similarity between these two proteins, they are immunologically indistinguishable. The purpose of examining the digestibility of the plant-produced AHAS proteins was to determine if the plant matrix had any impact on the digestibility and also to confirm that the plant-produced AHAS proteins had a similar digestibility to the *E. coli*-produced protein.

Samples of the AtAHAS-0107 test substance and CV127 leaf and seed extracts were incubated with SGF and SIF at 37˚C for 0.5, 2, 5, 10, 30 and 60 minutes. The 0 time point was taken by first inactivating the SGF or the SIF reaction mix (heating at 75˚C for 10 minutes) before adding the test protein. Controls consisted of SGF or SIF without any test protein added (to control for auto-digestion) as well as SGF or SIF containing test protein but without any added pepsin or pancreatin (to determine test protein stability). The control reactions were treated in exactly the same way as the test reactions except that only 0 and 60 minute time points were taken. Samples were subjected to SDS-PAGE and Coomassie Blue staining (for AtAHAS-0107) or Western blot analysis (for plant extracts) using a rabbit anti-AHAS peptide 2 polyclonal antibody. For reference and to aid in the interpretation of the SDS-PAGE and Western blot analysis results, digestibility studies were also undertaken with leaf and seed protein extracts from a non-transgenic soybean (expressing only endogenous AHAS).

The *E. coli*-produced AHAS protein was rapidly degraded in SGF. It was readily detected in the 0 time point and in the control with no added pepsin; however no full length (64 kDa) protein was visible after incubation for 0.5 minutes. Some low molecular weight bands (< 6 kDa) were observed at 0.5 minutes but these were degraded after 2 minutes. A low molecular weight band of approximately 6 kDa was apparent that was also observed in the SGF control without added test protein, suggesting that it may be derived from the pepsin enzyme preparation. In SIF, the *E. coli*-produced AHAS protein was also rapidly degraded, with no full length protein being visible after 0.5 minutes. Some degradation of the AHAS protein was observed in the SIF control without added pancreatin indicating the preparation was sensitive to degradation in the absence of pancreatin.

The plant-produced AHAS proteins were also rapidly digested in SGF with no full length AHAS being detectable after 2 minutes. An additional immunoreactive band at approximately 50 kDa was observed in leaf extracts. This band was stable in the absence of pepsin but was not visible during the digestion time course. Another immunoreactive band at approximately 36 kDa was observed in seed extracts and was rapidly degraded by pepsin within 0.5 minutes. The Applicant states these bands could either represent proteins that cross-react with the AHAS-specific antibody or smaller fragments of the AHAS protein.

In SIF, the AHAS protein from leaf extracts was rapidly digested, with no full length protein being visible after 0.5 minutes. This was also the case for full length AHAS in leaf extracts from the non-transgenic soybean. A lower molecular weight band of approximately 50 kDa, which was not present in any of the control reactions, was detected at 0.5 minutes and increased in intensity with time of digestion.

Another immunoreactive band of approximately 16 kDa was also detected which appeared from 0.5 minutes and increased in intensity with time of digestion. However, this band was also visible in the SIF control without added test protein, suggesting it could be a degradation product of pancreatin. A band of similar molecular weight was also visible in leaf and seed extracts of the non-transgenic soybean, including the SIF control without added test protein, confirming they most likely represent a degradation product of pancreatin. In seed extracts, full length AHAS is rapidly degraded and is not visible at 0.5 minutes. Full length AHAS was not detectable in SIF digests of seed extracts from the non-transgenic soybean. A lower molecular weight band of approximately 36 kDa, which was not present in the SIF control without added test protein, appears to accumulate with digestion time in the SIF digestion of seed extracts from both CV127 and the non-transgenic soybean. In addition, lower molecular weight bands of approximately 16 kDa and 6 kDa were also visible in the seed extract digestions as well as in the SIF control without added test protein suggesting they originate from pancreatin, rather than AHAS.

The 50 kDa and 36 kDa immunoreactive bands which were visible in the leaf and seed extract digests, respectively, and which appeared to increase in intensity with increasing time of digestion in SIF, were not present in the SGF and SIF digestions of the *E. coli*-produced AHAS protein. Because of this, the Applicant states these bands are unlikely to be a degradation product of the full length AHAS protein and speculates they are more likely to be immunoreactive proteins generated from digestion of a different protein present in leaf or seed extracts.

AtSEC61γ subunit

Digestibility studies were done with the *A. thaliana* SEC61γ subunit protein which had been produced in an *E. coli* expression system. The protein was produced as a fusion with glutathione S-transferase and then subsequently cleaved and purified and then used in digestion studies using SGF.

Samples of the *E. coli*-produced AtSEC61γ subunit protein were incubated with SGF at 37˚C for 0.5, 2, 5, 10, 30 and 60 minutes. The 0 time point was taken by first quenching the SGF reaction mix before adding the test protein. Controls consisted of SGF without any AtSEC61γ subunit protein added (to control for pepsin auto-digestion) as well as SGF containing the AtSEC61γ subunit protein but without any added pepsin (to determine the stability of the AtSEC61γ subunit protein preparation). The control reactions were treated in exactly the same way as the test reactions except that only 0 and 60 minute time points were taken. Samples were subjected to SDS-PAGE followed by Western blot analysis using an immunoaffinity purified rabbit anti- AtSEC61γ peptide polyclonal antibody.

The *E. coli*-produced AtSEC61γ subunit protein was rapidly degraded in SGF. The full length AtSEC61γ subunit protein (ca. 7 kDa) was readily detected at 0 minutes and in the SGF control without added pepsin but was no longer visible at 0.5 minutes.

4.6 Conclusion

Soybean line CV127 expresses the AHAS catalytic subunit from *A. thaliana*. This protein is immunologically indistinguishable from the endogenous imidazolinone-sensitive soybean AHAS therefore protein expression levels were measured as total AHAS (endogenous soybean AHAS plus *A. thaliana* AHAS). The highest AHAS levels were found in young leaves and plants but typically at levels that were too low to be quantified. The levels in soybean seed were also too low to be quantified and no AHAS protein was able to be detected in any processed soybean fraction.

Soybean line CV127 also contains the *AtSEC61γ* subunit gene from *A. thaliana* which was shown in the molecular characterisation to be weakly transcribed. No AtSEC61γ subunit protein was able to be detected in CV127 therefore, if it is expressed, it is below the level of detection (< 15 ppb in seed).

Several studies were done to confirm the identity and physicochemical and functional properties of AHAS expressed in CV127. These studies demonstrated that the AHAS protein expressed in CV127 is as expected in terms of its physicochemical and functional properties but the mature form of protein is slightly larger (by 34 amino acids) than anticipated due to the CTP cleavage site being different to what had been previously predicted. The AHAS protein expressed in CV127 is not glycosylated and exhibits the expected enzymatic activity.

An assessment was done to determine the potential toxicity and allergenicity of the AHAS protein as well as the AtSEC61γ subunit protein (should it be expressed). Both proteins are highly homologous to proteins that have been safely consumed in food. Bioinformatic analyses confirmed the lack of any significant amino acid sequence similarity of either protein to known protein toxins or allergens and digestibility studies demonstrated that both proteins would be rapidly digested in the gastrointestinal tract. The AHAS protein was also shown to be rapidly inactivated at temperatures > 60˚C and is unlikely to be present in processed products such as meal, protein isolate, protein concentrate and oil. Taken together, the evidence indicates that both proteins are unlikely to be toxic or allergenic to humans.

5. Other novel substances

As part of the safety assessment it is important to establish whether the expression of a novel protein is likely to result in the accumulation of other novel substances. If such substances are found to occur as a result of the genetic modification, then it is important to determine their potential toxicity.

In the case of herbicide tolerant plants one or more novel metabolites could accumulate as a result of the detoxification of a herbicide. However, herbicide detoxification is only one of a number of mechanisms that have been used to confer herbicide tolerance. Other mechanisms that have been used include:

- introduction of a gene encoding a herbicide insensitive form of an endogenous enzyme or modification of the gene encoding the endogenous enzyme such that the herbicide molecule can no longer bind;
- overexpression of a gene encoding a herbicide target enzyme such that sufficient amounts remain functional in the presence of herbicide;
- engineering of active herbicide efflux from plant cells.

These other mechanisms generally would not be expected to result in the accumulation of novel metabolites in the plant as they do not rely upon the conversion of the herbicide into herbicidally inactive forms.

In the case of soybean line CV127, herbicide tolerance is achieved by the introduction of a gene encoding a herbicide insensitive form of an endogenous enzyme, therefore no novel metabolites would be expected as a result of the genetic modification. The rationale for this is discussed below.

5.1 Mechanism of action of imidazolinone herbicides

The mechanism of action of imidazolinone herbicides on sensitive plants is to inhibit the AHAS enzyme, which interferes with branch-chain amino acid synthesis leading to the death of the plant. Crystallisation studies with the *A. thaliana* AHAS enzyme have shown that inhibition occurs by binding of the herbicide within the substrate access channel, which blocks substrate from gaining access to the enzyme's active site (McCourt et al 2006, Duggleby et al 2008). The herbicide itself is not altered in this interaction with the AHAS enzyme and the AHAS enzyme has no role in the metabolism of imidazolinone herbicides by plants.

5.2 Tolerance to imidazolinone herbicides

Studies have shown that a single amino acid substitution in the herbicide binding site of the AHAS enzyme can be sufficient to interfere with herbicide binding, producing an enzyme that is less sensitive to herbicide inhibition (Duggleby and Pang 2000).

The *A. thaliana* AHAS catalytic subunit expressed in CV127 contains a single amino acid substitution (asparagine for serine) at amino acid position 653 which produces the imidazolinone tolerance. This mutation does not affect feedback regulation by the branched chain amino acids or the normal catalytic function of the enzyme. The introduced enzyme is thus able to substitute for the endogenous soybean AHAS in the presence of imidazolinone herbicides.

The Applicant has confirmed through proprietary studies involving an imidazolinone-tolerant rice variety that the introduction of an imidazolinone-tolerant AHAS protein has no impact on the uptake, translocation and metabolism of the imidazolinone herbicides imazethapyr, imazapic and imazapyr by the plant (data not shown).

5.1 Conclusion

The expression of an imidazolinone-tolerant form of AHAS, derived from *A. thaliana*, in soybean line CV127 is unlikely to lead to the production or accumulation of any novel herbicide metabolites in the plant.

6. Compositional analysis

The main purpose of compositional analysis is to determine if any unexpected changes in composition have occurred to the food and to establish its nutritional adequacy. Compositional analysis can also be important for evaluating the intended effect where there has been a deliberate change to the composition of food.

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

6.1 Key components of soybean

For soybean intended for human food use, the key components considered important for compositional analysis include the proximates (moisture, crude protein, fat, ash, fibre), amino acids and fatty acids as well as the anti-nutrients phytic acid, trypsin inhibitors, lectins and isoflavones (OECD 2001). For this Application, additional analyses were also done on vitamins and minerals, stachyose and raffinose and phospholipid content.

Analyses for key components were done on seed, forage and on processed seed fractions. In general, soybean is cultivated for the production of seed, which is used as a source of both human food and animal feed, and is only infrequently used as a forage crop for livestock. As there are no human food products derived from forage, only the results of the compositional analyses for seed and its processed fractions are presented in this report. The compositional analyses for forage yielded comparable results to those obtained for seed.

Studies submitted:

BASF Report No BPS-015-07A (2009) Compositional Analysis of Grain from Imidazolinone-Tolerant Soybean BPS-CV127-9 Produced in Brazil in 2006/2007 in Comparison with that from Isoline Control and Conventional Soybean Varieties.

BASF Report No BPS-013-08 (2008) Compositional Analyses of Fractions Produced during Processing of Grain from Imidazolinone-Tolerant Soybean BPS-CV127-9 Produced in Brazil in 2006/2007 and Fate of AtAHAS in these Fractions.

BASF Report No BPS-011-08 (2008) Proximate and fibre composition of forage from Imidazolinone-Tolerant Soybean BPS-CV127-9 Produced in Brazil in 2007/2008 and Comparison with that from Isogenic Control and Commercial Soybean Varieties.

6.2 Study design and conduct

6.2.1 Seed

Imidazolinone treated and non-treated CV127 plants, together with the null segregant control and two other conventional soybean varieties, were each grown at six locations in Brazil during the 2006/2007 growing season and four locations in Brazil during the 2007 short season. The Applicant states the null segregant was used as the comparator for the compositional studies as it is the most closely related line to CV127, being identical in every aspect except that it lacks the introduced *csr1-2* gene. The null segregant used was obtained from the same stage of breeding as the CV127 line being evaluated. The two conventional soybean varieties grown alongside CV127 and the null segregant were used to establish a reference range of natural variability for each analyte against which values for CV127 and the null segregant could be compared. Values were also compared to data in the International Life Sciences Institute (ILSI) Crop Composition Database (ILSI 2009) which is a publicly available database providing compositional data from conventionally bred cotton, soybean and corn. A summary of the plant lines used for the compositional analyses is provided in Table 5.

Table 5: Plant lines used for the compositional analyses

Field Trial	CV127 Generation	Comparator	Reference Range
2006/2007 Growing Season	F5 CV127 line 127	F5 Null Segregant	Monsoy 8001 Coodetec 271
2007 Growing Season	F6 CV127 line 127	F6 Null Segregant	Monsoy 8001 Coodetec 271

The plants were grown using standard agronomic practices in a complete randomised block design with four replicate blocks per location. With the exception of the CV127 plants treated with imidazolinone herbicide at 70 g/ai/ha (CV127 + imi) all other plant lines in the study were treated with Bentazon + Acifluorfen-sodium (sold commercially as Volt) at the rate of 1.0 litres/ha. The seed was harvested at the conclusion of the growing season with approximately 2 kg of each replicate seed sample being removed for compositional analysis. For the two conventional soybean reference varieties, the same seed harvest and separation procedures were followed but approximately 500 g of seed was further sub-sampled from each 2 kg replicate seed sample. The four 500 g replicate samples for each reference variety from each field location were pooled to make a single sample for compositional analysis. Statistical analyses were only conducted for the CV127 and null segregant control lines. Data from the two conventional soybean varieties were used to establish a range of natural variability.

A total of 70 different components were analysed in the seed samples. These consisted of proximates (protein, fat, ash, carbohydrate, calories, total dietary fibre and moisture), crude fibre, acid detergent fibre (ADF), neutral detergent fibre (NDF), amino acids, fatty acids (C14-C22), minerals (calcium, iron, phosphorus, magnesium, and potassium), vitamins (α-, β-, γ-, δ- and total tocopherols, vitamins E and B_1 and folic acid), isoflavones (daidzin, malonyl-daidzin, daidzein, glycitin, malonyl-glycitein, genistin, malonyl-genistin, genistein), phospholipids (phosphatidyl ethanolamine, phosphatidic acid, phosphatidyl inositol, phosphatidyl choline), phytate, raffinose, stachyose, lectin, urease and trypsin inhibitor. Methods of analysis were based on internationally recognised procedures (e.g. those of the Association of Official Analytical Chemists), methods specified by the manufacturer of the equipment used for analysis, or other published methods.

Statistical analysis was done by analysis of variance using SAS Version 9.1 (SAS Institute Inc., Cary, NC) following two procedures: the General Linear Model and the Mixed Model procedures. With the exception of moisture content, all data were expressed on a dry weight basis for statistical analyses. Differences were assessed across location and by location. Means for varieties (averaged over locations) were compared using the Least Significant Difference (LSD) Procedure. In assessing the significance of any difference between means, a probability (p) value of 0.05 was used (i.e. a p-value of ≥0.05 was not significant).

Following the analysis by location, a second analysis was carried out to compare each of the sprayed and unsprayed CV127 treatments (CV127 + imi and CV127 treated with Volt) with the null segregant control. Differences were considered statistically significant at the 0.05 significance level.

6.2.2 Processed fractions

Seed harvested from four locations at the conclusion of the 2006/2007 field trials (see 6.2.1 above) was pooled to form a single sample for each treatment at each location and then shipped for processing and analysis. The seed was processed in a pilot scale processing apparatus using standard methods representative of those currently used to commercially process soybeans. The processing produced five fractions: defatted meal, toasted defatted meal, protein isolate, protein concentrate and refined oil. Proximate analysis was done on the toasted defatted meal, protein isolate and protein concentrate fractions. The toasted defatted meal was also analysed for raffinose, stachyose, trypsin inhibitor, urease, phytic acid, isoflavones and fibre (crude, ADF and NDF). Refined oil was analysed for fatty acids. Untoasted defatted meal was not analysed as its composition was expected to be very similar to toasted defatted meal.

For compositional analysis of processed fractions from seed, analysis of variance was carried out as described above for seed samples. Differences were only assessed across location and analyses were done to compare each treatment. Differences were considered statistically significant at the 0.05 confidence level.

6.3 Seed composition

6.3.1 Proximates

Results of the proximate analyses are shown in Table 6. For both growing seasons, there were no significant differences in ash, total dietary fibre or carbohydrate content between either of the CV127 treatments (+imi or -imi) and the control.

In the 2007 season there was also no significant difference in calorie content between either of the CV127 treatments and the control. A number of significant differences were found between the CV127 treatments and the control for the other analytes (moisture, protein, fat calories) but the mean levels were all within the ranges reported in the ILSI database. Calorie content is not reported in the ILSI database but the mean levels obtained were all within the reference range obtained for reference varieties grown at the same time.

Table 6: Proximate composition of seed

	2006/2007 Growing Season				2007 Growing Season				ILSI Database	
Analyte	**Control**	**CV127 -imi**	**CV127 + imi**	**Reference Standards**	**Control**	**CV127 -imi**	**CV127 +imi**	**Reference Standards**	**Global**	**Brazilian**
	Mean (range)									
Moisture (%)	10.1a* (9.2-10.9)	9.3b (8.7-10.2)	9.4b (8.8-9.9)	9.8 (9.4-10.5)	7.6a (7.1-8.2)	7.8a (7.4-8.2)	7.9b (7.6-8.2)	7.7 (7.0-8.2)	10.1 (4.7-34.4)	9.8 (7.6-11.2)
Ash (g/100g DW)	5.0a (4.6-5.3)	5.0a (4.6-5.5)	4.9a (4.5-5.4)	4.9 (4.5-5.3)	5.2a (4.8-5.7)	5.2a (4.8-5.7)	5.1a (4.7-5.6)	5.2 (4.9-5.5)	5.32 (3.89-6.99)	5.00 (4.58-5.47)
Protein (g/100g DW)	40.3a (38.1-42.2)	39.2b (37.0-42.0)	39.4b (37.3-41.9)	37.6 (36.4-39.6)	39.2a (36.4-41.6)	39.2ab (37.0-41.5)	39.7b (36.7-41.6)	39.2 (36.8-42.0)	39.47 (33.19-45.48)	40.15 (37.19-44.85)
Fat (g/100g DW)	21.7b (20.0-23.3)	22.6a (20.3-24.6)	22.7a (20.1-24.2)	22.8 (20.2-24.8)	20.2a (17.7-23.8)	20.7b (18.8-24.7)	20.5ab (18.9-24.1)	20.6 (16.9-25.1)	16.68 (8.10-23.56)	18.85 (14.44-23.56)
Total Dietary Fibre (g/100g DW)	24.90a (21.93-26.90)	24.55a (21.61-26.64)	24.70a (22.13-28.11)	25.43 (22.29-28.03)	24.54a (18.94-28.01)	24.73a (22.62-26.55)	24.66a (21.37-28.43)	25.0 (21.9-27.3)	NA[#]	NA
Carbohydrate[1] (g/100g DW)	33.1a (31.6-34.4)	33.2a (31.1-34.1)	32.9a (31.9-34.9)	34.7 (27.39-42.03)	35.4a (25.2-44.9)	34.8a (27.9-39.9)	34.7a (24.2-43.3)	35.0 (26.7-39.8)	38.2 (29.6-50.2)	36.0 (29.6-41.6)
Calories (g/100g DW)	390b (379-402)	395a (389-407)	395a (384-405)	393 (377-400)	382a (362-403)	377a (281-409)	383a (365-413)	382 (374-395)	NA	NA

*Numbers followed by the same letter are not statistically significantly different at p<0.05
[#]NA = not available
[1]Carbohydrates including total dietary fibre

6.3.2 Fibre

Results of the fibre analyses are shown in Table 7. In both seasons, there were no significant differences in crude fibre content between either of the CV127 treatments and the control. Statistically significant differences between the control and both CV127 treatments were observed for ADF and NDF in the 2006/2007 season and between the control and CV127 +imi for NDF in the 2007 season. The mean levels for NDF and ADF were all within the ranges reported in the ILSI database.

Table 7: Fibre composition of seed

	Crude Fibre	ADF	NDF
	Mean (range) g/100 g DW		
2006/07 Season			
Control	8.5a* (6.9-11.0)	11.39b (8.84-14.95)	14.98b (12.04-17.32)
CV127 -imi	8.3a (6.9-11.1)	13.14a (9.75-15.92)	17.46a (12.72-20.55)
CV127 +imi	8.1a (6.4-11.3)	13.76a (10.96-19.00)	17.52a (14.06-20.01)
Reference Standards	8.2 (6.7-9.3)	11.76 (9.32-14.43)	14.85 (10.63-16.92)
2007 Season			
Control	7.9a (6.7-10.6)	10.25ab (8.53-12.05)	14.08a (11.32-15.93)
CV127 -imi	8.0a (7.1-9.7)	9.66a (8.09-11.19)	14.24a (11.86-16.91)
CV127 +imi	8.2a (6.2-14.7)	10.49b (7.35-13.40)	15.13b (12.53-18.17)
Reference Standards	8.2 (7.2-12.1)	11.11 (8.89-12.59)	14.11 (12.26-16.18)
ILSI Database			
Global	7.81 (4.12-13.87)	11.97 (7.81-18.61)	12.33 (8.53-21.25)
Brazilian	8.46 (6.42-10.93)	11.34 (7.81-16.39)	12.39 (8.53-21.25)

* Numbers followed by the same letter are not statistically significantly different at $p<0.05$

6.3.3 Amino acids

Results of the amino acid analyses are shown in Table 8. For the 2006/2007 season, the only significant differences were lower levels of proline and tyrosine in seed from CV127 -imi compared to the control. There were no significant differences between CV127 +imi and the control. For the 2007 season, the only significant differences were higher levels of alanine and histidine and lower levels of tyrosine in the two CV127 treatments compared to the control. Higher levels of valine were also found in the CV127 -imi treatment compared to the control, but not the CV127 +imi treatment. The mean levels were however within the ranges reported in the ILSI database.

These results indicate that biosynthesis of the branched chain amino acids (leucine, isoleucine and valine) is largely unaffected by the expression of the *A. thaliana* AHAS catalytic subunit in CV127, suggesting this enzyme can readily substitute for the endogenous imidazolinone-sensitive AHAS enzyme in the presence of herbicide.

Table 8: Amino acid composition of seed

Amino Acid	2006/2007 Season				2007 Season				ILSI Database	
	Control	CV127	CV127+imi	Reference Standard	Control	CV127	CV127+imi	Reference Standard	Global	Brazilian
	Mean (range) g/100 g DW									
Alanine	1.64a (1.45-1.92)	1.59a (1.33-1.89)	1.63a (1.27-1.76)	1.56 (1.39-1.68)	1.51a (1.29-1.71)	1.58b (1.33-1.90)	1.62b (1.38-1.86)	1.58 (1.41-1.79)	1.72 (1.51-2.10)	1.73 (1.62-1.85)
Arginine	3.03a (2.64-3.09)	2.91a (2.35-3.51)	3.01a (2.36-3.33)	2.71 (2.34-2.97)	2.84a (2.37-3.34)	2.77a (2.44-3.22)	2.84a (2.53-3.15)	2.95 (1.49-3.45)	2.84 (2.29-3.40)	2.93 (2.59-3.28)
Aspartate	4.64a (4.02-5.37)	4.46a (3.67-5.24)	4.61a (3.66-5.17)	4.32 (3.94-4.72)	4.17a (3.50-4.66)	4.15a (3.32-4.68)	4.29a (3.85-4.65)	4.45 (3.70-5.06)	4.49 (3.81-5.12)	4.59 (4.23-5.12)
Cysteine	0.54a (0.51-0.58)	0.52a (0.48-0.56)	0.53a (0.50-0.56)	0.52 (0.49-0.56)	0.51b (0.43-0.55)	0.52a (0.43-0.56)	0.52a (0.46-0.56)	0.52 (0.44-0.59)	0.59 (0.37-0.81)	0.57 (0.50-0.81)
Glutamate	7.61a (6.59-8.92)	7.30a (6.10-8.46)	7.54a (5.83-8.45)	6.98 (6.16-7.68)	6.74a (5.72-7.41)	6.66a (5.49-7.52)	6.78a (6.17-7.46)	7.18 (5.90-8.55)	7.09 (5.84-8.20)	7.29 (6.58-8.09)
Glycine	1.65a (1.45-1.87)	1.62a (1.33-1.94)	1.65a (1.32-1.81)	1.56 (1.41-1.71)	1.50a (1.28-1.68)	1.47a (1.24-1.69)	1.50a (1.34-1.83)	1.56 (1.35-1.72)	1.69 (1.46-2.00)	1.70 (1.56-1.82)
Histidine	0.87a (0.78-1.03)	0.83a (0.69-1.04)	0.85a (0.67-0.97)	0.79 (0.72-0.86)	1.11a (0.80-1.32)	1.45b (0.93-2.11)	1.37b (0.91-2.13)	1.10 (0.73-1.34)	1.04 (0.88-1.18)	1.06 (0.98-1.18)
Isoleucine	1.61a (1.42-1.94)	1.56a (1.28-1.88)	1.61a (1.24-1.75)	1.54 (1.38-1.71)	1.46a (1.22-1.62)	1.41a (1.10-1.83)	1.43a (1.08-1.97)	1.50 (1.35-1.65)	1.81 (1.54-2.08)	1.85 (1.59-2.04)
Leucine	2.89a (2.51-3.40)	2.78a (2.31-3.36)	2.87a (2.24-3.16)	2.80 (2.45-3.65)	2.56a (2.19-2.86)	2.60a (2.30-2.86)	2.61a (2.44-2.85)	2.73 (2.39-3.15)	3.04 (2.59-3.62)	3.07 (2.81-3.38)
Lysine	2.48a (2.12-2.85)	2.40a (2.00-2.93)	2.46a (1.93-2.68)	2.33 (2.08-2.54)	2.24a (1.97-2.47)	2.29a (2.12-2.76)	2.27a (1.99-2.62)	2.32 (2.03-2.58)	2.56 (2.29-2.84)	2.58 (2.42-2.82)
Methionine	0.63a (0.60-0.67)	0.62a (0.59-0.68)	0.62a (0.59-0.64)	0.60 (0.55-0.65)	0.60a (0.53-0.66)	0.61a (0.51-0.66)	0.62a (0.58-0.66)	0.60 (0.53-0.69)	0.55 (0.43-0.68)	0.55 (0.50-0.68)
Phenylalanine	1.99a (1.77-2.37)	1.91a (1.59-2.29)	1.98a (1.52-2.19)	1.87 (1.64-2.09)	1.76a (1.52-1.94)	1.80a (1.55-2.16)	1.78a (1.65-1.90)	1.87 (1.64-2.12)	1.98 (1.63-2.35)	2.06 (1.82-2.24)
Proline	1.98a (1.74-2.32)	1.86b (1.56-2.31)	1.91ab (1.49-2.12)	1.82 (1.58-2.03)	1.74a (1.50-1.95)	1.77a (1.56-1.99)	1.73a (1.11-1.89)	1.86 (1.60-2.05)	2.00 (1.69-2.28)	2.06 (1.86-2.28)
Serine	2.09a (1.83-2.39)	2.02a (1.70-2.41)	2.07a (1.63-2.29)	1.96 (1.78-2.16)	1.82a (1.54-2.05)	1.81a (1.56-1.99)	1.83a (1.65-1.97)	1.92 (1.67-2.21)	2.02 (1.11-2.48)	2.17 (1.96-2.48)
Threonine	1.56a (1.34-1.80)	1.50a (1.25-1.81)	1.55a (1.25-1.72)	1.48 (1.36-1.55)	1.37a (1.21-1.50)	1.35a (1.14-1.62)	1.37a (1.17-1.67)	1.40 (1.23-1.550	1.47 (1.14-1.86)	1.40 (1.28-1.52)
Tryptophan	0.73a (0.59-0.84)	0.72a (0.60-0.92)	0.76a (0.60-1.03)	0.77 (0.57-1.16)	0.64a (0.53-0.74)	0.65a (0.50-0.79)	0.67a (0.51-0.84)	0.65 (0.56-0.73)	0.43 (0.36-0.50)	0.44 (0.38-0.49)
Tyrosine	1.34a (1.17-1.55)	1.27b (1.08-1.50)	1.31ab (1.06-1.42)	1.27 (1.12-1.39)	1.18a (0.98-1.33)	1.13b (1.01-1.33)	1.14b (1.02-1.38)	1.25 (1.14-1.45)	1.32 (1.02-1.61)	1.38 (1.27-1.56)
Valine	1.66a (1.36-1.98)	1.60a (1.34-1.94)	1.64a (1.26-1.79)	1.58 (1.40-1.67)	1.56a (1.35-1.830	1.64b (1.49-1.84)	1.62ab (1.41-1.79)	1.62 (1.31-1.87)	1.91 (1.60-2.20)	1.91 (1.63-2.08)

* Numbers followed by the same letter are not statistically significantly different at p<0.05.

6.3.4 Fatty acids

Seed samples were analysed for the complete spectrum of fatty acids (a total of 37 fatty acids). However, a number of fatty acids were either not detectable or present at levels below the limit of quantitation for the assay. Results are reported only for those fatty acids that had quantifiable levels. Results of the fatty acid analyses are shown in Table 9.

In seed from the 2006/2007 season there were no significant differences in mean levels of myristic, margaric, stearic, arachidic, and eicosenoic acids between the seed from either of the CV127 treatments and the control. Two fatty acids (palmitic and oleic) were significantly higher and two others (linoleic and linolenic) were significantly lower in seed from both CV127 treatments compared to the control. In the case of behenic acid, there was no significant difference between the CV127 +imi treatment and control, but levels in CV127 -imi were significantly lower compared to the control. The mean levels however were all within the ranges reported in the ILSI database.

In seed from the 2007 season, there were no significant differences in mean levels of myristic, palmitic, margaric, arachidic and behenic acids between the seed from either of the CV127 treatments and the control. There were also no significant differences in mean levels of any of the measured fatty acids between seed of the CV127 -imi treatment and the control, but significant differences in mean levels of stearic, oleic, linoleic, and linolenic fatty acids were observed between seed from the CV127 +imi and the control. The mean levels however were all within the ranges reported in the ILSI database.

6.3.5 Minerals

Results of the mineral analyses are shown in Table 10. In the 2006/2007 season, significant differences between the control and both CV127 treatments were observed for iron (lower) and magnesium (higher). In addition, calcium in the CV127 –imi treatment was significantly higher than the control and potassium in the CV127 +imi treatment was significantly lower. All the mean mineral levels obtained for the 2006/2007 season were within the ranges reported in the ILSI database. In the 2007 season, all the minerals analysed for both treatment were significantly different to the control except for iron in the CV127 +imi treatment. The mean mineral levels obtained for the 2007 season were within the ranges reported in the ILSI database except for potassium. Mean potassium levels obtained for the control, as well as the CV127 +imi treatment, were below the range reported in the ILSI database but within the range obtained for reference varieties grown at the same time.

Table 9: Fatty acid composition of seed

	2006/2007 Season				2007 Season				ILSI Database	
Fatty Acid	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Global**	**Brazilian**
	Mean (range) % Total FA									
Myristic 14:0	<0.09a* (nd-0.10)	<0.09a (nd-0.10)	<0.09a (nd-0.10)	0.10 (0.09-0.11)	0.11a (0.09-0.12)	0.11a (0.9-0.12)	0.11a (0.10-0.11)	0.09 (0.06-0.12)	NA[1]	NA
Palmitic 16:0	9.77b (9.12-10.44)	10.00a (9.23-10.42)	9.98a (9.35-10.59)	10.17 (9.08-11.16)	10.32a (9.67-10.82)	10.24a (9.85-10.93)	10.18a (9.75-10.64)	9.77 (8.31-10.67)	11.12 (9.55-15.77)	11.27 (10.28-12.73)
Margaric 17:0	<0.10a (nd-0.11)	<0.10a (nd-0.11)	<0.10a (nd-0.11)	<0.10 (nd-0.11)	0.11a (0.09-0.12)	0.11a (0.09-0.12)	0.11a (0.09-0.11)	0.10 (0.06-0.12)	0.114 (0.085-0.146)	0.106 (0.086-0.116)
Stearic 18:0	3.39a (2.84-3.87)	3.38a (2.93-3.86)	3.30a (2.77-3.61)	3.46 (2.97-3.92)	4.04b (3.14-5.05)	4.13ab (3.45-4.98)	4.21a (3.50-4.80)	3.99 (3.04-5.01)	4.01 (2.70-5.88)	3.95 (2.70-5.52)
Oleic 18:1	20.07c (18.47-21.02)	21.44b (20.34-22.74)	22.07a (20.10-25.76)	20.06 (18.54-21.38)	24.38b (21.72-33.15)	25.64b (22.81-32.45)	27.79a (22.65-43.63)	23.39 (20.13-28.06)	20.7 (14.3-32.2)	22.6 (18.7-28.9)
Linoleic 18:2	45.87b (44.01-47.68)	45.42a (43.05-46.70)	45.00a (42.60-46.62)	53.54 (52.36-54.40)	48.86a (42.54-50.83)	47.72a (43.25-50.50)	45.65b (31.97-49.57)	50.10 (46.78-52.80)	53.3 (42.3-58.8)	52.6 (48.2-55.5)
Linolenic 18:3	5.65a (5.05-6.10)	5.21b (4.80-5.67)	5.10c (4.59-5.72)	7.23 (6.31-8.15)	6.62a (4.20-8.11)	6.47ab (3.92-8.25)	6.32b (3.42-8.12)	7.06 (4.76-8.52)	8.34 (3.00-12.52)	7.06 (5.92-8.18)
Arachidic 20:0	0.37a (0.32-0.44)	0.35a (0.24-0.44)	0.34a (0.26-0.39)	0.31 (0.26-0.40)	0.39a (0.32-0.49)	0.42a (0.39-0.46)	0.42a (0.27-0.55)	0.33 (0.26-0.48)	0.32 (0.16-0.48)	0.37 (0.28-0.48)
Eicosenoic 20:1	0.13a (0.09-0.19)	0.12a (0.08-0.20)	0.13a (0.08-0.18)	0.14 (0.09-0.200	ND[2]	ND	ND	ND	0.20 (0.14-0.35)	0.22 (0.17-0.28)
Behenic 22:0	0.46a (0.37-0.52)	0.42b (0.34-0.46)	0.43ab (0.39-0.53)	0.40 (0.37-0.50)	0.51a (0.40-0.75)	0.49a (0.40-0.63)	0.50a (0.38-0.80)	0.44 (0.37-0.51)	0.40 (0.28-0.60)	0.45 (0.37-0.57)

* Numbers followed by the same letter are not statistically significantly different at p<0.05.
[1]NA = not available
[2]ND = not detectable

Table 10: Mineral composition of seed

	2006/2007 Season				2007 Season				ILSI Database
Mineral	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Global**
	Mean (range) mg/100 g DW								
Calcium	268b* (221-330)	277a (231-327)	266b (214-318)	254 (205-313)	214a (172-359)	274b (190-427)	286b (198-484)	234 (167-329)	217 (117-307)
Iron	8.50a (6.01-10.430	7.89b (5.56-10.91)	7.75b (5.79-10.48)	8.57 (6.54-10.35)	9.16a (7.98-10.90)	9.88b (8.23-13.00)	9.15a (7.80-11.40)	10.16 (7.93-14.13)	7.81 (5.54-10.95)
Magnesium	246b (204-266)	266a (218-326)	266a (227-304)	269 (225-308)	238a (211-266)	277b (258-309)	287c (268-309)	255 (223-293)	264 (219-313)
Phosphorus	687a (541-834)	677a (515-791)	667a (546-760)	670 (527-875)	768a (655-845)	717b (654-784)	732b (580-871)	745 (614-808)	714 (507-935)
Potassium	1928a (1782-2071)	1897ab (1720-2164)	1881b (1703-2069)	1961 (1772-2103)	1744a (1593-2061)	1877b (1713-2079)	1864b (1599-2021)	1758 (1595-2033)	2061 (1868-2316)

* Numbers followed by the same letter are not statistically significantly different at p<0.05.

6.3.6 Vitamins

Results of the vitamin analyses are shown in Table 11. In the 2006/2007 season, significant differences between the control and CV127 treatments were observed for all vitamins analysed except for total tocopherol. Increased levels were observed for α-, β- and δ-tocopherol, as well as vitamin E, and decreased levels were observed for folic acid, γ-tocopherol and vitamin B_1. The mean levels obtained for folic acid and α-tocopherol were within the ranges reported in the ILSI database. Data for β-, δ- and γ-tocopherol, total tocopherol and vitamin E content of soybeans were not available in the ILSI database. The mean levels obtained for Vitamin B_1, including for the control, were significantly higher than the maximum level reported in the ILSI database; however the mean levels for all vitamins analysed were within the ranges obtained for the reference varieties grown at the same time.

In the 2007 season, there were no significant differences between the CV127 treatments and the control for folic acid, total tocopherol and vitamin B_1 content. Significantly higher levels were observed for α-tocopherol (CV127 +imi only), β- and δ-tocopherol (CV127 -imi only), and vitamin E content (CV127 -imi only), with lower levels observed for γ-tocopherol (CV127 +imi only). The levels obtained for α-tocopherol were within the range reported in the ILSI database. All mean vitamin levels were within the ranges obtained for the reference varieties grown at the same time

6.3.7 Isoflavones

Results of the isoflavone analyses are shown in Table 12 . There was no significant difference in the mean levels for total glycitein between seed from both CV127 treatments and the control in the 2006/2007 season and between seed from CV127 -imi and the control in the 2007 season. Mean levels of total glycitein were however significantly lower in seed from CV127 +imi compared to the control in the 2007 season. Mean levels of total daidzein and total genistein were also significantly lower in seed of both CV127 treatments compared to the control in both seasons. However, for each of the isoflavones, the mean levels in CV127 were within the ranges reported in the ILSI database.

6.3.8 Phospholipids

Results of the phospholipid analyses are shown in Table 13. In the 2006/2007 season, the mean levels for all measured phospholipids in seed from both CV127 soybean treatments were significantly lower compared to the control. In contrast, in the 2007 season, there were no significant differences in mean phospholipid levels of seed from both CV127 treatments and the control, except for phosphatidic acid, which was significantly higher in seed from the CV127 + imi treatment compared to the control. Data for phospholipids are not available in the ILSI database however, with the exception of phosphatidyl inositol in the 2006/2007 season, all mean levels were within the ranges obtained for the reference varieties grown at the same time. The levels of phosphatidyl inositol obtained for both CV127 treatments were considerably lower than the reference range but as this result was not replicated in the 2007 season, it is not considered to be biologically meaningful.

Table 11: Vitamin composition of seed

	2006/2007 Season				2007 Season				ILSI Database	
Vitamin	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Global**	**Brazilian**
	Mean (range)									
Folic Acid µg/100 g DW	330a* (216-456)	267b (201-327)	270b (183-338)	291 (205-403)	365a (249-462)	363a (302-460)	374a (320-464)	323 (261-438)	360 (240-470)	NA[#]
α-tocopherol mg/100 g DW	3.04b (2.33-4.44)	3.44a (2.86-4.86)	3.49a (2.43-4.69)	3.22 (2.45-4.61)	3.21a (1.74-8.17)	3.53ab (1.96-6.94)	3.67b (1.96-9.230	3.54 (2.07-7.51)	1.91 (0.19-6.17)	3.44 (1.36-6.17)
β-tocopherol mg/100 g DW	0.60b (0.20-1.01)	0.90a (0.58-1.32)	0.90a (0.58-1.22)	0.70 (0.12-1.08)	0.84a (0.54-1.15)	1.04b (0.73-1.49)	0.96ab (0.68-1.37)	0.88 (0.53-1.34)	NA	NA
δ-tocopherol mg/100 g DW	6.18b (4.41-7.53)	6.49a (4.85-8.15)	6.56a (4.9-7.99)	6.09 (5.31-7.42)	7.65a (4.52-9.72)	8.68b (4.67-9.72)	8.15ab (3.31-11.35)	6.99 (5.16-8.94)	NA	NA
γ-tocopherol mg/100 g Dw	16.51a (11.62-20.83)	15.83b (12.50-17.69)	15.81b (12.31-18.84)	16.28 (12.68-20.88)	16.41a (13.79-21.63)	16.06a (13.91-21.05)	14.96b (12.28-18.46)	15.68 (11.65-22.27)	NA	NA
Total tocopherol mg/100 g DW	26.19a (19.26-31.54)	26.57a (21.59-29.61)	26.75a (21.95-31.12)	26.32 (21.68-31.03)	28.12ab (25.02-34.43)	29.31a (26.72-35.38)	27.41b (23.38-31.40)	27.09 (21.56-33.23)	NA	NA
Vitamin B1 mg/100 g Dw	0.65a (0.44-0.86)	0.51b (0.40-0.67)	0.52b (0.34-0.78)	0.58 (0.42-0.71)	0.52a (0.34-0.80)	0.55a (0.35-0.72)	0.55a (0.28-0.75)	0.48 (0.31-0.71)	0.20 (0.10-0.25)	NA
Vitamin E mg/100 g Dw	5.51b (4.11-7.21)	5.88a (5.05-7.30)	5.91a (4.68-7.16)	5.70 (4.95-7.05)	5.75a (3.94-10.79)	6.08b (4.35-10.09)	6.05ab (4.35-11.21)	5.98 (4.64-9.92)	NA	NA

* Numbers followed by the same letter are not statistically significantly different at p<0.05
[#]NA = not available

Table 12: Isoflavone composition of seed

	2006/2007 Season				2007 Season				ILSI Database	
Isoflavone	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Global**	**Brazilian**
	Mean (range) mg/100 g DW									
Total Daidzein	72.2b* (48.4-106.6)	52.8a (41.2-67.7)	52.4a (41.4-72.4)	81.3 (66.2-96.2)	114.6a (19.7-237.8)	77.9b (22.6-162.2)	79.0b (14.7-161.0)	100.2 (30.8-186.4)	86.3 (6.0-245.4)	51.0 (6.0-112.9)
Total Genistein	101.7b (57.1-153.4)	86.1a (56.9-138.0)	83.4a (60.2-121.0)	134.5 (102.8-166.6)	144.5a (26.1-270.1)	115.9b (26.8-232.4)	114.6b (11.8-246.8)	144.6 (54.3-255.1)	97.9 (14.4-283.7)	65.2 (14.4-135.7)
Total Glycitein	22.3a (14.9-31.4)	21.7a (16.6-27.5)	21.7a (14.6-30.2)	49.2 (36.3-75.4)	19.0a (7.3-30.2)	17.5ab (11.5-25.7)	16.5b (7.4-23.9)	35.3 (27.9-43.0)	16.1 (1.5-31.0)	13.3 (1.5-26.4)

* Numbers followed by the same letter are not statistically significantly different at p<0.05

Table 13: Phospholipid composition of seed

	2006/2007 Season				2007 Season				ILSI Database	
Phospholipid	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Global**	**Brazilian**
	Mean (range) mg/g fat									
Phosphatidyl ethanolamine	101.9a* (89.7-127.9)	92.6b (51.0-110.3)	90.9b (57.9-108.9)	98.2 (79.4-113.7)	109.4a (65.7-156.8)	109.2a (61.7-155.5)	106.0a (52.1-139.5)	100.5 (51.1-133.7)	NA[#]	NA
Phosphatidic acid	4.0a (1.8-6.9)	2.8b (0.8-4.6)	2.9b (1.0-4.7)	4.1 (2.1-7.5)	2.1a (0.7-6.7)	2.1a (0.6-6.1)	2.6b (0.5-9.9)	3.0 (0.8-7.2)	NA	NA
Phosphatidyl inositol	11.8a (10.1-14.20	9.8b (6.8-11.0)	9.6b (6.1-11.20	12.3 (11.4-13.50	10.4a (8.6-14.80	10.0a (8.1-13.20	9.7a (8.0-11.0)	10.8 (8.1-12.4)	NA	NA
Phosphatidyl choline	29.3a (24.9-38.5)	27.0b (14.9-34.1)	26.9b (15.9-32.20	30.7 (25.6-38.2)	32.9a (20.6-45.0)	32.2a (19.8-39.4	32.5a (17.5-40.2)	33.4 (17.0-42.0)	NA	NA

* Numbers followed by the same letter are not statistically significantly different at p<0.05
[#]NA = not available

6.3.9 Anti-nutrients

Results of the anti-nutrient analyses are shown in Table 14 . There were no significant differences in the levels of phytic acid, urease, and trypsin inhibitor between seed from either of the CV127 treatments and the control in either season. The measured levels for phytic acid and trypsin inhibitor, including for the controls, were considerably lower than the minimum level reported for these constituents in the ILSI database but were within the ranges obtained for the reference varieties grown at the same time. The Applicant states this most likely reflects a germplasm characteristic of soybean varieties adapted for tropical cultivation conditions in Brazil.

The mean levels for raffinose and stachyose for both CV127 soybean treatments produced in the 2006/2007 season were significantly different to levels in the control. In the 2007 season, statistically significant differences were also observed but in the case of raffinose this was confined to the CV127 -imi treatment only. The measured levels for raffinose and stachyose, including for the controls, were higher than the maximum level reported for these constituents in the ILSI database but were within the ranges obtained for the reference varieties grown at the same time.

The mean levels for lectins in the 2006/2007 season were not significantly different to levels in the control. This contrasts with the 2007 season where mean lectin levels were significantly lower in the two CV127 treatments compared to the control. The mean levels were within the range reported in the ILSI database.

6.4 Processed fraction composition

6.4.1 Toasted defatted soybean meal

Results of the toasted defatted soybean meal analyses are shown in Table 15. The only significant differences observed in meal from CV127 compared to the control were decreased levels for moisture, total daidzein, and total genistein and increased levels for trypsin inhibitor.

No literature ranges were available for moisture content of soybean meal; however the mean level was comparable to that obtained for meal from the two reference varieties grown at the same time. The Applicant states the lower moisture level in meal obtained from CV127 is most likely due to slight variations in toasting conditions between treatments.

No literature ranges were available for isoflavone content of soybean meal; however the results obtained were consistent with that obtained for seed, i.e. significantly lower levels compared to the control. In the case of the seed, the mean isoflavone levels obtained were all within the ranges reported in the ILSI database.

While the mean level for trypsin inhibitor from meal derived from CV127 was significantly higher compared to the control, as well as the two reference varieties grown at the same time, it was still well below the minimum level reported for trypsin inhibitor in the literature.

Table 14: Anti-nutrient composition of seed

	2006/2007 Season				2007 Season				ILSI Database	
Anti-nutrient	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Control**	**CV127 –imi**	**CV127 +imi**	**Reference Varieties**	**Global**	**Brazilian**
	Mean (range)									
Phytic Acid (mg/g DW)	2.95a* (1.43-6.09)	2.75a (1.25-4.52)	2.54a (0.71-5.27)	2.89 (1.47-7.39)	3.36a (1.89-6.00)	3.96a (2.57-6.22)	3.81a (2.63-4.770	4.06 (2.37-5.44)	11.21 (6.34-19.60)	NA[#]
Raffinose (g/100 g DW)	1.1c (0.9-1.5)	1.4a (1.0-1.8)	1.3b (1.0-1.7)	1.1 (0.8-1.6)	1.30a (1.00-1.50)	1.2b (0.9-1.30	1.2ab (0.9-1.7)	1.3 (1.1-1.4)	0.355 (0.212-0.661)	NA
Stachyose (g/100 g DW)	3.7a (3.0-4.2)	3.6b (2.9-4.0)	3.6b (3.1-4.1)	3.8 (3.1-4.6)	4.0a (3.0-4.6)	3.7b (3.0-4.2)	3.6b (2.4-4.1)	4.1 (3.1-4.80	2.19 (1.21-3.50)	NA
Lectins (HU^1/mg DW)	2.24a (1.35-3.35)	2.23a (1.43-3.70)	2.20a (1.27-3.49)	2.03 (1.38-3.53)	1.70a (0.85-3.43)	0.84b (0.17-1.710	0.92b (0.11-3.430	1.65 (0.67-2.67)	1.718 (0.105-9.038)	0.815 (0.299-1.892)
Urease (ΔpH)	1.93a (1.27-2.12)	1.85a (0.67-2.18)	1.91a (0.90-2.21)	1.93 (1.43-2.16)	1.53a (0.28-2.06)	1.57a (0.41-2.00)	1.58a (0.49-2.03)	1.63 (0.27-2.02)	NA	NA
Trypsin Inhibitor (TIU^2/mg DW)	12.29a (6.03-16.4)	12.38a (8.69-16.58)	12.01a (9.14-15.13)	11.45 (5.03-19.64)	13.16a (8.48-17.97)	13.72a (8.16-18.20)	13.80a (7.82-18.03)	14.02 (9.84-16.76)	48.33 (19.59-118.68)	NA

* Numbers followed by the same letter are not statistically significantly different at $p<0.05$

[#]NA = not available

[1] Haemagglutinating Units

[2] Trypsin Inhibitor Units

Table 15: Defatted toasted soybean meal composition

Analyte	Control	CV127 +imi	Reference Line 1	Reference Line 2	Literature Range
	Mean (range)				
Moisture (g/100 g FW)	4.3a* (3.8-4.8)	3.4b (3.2-3.8)	3.9ab (3.6-4.1)	4.5a (3.6-5.5)	NA
Ash (g/100 g DW)	6.3ab (6.0-6.6)	6.3ab (6.1-6.6)	6.2b (5.8-6.5)	6.3ab (5.8-6.6)	5.5-6.5[1]
Fat (g/100 g DW)	1.2a (0.4-1.9)	1.2a (0.9-1.9)	1.0a (0.6-1.4)	1.2a (0.7-2.4)	0.5-2.40[2]
Protein (g/100 g DW)	51.1a (49.8-53.0)	50.6a (48.8-51.4)	48.2b (46.4-49.3)	48.1b (47.3-49.1)	44-61.4[3]
Carbohydrates (g/100 g DW)	41.5b (40.1-42.1)	41.9b (41.3-42.7)	44.6a (43.4-45.7)	44.4a (43.8-45.4)	32.0-38.0[4]
Calories (kcal/100 g DW)	381a (377-384)	381a (380-383)	381a (378-382)	381a (377-389)	NA
Crude Fibre (g/100 g DW)	10.1a (9.5-10.7)	10.0a (9.7-10.2)	10.5a (9.9-11.30	10.5a (8.9-12.4)	NA
ADF (g/100 g DW)	8.41bc (7.87-8.94)	7.74c (7.31-7.92)	9.44a (8.72-9.88)	9.11ab (8.41-9.83)	NA
NDF (g/100 g DW)	16.49ab (15.01-17.26)	14.86b (13.17-17.170	16.4ab (15.43-17.50)	18.01a (16.50-20.46)	NA
Raffinose (g/100 g DW)	1.9ab (1.6-2.3)	1.9b (1.5-2.2)	2.1a (1.7-2.5)	1.8b (1.5-2.1)	1.0-2.0[5]
Stachyose (g/100 g DW)	5.1a (4.4-5.9)	5.0a (4.5-5.4)	5.4a (5.0-5.8)	4.8a (4.0-5.5)	4.0-5.3[6]
Trypsin Inhibitor (TIU/mg)	1.24b (0.84-1.56)	2.03a (1.69-2.56)	1.16b (0.52-1.64)	1.16b (1.01-1.23)	3.8-17.9[7]
Urease (ΔpH)	0.02a (0.01-0.04)	0.04a (0.02-0.06)	0.05a (0.02-0.09)	0.04a (0.03-0.05)	0.05-0.20[8]
Phytic Acid (mg/g DW)	4.32ab (4.0-4.8)	4.07b (2.9-5.0)	3.78b (3.3-4.7)	4.51ab (3.6-5.2)	1.3-4.1[9]
Total Daidzein (mg/100 g DW)	99.0a (85.1-123.6)	74.8b (67.4-82.7)	103.3a (95.2-118.5)	100.6a (95.2-118.5)	NA
Total Glycitein (mg/100 g DW)	25.1b (24.1-26.5)	22.9b (20.9-27.9)	48.3a (42.8-52.4)	52.4a (45.9-57.9)	NA
Total Genistein (mg/100 g DW)	133.5b (114.5-167.9)	115.7c (97.7-132.6)	166.8a (149.7-199.1)	161.7a (138.7-203.4)	NA

* Values in the same row followed by the same letter are not significantly different at $p<0.05$.
[1]Fulmer (1988), Orthoefer (1978); [2]Han et al (1991), Orthoefer (1978); [3]Orthoefer (1978), Smith and Circle (1972); [4]Waggle and Kolar (1979); [5]Rackis (1974); [6]Coon et al (1988), Kuo et al (1988), Rackis (1974); [7]Anderson and Wolf (1995), Rackis (1974); [8]Lee and Garlich (1992); [9]Anderson and Wolf (1995), Mohamed et al (1991).

6.4.2 Protein isolate and concentrate

Results of the protein isolate and concentrate analyses are shown in Table 16. The only significant differences observed were for ash in protein isolate which was decreased in CV127 compared to the control. The mean levels for ash in CV127 were however within the literature range for other soybean varieties. No significant differences were observed for any of the other analytes in either protein isolate or concentrate.

Table 16: Proximate composition of protein isolate and protein concentrate

Analyte	Control	CV127 +imi	Reference Line 1	Reference Line 2	Literature Range
Protein Isolate	**Mean (range)**				
Ash (g/100 g DW)	3.3a* (2.9-4.0)	2.6b (2.2-3.2)	3.2ab (2.7-3.8)	2.6ab (1.9-3.1)	2.3-7.6[1]
Fat (g/100 g DW)	5.4a (4.1-6.6)	7.0a (5.8-9.0)	5.2a (3.9-6.1)	5.6a (4.6-6.9)	0.1-2.5[2]
Protein (g/100 g DW)	90.4ab (89.0-91.0)	90.4ab (88.1-92.7)	89.6b (88.4-91.1)	91.5a (90.6-92.5)	85.2-92.0[3]
Carbohydrates (g/100 g DW)	0.9ab (0.1-1.8)	0.7ab (0.0-2.1)	2.1a (1.0-3.5)	0.8ab (0.0-1.7)	0.3-0.6[4]
Calories (kcal/100 g DW)	414a (407-417)	428a (417-436)	414a (408-418)	420a (411-427)	NA[#]
Protein Concentrate	**Mean (range)**				
Ash (g/100 g DW)	4.0ab (3.6-4.1)	4.4a (4.0-4.9)	3.7b (3.2-4.4)	4.0ab (3.8-4.2)	4.7-6.5[5]
Fat (g/100 g DW)	5.1a (3.7-7.4)	6.2a (4.5-9.2)	5.1a (3.7-6.8)	4.4a (3.8-5.7)	0.9-2.0[6]
Protein (g/100 g DW)	79.5a (76.7-81.1)	78.2a (70.9-85.6)	80.9a (77.1-82.8)	78.2a (77.0-79.3)	66.2-78.1[5]
Carbohydrates (g/100 g DW)	11.5a (7.4-14.2)	11.4a (1.0-18.6)	10.4a (7.3-13.5)	13.3a (11.5-14.9)	17.1-25.0[7]
Calories (kcal/100 g DW)	410a (403-421)	414a (406-429)	411a (404-421)	406a (402-412)	NA

* Values in the same row followed by the same letter are not significantly different at p<0.05.
[#] Not available.
[1]Smith and Circle (1972), Wolf (1983); [2]Horan (1974), Wolf (1983); [3]Torun (1979), Waggle and Kolar (1979); [4]Waggle and Kolar (1979), Wolf (1983); [5]Bookwalter (1978), Smith and Circle (1972); [6]O'Dell (1979), Wolf (1983); [7]Mattil (1974), Smith and Circle (1972).

6.4.3 Refined oil

Results of the refined oil analyses are shown in Table 17. Myristic acid (14:0) was detected in refined oil but was present below the level of quantification. There were no significant differences between CV127 and the control for palmitic, stearic, arachidic, eicosenoic and tetracosanoic acid. Refined oil from CV127 contained higher levels of oleic acid and decreased levels of linoleic, linolenic and behenic acid compared to the control. These levels were all within the literature ranges except for behenic acid, which was slightly higher than the range reported in the literature. The differences observed for refined oil were consistent with the differences identified in the seed analyses for fatty acid content.

Table 20: Fatty acid composition of refined oil

Fatty Acid	Control	CV127 +imi	Reference Line 1	Reference Line 2	Literature Range[1]
	Mean (range) g/100 g oil				
Palmitic 16:0	10.33ab* (9.66-10.90)	9.94b (9.61-10.42)	10.38a (9.66-10.76)	9.19c (8.94-9.32)	7-12
Stearic 18:0	3.84a (3.54-4.21)	3.75a (3.25-4.21)	3.67a (3.44-3.97)	3.36b (3.15-3.59)	2-5.5
Oleic 18:1	21.96c (21.03-22.66)	24.28a (22.56-25.62)	18.88d (17.78-19.74)	21.24c (20.65-21.94)	20-50
Linoleic 18:2	51.52b (50.95-52.29)	50.45c (49.43-51.67)	53.66a (53.20-54.06)	53.99a (53.58-54.30)	35-60
Linolenic 18:3	6.19b (5.93-6.41)	5.64c (5.26-6.02)	7.75a (7.27-8.27)	6.49b (6.31-6.64)	2-13
Arachidic 20:0	0.42a (0.38-0.48)	0.42a (0.38-0.48)	0.35b (0.29-0.38)	0.29c (0.29)	0.2-1.0
Eicosenoic 20:1	0.23a (0.14-0.29)	0.22a (0.19-0.29)	0.18a (0.14-0.19)	0.19a (0.19)	<1.0
Behenic 22:0	0.60a (0.57-0.67)	0.54b (0.48-0.57)	0.48c (0.43-0.53)	0.46c (0.43-0.48)	<0.5
Tetracosanoic 24:0	0.19a (0.19)	0.19a (0.19)	0.13b (0.10-0.19)	0.19a (0.19)	NA

* Values in the same row followed by the same letter are not significantly different at p<0.05.
# Not available.
[1]Pryde (1990)

6.5 Endogenous allergens

Soybean naturally contains allergenic proteins and is one of a group of known allergenic foods including milk, eggs, fish, shellfish, wheat, peanuts, tree nuts and sesame. This group of foods accounts for approximately 90% of all food allergies (Metcalfe *et al*., 1996). The FARRP Allergen Protein Database lists 33 allergenic proteins for soybean. These include the β-conglycinin family of proteins, the glycinin family, the trypsin inhibitor family and other miscellaneous allergens. The major allergens are among the major proteins present in soybean seed.

Since soybean is associated with allergic effects in susceptible individuals, a study was done to determine if there were any significant differences in the levels of known allergens in soybean line CV127 compared to the non-GM parent line (Conquista). This was done using proteomic techniques to directly monitor the levels of specific allergens in the seed.

Studies submitted:

BASF Report No 16331 (2007) Identification and Quantification of Allergens in Different Soybean Lines, Comparison between Cultivance Soybean Event 127 and its Conventional Counterpart, Conquista.

The Applicant states that several publications describing proteomic studies in soybeans are available in the literature (Herman et al 2003, Hajduch et al 2005, Natarajan et al 2006a, Natarajan et al 2006b, Natarajan et al 2007), with one of the authors also publishing an interactive web database containing clickable two dimensional (2D)-protein maps of proteins from soybean seeds with multiple annotated spots (Hajduch et al 2005).

The availability of these data means that 2D-PAGE can be used to compare the allergen content of seed from soybean line CV127 and control lines.

A proteomic study was therefore done using protein extracts from soybean seed obtained from Conquista and soybean line CV127 (F_8 generation) grown in the glasshouse in Brazil, together with Conquista seed from glasshouse grown plants in the US. This enabled a comparison between the GM line and the parental non-GM control, grown at the same location, as well as a comparison between two identical lines grown at different locations to determine the variability in protein patterns. The study focussed on the major soybean allergens (Kunitz trypsin inhibitor, dehydrin-like protein, β-conglycinin alpha subunit, glycinin A5A4D3 subunit precursor, glycinin precursor,β-conglycinin alpha prime subunit) and the major lectin, agglutinin. Two other proteins, lipoxygenase and sucrose binding protein, were monitored as controls.

Protein extracts from each of the soybean lines were first separated according to isoelectric point using isoelectric focussing, and then separated according to migration in SDS-PAGE. Proteins were visualised using silver staining. Image analysis and comparison of the relative quantities of identical proteins in the samples was done using commercially available software (ImageMaster Platinum™). For each soybean line, three replicate gels were prepared and analysed.

The results indicate a high degree of similarity in protein pattern between Conquista and CV127 soybean and suggest there are no significant differences in the levels of major allergens between the two lines.

6.6 Conclusion

The compositional analyses do not indicate any differences of biological significance between the seed from CV127 soybean and the null segregant control. Significant differences were noted in a number of constituents. However the differences were typically small and almost all mean values were within the range reported in the ILSI database and/or were within the range obtained for reference varieties grown at the same time. Any observed differences are therefore considered to represent the natural variability that exists within soybean. The spraying of CV127 soybean with imidazolinone herbicides did not have a significant effect on seed composition.

In addition, no significant differences in endogenous allergen content of seed from CV127 soybean compared to the non-GM parental line were identified.

7. Nutritional impact

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

If the compositional analysis indicates biologically significant changes to the levels of certain nutrients in the GM food, additional nutritional assessment should be undertaken to assess the consequences of the changes and determine whether nutrient intakes are likely to be altered by the introduction of such foods into the food supply.

Where a GM food has been shown to be compositionally equivalent to conventional varieties, as is the case for soybean line CV127, the evidence to date indicates that feeding studies using target livestock species will add little value to the safety assessment and generally are not warranted (OECD, 2003; EFSA, 2008).

Studies submitted:

BASF Project No 10.06.21.033-01 (2008) Performance of chickens fed with feed containing soybean meal from genetically modified imidazolinone-tolerant CV127 soybean as compared to conventional soybean meals.

Soybean line CV127 is the result of a simple genetic modification to confer herbicide tolerance with no intention to significantly alter nutritional parameters in the food. In addition, the extensive compositional analyses of seed that have been undertaken to demonstrate the nutritional adequacy of CV127 indicate it is equivalent in composition to conventional soybean varieties. The introduction of food from soybean line CV127 into the food supply is therefore expected to have little nutritional impact.

The Applicant did however undertake a feeding study in chickens to assess their growth performance and general health when fed diets containing soybean meal from CV127 compared to diets containing soybean meal from non-GM soybean. The results of this study, which did not show any significant differences between meal from CV127 compared to the non-GM soybean, were consistent with the findings from the compositional analyses (Section 6).

References

American Soybean Association (2011) 2011 Soystats. A reference guide to important soybean facts and figures. American Soybean Association. Available from http://www.soystats.com/2011/default.htm [Accessed on 16 December 2011]

Anderson RL, Wolf WJ (1995) Compositional changes in trypsin inhibitors, phytic acid, saponins, and isoflavones related to soybean processing. Journal of Nutrition 125: 518S-588S

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815

Aragáo FJL, Barros LMG, Brasileiro ACM, Ribeiro SG, Smith FD, Sanford JC, Faria JC, Rech EL (1996) Inheritance of foreign genes in transgenic bean (*Phaseolus vulgaris* L.) co-transformed via particle bombardment. Theoretical and Applied Genetics 93: 142-150

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

Bannon GA, Goodman RE, Leach JN, Rice E, Fuchs RL, Astwood JD (2002) Digestive stability in the context of assessing the potential allergenicity of food proteins. Comments on Toxicology 8(3): 271-285

Beckmann R, Bubeck D, Grassucci R, Penczek P, Verschoor A, Blobel G, Frank J (1997) Alignments of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. Science 278: 2123-2126

Bookwalter GN (1978) Soy protein utilization in food systems. In: Friedman M (ed) Nutritional improvement of food and feed proteins. Plenum Press, New York, NY, p. 749-766

Coon C, Akavanichan O, Cheng T (1988) The effect of oligosaccharides on the nutritive value of soybean meal. In: McCann L (ed) Soybean utilization alternatives. Centre for Alternative Crops and Products, University of Minnesota., St Paul, MN, p. 203-213

Davies, ME (1964) Acetolactate and acetoin synthesis in ripening peas. Plant Physiology 39: 53-59

Delaney B, Astwood JD, Cunny H, Conn RE, Herouet-Guicheney C, MacIntosh S, Meyer, LS, privalle L, Gao, Y, Mattsson J, Levine M, ILSI International Food Biotechnology Committee Task Force on Protein Safety (2008) Evaluation of protein safety in the context of agricultural biotechnology. Food and Chemical Toxicology 46: S71-S97

Duggleby RG, Pang SS (2000) Acetohydroxyacid synthase. Journal of Biochemistry and Molecular Biology 33(1): 1-36

Duggleby RG (2006) Domain relationships in thiamine diphosphate-dependent enzymes. Accounts of Chemical Research 39: 550-557

Duggleby RG, McCourt JA, Guddat LW (2008) Structure and mechanism of inhibition of plant acetohydroxyacid synthase. Plant Physiology and Biochemistry 46: 309-324

EFSA (2008) Safety and nutritional assessment of GM plants and derived food and feed: The role of animal feeding trials. Food and Chemical Toxicology 46:S1-S70

Fulmer RW (1988) The soybean as a chemical factory. In: McCann L (ed) Soybean utilization alternatives. Centre for Alternative Crops and Products, University of Minnesota., St Paul, MN, p. 1-12

Grey, D (2006) Growing Soybeans in Northern Victoria. State of Victoria, Department of Primary Industries

Hajduch M, Ganapathy A, Stein JW, Thelen JJ (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution, two-dimenstional reference maps, expression profiles, and an interactive proteome database. Plant Physiology 137: 1397-1419

Han Y, Parsons CM, Hymowitz T (1991) Nutritional evaluation of soybeans varying in trypsin inhibitor content. Poultry Science 70: 896-906

Hanein D, Matlack KE, Jungnickel B, Plath K, Kalies KU, Miller KR, Rapoport TA, Akey CW (1996) Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. Cell 87:721-732

Hartmann E, Sommer T, Prehn S, Görlich D, Jentsch S, Rapoport TA (1996) Evolutionary conservation of components of the protein translocation complex. Nature 367: 654-657

Haughn GW, Somerville C (1990) A mutation causing imidazolinone resistance maps to the *Csr1* locus of *Arabidopsis thaliana*. Plant Physiology 92: 1081-1085

Herman EM, Helm, RM, Jung R, Kinney AJ (2003) Genetic modification removes an immunodominant allergen from soybean. Plant Physiology 132: 36-43

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

Horan FE (1974) Soy protein products and their production. Journal of the American Oil Chemists' Society 51: 67A-73A

ILSI (2009) Crop Composition Database. International Life Science Institute, Washington DC. http://www.cropcomposition.org/query/index.html

James AT, Rose IA (2004) Integrating crop physiology into the Australian soybean improvement program. In: 4th International Crop Science Congress, Brisbane, Australia, September 2004.

Kimber I, Kerkvliet NI, Taylor SL, Astwood JD, Sarlo K, Dearman RJ (1999) Toxicology of protein allergenicity: prediction and characterization. Toxicological Sciences 48(2):157-162

Kuo TM, VanMiddlesworth JF, Wolf WJ (1988) Content of raffinose oligosaccharides and sucrose in various plant seeds. Journal of Agriculture and Food Chemistry 36: 32-36

Lee Y-T, Duggleby RG (2001) Identification of the regulatory subunit of *Arabidopsis thaliana* acetohydroxyacid synthase and reconstitution with its catalytic subunit. Biochemistry 40: 6836-6844

Lee H, Garlich JD (1992) Effect of overcooked soybean meal on chicken performance and amino acid availability. Poultry Science 71: 499-508

Mattil KF (1974) Composition, nutritional, and functional properties, and quality criteria of soy protein concentrates and soy protein isolates. Journal of the American Oil Chemists' Society 51: 81A-84A

Mazur BJ, Chok-Fun C, Smith JK (1987) Isolation and characterization of plant genes coding for acetolactate synthase, the target enzyme for two classes of herbicide. Plant Physiology 85: 1110-1117

McCourt JA, Pang SS, King-Scott J, Guddat LW, Duggleby RG (2006) Herbicide-binding sites revealed in the structure of plant acetohydroxyacid synthase. Proceedings of the National Academy of Sciences (USA) 103(3): 569-573

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

Mohamed AI, Mebrahtu T, Rangappa M (1991) Nutrient composition and anti-nutritional factors in selected vegetable soybean (Glycine max [L.] Merr.). Plant Foods for Human Nutrition 41: 89-100

Natarajan S, Xu C, Bae H, Caperna TJ, Garrett WM (2006a) Proteomic analysis of allergen and antinutritional proteins in wild and cultivated soybean seeds. Journal of Plant Biochemistry and Biotechnology 15: 103-108

Natarajan S, Xu C, Bae H, Caperna TJ, Garrett WM (2006b) Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic analysis. Journal of Agricultural and Food Chemistry 54: 3114-3120

Natarajan S, Xu C, Bae H, Bailey BA (2007) Proteomic and genomic characterization of Kunitz trypsin inhibitors in wild and cultivated soybean genotypes. Journal of Plant Physiology 164: 756-763

Newhouse KE, Smith WA, Starrett MA, Schaefer TJ, Singh BK (1992) Tolerance to imidazolinone herbicides in wheat. Plant Physiology 100: 882-886

O'Dell BL (1979) Effect of soy protein on trace mineral availability. In: Wilke HL, Hopkins DT, Waggle DH (eds) Soy protein and human nutrition. Academic Press, New York, NY, p. 187-204

OECD (2001) Consensus document on compositional considerations for new varieties of soybean: key food and feed nutrients and anti-nutrients. Series on the Safety of Novel Foods and Feeds, No. 2. Organisation for Economic Co-operation and Development, Paris

OECD (2003) Considerations for the Safety Assessment of Animal Feedstuffs Derived from Genetically Modified Plants. Series on the Safety of Novel Foods and Feeds, No. 9. Organisation for Economic Cooperation and Development, Paris

Orthoefer FT (1978) Processing and utilization. In: Norman AG (ed) Soybean physiology, agronomy and utilzation. Academic Press., New York, NY, p. 219-246

Osborne AR, Rapoport TA, van den Berg B (2005) Protein translocation by the Sec61/SecY channel. Annual Review of Cell and Developmental Biology 21: 529-550

Pryde EH (1990) Composition of soybean oil. In: Erickson DR, Pryde EH, Brekke OL, Mounts TL, Falb RA (eds) Handbook of soy oil processing and utilization. 5th Edition, American Soybean Association and American Oil Chemists' Society, St. Louis, MO, and Champaign, IL, p. 13-31.

Rackis JJ (1974) Biological and physiological factors in soybeans. Journal of the American Oil Chemists' Society 51: 161A-174A
Rapoport TA, Jungnickel B, Kutay U (1996) Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. Annual Review of Biochemistry 65: 271-303

Sambrook J, Fritsch J, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd Edition. Cold Spring harbour Laboratory Press, Cold Spring Harbour, New York.

Sathasivan K, Haughn GW, Murai N (1990) Nucleotide sequence of a mutant acetolactate synthase gene from an imidazolinone-resistant *Arabidopsis thaliana* var. Columbia. Nucleic Acids Research 18(8): 2188

Short JM, Fernandez JM, Sorge JA, Huse WD (1988) λ Zap: a bacteriophage λ expression vector with in vivo excision properties. Nucleic Acids Research 16(15): 7583-7600

Shurtleff W, Aoyagi A (2007) History of soybean crushing: Soy oil and soybean meal. In: History of soybeans and soyfoods: 1100 B.C. to the 1980s. SoyInfo Center, Lafayette, California

Singh BK, Stidham MA, Shaner DL (1988) Assay of acetohydroxyacid from plants. Analytical Biochemistry 171: 173-179

Smith AK, Circle SJ (1972) Chemical composition of the seed. In: Smith AK, Circle SJ (eds) Soybeans: chemistry and technology. Vol 1, Proteins. Avi Publishing, Wesport, CT., p. 61-92

Stidham MA, Singh BK (1991) Imidazolinone-acetohydroxyacid synthase interactions. Ch 6 In: Shaner D, O'Connor S, (eds) The imidazolinone herbicides. CRC Press, Boca Raton, FL., p. 71-90.

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

Thomas K, MacIntosh S, Bannon G, Herouet-Guicheney C, Holsapple M, Ladics G, McClain S, Vieths S, Woolhiser M, Privalle L (2009) Scientific advancement of novel protein allergenicity evaluation: An overview of work from the HESI Protein Allergenicity Technical Committee (2000 - 2008). Food and Chemical Toxicology 47:1041-1050

Torun B (1979) Nutritional quality of soybean protein isolates: studies in children of preschool age. In: Wilke HL, Hopkins DT, Waggle DH (eds) Soy protein and human nutrition. Academic Press, New York, NY, p. 101-119

USDA (2009) Oilseeds: World markets and trade. Circular Series FOP1-09. Foreign Agricultural Service, United States Department of Agriculture

Waggle DH, Kolar CW (1979) Types of soy protein products. In: Wilke HL, Hopkins DT, Waggle DH (eds) Soy protein and human nutrition. Academic Press, New York, NY, p. 19-51

Wolf WJ (1983) Edible soybean protein products. In: Wolff IA (ed) CRC handbook of processing and utilization in agriculture. Vol II, Part 2. CRC Press, Boca Raton, FL, p. 23-55