Centre for Integrated Research in Biosafety

Tel: +64 3 364 2500, Fax: + 64 3 364 2590
Email: brigitta.kurenbacj@canterbury.ac.nz

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Submission II on the Assessment Report for **APPLICATION A1018 FOOD DERIVED FROM HIGH OLEIC ACID SOYBEAN DP-DP-305423-1-1**

Submitted to Food Standards Australia/New Zealand (FSANZ)

by

Submitter: Centre for Integrated Research in Biosafety
23 October 2009

Authors: Brigitta Kurenbach*, Dorien S. Coray, Jack A. Heinemann, Ryan J. Catchpole and Leighton A. Turner

*Contact address: Dr. Brigitta Kurenbach. University of Canterbury, Private Bag 4800, Christchurch, New Zealand
voice: + 64 3 364 2500; fax + 64 3 364 2590; email: brigitta.kurenbach@canterbury.ac.nz

Summary of key findings

This submission evaluates several aspects of FSANZ's Assessment Report of soybean DP-305423-1 and its Supplementary Document 1. It is structured in three parts.

Part 1 outlines our concerns with the inconsistent and inappropriate use of controls and test material described in the Assessment Report. Different "comparators" were used in the various parts of the molecular and compositional analysis. Apart from the proper comparator variety 'Jack', null-segregants of GM varieties with the event DP-305423-1 (BC1F) were used. Codex Alimentarius guidelines allow a regulator to insist that the comparator not be a GMO because the use of a GMO as a comparator undermines the baseline data that makes a comparative analysis possible. Equally, the use of test material is not consistent throughout the supplied studies. T-line plants, derived from the initial transformed plant, were used for some analyses, and BC1F-line plants, derived after repeated outcrossing to unspecified 'elite varieties' were used in other studies. Additionally, some parts of the Assessment Report come without the specification of the genetic background of test and control material, being only referred to as "soybean DP-305423-1" or "control line". This makes data analysis or correlation of data across assays impossible. Furthermore, we object to the use of pooled data in the GM-HRA protein expression study and the analysis of key components. This decreases the likelihood of detecting statistically significant differences between sites and between test plant and control. Still, significant differences were detected for several components but we are concerned that these were not further addressed by FSANZ.

Part 2 discusses FSANZ's assessment of new open reading frames (ORFs) that were or may have been created in soybean DP-305423-1. The Applicant identified two fusion-ORFs at the junctions of two of the four insertion sites. These were examined only using sequence data (that is, "bioinformatically"), and FSANZ assumed that the ORFs are of "very low safety concern" without requiring any experimental proof. Against FSANZ guidelines, the possibility that new ORFs that might have arisen within or between the four insertion sites was not examined. Likewise, potential changes in the chromosome or the wider genome of soybean DP-305423-1 and consequent unintended effects were not addressed.

Part 3 discusses the expression of the transgenic *fad2-1* sequence. FSANZ assumed that the introduced *fad2-1* sequence, which comprises about 40% of the original DNA sequence of the gene, will not be expressed because it was introduced to induce silencing of the endogenous *fad2-1* gene. This assumption is not supported by experimental data showing that indeed the transcript will not be translated or is part of some other regulatory process. This is against FSANZ guidelines (2.e p.89 of FSANZ, 2008), which state that "*evidence of non-expression of a gene, in the case where a transferred gene is not expected to express any novel substances (e.g. because it has a 'silencing' role…)*" must be included in the application.

Part 4 discusses FSANZ's conclusion that the GM-HRA protein derived from bacteria is a suitable surrogate in safety studies. The source of the protein used to obtain mouse antiserum is unclear from FSANZ's Assessment Report. If it was also derived from bacteria, isoforms that are specific to the plant (e.g. by post-translational modifications) might not be recognised. The potential for adverse effects of five added amino acids at the N-terminus of GM-HRA were not discussed by FSANZ. The assumption that this change will not result in differences in post-translational processing, including cleavage during translocation, cannot be drawn from the presented data. Both have implications for the detection of potential allergens.

Thus, we encourage FSANZ to change their preference from Option 2: accept to Option 1: reject unless and until the concerns of the INBI submissions are satisfied using appropriate data from economic and safety studies.

Introduction

This submission from the Centre for Integrated Research in Biosafety (INBI) is one of two. Our submissions were prepared in response to an invitation from Food Standards Australia/New Zealand to comment on application A1018. A1018 is an application to amend the Australia New Zealand Food Standards Code to allow foods derived from soybean line DP-DP-305423-1-1 into the human food supply. In this case, the *fad2-1* gene is silenced by expressing a partial duplicate (*gm-fad2-1*) which is sufficient to evoke an RNAi effect. Silencing of *fad2-1* inhibits conversion of oleic acid into linoleic acid, resulting in an accumulation of the former and lower levels of the latter.

Our submissions were built in large part using the Biosafety Assessment Tool (https://bat.genok.org/bat/) produced by the University of Canterbury and GenØk – Centre for Biosafety. This is a free-to-the-public resource for hazard identification and risk assessment of genetically modified organisms.

This submission is based on the Executive Summary of Application A1018 Food derived from high oleic acid soybean line DP-DP-305423-1 Assessment Report (FSANZ 2009a) and its Supporting Document 1 (referred to hereafter as SD-1 and cited as FSANZ 2009b), prepared by FSANZ. It is in four parts which follow a list of recommendations in brief. The first part evaluates the use of controls, comparators and test material throughout the Assessment Report. Part two discusses detected and potential new open reading frames resulting from the introduction of the transgenic material. The third part discusses FSANZ's assumption that the transgenic *fad2-1* sequence will not be expressed in soybean DP-305423-1. Part four raises the issue of potential isoforms of the GM-HRA protein, which may have gone unnoticed if the antiserum was obtained using bacterially derived protein.

Summary of recommendations

Following this summary is a detailed explanation of the recommendations.

1. FSANZ should require the Applicant to submit all safety studies using the appropriate, closely related non-GM parent 'Jack' as comparator with the *most closely related GM test variety*.
2. FSANZ should require the Applicant to submit all safety data using material from the 'T series' plants (progeny derived by selfing of the original transformed plant) as test material rather than the 'BC1F series' which are too distantly related to a conventional parent for proper comparisons to be made.
3. FSANZ should require the Applicant to provide full genotypes and breeding histories of the 'elite cultivars' used in crosses and backcrosses with T plants, which were then used as test material in several studies.
4. We recommend that FSANZ incorporate the South America field data into the Assessment Report.
5. FSANZ should require the Applicant to provide experimental data demonstrating that each statistically significant difference between soybean DP-305423-1 and 'Jack' in the analysis of key components and in respect to the broad range of values obtained for GM-HRA expression in different environments raises no safety concerns.

6. FSANZ should require the Applicant to submit data on the potential immunostimulatory effects of potential novel peptides that could be produced from the two identified new open reading frames. These data should comply with FAO/WHO guidelines.
7. FSANZ should require the Applicant to demonstrate that the two identified new ORFs are indeed not expressed to a specified level of expression or, if expressed, create no new safety risk.
8. FSANZ should require the Applicant to describe the experiments and the limits of detection in which potential unintended effects resulting from the disruption of endogenous genes were identified.
9. FSANZ should require the Applicant to submit data to verify the absence of additional new ORFs arising by the insertion of the rDNA within or outside the four insertion loci.
10. FSANZ should require the Applicant to demonstrate that the transformation process did not result in changes to expression or sequence of endogenous genes surrounding the four insertions.
11. FSANZ should require the Applicant to submit proof that no peptides are translated from any species of RNA that arises from transcription over the transgenic *gm-fad2-1* gene.
12. FSANZ should confirm that the source of GM-HRA used to produce the anti-HRA antiserum used in several assays originated from seeds of DP-305423-1. If this cannot be established, FSANZ should require the Applicant to resubmit data using an appropriate antiserum.
13. FSANZ should require the Applicant to perform the relevant experiments to confirm that there are no unprocessed or partly processed isoforms of GM-HRA in cells. If that assurance is not possible, FSANZ should require the Applicant to verify that these isoforms raise no safety concerns

Part 1: Inconsistent and inappropriate controls, test material and treatment of data

1.1 Choice of Test substance and Controls for Comparative Analysis

The Assessment Report makes reference to more than one 'comparator' as a source of control material for safety analyses. Inconsistent use of a single, common and proper comparator as a control fundamentally undermines the validity of the safety assessment. The purpose of the *comparator* is to provide the standard baseline for all measurements, and be the single common element in all experiments using material grown in multiple locations and years. The ad hoc mixing of control sources therefore invalidates any attempt to draw sound scientific conclusions across the experiments in this dossier.

Codex Alimentarius guidelines define the conventional counterpart (which is the proper comparator) as "*a related plant variety, its components and/or products for which there is experience of establishing safety based on common use as food"* (Codex, 2003). Transgenic soybean DP-305423-1 was derived by bioballistic transformation of soybean cultivar 'Jack' and thus is the most appropriate conventional parent to serve as the comparator.

On p. 5 of SD-1 FSANZ apparently agree, stating that:

> "*The soybean cultivar 'Jack' has been used as the parental variety for the high oleic acid trait described in this application, and thus is regarded as the near-isogenic line for the purposes of comparative assessment*" (emphasis added).

However, throughout the report FSANZ then allow the Applicant to mix and match controls rather than consistently use the proper comparator as the baseline for comparisons. For some analyses, controls are not specified at all (see Table 1 and section 1.1.3 below for details).
According to the breeding information, the T generations (T1-T5) were derived from selfed plants originating from the initial transformed plant (T0). They were used in several assays (see Table 1). In these instances, using the parental line 'Jack' as comparator was the appropriate choice. The Applicant used F plants, which are crosses between T line plants and an elite line, in the analyses of transgene generational stability. In this case, both 'Jack' and the parental 'elite line' were used as controls in Southern blots.

While Codex does not preclude the use of control lines in addition to the proper comparator, these additional lines should not substitute for the comparator. Studies where this occurred should be repeated and resubmitted to FSANZ for evaluation.

1.1.1 Some analyses lack proper conventional comparator

In assays used to measure GM-HRA expression and in parts of the compositional analysis, different lines were used as comparators. These controls were null-segregants of BC (backcross) lines derived from repeated breeding of the T3 generation with non-GM 'elite varieties' followed by selfing.

BC null-segregants are derived from crosses of a non-GM line with DP-305423-1 lines, and therefore are also GMOs. Their use is inconsistent with Codex Alimentarius guidelines which say that it "*is recognized that for the foreseeable future, foods derived from modern biotechnology will not be used as conventional counterparts*" (footnote p. 2 of both Codex, 2003, Codex, 2008). This is because GM lines do not have a history of safe use which is necessary for comparison.

Table 1: Transgenic organisms and controls used in the safety assessment of soybean DP-305423-1

Analysis	DP-305423-1 soybean	Control used
transgene copy number and insertion integrity	T4	Comparator ('Jack')
Plasmid backbone analysis	T4 and T5	Comparator ('Jack')
Physical map of inserted DNA	Not reported	Not reported
Open reading frame analysis	Not reported	Not reported
Generational stability	T4, T5, F2	Comparator ('Jack') and elite line PHSB02
Segregational analysis	F2, F3, BC1F2 (2 elite backgrounds of each)	No controls reported for any method used[a]
fad2-1 gene silencing	Not reported	Jack
GM-HRA protein characterization (equivalence assessment)	Not reported	*E. coli* derived HRA
Glycosylation analysis	Not reported	Known glycosylated and non-glycosylated proteins
GM-HRA expression analysis	BC1F5	BC1F5 null
Compositional analysis[#]	BC1F5 or unidentified variety	BC1F5 null and 4 commercial varieties[b] or comparator ('Jack')

a) Plants were analysed using gas chromatography, Southern analysis and/or PCR.
b) BC1F5 and the null segregant were grown in 6 locations in North America, the four commercial lines in 2 of the above locations plus four other locations.
see section 1.1.3 below for ambiguous data in this part of the Assessment Report

Without the proper conventional parent as a control, additional effects of the transformation process might be overlooked in the safety assessment. For example, bioballistics transformation often leads to chromosomal rearrangements, which do not necessarily have to be at the locus of integration of the transgenes (discussed in more detail in Part 2) (p. 20 Latham et al., 2006). FSANZ do not mention analysis undertaken by the Applicant to determine whether DNA of soybean DP-305423-1 has been altered outside the insertion locus. If other genomic changes occurred during the transformation or tissue regeneration process and they disrupted or altered genes located on another chromosome, the changes might still be present in the null-segregants.

FSANZ need look no further than at the elevated levels of the 9,15 isomer of linoleic acid in the null segregrants, which no longer have the transgene but continue to display traits unique to the GM plants. The four conventional varieties grown as reference lines did not contain the 9,15 isomer in detectable amounts (p. 33 FSANZ, 2009b). This is consistent with previous findings (FSANZ, 2000, Kinney and Knowlton, 1997) showing that this isomer does not occur in conventional soybeans and is unique to GM varieties. No explanation is given by FSANZ for the presence of the 9,15 isomer in the GM plants or the null-segregants.

The presence of the 9,15 isomer is an example of why GMOs should not be used to substitute for the proper, non-GM comparator. Both the test material and the null-segregant exhibit changes to their fatty acid profiles compared to any non-GM soybean, and which persist even when the transgene event is removed from the plant through breeding. The proper non-GM comparator (in this case only as used by Kinney and Knowlton (1997) and FSANZ (2000)), must be used to detect these unintended effects.

FSANZ list several food sources that naturally contain the 9,15 isomer (p. 33 FSANZ, 2009b). We counsel against this kind of comparison.[1] Foods such as mango pulp, cheese and beef are compositionally and nutritionally not comparable to soybean or soybean oil, are not cooked or mixed with the same range of ingredients, and thus are not predictive of the safety or otherwise of soybeans that make high levels of the 9,15 isomer.

The use of a GM comparator is not the current best practice standard set by Codex Alimentarius and all studies using these controls should be redone with the proper comparator instead. We recommend that FSANZ require the Applicant to submit all safety studies using GM test material derived from plants that are closely related to appropriate the appropriate parent non-GM comparator. The most robust experiments would use material with event DP-305423-1 from the earliest possible derivative of the first regenerated plant to compare with 'Jack' so as to ensure that 'Jack' was still the closest conventional relative to the transgenic plant being examined. (This is why, below, we argue that the testing should have been done using the T varieties as the test material.)

1.1.2: Inconsistent use of test material

The Assessment Report indicates that different lines containing event DP-305423-1 were used as test material. In addition to material from T line (see above), several studies used crosses between T and elite lines (F plants) or backcrosses of F to the "appropriate elite line" (BC). Crosses into several commercial lines were made by the Applicant for reasons other than optimizing the safety assessment, because they are both unnecessary for the safety assessment and they reduce the power of the experiment to reveal potential harms. This is confirmed by FSANZ who say:

> *a breeding programme was undertaken for the purposes of:*
> *…*

[1] See extended discussion in Part 1 of accompanying INBI Submission by Heinemann et al.

• *ensuring that the DP-305423-1 germplasm is incorporated into a wide cultivar base for commercialisation of GM high oleic acid soybean* (p.8 FSANZ, 2009b).

F and BC lines should not have been used for safety testing, since it is obviously possible to obtain enough material from T-plants for testing. This is clear both because this material was available for the analyses described in rows 1-3 of Table 1, and the T line should be as easy to cultivate for material as F and BC lines.

Once an event is approved in one variety, it is approved for all subsequent varieties derived from conventional breeding, so safety testing by the developer does and should not have to be delayed to this late stage of product development. Even if there were compelling reasons to test backcrosses, these tests should be done in addition to tests on material from T generations and comparator 'Jack'. We recommend FSANZ require the Applicant to submit safety data using material from T plants as test material. As discussed in section 1.1.1, all experiments should be conducted with comparator 'Jack', which was identified by FSANZ as the proper comparator.

Table 2: Summary of instances where incorrect controls were used or controls cannot be deciphered

Section	**FSANZ statement**	**Comment**
3.4.3 Physical map of the inserted DNA	*genomic DNA from DP-305423-1 soybean*" as sample and "*control soybean samples"*	The introduction to section 3.4 mentions T4, T5 and 'Jack', but this only refers to the "*(e)valuation of insert copy number, insert integrity and presence/absence of plasmid backbone sequences"*, not the physical map (p.13 of SD-1)
3.3.4 Open reading frame (ORF) analysis	*soybean DP-305423-1*	
3.6 The *fad2-1* partial sequence and gene silencing	*soybean DP-305423-1*	
4.3 GM-HRA protein characterization	*plant-derived GM-HRA protein isolated from soybean DP-305423-1*	This includes the glycosylation analysis
Table 8 (p.37 of SD-1)	*soybean DP-305423-1 oil*	Neither the legend (p. 37 of SD-1) nor the paragraph referring to Table 8 (p.36 SD-1) defines which organism was used to obtain presented data.
5.4 Assessment of endogenous allergenic potential	*soybean DP-305423-1 and the non-GM parent 'Jack'*	

Additionally, FSANZ's report provides no genetic description of the 'elite lines' used in the crosses and backcrosses, or even if backcrosses (F → BC generations) were performed with the same elite line(s) used to create the F generations. Without this information, it is impossible to correlate results obtained in different parts of the assessment. We recommend that FSANZ require the Applicant to provide full genotypes and breeding histories of the 'elite cultivars' used in crosses and backcrosses with T plants, which were then used as test material in several studies.

1.1.3 Lacking identification of test material and controls

Overall, the selective use and identification of controls and GM plants by the Applicant is confusing, unnecessary and damaging to the validity of a subsequent safety assessment. The term 'soybean DP-305423-1' is used to refer to both T and BC generation plants containing the DP-

305423-1 event, and frequently it is not clear which one is in fact being analyzed. These cases are summarized in Table 2.

Additional to the ambiguous statements listed in Table 2, Section 5 (Compositional analysis) contains contradictory statements regarding the test material and the comparator used. In the introductory paragraphs of 5.3.2 Fatty acids (p. 33 FSANZ, 2009b) it is stated that:

> "*The levels of 25 fatty acids in soybean DP-305423-1 and null segregant control seed were measured (…)"* and in the next paragraph that "*[r]esults [for the measurements] are given in Table 6…*".

"Table 6: Percentage composition, relative to total fat, of major fatty acids in soybean DP-305423-1 and control seed", does not further specify the source of the test material or the control. We are confused as to what were used as controls for each of these measurements, because FSANZ then go on to say that:

> "*[t]ogether, these two fatty acids [heptadecanonic and heptadecenoic] constitute approximately 2% of the total fatty acid content in soybean DP-305423-1, compared to approximately 0.17% in the control line 'Jack'* (see Table 6)" (emphasis added to p. 36 FSANZ, 2009b).

Since 'Jack' is not the null-segregant, the control for these different data could not be the same in all measurements. Additionally, the test material may not have been the same. If the Applicant saw a reason to use 'Jack' as a control, the test material may have been derived from T plants. Without further information, it is impossible to say which test and control material was used in this section. Given this kind of ambiguity about test material and controls used in safety studies, it is difficult to see how FSANZ came to conclude that they had adequate information to advocate for Option 2 (accept).

1.2 Data from field trials were pooled before analysis

The rationale for testing the GMO under different conditions (locations and years) is that both parentage and environment contribute to the variance in phenotypes measured for the compositional analyses (Reynolds et al., 2005). It requires careful experimental preparation, execution and data analysis to identify unintended effects that may appear only under some environmental conditions (Zolla et al., 2008).

For the GM-HRA protein expression analysis and the analyses of other key components of soybean DP-305423-1, samples were taken from plants grown in six different locations in North America, and results are presented as pooled data. Pooling data from different field sites to make a range hides individual differences between the GMO and its control at any specific site. Instead, all significant differences between test and comparator in each site and between sites should have been identified and evaluated. Statistically significant differences between test and comparator were potentially hidden in pooled data. Additionally, the four non-GM commercial lines used as reference lines in the key component analysis (to create a 'population tolerance interval') were not grown in the same locations as the GMO. Adding further values derived from the literature, which are often obtained under incomparable conditions, increases the variability range of the data but does not help to identify legitimate safety issues that should be investigated.

Moreover, the Applicant collected agronomic data from Chile and Argentina (EFSA, 2007) but FSANZ only comment on data from North America. It is difficult to believe that the data from South America is not relevant with the data from North America is. According to the FSANZ guidelines:

"An applicant should submit all information relevant to the consideration of the safety of a substance, whether the information is an explicit requirement of the Handbook or not" (p. 9 FSANZ, 2009c).

Has the Applicant omitted this data from the dossier supplied to FSANZ or has FSANZ overlooked the data? We recommend that FSANZ incorporate the South America field data into the Assessment Report.

Despite the poor use of statistics, several statistically significant differences were detected between soybean DP-305423-1 and the null-segregant in the compositional analysis of key components. These differences were in the fat and ash content and the amount of several fatty acids, namely myristic, palmitoleic, arachidic, lignoceric, eicosenoic, heptadecanoic and heptadecenoic acids. Expression levels of the transgenic GM-HRA protein also seem to vary considerably between the locations. For example, measured concentrations of that protein varied between 0 and 4.9 μg/mg dry weight in seed samples, or between 0.78 and 51 μg/mg dry weight in forage. The range in these measurements is extremely large. FSANZ's assessment did not include any follow up on the potential biological significance or reasons for these changes. Only for heptadecanoic and heptadecenoic acids do they acknowledge that the Applicant has at least speculated on a cause.

FSANZ should require the Applicant to provide experimental data demonstrating that each statistically significant difference between soybean DP-305423-1 and 'Jack' in the analysis of key components and in respect to the broad range of values obtained for GM-HRA expression in different environments raises no safety concerns.

Part 2: Presence and expression of potential fusion proteins

2.1: Characterisation of identified new open reading frames (ORFs)

As part of the molecular characterization of event DP-305423-1, the Applicant examined the 5' and 3' junctions of each of the four insertion sites for the presence of new ORFs. Two fusion-ORFs that span the junctions were identified by the Applicant. This is in accordance with Codex Alimentarius guidelines:

§31
Information […] should include: …
D) identification of any open reading frames within the inserted DNA or created by the insertions with contiguous plant genomic DNA including those that could result in fusion proteins (emphasis added to p. 14 Codex, 2003).

Further, Codex Alimentarius states that:

§33
In addition, information should be provided: …
E) to indicate whether there is any evidence to suggest that one or several genes in the host plant has been affected by the transformation process; and
F) to confirm the identity and expression pattern of any new fusion proteins (p. 15 Codex, 2003).

#1 at the 5' insert border of Insertion 2 (106 amino acids, nine contributed by the recombinant DNA)

FSANZ conclude that this ORF has a "*low degree of novelty*", due to having only nine amino acids contributed from the introduced recombinant DNA (rDNA), and it is unlikely to be expressed due to "*the absence of upstream and adjacent transcriptional elements*".

To conclude that a novel protein is likely to be of no safety concern because of the addition of 'only' nine amino acids is not a research-based conclusion. In fact, the change of only two amino acids (P183A and W560L) in the *gm-hra* gene used in soybean DP-305423-1 is enough to confer tolerance to ALS-inhibiting herbicides. Changes of single amino acids can drastically alter the characteristics of proteins (e.g. Doyle and Amasino, 2009, Hanzawa et al., 2005, Zubieta et al., 2008), a fact that underpins the field of directed evolution (reviewed in e.g. Bloom and Arnold, 2009, Tracewell and Arnold, 2009). One of the characteristics that can be changed is immunogenicity. For example, several groups reported significant decreases of IgE binding to a major peanut allergen after mutating single nucleotides (Glaspole et al., 2005, King et al., 2005, Ramos et al., 2009). Even more surprising, in some cases not even an amino acid change is necessary to alter the characteristics of a protein! Kimchi-Sarfaty et al. demonstrated that even synonymous single nucleotide polymorphisms (i.e. differences in the nucleotide sequence of a gene that do not alter the resulting amino acid sequence) can change the substrate specificity of the resulting protein, potentially by affecting its folding patterns during translation (Kimchi-Sarfaty et al., 2007). Changes in the tertiary structure alone can turn benign proteins into toxins (Bucciantini et al., 2002, Ellis and Pinheiro, 2002, Ross and Poirier, 2005), as demonstrated for the Prp proteins causing Creutzfeld-Jacob disease and mad cow disease (Caughey and Baron, 2006). Nine new amino acids are therefore more than enough to cause biological effects. It is only through proper scientific testing that FSANZ can rule out unintended or unanticipated effects.

#2 at the 5' genomic border of Insert 3 (235 amino acids, 54 contributed by rDNA)
FSANZ conclude that this ORF "*has a low likelihood of transcription because the truncated KTi3 promoter upstream is missing the elements necessary for transcription*". The absence of *known* transcriptional elements, i.e. promoters or enhancers, is neither proof of, nor sufficient evidence for, the conclusion that the DNA will not be transcribed. In current guidance documents, the European Food Safety Authority (EFSA) state that:

> *"It is also clear that not all functions and/or sequence patterns of plant genes and non-coding sequences (like promoters and enhancers) are known. Thus flanking sequence information will not provide unequivocal evidence for safety but will support the risk assessment substantially"* (p. 20 EFSA, 2006).

Only bioinformatics analysis of the derived amino acid sequences of the two ORFs was used to build an assessment of their likelihood to cause an immune response based on their similarity to known or putative allergens (Section 4.1 of SD-1). As one parameter, the predicted similarity to epitopes was determined using a window of eight consecutive amino acids. This is a conservative window that favours false negative results because linear epitopes have been identified that are as short as 5 amino acids (Banerjee et al., 1999, Beezhold et al., 1999). To reduce the probability for false negatives, FAO/WHO guidelines recommend using a window size of six amino acids for this analysis (FAO/WHO, 2001). We recommend that FSANZ require the Applicant to submit data using these parameters, to provide a comparison between search standards.

From sequence and bioinformatical analyses FSANZ conclude

> "*that there are very low safety concerns relating to the two ORFs created by the transformation procedure used to generate soybean DP-305423-1*" (p. 19 FSANZ, 2009b).

This conclusion is based only on assumptions and not supported by any experimental data showing that the ORFs do not produce novel RNA molecules[2] or proteins or, if the ORFs are expressed, that the products have no new toxicity/allergenicity potential. Appropriate laboratory experiments can be conducted, making it unnecessary for FSANZ to rely on assumptions. FSANZ should require the Applicant to demonstrate that the two identified new ORFs are indeed not expressed to a specified level of detection or, if expressed, create no new safety risk. Ultimately, only properly designed and conducted feeding and allergenicity trials would be able to capture unanticipated and unintended toxins/allergens. This is consistent with EFSA guidelines especially in cases such as this where there are other unintended effects of the modification:

> ''*… if there are any indications for the potential occurrence of unintended effects, based on the preceding molecular, compositional, phenotypic or agronomic analysis, not only new constituents, but also the whole GM food and feed should be tested*" (Section III 7.8.4 EFSA, 2006).

Furthermore, it is not apparent from the Assessment Report what genomic sequences contribute to the two new ORFs. Without this information it is not possible to exclude unintended effects due to the potential disruption or alteration of functional endogenous genes.

FSANZ should require the Applicant to describe the experiments and the limits of detection in which potential unintended effects resulting from the disruption of endogenous genes were identified.

2.2: Unidentified new ORFs

Paragraph 31 D of Codex Alimentarius guidelines (see above) allows regulators to ask for studies that go beyond sequencing the genome-insert borders. This is also in accordance with FSANZ guidelines, which require the Applicant to provide a full molecular characterization, including:

> *"…*
> *(v) the identification and characterisation of any unexpected open reading frames <u>within the inserted DNA</u> or created by insertion with contiguous genomic DNA, including those that could result in fusion proteins or unexpected protein expression products"* (emphasis added, p. 88 of FSANZ, 2008).

The DP-305423-1 event consists of multiple fragments integrated into at least four locations in the genome, presumably within the same chromosome[3]. These four integration sites are separated by an unspecified length of genomic DNA. Given that three of the four insertions consist of multiple integrations of rDNA, either consecutively or separated by non-rDNA (which may be either genomic DNA from the soy or filler DNA), it is necessary to investigate if further unanticipated fusion ORFs were created <u>within or between</u> the four insertions. FSANZ's Assessment Report does not contain any such evaluation of such data, if indeed they exist.

We recommend that FSANZ require the Applicant to submit data to verify the absence of new ORFs arising by all insertions of the rDNA.

Transformation using bioballistics results in changes in the chromosome, a fact FSANZ seem to agree with:

[2] See INBI submission by Heinemann et al for discussion of novel RNA molecules.
[3] FSANZ conclude that the four insertion sites are linked on one chromosome based on the inheritance patterns alone. Since the Applicant determined the genomic sequences around each site, data should be provided that unambiguously prove that the inserts are actually situated in the same region of one specific chromosome.

> *The nature of the process by which genetic material is introduced into an organism may result in unintended effects that include the creation of new open reading frames (ORFs) in the genome of the organism. These ORFs may, in turn, lead to the production of novel proteins which could have implications for human safety, particularly regarding toxicity and allergenicity, if ingested* (emphasis added, p.15 of FSANZ, 2009b).

These unintended effects may be due in part to the disruption of endogenous genes or due to the mechanism of transgene integration itself (e.g. Kohli et al., 2003, Wilson et al., 2004). The latter can result in rearrangements (scrambling) of endogenous sequences, mutations or insertion of filler DNA (reviewed e.g. in Filipecki and Malepszy, 2006, Kohli et al., 2003, Latham et al., 2006, Wilson et al., 2004). In a study of two oat lines which were co-transformed with one or two plasmids via bioballistics, the authors found that in both cases the integration had resulted in "*multiple transgene and genomic DNA rearrangements and regions of scrambling […]*" (Makarevitch et al., 2003). These changes can create new or altered ORFs in chromosomal sequences outside the locus of transgene insertion. Zolla et al (2008) conclude that:

> *[I]t is also evident that the insertion of a single gene does not result in a unique newly expressed protein, but rather in many differently expressed genes with respect to the control. This could be due to the fact that, when the transgene enters the nucleus, many genetic loci are randomly affected by the insertion procedure.* (p. 1854 Zolla et al., 2008).

No data are presented that demonstrate the absence of changes to endogenous genes. FSANZ should require the Applicant to demonstrate that no endogenous genes surrounding the four insertions were altered by the transformation process.

Part 3: Expression of the transgenic *fad2-1* sequence

FSANZ report that northern blots conducted by the Applicant confirmed the silencing of the endogenous *gm-fad2-1* and (to a lesser degree) *gm-fad2-2* genes in soybean line DP-305423-1. What the Assessment Report does not contain is proof that the introduced *fad2-1* gene sequence does not have any other functions. Specifically, the proportion of recombinant *fad2-1* transcripts actually participating in silencing of the endogenous *gm-fad2-1* and *gm-fad2-2* genes is not reported. It is likely that not all transcripts will be converted into siRNA species and that a proportion might be acting as truncated mRNAs translated into protein.

According to FSANZ guidelines, this possibility needs to be taken into account:

> *2. The characterization of novel proteins or other novel substances*
>
> *This part includes all of the following:*
> *(a) A full description of the biochemical function and phenotypic effects of all novel substances (e.g. a protein or an untranslated RNA) that could potentially be expressed in the new GM organism,[…]* (emphasis added, p. 88 of FSANZ, 2009c).

The silencing as documented by Northern blots certainly confirms that the target mRNA is reduced. However, this does not in itself prove the absence of harm from translation or other regulatory functions caused by the introduced *fad2-1* sequence.

FSANZ state that:

> *"The gm-fad2-1 partial sequence is not expected to give rise to any protein product and the intent of its transcription in soybean DP-305423-1 is to decrease the expression of one of the endogenous soybean fatty acid desaturase genes […]. No novel protein is therefore expected to be produced from transcript arising from the partial gm-fad2-1 sequence"* (emphasis added, p.17 of FSANZ, 2009b)

and even more assertively that:

> *"Since the introduced gm-fad2-1 element is a partial sequence rather than a complete gene, a functional protein is not produced during its transcription in cells of soybean DP-305423-1"* (emphasis added, p.5 of FSANZ, 2009a)

but FSANZ have not provided any evidence that this expectation is indeed true. FSANZ guidelines, which state that the characterisation of novel proteins or other novel substances must include:

> *2. …*
> *(e) Evidence of non-expression of a gene, in the case where a transferred gene is not expected to express any novel substances (e.g., because it has a 'silencing' role […])* (emphasis added, p. 89 of FSANZ, 2009c)

are being overruled in this assessment but no justification of why these guidelines do not apply was given. The experiments necessary to verify the absence (or otherwise) of a novel protein as required in 2.e above are within the ability of the Applicant and should not be left to assumption. We recommend FSANZ require the Applicant to submit proof that no peptides are translated from any species of RNA that arises from transcription over the transgenic *gm-fad2-1* gene.

Part 4: Characterisation of protein GM-HRA

4.1 Potential non-identification of novel GM-HRA isoforms

The Applicant used a range of methods to confirm the identity of GM-HRA protein produced in both a bacterial expression system and soybean DP-305423-1. In addition, the Applicant showed that they could isolate non-glycosylated isoforms of GM-HRA. The predominant isoforms of both *in bacteria* and *in planta*-derived GM-HRA were shown to be of an equivalent mass, N-terminal sequencing confirmed the proteins had similar amino acid sequences, and glycoprotein staining indicated that neither protein extracts contained glycoforms. In light of this evidence, FSANZ concluded that "*microbially-derived GM-HRA protein is a suitable surrogate for use in safety assessment studies*" (p. 22 FSANZ, 2009b).

However, the antiserum used in the western blot and immunoaffinity chromotography lacks sufficient description to draw the conclusion that all *in planta*-produced isoforms would be detected. Post-translational modifications vary by species, tissue and time of development, and important epitopes can be created by post-translational modifications (Kuster et al., 2001). Unless FSANZ can verify that the mouse anti-GM-HRA antiserum was raised to protein isolated from the seeds soybean DP-305423-1 FSANZ cannot know that the antiserum contains antibodies that would detect minority glycoforms, or other isoforms. FSANZ have not commented on either the source or the appropriateness of that protein source used to raise mouse antiserum for use in western blots and immunoaffinity chromotography purification processes.

Knowing the source of the protein used to elicit the immune response in mice is critical for reaching the appropriate conclusions regarding the ability of immunoaffinity chromatography to draw a representative sample of GM-HRA isoforms from soybean DP-305423-1. If immmunoaffinity

chromotography were only capable of capturing bacteria-derived isoforms, then novel GM-HRA isoforms (for example certain glycoforms) might be inadvertently and wrongly lost during the purification process. Unless FSANZ can confirm that GM-HRA extracted from the seeds of soybean DP-305423-1 was used to obtain mouse antiserum, they may overstate the data upon which they base the claim that protein made by bacteria is equivalent to the protein made by plants.

4.2 The presence of a novel chloroplast transit sequence at the N-terminal of GM-HRA

> *"Additionally, the coding region of the gm-hra gene differs from the endogenous als gene by having 5 artificial codons ... These codons are adjacent to a chloroplast transit peptide protein sequence (see Section 4.2) and are presumably removed with it during targeting of the HRA protein to the chloroplast"* (p. 12 FSANZ, 2009b)

It is not obvious to us upon which factual basis FSANZ presume that the addition of 5 amino acids adjacent to the N-terminal end of GM-HRA will pose no new hazards. Can FSANZ confirm that this novel protein still undergoes "normal" post-translational cleavage liberating the novel chloroplast targeting sequence upon translocation? Incomplete or inefficient processing of the transit sequence could produce novel GM-HRA isoforms in soybean DP-305423-1. These novel isoforms may not be detected using the antiserum prepared by the Applicant and thus may have gone unnoticed. Their potential existence remains a safety issue, e.g., as possible allergens.

We recommend that FSANZ require the Applicant to perform the relevant experiments to confirm that there are no unprocessed or partly processed isoforms of GM-HRA in cells. If that assurance is not possible, FSANZ should require the Applicant to verify that these isoforms raise no safety concerns.

Thus, we encourage FSANZ to change their preference from Option 2: accept to Option 1: reject unless and until the concerns of the INBI submissions are satisfied using appropriate data from economic and safety studies.

Respectfully submitted on behalf of the Centre,

Dr. Brigitta Kurenbach

References:

Banerjee, B., Greenberger, P. A., Fink, J. N. and Kurup, V. P. (1999). Conformational and linear B-cell epitopes of Asp f 2, a major allergen of Aspergillus fumigatus, bind differently to immunoglobulin E antibody in the sera of allergic bronchopulmonary aspergillosis patients. Infect. Immun. *67*, 2284-91.

Beezhold, D. H., Hickey, V. L., Slater, J. E. and Sussman, G. L. (1999). Human IgE-binding epitopes of the latex allergen Hev b 5. J. Allergy Clin. Immunol. *103*, 1166-72.

Bloom, J. D. and Arnold, F. H. (2009). In the light of directed evolution: Pathways of adaptive protein evolution. Proc Natl Acad Sci USA *106*, 9995-10000.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M. and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature *416*, 507-11.

Caughey, B. and Baron, G. S. (2006). Prions and their partners in crime. Nature *443*, 803-810.

Codex (2003). Codex Work on Foods Derived from Biotechnology. CAC/GL 45-2003. Codex. http://www.who.int/foodsafety/biotech/codex_taskforce/en/.

Codex (2008). Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA animals. CAC/GL 68-2008. Codex Alimentarius Commission.

Doyle, M. R. and Amasino, R. M. (2009). A single amino acid change in the Enhancer of Zeste ortholog CURLY LEAF results in vernalization-independent, rapid-flowering in Arabidopsis. Plant Physiol.

EFSA (2006). Guidance Document of the Scientific Panel on Genetically Modified Organisms for the Risk Assessment of Genetically Modified Plants and Derived Food and Feed. EFSA J. *99*, 1-100.

EFSA (2007). Summary of the Application for Authorisation of Genetically Modified 305423 Soybean and Derived Food and Feed in Accordance with Regulation (EC) 1829/2003. Application EFSA-GMO-NL-2007-45. European Food Safety Authority.

Ellis, R. J. and Pinheiro, T. J. (2002). Medicine: danger--misfolding proteins. Nature *416*, 483-4.

FAO/WHO (2001). Allergenicity of Genetically Modified Foods. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology. Food and Agriculture Organisation of the United Nations. http://www.who.int/fsf/Documents/Biotech_Consult_Jan2001/report20.pdf.

Filipecki, M. and Malepszy, S. (2006). Unintended consequences of plant transformation: a molecular insight. J Appl Genet *47*, 277-286.

FSANZ (2000). Application A387 Food derived from high oleic acid soybean lines G94-1, G94-19 and G168 Draft Risk Analysis Report. Australia New Zealand Food Authority (now Food Standards Australia New Zealand). Food Standards Australia New Zealand.

FSANZ (2008). Application Handbook. Food Standards Australia New Zealand.

FSANZ (2009a). Application A1018 Food Derived From High Oleic Acid Soybean Line DP-305423-1 Assessment Report. Food Standards Australia New Zealand.

FSANZ (2009b). Application A1018 Food Derived From High Oleic Acid Soybean Line DP-305423-1 Supporting Document 1 (Food Standards Australia New Zealand).

FSANZ (2009c). Application Handbook. Food Standards Australia New Zealand.

Glaspole, I. N., de Leon, M. P., Rolland, J. M. and O'Hehir, R. E. (2005). Characterization of the T-cell epitopes of a major peanut allergen, Ara h 2. Allergy *60*, 35-40.

Hanzawa, Y., Money, T. and Bradley, D. (2005). A single amino acid converts a repressor to an activator of flowering. Proc Natl Acad Sci USA *102*, 7748-7753.

Kimchi-Sarfaty, C., Oh, J. M., Kim, I. W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V. and Gottesman, M. M. (2007). A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science *315*, 525-528.

King, N., Helm, R., Stanley, J. S., Vieths, S., Luttkopf, D., Hatahet, L., Sampson, H., Pons, L., Burks, W. and Bannon, G. A. (2005). Allergenic characteristics of a modified peanut allergen. Mol Nutr Food Res *49*, 963-971.

Kinney, A. J. and Knowlton, S. (1997). Desinger Oils for Food Applications. In Genetic Engineering for Food Industry: A Strategy for Food Quality Improvement, S. Harlander, and S. Roller, eds. (London, Blackie Academic), pp. 193-213.

Kohli, A., Twyman, R. M., Abranches, R., Wegel, E., Stoger, E. and Christou, P. (2003). Transgene integration, organization and interaction in plants. Plant Mol. Biol. *52*, 247–258.

Kuster, B., Krogh, T. N., Mortz, E. and Harvey, D. J. (2001). Glycosylation analysis of gel-separated proteins. Proteomics *1*, 350-61.

Latham, J. R., Wilson, A. K. and Steinbrecher, R. A. (2006). The mutational consequences of plant transformation. J Biomed Biotechnol.

Makarevitch, I., Svitashev, S. K. and Somers, D. A. (2003). Complete sequence analysis of transgene loci from plants transformed via microprojectile bombardment. Plant Mol Biol *52*, 421-432.

Ramos, M. L., Huntley, J. J., Maleki, S. J. and Ozias-Akins, P. (2009). Identification and characterization of a hypoallergenic ortholog of Ara h 2.01. Plant Mol Biol *69*, 325-335.

Reynolds, T. L., Nemeth, M. A., Glenn, K. C., Ridley, W. P. and Astwood, J. D. (2005). Natural Variability of Metabolites in Maize Grain: Differences due to Genetic Background. J. Agrig. Food Chem. *53*, 10061-10067.

Ross, C. A. and Poirier, M. A. (2005). What is the role of protein aggregation in neurodegeneration? Nat. Rev. Mol. Cell Biol. *6*, 891-898.

Tracewell, C. A. and Arnold, F. H. (2009). Directed enzyme evolution: climbing fitness peaks one amino acid at a time. Curr Opin Chem Biol *13*, 3-9.

Wilson, A., Latham, J. and Steinbrecher, R. (2004). Genome Scrambling - myth or reality? Technical Report: October 2004. EcoNexus. www.econexus.info.

Zolla, L., Rinalducci, S., Antonioli, P. and Righetti, P. G. (2008). Proteomics as a complementary tool for identifying unintended side effects occurring in transgenic maize seeds as a result of genetic modifications. J Proteome Res ***7*****, 1850-61.**

Zubieta, C., Arkus, K. A. J., Cahoon, R. E. and Jez, J. M. (2008). A single amino acid change is responsible for evolution of acyltransferase specificity in bacterial methionine biosynthesis. J Biol Chem ***283*****, 7561-7567.**