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

20 December 2017
[35-17]

Supporting document 1

Safety Assessment Report (at Approval) – Application A1143

Food derived from DHA Canola Line NS-B50027-4

Summary and conclusions

Background

A genetically modified (GM) canola line with OECD Unique Identifier NS-B50027-4 (herein referred to as DHA canola) has been developed by Nuseed Pty Ltd. This canola line has been genetically modified to introduce, into the seed, the pathway for production of the omega-3 long chain polyunsaturated fatty acid (n-3 LC PUFA) docosahexaenoic acid (DHA) from oleic acid (OA). Other n-3 LC PUFAs in the DHA synthesis pathway, particularly eicosapentaenoic acid, (EPA) would also be present.

Coding sequences from seven genes in the DHA pathway have been introduced as follows:
- Δ12 desaturase *(Lackl-Δ12D)* from the yeast *Lachancea kluyveri*
- Δ15-/ ω3 desaturase *(Picpa- ω3D)* from the yeast *Pichia pastoris*
- Δ6 elongase *(Pyrco-Δ6E)* from the marine microalga *Pyramimonas cordata*
- Δ6 desaturase *(Micpu-Δ6D)* from the marine microalga *Micromonas pusilla*
- Δ5 elongase *(Pyrco-Δ5E)* from the marine microalga *Pyramimonas cordata*
- Δ5 desaturase *(Pavsa-Δ5D)* from the marine microalga *Pavlova salina*
- Δ4 desaturase *(Pavsa-Δ4D)* from the marine microalga *Pavlova salina*

In addition, DHA canola also contains the phosphinothricin N-acetyltransferase (*pat*) gene from *Streptomyces viridochromogenes* that confers tolerance to the herbicide phosphinothricin – also known as glufosinate ammonium (glufosinate). The glufosinate tolerance was used for selection of putative transformants during the transformation stage and was not subsequently selected for during the breeding of the final DHA canola line.

This safety assessment report addresses only food safety and nutritional issues of the GM food *per se*. 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

Canola is rapeseed (*Brassica napus, B. rapa* or *B. juncea*) which has been conventionally bred to contain less than 2% erucic acid and less than 30 micromoles of glucosinolates per gram of seed solids, by definition. Rapeseed is the second largest oilseed crop in the world behind soybean, although annual production is around 25% of that of soybean.

Canola seeds are processed into two major products, oil and meal. The oil is the major product for human consumption, being used directly for cooking and as an ingredient in a variety of manufactured food products including salad and cooking oil, margarine, shortening and a range of prepared foods such as mayonnaise, sandwich spreads, creamers and coffee whiteners. Canola oil is the third largest source of vegetable oil in the world after soybean oil and palm oil. Whole canola seeds are being used increasingly in products such as breads. More recently the meal has been identified as a potential alternative source of protein for human consumption.

Molecular characterisation

DHA canola was generated through *Agrobacterium*-mediated transformation with a single T-DNA containing eight expression cassettes. Comprehensive molecular analyses indicate there are two insertion sites on different chromosomes. Specific information about the nature of the insertions was supplied by the Applicant but is confidential commercial information (CCI). One insertion site comprises the coding regions and complete associated regulatory elements of less than eight of the gene cassettes; the sequences all perfectly match the corresponding sequences in the T-DNA used to introduce the genetic material. The insert is located within one native gene, the expression of which was subsequently shown not to have been disrupted. The other insertion site contains more than one copy of various cassettes, again all perfectly matching the corresponding cassettes in the T-DNA, although not necessarily in the same orientation. While expression of the native gene into which the cassettes have inserted has been disrupted, no negative phenotypic effect was documented in the subsequent breeding programme.

The introduced genes are stably inherited across multiple generations and in different genetic backgrounds. No plasmid backbone (and hence, no antibiotic resistant marker) has been incorporated into either of the transgenic loci.

Characterisation and safety assessment of new substances

Newly expressed proteins

DHA canola expresses eight novel proteins, Lackl-Δ12D, Picpa-ω3D, Pyrco-Δ6E, Pyrco Δ5E, Micpu-Δ6D, Pavsa-Δ5D, Pavsa-Δ4D and PAT.

The seven proteins in the n-3 LC-PUFA pathway were only detected in developing and mature seed. This is expected since the genes all have seed-specific promoters. The levels for all seven proteins were low with Pyrco-Δ5E having the lowest and Pavsa-Δ4D the highest expression. Given that the range of total protein in seeds was 5.95 – 18.0 mg/g of seed, the levels of the seven proteins (ranging from 20 – 740 ng/mg total protein) represent a negligible proportion of the total protein. The PAT protein levels were too low to be quantified in any of the plant parts analysed. Thus, dietary exposure to any of the proteins would be very low.

Previous safety assessments of PAT indicate the protein would be rapidly degraded in the digestive system following ingestion and would be inactivated by heating. Additionally, updated bioinformatic studies considered in this assessment confirm the lack of any

significant amino acid sequence similarity to known protein toxins or allergens. The low levels of the seven proteins in the n-3 LC-PUFA pathway that were expressed in the DHA canola precluded their direct characterisation. An indirect approach was therefore used which relied upon: a) translation of the known DNA sequences of the seven introduced genes present in DHA canola; b) the fact the proteins function as predicted in the plant; and c) the functionality of the seven proteins expressed in heterologous systems.

For these seven proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to know protein toxins or allergens. *In vitro* digestibility studies indicated that the proteins would be rapidly degraded in the gastro-intestinal tract following ingestion. Thermolability studies for the seven proteins could not be done but evidence from the literature suggests integral membrane proteins are generally not thermally stable and are functionally inactivated following heating.

Taken together, the evidence indicates that should Lackl-Δ12D, Picpa-ω3D, Pyrco-Δ6E, Pyrco Δ5E, Micpu-Δ6D, Pavsa-Δ5D, Pavsa-Δ4D or PAT be present in the diet they are unlikely to be toxic or allergenic in humans.

Herbicide metabolites

The presence of the PAT protein in DHA canola was intended for use in the early selection stages of product development and is not intended to confer agronomically useful tolerance to glufosinate ammonium. The trait was not selected for during product development and the PAT protein is hardly detectable in plant parts. Nonetheless, should DHA canola be sprayed with glufosinate ammonium and survive the spraying, the metabolic profiles resulting from any novel protein x herbicide interaction have already been established through a significant history of use. The glufosinate-tolerance trait is present in lines from 24 previous applications considered by FSANZ. There are no concerns that the spraying of DHA canola with glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicide and already used in the food supply.

Compositional analyses

Detailed compositional analyses were undertaken on seed from DHA canola. Seed samples harvested from DHA canola grown in eight different locations were analysed for proximates, fibre, fatty acids, amino acids, minerals, vitamins, tocopherols/sterols and anti-nutrients (glucosinolates, phytic acid, and phenolics). The levels of each analyte in DHA canola were compared to levels in: a) the non-GM parental line, AV Jade; b) eight non-GM commercial reference lines grown at the same locations; and c) levels recorded in the literature.

In total, 106 individual analytes and some groupings were analysed. A total of 15 individual analytes (including eight fatty acids) had more than 33% of the observations below the assay limit of quantitation (LOQ) for both DHA canola and the AV Jade control and were excluded from the statistical analysis. For a further nine individual fatty acids, statistical analysis could not be performed because levels were at, or close to, zero in AV Jade but were detectable in DHA canola. The data for 91 individual analytes (including these nine fatty acids) were therefore considered.

Of these 91 individual analytes, a total of 48 (including 14 of the statistically analysed individual fatty acids) differed significantly between DHA canola and AV Jade. The changes in fatty acid profile were consistent with those expected as a result of the introduction of seven enzymes from the n-3 LC-PUFA synthesis pathway leading to the production of DHA from OA.

For the 34 non-fatty acid analytes that showed significant differences, the levels of all but δ_5_avenasterol were within the reference range generated from the non-GM varieties. Comparative information in the literature was not available for many of the analytes but, where it was available, it is noted in five cases where the mean level of the DHA canola analyte fell outside the literature range it also fell outside for AV Jade. This highlights the limitations of the literature in capturing the range of variability found in canola.

DHA canola seeds were found to have a significantly higher level of total *trans* fatty acids than that found in the parental non-GM control and other commercial non-GM canola lines. Although the level was increased, the total *trans* fatty acids level in the DHA canola seed was below 1%. As canola oils are diluted when used, the overall level of *trans* fatty acids consumed would also be further reduced. Furthermore, *trans* fatty acids are present in other refined non-GM vegetable oils, including soybean, sunflower and rice oils, and the *trans* fatty acid content of DHA canola is not expected to vary significantly from these other retail vegetable oils. Finally, a consideration of the consumption data of *trans* fatty acids in the Australian and New Zealand diets (see section 6.1 of SD1) indicates consumption of food derived from DHA canola does not pose a public health concern.

Apart from the intended change to fatty acid profile and a small increase in *trans* fatty acids, seed from DHA canola is otherwise compositionally equivalent to seed from conventional canola varieties.

Conclusion

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

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACP	acyl carrier protein
BLAST	Basic Local Alignment Search Tool
bp	base pairs
Code	*Australia New Zealand Food Standards Code*
ddPCR	droplet digital polymerase chain reaction
DMPT	demethylphosphinothricin
DNA	deoxyribonucleic acid
dw	dry weight
ELISA	enzyme linked immunosorbent assay
FADs	fatty acid desaturases
FAO	Food & Agricultural Organization of the United Nations
FASTA	Fast Alignment Search Tool – All
FRET	fluorescence resonant energy transfer
FSANZ	Food Standards Australia New Zealand
fw	fresh weight
g	gram
GC	gas chromatography
GC-FID	gas chromatography-flame ionisation
GFP	green fluorescent protein
GLA	γ-linolenic acid
GM	genetically modified
His	histidine
IgE	Immunoglobulin E
IS	intervening sequence(s)
KASP	Kompetitive Allele Specific PCR
kDa	kilo Dalton
kg	kilogram
Lackl-Δ12D	Δ12 desaturase from the yeast *Lachancea kluyveri*
LB	Left Border of T-DNA (*Agrobacterium tumefaciens*)
LC/MRM/MS	liquid chromatography/multiple reaction monitoring/mass spectrometry
LOQ	limit of quantification
MS/MS	tandem mass spectrometry
mg	milligram
Micpu-Δ6D	Δ6 desaturase from the marine microalga *Micromonas pusilla*
MRL	maximum residue limit
MT	million tonnes
n-3 LC-PUFAs	omega-3 long chain polyunsaturated fatty acids
NCBI	National Centre for Biotechnology Information
ng	nannogram
NMR	nuclear magnetic resonance
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PAS	polymerase chain reaction amplicon sequencing

PAT	phosphinothricin N-acetyltransferase
Pavsa-Δ4D	Δ4 desaturase from the marine microalga *Pavlova salina*
Pavsa-Δ5D	Δ5 desaturase from the marine microalga *Pavlova salina*
PCR	polymerase chain reaction
Picpa- ω3D	Δ15-/ ω3 desaturase from the yeast *Pichia pastoris*
PPT	phosphinothricin
PUFA	polyunsaturated fatty acid
Pyrco-Δ5E	Δ5 elongase from the marine microalga *Pyramimonas cordata*
Pyrco-Δ6E	Δ6 elongase from the marine microalga *Pyramimonas cordata*
qPCR	quantitative (real time) PCR
RB	Right Border of T-DNA (*Agrobacterium tumefaciens*)
RBD (oil)	Refined, deodorised and bleached
SAS	Statistical Analysis Software
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
TAGs	triacylglycerols
T-DNA	transfer DNA
Ti	tumour inducing
µg	microgram
µm	micrometres
U.S.	United States of America
UTR	untranslated region
WGS	whole genome sequencing

Fatty acid abbreviations

AA	arachidonic acid (20:4)
ALA	α-linolenic acid (18:3 n-3)
DHA	docosahexaenoic acid (22:6)
DPA	docosapentaenoic acid (22:5)
EPA	eicosapentaenoic acid (20:5)
ETA	eicosatetraenoic acid (20:4)
GLA	γ-linolenic acid (18:3 n-6)
LA	linoleic acid (18:2)
OA	oleic acid (18:1)
SDA	stearidonic acid (18:4)

Greek symbols used in the text

α	alpha
γ	gamma
Δ	delta
χ	chi
µ	mu (micro)
ω	omega

1   Introduction

Nuseed Pty Ltd (Nuseed) has submitted an application to FSANZ to vary Schedule 26 in the *Australia New Zealand Food Standards Code* (the Code) to include food from a new genetically modified (GM) canola (*Brassica napus*) line with OECD Unique Identifier NS-B50027-4 (herein referred to as DHA canola). This canola line has been genetically modified to introduce, into the seed, the pathway for production of the omega-3 long chain polyunsaturated fatty acid (n-3 LC PUFA) docosahexaenoic acid (DHA) from oleic acid. Other n-3 LC PUFAs in the DHA synthesis pathway, particularly eicosapentaenoic acid, (EPA) may also be present.

Coding sequences from seven genes in the DHA pathway have been introduced as follows:

- Δ12 desaturase *(Lackl-Δ12D)* from the yeast *Lachancea kluyveri*
- Δ15-/ ω3 desaturase *(Picpa- ω3D*) from the yeast *Pichia pastoris*
- Δ6 elongase *(Pyrco-Δ6E)* from the marine microalga *Pyramimonas cordata*
- Δ6 desaturase *(Micpu-Δ6D)* from the marine microalga *Micromonas pusilla*
- Δ5 elongase *(Pyrco-Δ5E)* from the marine microalga *Pyramimonas cordata*
- Δ5 desaturase *(Pavsa-Δ5D)* from the marine microalga *Pavlova salina*
- Δ4 desaturase *(Pavsa-Δ4D)* from the marine microalga *Pavlova salina*

In addition, DHA canola also contains the phosphinothricin N-acetyltransferase (*pat*) gene from *Streptomyces viridochromogenes* that confers tolerance to the herbicide phosphinothricin – also known as glufosinate ammonium (glufosinate). The glufosinate tolerance was used for selection of putative transformants during the transformation stage and was not subsequently selected for during the breeding of the final DHA canola line. The PAT protein has been assessed by FSANZ in 24 previous FSANZ applications, and globally is represented in six major crop species and over 30 approved GM single plant events (CERA 2011).

The Applicant states the purpose of DHA canola is to provide a sustainable and land-based source of n-3 LC PUFAs, particularly EPA and DHA, to meet increased human consumption and demand from aquaculture.

It is the Applicant's intention to commercially cultivate DHA canola initially in Australia. An application for commercial release has been submitted to the Office of the Gene Technology Regulator (OGTR) (DIR 155[1]). DHA canola has been grown under limited and controlled conditions in Australia under OGTR Licence DIR123[2]. It is also the Applicant's intention to pursue commercial plantings in other canola-growing countries such as Canada and the U.S. It is therefore anticipated food products derived from DHA canola will enter the Australian and New Zealand food supplies via both local production and imports from major canola-producing countries.

---

[1] http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/dir155
[2] DIR 123 documents - http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/DIR123

2   History of use

2.1   Host organism

Canola (a trade name purported to be derived from 'Canadian oil low acid') is the name used for rapeseed (*Brassica napus, Brassica rapa* or *Brassica juncea*) crops that have less than 2% erucic acid (a fatty acid)[3] and less than 30 micromoles of glucosinolates per gram of seed solids (OECD 2001). Canola varieties were first developed in Canada in the 1950s, using traditional breeding techniques, in response to a demand for food-grade rapeseed products and animal feed with improved palatability. Rapeseed-derived products that do not meet the compositional standard cannot use the trademarked term, 'canola'. In some countries, the term canola is not used and the low erucic acid crop is known generically as rapeseed.

The Brassicaceae family, to which canola belongs, contains a number of vegetable species including broccoli, cabbage, turnip, and radish. The family is also generically referred to as the Mustard family and seeds of a number of species within the family, particularly *B. nigra* (black mustard), *B. juncea* (brown Indian mustard) and *B. hirta* (white mustard) are ground to produce the spice known as mustard. Mustard allergy (from ingestion of a variety of plant parts) can be a serious problem (see e.g. Health Canada 2016).

Canola is one of the most important oilseed crops in the world with its oil being a source of the LC monounsaturated oleic acid (OA), omega-6 (n-6) LC PUFA linoleic acid (LA), and omega-3 (n-3) LC PUFA α-linolenic acid (ALA) – see Section 3.2.9. It does not, however, produce other important LC-PUFAs such as EPA and DHA.

Rapeseed is the second largest oilseed crop in the world behind soybean. In 2014, world production was 73.8 MT and the major oilseed-rape-producing countries were Canada (15.6 MT), China (14.8 MT) and India (7.8 MT); Australia ranked 6th at 3.8 MT (FAOSTAT3 2015). In the case of China and India, a significant amount of non-canola quality rapeseed, is included in the term 'rapeseed'. All of Australia's production is canola. New Zealand canola production was minor at approximately 3,000 T. In 2013, Canada was the largest exporter of canola seed (6.9 MT), while Australia was the second largest exporter at 3.8 MT. Australia's major export destinations in 2015/16 were Europe, Japan and Bangladesh (AOF 2016) and represented around 15% of the world's canola export trade.

In Australia, canola is the third largest broad acre crop behind wheat and barley and the growing area extends from south-western Western Australia to south-eastern Australia and northern New South Wales. GM canola now accounts for approximately 20% of the total Australian crop.

Canola seeds are processed into two major products, oil and meal. Very briefly, the processes involved in the production of oil and meal (CCC 2012) involve seed cleaning, seed pre-conditioning and flaking, seed cooking, pressing the flake to mechanically remove a portion of the oil, solvent extraction of the press-cake to remove the remainder of the oil, and desolventising and toasting of the meal. The oil is the major product for human consumption, being used in a variety of manufactured food products including salad and cooking oil, margarine, shortening and a range of prepared foods such as mayonnaise, sandwich spreads, creamers and coffee whiteners. The meal provides a good protein source in stock feed for a variety of animals, primarily pigs, poultry and dairy cattle (Bonnardeaux 2007). More recently, it has been identified as a potential alternative source of protein isolate for

---

[3] Codex Standard for Named Vegetable Oils (CODEX-STAN 210-1999) - http://www.fao.org/docrep/004/y2774e/y2774e04.htm

aquaculture[4] and human consumption (Bos et al. 2007; Campbell et al. 2016); the incorporation of whole meal into food for humans is restricted by anti-nutritional factors such as sinapine, glucosinolates, phytic acid and tannins (Yoshie-Stark et al. 2008) although these can be reduced by appropriate extraction (Tan et al. 2011a) Whole canola seeds are being used increasingly in products such as breads.

Another possible food product that can be derived from the canola plant is bee pollen (Bogdanov 2016).

The canola variety used as the host for the transformation generating DHA canola was AV Jade. This variety was developed in Australia by the Department of Primary Industries, Victoria and the Grains Research and Development Corporation as part of the National Brassica Improvement Program (Gororo and Burton 2006). It is an early to mid-maturing, high-yielding, open-pollinated, conventional variety suited to areas receiving an annual rainfall between 350 – 500 mm (Dovuro Seeds 2008). It also shows good transformation efficiency.

2.2   Donor organisms

In addition to specific information provided in the sections below, the Applicant conducted a search of the PubMed[5] database, maintained by the U.S. National Library of Medicine, for scientific literature on potential allergy or toxicity concerns with the six donor organisms. The species names were used in searches without and with keywords (allergen, allergy, toxicity and toxin).

Unpublished study submitted
2016. Bioinformatics analysis of the potential allergenicity and toxicity of proteins encoded by genes inserted in canola (*Brassica napus*) for production of omega 3 fatty acids. Internal Nuseed report. Nuseed Pty Ltd. Report No. 2016-017

No publications were identified that showed the source organisms are associated with human allergenicity or toxicity.

2.2.1   Microalgae

Eukaryotic algae in general account for half the primary productivity at the base of the food chain and marine microalgae in particular contribute to natural marine productivity as well as being essential components of aquaculture feed and being cultivated for human nutritional products (Guschina and Harwood 2006; Barakoni et al. 2015). Specifically, many contain n-3 LC PUFAs which make them commercially important both as dietary and industrial sources of the fatty acids and as a source of fatty acid synthesis genes that can be used in the genetic modification of other organisms, particularly plants (Zhou et al. 2007; Harwood and Guschina 2009; Greenwell et al. 2010; Petrie and Singh 2011)

2.2.1.1   *Micromonas pusilla*

Micromonas pusilla, the source of the *Micpu-Δ6D* gene (Petrie et al. 2010a) is a pear-shaped, unicellular alga which bears one flagellum and contains one mitochondrion and one chloroplast. It is a significant primary producer in the marine ecosystem. *Micromonas* belongs to the ancient lineage giving rise to higher plants.

---

[4] https://grdc.com.au/Media-Centre/Ground-Cover/Ground-Cover-Issue-108-Jan-Feb-2014/Canola-role-for-aquaculture
[5] PubMed – http://www.ncbi.nlm.nih.gov/pubmed

2.2.1.2 *Pyramimonas cordata*

The *Pyrco-Δ5E* and *Pyrco-Δ6E* genes are sourced from the marine microalga *Pyramimonas cordata* (Petrie et al. 2010b) which is also a unicellular green flagellate alga with ubiquitous marine distribution.

2.2.1.3 *Pavlova salina (now [Rebecca salina](#)[6])*

The *Pavsa-Δ5D* and *Pavsa-Δ4D* genes are sourced from *Pavlova salina*, a golden-brown flagellate (Zhou et al. 2007). This marine unicellular microalga produces lipids containing approximately 50% n-3 LC PUFAs and is regarded as a useful mariculture feedstock, particularly in tropical hatcheries (Volkman et al. 1991).

2.2.2 Yeasts

Yeasts are single-celled organisms of the Fungi kingdom that have been used in food production for many years, particularly in baking and making alcoholic beverages (Fleet 2006). Yeast extract, a by-product of the brewing industry, provides spreads such as Marmite, Vegemite and Promite. Several species of yeast, and particularly *Saccharomyces* spp., have been genetically modified to efficiently produce food processing proteins.

2.2.2.1 *Pichia pastoris (now [Komagataella pastoris](#)[7])*

Pichia pastoris is the source of the Picpa- ω3D gene (Zhang et al. 2008). It is a methylotroph (an organism that can grow using methanol as its only energy source) and is used widely in research and industry as a heterologous production organism for proteins using recombinant DNA techniques (Ahmad et al. 2014)

2.2.2.2 *Lachancea kluyveri*

This budding yeast, the source of the Lackl-Δ12D gene (Petrie et al. 2012), is a PUFA-producing yeast located phylogenetically between the Saccharomyces and Kluyveromyces lineages and was previously known as *Saccharomyces kluyveri*. It is thought to be widespread in the environment and, for example, has been found on cheese (Tzanetakis et al. 1998; Wojtatowicz et al. 2001) and has been reported to be involved in solubilisation of pectin such as occurs during the softening of olives and cherries in brine (Considine and Considine 1982). A closely related species *Kluyveromyces lactis* is used widely for the production of metabolites and heterologous proteins that are used, for example in infant nutrition products and the fermented drink kefir (Spohner et al. 2016).

2.2.3 Other organisms

2.2.3.1 *Streptomyces viridochromogenes*

The source of the *pat* gene is the bacterial species *Streptomyces viridochromogenes*. The *Streptomycetae* bacteria were first described in the early 1900's. These organisms are generally soil-borne, although they may also be isolated from water. They are not typically pathogenic to animals including humans, and few species have been shown to be phytopathogenic (Kützner 1981; Bradbury 1986).

Although this organism is not used in the food industry directly, the *pat* gene from *S. viridochromogenes* has been used to confer glufosinate ammonium-tolerance in food

---

[6] Algaebase - http://www.algaebase.org/search/species/detail/?species_id=66781
[7] EnsemblFungi - http://fungi.ensembl.org/Komagataella_pastoris/Info/Index

producing crops over the past decade. The *bar* gene from the closely related species
S. hygroscopicus produces a protein that is structurally and functionally equivalent to the
protein encoded by the *pat* gene (Wehrmann et al. 1996).

2.2.3.2   *Regulatory elements from other organisms*

Genetic elements from seven other organisms not mentioned above (plant species - *Linum
usitatissimum* (flax), *Arabidopsis thaliana, Glycine max* (soybean) and *Nicotiana tabacum*
(tobacco); viruses - Tobacco mosaic virus and Cauliflower mosaic virus; bacterium –
Agrobacterium tumefaciens) have been used in the genetic modification of DHA canola (refer
to Table 1). These sequences are used to drive, terminate or enhance expression of the
novel genetic material. None of the sources of these genetic elements is associated with
toxic or allergenic responses in humans. The genetic elements derived from the plant viral
and bacterial pathogens are not pathogenic per se and do not cause pathogenic symptoms
in DHA canola.

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
- the genetic stability of the inserted DNA and any accompanying expressed traits.

Unpublished Studies submitted:

2016. Molecular Characterization of Genetically Modified Canola NSB50027-4 Producing High
    Percentage of Long-Chain Omega-3 (LC-ω3) Fatty Acids in Seed. Internal Nuseed/CSIRO
    report, Nuseed Pty Ltd. Report N° 2016-002.
2016. Supplemental File 1, Report N° 2016-002
2016. Supplemental File 2, Report N° 2016-002 (all CCI)
2016. Supplemental File 3, Report N° 2016-002
2016. Supplemental File 4, Report N° 2016-002 (all CCI)
2016. Supplemental File 5, Report N° 2016-002 (all CCI)
2016. Supplemental File 6, Report N° 2016-002 (all CCI)
2017. Bioinformatics Analysis of the Potential Fusion Proteins at DNA Junctions in Canola (*Brassica
    napus*) for Omega 3 Fatty Acids: Identity Comparison to Allergens and Toxins. Internal Nuseed
    report, Nuseed Pty Ltd. Report N° 2016-004
2017. Bioinformatics analysis of hypothetical open reading frame proteins throughout the inserts in
    canola (Brassica napus) for omega-3 fatty acids: identity comparison to allergens and toxins.
    Internal Nuseed report, Nuseed Pty Ltd. Report N° 2017-026
2016. Inheritance of the omega-3 trait – DHA canola (OECD ID NS-B50027-4). Internal Nuseed
    report, Nuseed Pty Ltd. Report N° 2016-019

3.1   Method used in the genetic modification

The method was based on that from two publications (Bhalla and Singh 2008; Belide et al.
2013). Cotyledonary petioles from seedlings of cultivar AV Jade were co-cultured with
disarmed *Agrobacterium tumefaciens* strain AGL1 (Lazo et al. 1991) harbouring a binary
vector system (Deblaere et al. 1987). The binary vector, plasmid pJP3416_GA7-ModB
(Figure 1), contained the eight genes of interest between right border (RB) and left border
(LB) T-DNA sequences.

Following co-culture, the petioles were transferred to a callus initiation medium containing cefotaxime and timentin (to inhibit the growth of excess *Agrobacterium)*, and phosphinothricin (for selection of putative transformants). Surviving green calli were transferred to a shoot initiation medium and resulting shoots were isolated and transferred to shoot elongation medium and later rooting medium. Leaves from these shoots were used for testing of T-DNA presence by polymerase chain reaction (PCR) amplification of the seven fatty acid synthesis genes and *pat* selection marker. Positive plantlets (generation T_0) were transferred to the glasshouse. DHA canola (derived from the T_0 generation event B0050 – see Figure 4) was ultimately chosen as the lead event based on superior agronomic, biochemical, genetic and molecular characteristics. Specifically, this selection and advancement was based on:

- copy number and integrity of the T-DNA insert
- segregation and homozygosity of transgenic insert monitored by digital droplet PCR (ddPCR) and zygosity testing
- segregation pattern and production of DHA or other n-3 LC PUFAs in the seed measured by gas chromatography (GC)-flame ionisation detector (FID) (GC-FID) and nuclear magnetic resonance (NMR) spectrometry
- phosphinothricin resistance (in the early stages of selection)
- genetic stability of the event in different generations and genetic backgrounds
- expression of ω3 genes monitored by qPCR, ddPCR and ELISA
- suitable agronomic traits (e.g. grain yield, oil content, blackleg resistance, plant emergence and vigour, flowering time and duration, maturity, plant survival, height and lodging at harvest, seed shattering) for crop production through field testing at different locations over winter and summer plantings.

Figure 1: Genes and regulatory elements contained in plasmid pJP3416_GA7-ModB

3.2   Function and regulation of introduced gene fragments

Information on the genetic elements in the T-DNA used for transformation is summarised in Table 1. The complete plasmid is 31,564 bp comprising 8,052 bp vector backbone and 23,512 bp T-DNA (Figure 1). The T-DNA comprises eight cassettes located between a 163 bp RB and 161 bp LB. Intervening sequences, where present, have assisted with the cloning of the various components of each cassette.

Table 1: Description of the genetic elements contained in the T-DNA of pJP3416_GA7-ModB

IS = intervening sequence

Genetic element	Relative bp location on plasmid	Size (bp)	Source	Orient.	Description &Function	Reference
Right Border	3398 - 3560	163				
Multiple cloning site	3561 - 3628	68				
Micpu-Δ6D cassette						
Linus-Cn12	3629 - 4166	538	*Linum usitatissimum* (flax)	anti-clockwise	• Terminator region of the conlinin2 gene • Directs polyadenylation of the *Micpu-Δ6D* gene	Chaudhary et al. (2001)
IS	4167 - 4174	8				
Micpu-Δ6D	4175 - 5569	1395	*Micromonas pusilla*	anti-clockwise	• Coding sequence of the *Δ6 desaturase* gene • NCBI accession XM_003056946	Petrie et al. (2010a)
IS	5570 - 5572	3				
5′ UTR leader	5573 - 5637	65	Tobacco mosaic virus (TMV)	anti-clockwise	• Leader sequence • Enhances expression of the Micpu-Δ6D protein.	Gallie et al. (1987)
IS	5638 - 5645	8			•	
Linus-Cn12	5646 - 7678	2033	*Linum usitatissimum* (flax)	anti-clockwise	• Promoter region of the conlinin2 gene • Directs transcription of the *Micpu-Δ6D* gene	Chaudhary et al. (2001)
IS	7678 - 7684	7				
Pyrco-Δ5E cassette						
Arath-FAE1	7685 - 8618	934	*Arabidopsis thaliana*	clockwise	• Promoter region from the fatty acid *elongase* gene • Directs transcription of the *Pyrco-Δ5E* gene	Rossak et al. (2001)
5′ UTR leader	8619 - 8683	65	Tobacco mosaic virus (TMV)	clockwise	• Leader sequence • Enhances expression of the Pyrco-Δ5E protein	Gallie et al. (1987)
IS	8684 - 8686	3				
Pyrco-Δ5E	8687 - 9493	807	*Pyramimonas cordata*	clockwise	• Coding sequence of the *Δ5 elongase* gene • NCBI accession GQ202035	Petrie et al. (2010b)
IS	9494 - 9508	15				
Glyma-lectin	9509 - 9842	334	*Glycine max* (soybean)	clockwise	• Terminator region of the *lectin* gene • Directs polyadenylation of the *Pyrco-Δ5E* gene	Vodkin et al. (1983); Cho et al. (1995)
Pavsa-Δ5D cassette						
Brana-FPI	9843 - 10200	358	*Brassica napus* (oilseed rape)	clockwise	• Promoter region of the napin gene • Directs transcription of the *Pavsa-Δ5D* gene	Stålberg et al. (1993)

Genetic element	Relative bp location on plasmid	Size (bp)	Source	Orient.	Description &Function	Reference
5′ UTR leader	10201 - 10265	65	Tobacco mosaic virus (TMV)	clockwise	• Leader sequence • Enhances expression of the Pavsa-Δ5D protein	Gallie et al. (1987)
IS	10266 - 10268	3				
Pavsa-Δ5D	10269 - 11549	1281	*Pavlova salina*	clockwise	• Coding sequence of the *Pavsa-Δ5D* gene	Zhou et al. (2007)
Agrtu-NOS	11550 - 11804	255	*Agrobacterium tumefaciens*	clockwise	• Terminator region of the nopaline synthase gene • Directs polyadenylation of the *Pavsa-Δ5D* gene	Bevan et al. (1983); Dhaese et al. (1983)
Nicta-RB7	11805 - 12972	1168	*Nicotiana tabacum* (tobacco)	clockwise	• AT-rich matrix attachment region of the root-specific *Rb7* gene • Increases expression of the *Pavsa-Δ5D* gene	Hall et al. (1991); Halweg et al. (2005)
Picpa- ω3D cassette						
Linus-Cn11	12973 - 13706	734	*Linum usitatissimum* (flax)	anti-clockwise	• Terminator region of the conlinin1 gene • Directs polyadenylation of the *Picpa- ω3D* gene	Chaudhary et al. (2001)
Picpa- ω3D	13707 - 14957	1251	*Pichia pastoris*	anti-clockwise	• Coding sequence of the Δ15-/ ω3 desaturase gene • NCBI Accession EF116884	Zhang et al. (2008)
IS	14958 - 14960	3				
5′ UTR leader	14961 - 15025	65	Tobacco mosaic virus (TMV)	anti-clockwise	• Leader sequence • Enhances expression of the Picpa- ω3D protein	Gallie et al. (1987)
Linus-Cn11	15026 - 15475	450	*Linum usitatissimum* (flax)	anti-clockwise	• Promoter region of the conlinin1 gene • Directs transcription of the *Picpa- ω3D* gene	Chaudhary et al. (2001)
Pavsa-Δ4D cassette						
Linus-Cn12	1476 - 17508	2033	*Linum usitatissimum* (flax)	clockwise	• Promoter region of the conlinin2 gene • Directs transcription of the *Pavsa-Δ4D* gene	Chaudhary et al. (2001)
5′ UTR leader	17509 - 17573	65	Tobacco mosaic virus (TMV)	clockwise	• Leader sequence • Enhances expression of the Pavsa-Δ4D protein	Gallie et al. (1987)
IS	17574 - 17576	3				
Pavsa-Δ4D	17577 - 18923	1347	*Pavlova salina*	clockwise	• Coding sequence of the Δ4 desaturase gene • NCBI accession DQ995517	Zhou et al. (2007)
Linus-Cn12	18924 - 19461	538	*Linum usitatissimum* (flax)	clockwise	• Terminator region of the conlinin2 gene • Directs polyadenylation of the *Pavsa-Δ4D* gene	Chaudhary et al. (2001)
Lackl-Δ12D cassette						
Linus-Cn11	19462 - 19911	450	*Linum usitatissimum* (flax)	clockwise	• Promoter region of the conlinin1 gene • Directs transcription of the *Lackl-Δ12D* gene	Chaudhary et al. (2001)
5′ UTR leader	19912 - 19976	65	Tobacco mosaic virus (TMV)	clockwise	• Leader sequence • Enhances expression of the Lackl-Δ12D protein	Gallie et al. (1987)
IS	19977 - 19979	3				
Lackl-Δ12D	19980 - 21233	1254	*Lachancea kluyveri*	clockwise	• Coding sequence of the Δ12 desaturase gene • NBCI accession AB115968	Petrie et al. (2012); Watanabe et al. (2014)
Linus-Cn11	21234 - 21967	734	*Linum usitatissimum* (flax)	clockwise	• Terminator region of the conlinin1 gene	Chaudhary et al. (2001)

Genetic element	Relative bp location on plasmid	Size (bp)	Source	Orient.	Description &Function	Reference
					• Directs polyadenylation of the *Lackl-Δ12D* gene	
Nicta-RB7	21968 - 23135	1168	*Nicotiana tabacum* (tobacco)	clockwise	• AT-rich matrix attachment region of the root-specific *Rb7* gene • Increases expression of the *Lackl-Δ12D* protein	Hall et al. (1991); Halweg et al. (2005)
IS	23136 - 23143	8				
Pyrco-Δ6E cassette						
Arath-FAE1	23144 - 24077	934		clockwise	• Promoter region from the fatty acid elongase gene • Directs transcription of the *Pyrco-Δ6E* gene	Rossak et al. (2001)
5′ UTR leader	24078 - 24142	65	Tobacco mosaic virus (TMV)	clockwise	• Leader sequence • Enhances expression of the Pyrco-Δ6E protein	Gallie et al. (1987)
IS	24143 - 24145	3				
Pyrco-Δ6E	24146 - 25015	870	*Pyramimonas cordata*	clockwise	• Coding sequence of the *Δ6 elongase* gene • NCBI accession GQ202034	Petrie et al. (2010b)
Glyma-lectin	25016 - 25349	334	*Glycine max* (soybean)	clockwise	• Terminator region of the lectin gene • Directs polyadenylation of the *Pyrco-Δ6E* gene	Vodkin et al. (1983); Cho et al. (1995)
IS	25350 - 25371	22				
pat cassette						
35S	25372 - 25909	538	Cauliflower Mosaic virus (CMV)	clockwise	• Constitutive promoter of the 35S RNA • Directs transcription of the *pat* gene	Kay et al. (1987); Coutu et al. (2007)
IS	25910 - 25918	9				
pat	25919 - 26470	552	*Streptomyces viridochromogenes*	clockwise	• Coding sequence of the *phosphinothricin acetyltransferase* gene	Dröge et al. (1992)
IS	26471 - 26478	8			•	
Agrtu-NOS	26479 - 26731	253	*Agrobacterium tumefaciens*	clockwise	• Terminator region of the *nopaline synthase* gene • Directs polyadenylation of the *pat* gene	Bevan et al. (1983); Dhaese et al. (1983)
IS	26732 - 26748	17				
Left Border	26749 - 26909	161				

All fatty acid enzyme genes were synthesised at Gene Art[8], according to a sequence in the National Centre for Biotechnology Information (NCBI[9]) database (accession number indicated in Table 1).

The transcription of the fatty acid enzyme genes is initiated by seed-specific promoters that therefore confine the synthesis pathway to the seed. The 5′ UTR leader sequence from TMV is present in all of the fatty acid cassettes and functions to enhance the translation of the mRNA transcribed from each of the genes thereby enhancing expression of each protein. The presence, as a spacer between two of the genes (*Pavsa-Δ5D* and *Lackl-Δ12D)*, of the

---

[8] GeneArt – formerly Life Science Technologies, now Thermo Fisher Scientific - https://www.thermofisher.com/au/en/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis.html
[9] NCBI - https://www.ncbi.nlm.nih.gov/

matrix attachment region from the Rb7 gene of tobacco (Hall Jr et al. 1991; Halweg et al. 2005) further enhances expression and stability, and thereby assists in maximising production of substrates in the pathway.

The different orientations (clockwise/anticlockwise) of the genetic elements assist in improving transcription and increasing expression by including some spacing in protein transcription.

3.2.1   *Micpu-Δ6D* cassette

Delta-6 desaturases catalyse the addition of a double bond at the sixth carbon-carbon bond position from the carboxylic acid end (α end – see Figure 3) in fatty acids. The *Micpu-Δ6D* gene is required for the conversion of the n-3 LC PUFA, α-linolenic acid (ALA), to the n-3 LC PUFA, stearidonic acid (SDA) – see Figure 3. The gene is regulated by the seed-specific promoter and terminator from the conlinin2 (legumin-like seed storage protein) gene of flax.

3.2.2   *Pyrco-Δ5E* cassette

The delta-5 elongase from *P. cordata* (Petrie et al. 2010b) shows high native efficiency in the conversion of EPA to docosapentaenoic acid (DPA) (Figure 3) and hence the *Pyrco-Δ5E* gene was chosen for use in DHA canola. It is driven by the embryo-specific fatty acid elongase gene promoter from Arabidopsis that shows maximal activity during the period of major storage lipid accumulation (Rossak et al. 2001). *Pyrco-Δ5E* is terminated by the lectin gene from soybean.

3.2.3   *Pavsa-Δ5D* cassette

Delta-5 desaturases add a double bond at the 5th carbon-carbon bond from the α end in fatty acids. The Δ5D from *P. salina* shows greater than 95% efficiency in desaturating its substrate (Zhou et al. 2007). In DHA canola the double bond is added to eicosatetraenoic acid (ETA) to give EPA (Figure 3). The *Pavsa-Δ5D* gene is driven by the seed-specific promoter from the napin gene of *B. napus*. Polyadenylation is regulated by the nos terminator from *A. tumefaciens.*

3.2.4   *Picpa- ω3D* cassette

Omega-3 desaturases insert a double bond between the third and fourth carbon from the methyl end (ω end) of a fatty acid. In DHA canola, this double bond is added to linoleic acid (LA) to give ALA (Figure 3). The *Picpa- ω3D* gene is driven by the promoter from the seed-specific conlinin1 gene of flax. The terminator from the same flax gene directs polyadenylation.

3.2.5   *Pavsa-Δ4D* cassette

The Pavsa-Δ4 desaturase catalyses the addition of a double bond at the fourth carbon-carbon bond from the α end of DPA to give DHA. As for the *Micpu-Δ6D* cassette, the promoter and terminator for the *Pavsa-Δ4D* gene are both from the conlinin2 gene of flax.

3.2.6   *Lackl-Δ12D* cassette

The seed-specific Lackl-Δ12D inserts a double bond at the delta-12 (omega-6) position of OA to give LA (Figure 3). As for the *Picpa- ω3D* cassette, the promoter and terminator for the *Lackl-Δ12D* gene are both from the conlinin1 gene of flax.

3.2.7 *Pyrco-Δ6E* cassette

The *Pyrco-Δ6E* gene from *P. cordata* codes for a highly efficient enzyme that converts SDA to ETA (Petrie et al. 2010b). The gene is regulated by the same fatty acid elongase gene promoter from Arabidopsis and soybean lectin gene terminator as used to regulate the *Pyrco-Δ5E* gene.

3.2.8 *pat* cassette

The *pat* gene from *S. viridochromogenes* and the *bar* gene from *S. hygroscopicus* both confer tolerance to herbicides containing glufosinate ammonium. Both genes code for polypeptides of 183 amino acids and share 87% homology at the sequence level (Wehrmann et al. 1996). Both genes have been widely used for genetic modification of food species.

Both the *bar* and *pat* genes, like other bacterial; genes, have relatively high G:C content when compared to plant genes, and as a consequence the native microbial genes are inefficiently expressed in plants. In order to improve this expression, the codon usage pattern is modified but the resultant protein sequence is not altered (OECD 1999)

The *pat* gene is driven by the constitutive CaMV 35S promoter. Transcription is terminated by the polyadenylation signal from the *nos* gene of *A. tumefaciens*.

3.2.9 Function of the introduced fatty acid synthesis genes

Plant oils are primarily composed of triacylglycerols (TAGs) that in turn comprise three fatty acid chains usually 16 or 18 carbons long (Durrett et al. 2008). The synthesis of these fatty acids in plants occurs in plastids and essentially results in the formation of 16:0 (palmitic), 18:0 (stearic) and 18:1 (oleic) fatty acids. These fatty acids are esterified to acyl carrier protein (ACP). Thioesterases then release them from ACP so that they may be exported to the endoplasmic reticulum for desaturation (addition of double bonds) by fatty acid desaturases (FADs) and assembly into TAGs. The most abundant plant fatty acids are LA and ALA, produced by desaturation of OA.

LC-PUFAs are fatty acids with 18–20 carbons or more, and more than one double bond, which can be categorised into two main families — n-6 and n-3 — depending on the position of the first double bond from the methyl ('omega' or 'ω') end group of the fatty acid. Common examples include:

- n-3: ALA, SDA, EPA and DHA
- n-6: LA, γ-linolenic acid (GLA) and arachidonic acid (AA),

While plant fatty acid synthesis generally ends at the saturated fatty acids C16:0 or C18:0 and the monounsaturated C18:1, some plants do produce LC-PUFAs but not at levels that are commercially viable (Abedi and Sahari 2014). In order to facilitate a LC-PUFA biosynthesis pathway, a series of aerobic desaturations and elongations are required. The enzymes needed to do this are introduced via genetic modification (Vanegas-Calerón et al. 2010) and can result in parallel pathways that yield both *n*-6 and *n*-3 LC-PUFAs (Petrie et al. 2010a) – see Figure 2. Elongases and desaturases play critical roles in regulating the length and degree of unsaturation of fatty acids. Desaturases essentially remove two hydrogen atoms from a fatty acid, thereby creating a carbon double bond, a process known as desaturation. Elongases facilitate the addition of two carbon atoms via a sequential 4-stage process.
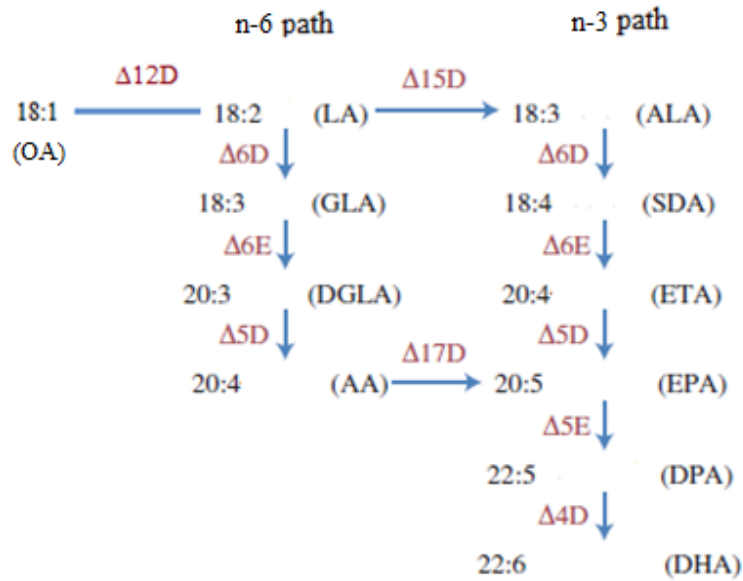
Figure 2: The parallel n-6 and n-3 pathways for LC-PUFA biosynthesis
Diagram modified from Petrie et al. (Petrie et al. 2010a)

Metabolic engineering of a pathway through to EPA has been reported by a number of groups but the conversion of this fatty acid to DHA has been problematic because of the undesirable n-6 co-production, the ability to achieve a continuous flux of substrates through the entire pathway and the inefficiency of the critical Δ-5 elongase-catalysed conversion of EPA to DPA (Petrie et al. 2012)

In DHA canola, the genetic modification results in the favouring of the *n*-3 pathway (Figure 3) from the outset. This is because the Δ15D encoded by the *Picpa- ω3D* gene converts the *n*-6 substrate (LA) to the corresponding *n*-3 fatty acid (ALA), thereby increasing the flow of the *n*-6 intermediate into its *n*-3 derivative (Zhang et al. 2008), and the Δ6D encoded by the *Micpu-Δ6D* gene has a stronger preference for *n*-3 substrate over *n*-6 substrate (Petrie et al. 2010a). Expression of fatty acid intermediates is maximised by the use of genes with high native expression, and enhancers (TMV 5′ UTR and tobacco Rb7) to ensure the desired flux of substrates.
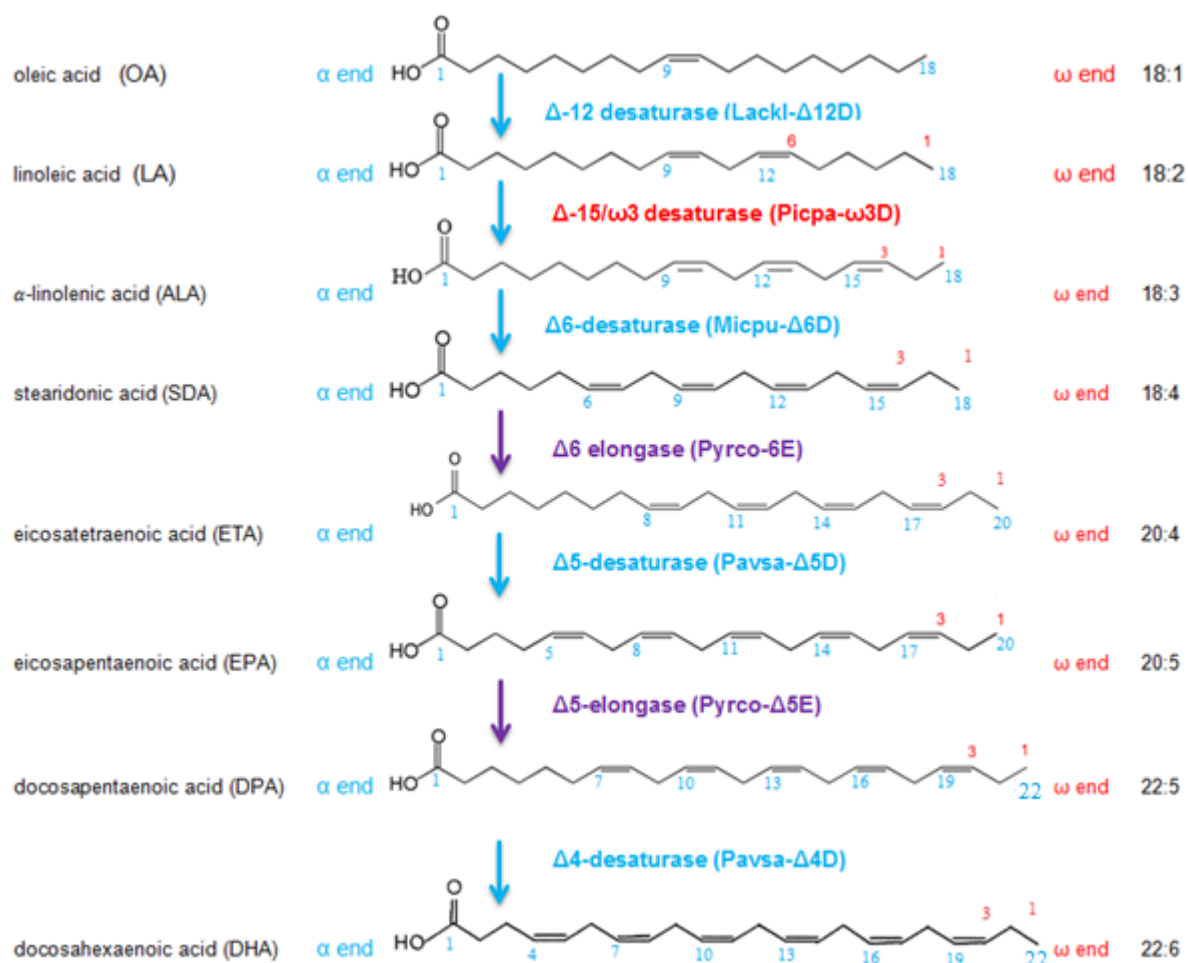
Figure 3: DHA biosynthesis pathway engineered into DHA canola

3.3 Breeding of DHA canola

T_1 seeds from the T_0 generation were advanced to the T_7 generation through self fertilisation and single seed descent. Two further breeding strategies (Figure 4) were used for the introgression of the DHA canola event into other elite canola lines:

- A single T_3 hemizygous DHA canola line – B0050-027-18-20 - was cross-pollinated with six non-genetically modified (non-GM) elite lines[10] to produce six F_1 generation populations that were then backcrossed to the non-GM (recurrent) parent to produce six BC_1F_1 populations. Those populations that were heterozygous for the T-DNA inserts were then selfed to yield BC_1F_2 and BC_1F_3 populations.
- A single T_5 hemizygous DHA canola line – B0050-027-18-20-12-19 - was cross-pollinated with six non-GM elite lines[11] to produce six F_1 generation populations each of which was then selfed to produce six F_2 populations.

---

[10] 04CC-024*5026W-41-12, CC05004*507L-03-4-X-12, CC06026-13-11, NC0013-X-X-10, ND0004B6, NX0052-10

[11] ATR Wahoo, ATR Stingray, ATR Bonito, Jackpot TT, NT0252 and NT0272
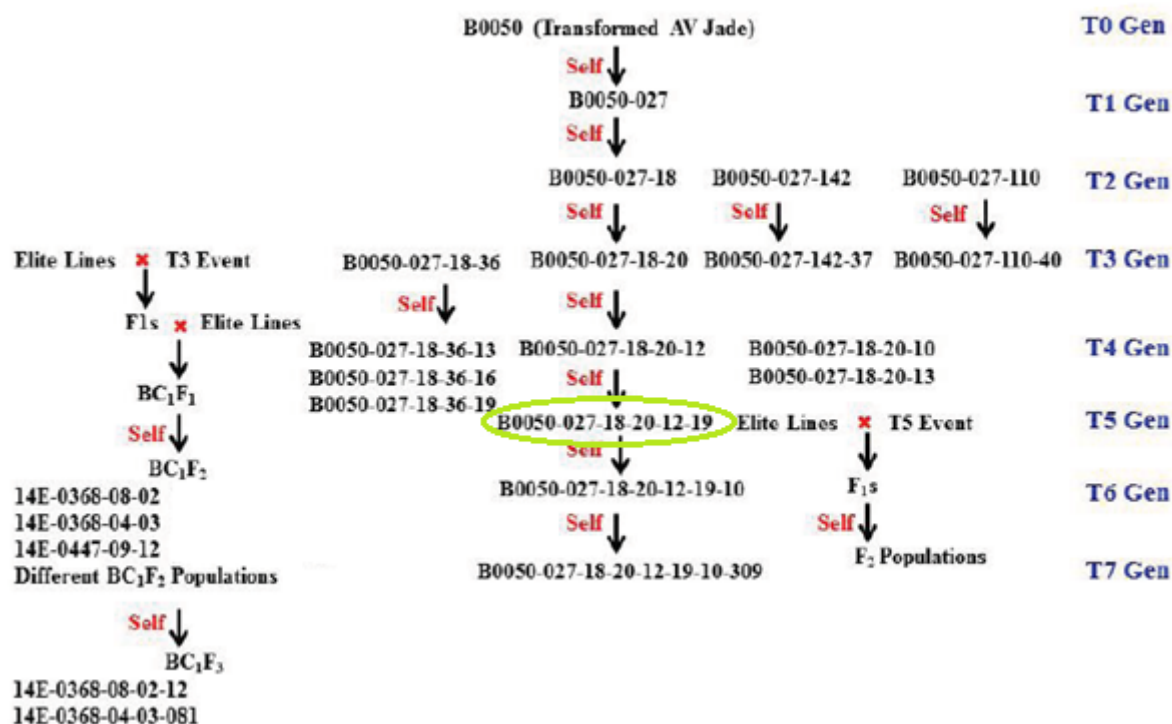
Figure 4: Breeding tree for DHA canola
The circled line in T_5 Gen was the line used for sequencing and hence became the event specific control in subsequent analyses.

The generations and controls used for various analyses described in this report are given in Table 2.

Table 2: DHA canola generations used for various analyses

Analysis	DHA canola generation(s) used	Control(s) used	Reference material
Molecular characterisation (Section 3.4.)	T_3, T_4, T5, BC_1F_2 and BC_1F_3		*B. napus*, *B. rapa*, *B. oleracea* reference genomes
Genetic stability (Section 3.5.1)	T_3, T_4, T_5, T_6, T_7, BC_1F_2, BC_1F_3, F_2	• AV Jade • Dwarf Essex • B0050-027-18-20-12-19	N/A
Mendelian inheritance (Section 3.5.2)	BC_1F_2, F_2	N/A	N/A
Phenotypic stability (Section 3.5.2)	T_1, T_2, T_3, T_4, T_5, T_6, T_7	N/A	
Protein expression levels in plant parts (Section 4.1.3)	Plants grown from T_3 (B0050-027-18-20) seed	AV Jade	
Compositional analyses (Section 5)	Plants grown from T_3 (B0050-027-18-20) seed	AV Jade	AV Zircon, AV Garnet, ATR Bonito, ATR Gem, ATR Wahoo, ATR Stingray, Monola 515TT

23

3.4   Characterisation of the genetic modification in the plant

A range of analyses were undertaken to characterise the genetic modification in DHA canola. These analyses focussed on the nature of the insertion of the introduced genetic elements and whether any unintended re-arrangements may have occurred as a consequence of the transformation procedure.

3.4.1   Next Generation Sequencing: insert number, insert integrity and plasmid backbone

3.4.1.1   Insert number and insertion site

Genomic DNA was isolated from seed or leaf samples from glasshouse-grown plants of eight lines from the T_3 and T_4 generations (Figure 4 and Table 2). The DNA was sheared, using Covaris technology[12], processed for deep sequencing, and enriched through a number of cycles of PCR. The PCR products from the eight lines were then pooled, quantified and sequenced using Illumina MiSeq®[13] technology that yielded raw, 250 bp paired-end reads[14]. An *in silico* analysis then followed in which only those reads aligning to sequences in the 31,564 bp plasmid/vector pJP3416_GA7-ModB (known as vector-targeted sequencing) were selected (the non-aligning reads represent off-target amplification). In the majority (*ca.* 98%) of cases, both ends (reads) aligned with a plasmid sequence.

In the remaining cases, where there was only alignment of one end (read), a further comparison was done with sequences in reference genomes of *B. napus (*cv Darmor-*bzh* (Chalhoub et al. 2014)), *B. rapa* (accession Chiifu (Wang et al. 2011)) and *B. oleracea* (Liu et al. 2014). Those reads having significant matches to both the plasmid and one of the Brassica genomes were used to define the copy number and integration sites of the T-DNA in DHA canola since these reads represent sequences in potential integration sites. The reads were re-assembled into four contigs which could then be mapped to locations in the reference genomes.

For all eight lines, two T-DNA inserts at specific sites on different chromosomes were identified (and confirmed the genetic stability of the inserts). An evaluation of the functioning of each insert was made by crossing the DHA canola T_3 lines with elite non-GM lines and then developing lines (from BC_1F_2 and BC_1F_3) with single T-DNA inserts through backcrossing and selfing (see Figure 4). Analysis of DHA composition of the seed in these lines indicated that while one of the inserts alone produced no DHA, and the other insert produced some DHA, the presence of both inserts was required for maximal DHA formation.

3.4.1.2   Insert integrity and sequencing

The integrity of the two inserts and characterisation of adjacent genomic sequences was assessed through combining a) whole genome sequencing (WGS) of representative lines from BC_1F_2 and BC_1F_3 containing one or other insert and b) PCR-amplicon sequencing (PAS) of the T_5 line containing both inserts.

For WGS, the sheared and processed DNA fragments were sequenced using both Illumina HiSeq® (yielding 150 bp paired sequence reads) and Illumina MiSeq® (yielding 250 bp paired sequence reads) technology. The trimmed and cleaned reads were then assembled into scaffolds[15] that were compared by *in silico* analysis to plasmid pJP3416_GA7-ModB.

---

[12] Covaris - http://covarisinc.com/applications/dnarna-shearing-for-ngs/
[13] Illumina MiSeq - https://www.illumina.com/systems/sequencing-platforms/miseq.html
[14] Paired-end sequencing sequences both ends of a single-strand fragment for a specified distance along the sequence i.e. there is an unsequenced area in the centre of the fragment.
[15] A scaffold is a series of contigs that are in the right order but not necessarily connected.

Those sequences with significant matches were further re-assembled. The depth of coverage (i.e. the average number of times any base of the genome is expected to be independently sequenced) was 30X.

For PAS, 20 primer pairs (based on the sequence obtained from WGS) were designed to amplify both T-DNA inserts and adjacent genomic sequences in the total DNA isolated from the T_5 line (B0050-027-18-20-12-19 – see Figure 3). The resulting amplicons were then Sanger sequenced and the contig sequences were compared to those obtained from WGS, and the vector-targeted sequencing described in Section 3.4.1.1.

Specific information about the nature of the insertions was supplied by the Applicant but is CCI. The sequencing confirmed that:

- One insert is a partial insert containing some of the gene cassettes all complete and perfectly matching the same cassettes in the T-DNA of plasmid pJP3416_GA7-ModB. A translation of the genes yielded proteins with the expected amino acid sequences. The insert is located within one native gene, the expression of which was subsequently shown not to have been interrupted.
- One insert contains more than one copy of various gene cassettes, again all perfectly matching the corresponding cassettes in the T-DNA of plasmid pJP3416_GA7-ModB (although not necessarily in the same orientation) and flanked by partial RB and LB sequences. A translation of the genes yielded proteins with the expected amino acid sequences. While expression of the native gene into which the cassettes had inserted had been disrupted, no negative phenotypic effect was documented in the subsequent breeding programme.
- The inserts are stably inherited.

3.4.1.3   Plasmid backbone

The vector-targeted 250 bp paired-end reads, obtained for the eight T_3 and T_4 lines indicated in Section 3.4.1.1 and aligning to the plasmid reference, all aligned to the T-DNA region. No reads aligned to the backbone region. The results from WGS (Section 3.4.1.2) with BC_1F_2 and BC_1F_3 lines containing only one T-DNA insert similarly confirmed there was no plasmid backbone present. Taken together, the results also confirmed the stability of the DHA event in different generations and lines.

3.4.2   Open reading frame (ORF) analysis

Sequences of potential ORFs were predicted *in silico* from the translation of DNA from a) the two inserts b) the four junction regions between plant genomic DNA and the two DNA inserts and c) any joined segments of sequences within one of the inserts. For the inserts, potential ORFs were defined by both start-to-stop (longest continuous predicted amino acid sequence from methionine (ATG) to stop codons (TAA, TAG, or TGA)), in each of the six reading frames, as well as the region between two stop codons. For the junction regions and joined insert segments, potential ORFs were defined by stop-to-stop codons.

Given the length of the two inserts, and the *in silico* analyses consider each of the six reading frames, many hundreds of potential ORFs would be expected in the two inserts. However few, if any of these are likely to result in an expressed product. ORFs represent only hypothetical coding regions. In order to be translated into a protein, there is an additional requirement for the presence of a promoter, splice sites and/or a terminator (Young et al. 2012). In addition, in eukaryotes, proteins are not usually expressed from the same linear segment of DNA and especially not from tightly packed sequences such as found in the DNA introduced into GM organisms. Since it is necessary for the reading frames to have appropriate ribosomal binding sites in order to be translated into proteins, linear DNA

sequences will significantly limit the potential for translation. Over 700 ORFs were considered in total with the number generated from stop-to-stop being much higher than the number generated from start-to-stop.

The junction region/joined insert segments search identified 25 ORFs in total – 9 associated with one insert and 16 associated with the other insert – ranging in size from 7 – 126 amino acids.

All putative polypeptides were analysed using a bioinformatics strategy to determine their similarity to known protein toxins or allergens (refer to Section 4.1.6).

3.5   Stability of the genetic change in DHA canola

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 (as produced in the initial transformation event) over successive generations. It is best assessed by molecular techniques. 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).

3.5.1   Genetic stability

In addition to evidence provided from analyses described in Sections 3.4.1.1 and 3.4.1.2, genetic stability was assessed through two other approaches.

3.5.1.1   Amplification of junction regions

Total DNA was extracted from each of 20 seeds from each of generations T_3 to T_7 (100 seeds in total) and from seeds of the characterised T_5 line (event positive control) and AV Jade (non-GM control). Four separate PCR assays, each with unique primers (giving four distinct amplicon sizes) targeting the four junction regions of the two inserts, were run for each seed sample. The amplicons from each of the four assays were then mixed and electrophoresed on agarose gels. No bands were detected for the AV Jade control. Identical banding patterns were observed for all seeds from all generations including the event positive control thereby indicating the generational stability of DHA canola.

3.5.1.2   KASP genotyping

Kompetitive Allele-Specific PCR[16] (KASP™) assays, which targeted the four junctions (two assays/junction) of the two T-DNA inserts, were developed mainly for the purpose of trait purity testing and detection of DHA canola. Essentially, this type of PCR allows for the generation of fluorescence signals that are detected, and can be analysed, using any cluster analysis viewing software. If the genotype at a given allele is homozygous, only one of two possible fluorescent signals will be generated. If the genotype is heterozygous, a mixed fluorescent signal will be generated. Detected signals are plotted as a graph (Figure 5) with samples of the same genotype clustering together.

A total of 180 T_5 (event specific positive control) seeds, 313 T_6 seeds, and 20 seeds from each of the T_3 to T_7 generations were tested (along with seeds from the event positive control - B0050-027-18-20-12-19, non-GM controls (Dwarf Essex and AV Jade), hemizygous controls (Dwarf Essex or AV Jade crossed with the event positive line; and DNA from a non-

---

[16] KASP genotyping chemistry - http://www.lgcgroup.com/products/kasp-genotyping-chemistry/#.WOxgCE1MRaQ

template control) using the KASP genotyping. Homozygous T-DNA alleles were observed in seeds from all generations (Figure 5) thereby confirming genetic stability.
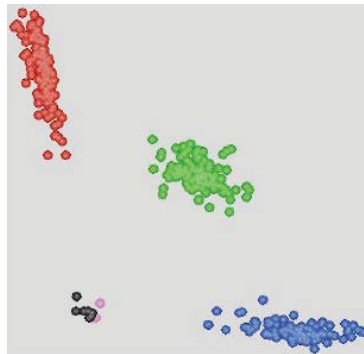
Figure 5: Example of a printout from a downstream KASP assay of one of the inserts
Red dots indicate samples with homozygous non-GM wildtype alleles; green dots indicate samples with hemizygous alleles; blue dots indicate samples with homozygous T-DNA alleles; black dots indicate non-template controls; pink dots indicate failed KASP reactions.

3.5.2   Phenotypic stability

As for the genetic stability two approaches were taken to assess phenotypic stability.

3.5.2.1   Mendelian inheritance

This was calculated for the various populations generated by crossing DHA canola lines with elite non-GM lines as described in the two further breeding strategies outlined in Section 3.3). For the purposes of descriptive annotation, the alleles for the presence of the two T-DNA inserts were designated A and B while those for the wild-type counterpart were designated a and b. The two breeding strategies could then be represented as shown in Figures 6A and 6B.
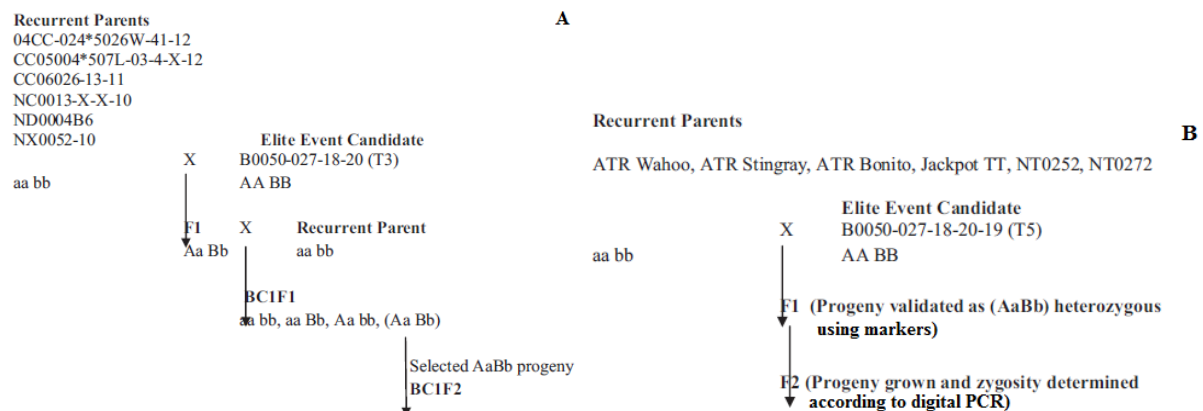
Figure 6: Schematic crosses to produce BC_1F_2 (A) and F_2 (B) progeny

For the BC_1F_2 progeny produced by the breeding strategy represented in Figure 6A, Mendelian inheritance was measured (through use of locus-specific digital PCR and KASP genotyping) by:
- % LC-PUFA phenotype i.e. (five phenotypes expected – see Table 3A)
- Genotype (nine genotypes expected – see Table 3B)

The Chi-squared test ($\chi2$) was used to check whether segregation of the Locus A and Locus B T-DNA inserts fitted Mendelian patterns. Table 3 presents the results for the two analyses.

Table 3: Chi-squared tests for hypotheses that the %LC-PUFA phenotype (A) and two loci genotypes (B) fit the expected Mendelian segregation pattern in BC_1F_2 progeny

A Genotype	Locus A	Locus B	Expected % of progeny	Expected Observable Phenotype % LC PUFA	Expected number of progeny	Observed number of progeny
aabb	wt	wt	0.0625	0	7.4	15
aaBB	wt	Hom	0.0625	<1	7.4	7
aaBb	wt	Het	0.125	<1	14.9	15
Aabb	Het	wt	0.125	>1 -7.5		
AaBb	Het	Het	0.25	>1 -7.5		
AaBB	Het	Hom	0.125	>1 -7.5	81.8	75
AAbb	Hom	wt	0.0625	>1 -7.5		
AABb	Hom	Het	0.125	>1 -7.5		
AABB	Hom	Hom	0.0625	> 7.5	7.4	7

Chi-squared Value $\chi^2 = 8.31$

To be significant ($p<0.05$), the χ^2 value would need to be >9.49

B Genotype	Locus A	Locus B	Expected % of progeny	Expected Observable Phenotype % LC PUFA	Expected number of progeny	Observed number of progeny
aabb	wt	wt	0.0625	0	7.1	12
aaBB	wt	Hom	0.0625	<1	7.1	7
aaBb	wt	Het	0.125	<1	14.1	15
Aabb	Het	wt	0.125	>1 -7.5	14.1	17
AaBb	Het	Het	0.25	>1 -7.5	28.3	35
AaBB	Het	Hom	0.125	>1 -7.5	14.1	14
AAbb	Hom	wt	0.0625	>1 -7.5	7.1	3
AABb	Hom	Het	0.125	>1 -7.5	14.1	6
AABB	Hom	Hom	0.0625	> 7.5	7.1	4

Chi Squared Value $\chi^2 = 14.04$

To be significant ($p<0.05$), the χ^2 value would need to be >15.51

The χ^2 values for both analyses are not significant and therefore show the observed segregation matches the expected Mendelian inheritance pattern.

For the 12 F_2 progeny lines (six recurrent parents by each of the two loci) produced by the breeding strategy represented in Figure 6B, χ^2 values were obtained to test the hypothesis that at each locus the genotype fitted the expected Mendelian inheritance segregation pattern (which was 1 homozygous T-DNA: 2 heterozygous T-DNA: 1 wild type). A representative result (for progeny derived using the ATR Stingray parent) is given in Table 4. Results for all other progeny from the five other parent lines were similar, with no χ^2 value being significant i.e. the observed segregation matched the expected segregation.

Table 4: Chi-squared tests for hypotheses that Locus A and Locus B genotypes fit the expected Mendelian segregation pattern in F_2 progeny

Pedigree	Zygosity	F_2 Locus A Observed	F_2 Locus A Expected 1:2:1	F_2 Locus B Observed	F_2 Locus B Expected 1:2:1
ATR Wahoo / B0050-027-18-20-12-19	Hom	9	16	13	16
	Het	38	31	38	31
	wt	15	16	13	16
		62	62	64	62

$$\chi^2 = 4.32 \qquad \chi^2 = 4.42$$

To be significant ($p<0.05$), the χ^2 value would need to be >5.99

3.5.2.2   *Seed oil composition*

Oil was extracted from seeds (some from single glasshouse-grown plants, some bulked from field-grown plants grown in different locations) from generations $T_1 - T_7$, and analysed for % DHA of the total fatty acids. Overall, the studies showed significant quantities of DHA were produced in all lines, thereby indicating the trait is stable across different generations grown in different environments.

3.6   Conclusion

DHA canola was generated through *Agrobacterium*-mediated transformation with a single T-DNA containing eight expression cassettes. Comprehensive molecular analyses indicate there are two insertion sites on different chromosomes. Specific information about the nature of the insertions was supplied by the Applicant but is CCI. One insertion site comprises the coding regions and complete associated regulatory elements of less than eight of the gene cassettes; the sequences all perfectly match the corresponding sequences in the T-DNA used to introduce the genetic material. The insert is located within one native gene, the expression of which was subsequently shown not to have been disrupted. The other insertion site contains more than one copy of various cassettes, again all perfectly matching the corresponding cassettes in the T-DNA, although not necessarily in the same orientation. While expression of the native gene into which the cassettes have inserted has been disrupted, no negative phenotypic effect was documented in the subsequent breeding programme.

The introduced genes are stably inherited across multiple generations and in different genetic backgrounds. No plasmid backbone (and hence, no antibiotic resistant marker) has been incorporated into either of the transgenic loci.

4   Characterisation and safety assessment of new substances

The main purpose of the characterisation is to describe the nature of any new substances and their phenotypic and biochemical effects on the organism in which they are expressed, particularly in the parts of the organism consumed as food. Typically, the main focus of the characterisation is on newly expressed (or potentially expressed) proteins, but other (non-protein) substances may need to be considered.

4.1 Newly expressed proteins

In considering the safety of newly expressed proteins it is important to note 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 newly expressed proteins in the organism as well as a detailed understanding of their biochemical function and phenotypic effects. It is also important to determine if the newly expressed proteins are expressed as expected, including whether any post-translational modifications have occurred.

Two types of proteins were considered:

- The proteins expected to be directly produced as a result of the translation of the introduced genes. A number of different analyses were done to characterise these proteins and determine their *in planta* expression.
- Those that may be potentially translated as a result of the creation of ORFs during the transformation process (see Section 3.4.2).

4.1.1   PAT protein

The *pat* gene from *S. viridochromogenes* confers tolerance to the antibiotic (natural herbicide) called bialaphos (Murakami et al. 1986) that is also produced by *S. viridochromogenes* i.e. the bacterium has evolved a mechanism to avoid the toxicity of its own product. In nature, bialaphos is produced by the host to help it eliminate other bacteria that may be competitors for the same food source. Bialaphos, now also used as a non-selective herbicide, is a tripeptide composed of two L-alanine residues and an analogue of glutamate known as L-phosphinothricin (L-PPT) (see Thompson et al. 1987) more recently known also as glufosinate ammonium. Free L-PPT released from bialaphos by peptidases (or applied directly as a synthetic herbicide) inhibits glutamine synthetase which in turn leads to rapid accumulation of ammonia and subsequent cell death.

The homologous polypeptide produced by the *pat* and *bar* genes (see Section 2.2.3.1) is known as phosphinothricin acetyltransferase (PAT); it is an acetyl transferase with enzyme specificity for both L-PPT and demethylphosphinothricin (DMPT) in the acetylation reaction (Thompson et al. 1987). In the presence of acetyl-CoA, PAT catalyses the acetylation of the free amino group of L-PPT to N-acetyl-L-PPT, a herbicidally-inactive compound. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for L-PPT and has been shown to have a very low affinity to related compounds and amino acids; even excess glutamate is unable to block the L-PPT-acetyltransferase reaction (Thompson et al. 1987). The proteins from the two different sources have a sequence identity of 85%.

A translation of the DNA sequence of the *pat* gene in DHA canola gives a protein comprising 183 amino acids with a calculated molecular weight of 20.67 kDa (Figure 7).

```
  1  MSPERRPVEI  RPATAADMAA  VCDIVNHYIE  TSTVNFRTEP  QTPQEWIDDL
 51  ERLQDRYPWL  VAEVEGVVAG  IAYAGPWKAR  NAYDWTVEST  VYVSHRHQRL
101  GLGSTLYTHL  LKSMEAQGFK  SVVAVIGLPN  DPSVRLHEAL  GYTARGTLRA
151  AGYKHGGWHD  VGFWQRDFEL  PAPPRPVRPV  TQI
```

Figure 7: Translated amino acid sequence of the PAT protein

4.1.2   Fatty acid desaturases and elongases

Unpublished studies submitted:

2016. Characterisation of *Lachancea kluyveri* Δ12-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-005.

2016. Characterisation of *Pichia pastoris* ω3-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-006.

2016. Characterisation of *Micromonas pusilla* Δ6-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-007.

2016. Characterisation of *Pyramimonas cordata* Δ6-Elongase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-008.

2016. Characterisation of *Pavlova salina* Δ5-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-009.

2016. Characterisation of *Pyramimonas cordata* Δ5-Elongase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-010.

2016. Characterisation of *Pavlova salina* Δ4-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-011.

A translation of the DNA sequences of the seven desaturase/elongase genes in DHA canola (together with calculated molecular weight) is provided in Figures 8 – 14.

In silico analyses using Vector NTI[17] software were conducted, where the sequence of each protein was used to search for homologous sequences present in organisms used in food, food production or in animal feeds. The purpose of this search was to identify the similarity (sequence identity) of each protein to other proteins present in consumed foods or used in food production or animal feeds. Results from this search are provided in Appendix 1 and indicate that similar proteins may be consumed in the diet or are already present in humans.

```
MSAVTVTGSDPKNRGSSSNTEQEVPKVAIDTNGNVFSVPDFTIKDILGAIPHECYERRLA
TSLYYVFRDIFCMLTTGYLTHKILYPLLISYTSNSIIKFTFWALYTYVQGLFGTGIWVLA
HECGHQAFSDYGIVNDFVGWTLHSYLMVPYFSWKYSHGKHHKATGHMTRDMVFVPATKEE
FKKSRNFFGNLAEYSEDSPLRTLYELLVQQLGGWIAYLFVNVTGQPYPDVPSWKWNHFWL
TSPLFEQRDALYIFLSDLGILTQGIVLTLWYKKFGGWSLFINWFVPYIWVNHWLVFITFL
QHTDPTMPHYNAEEWTFAKGAAATIDRKFGFIGPHIFHDIIETHVLHHYCSRIPFYNARP
ASEAIKKVMGKHYRSSDENMWKSLWKSFRSCQYVDGDNGVLMFRNINNCGVGAAEK
```

Figure 8: Translated amino acid sequence (plus potential glycosylation sites – green text) of the LackI-Δ12D protein (416 amino acids; 48.2 kDa)

```
MSKVTVSGSEILEGSTKTVRRSGNVASFKQQKTAIDTFGNVFKVPDYTIKDILDAIPKHC
YERSLVKSMSYVVRDIVAISAIAYVGLTYIPLLPNEFLRFAAWSAYVFSISCFGFGIWIL
GHECGHSAFSNYGWVNDTVGWVLHSLMVPYFSWKFSHAKHHKATGHMTRDMVFVPYTAE
EFKEKHQVTSLHDIAEETPIYSVFALLFQQLGGLSLYLATNATGQPYPGVSKFFKSHYWP
SSPVFDKKDYWYIVLSDLGILATLTSVYTAYKVFGFWPTFITWFCPWILVNHWLVFVTFL
QHTDSSMPHYDAQEWTFAKGAAATIDREFGILGIIFHDIIETHVLHHYVSRIPFYHAREA
TECIKKVMGEHYRHTDENMWVSLWKTWRSCQFVENHDGVYMFRNCNNVGVKPKDT
```

Figure 9: Translated amino acid sequence (plus potential glycosylation sites – green text) of the Picpa-ω3D protein (415 amino acids; 47.8 kDa)

---

[17] Vector NTI - https://www.thermofisher.com/au/en/home/life-science/cloning/vector-nti-software.html

```
MCPPKTDGRSSPRSPLTRSKSSAEALDAKDASTAPVDLKTLEPHELAATFETRWVRVED
VEYDVTNFKHPGGSVIFYMLANTGADATEAFKEFHMRSLKAWKMLRALPSRPAEIKRSE
SEDAPMLEDFARWRAELERDGFFKPSITHVAYRLLELLATFALGTALMYAGYPIIASVV
YGAFFGARCGWVQHEGGHNSLTGSVYVDKRLQAMTCGFGLSTSGEMWNQMHNKHHATPQ
KVRHDMDLDTTPAVAFFNTAVEDNRPRGFSRAWARLQAWTFVPVTSGLLVQAFWIYVLH
PRQVLRKKNYEEASWMLVSHVVRTAVIKLATGYSWPVAYWWFTFGNWIAYMYLFAHFST
SHTHLPVVPSDKHLSWVNYAVDHTVDIDPSRGYVNWLMGYLNCQVIHHLFPDMPQFRQP
EVSRRFVPFAKKWGLNYKVLSYYGAWKATFSNLDKVGQHYYVNGKAEKAH
```

Figure 10: Translated amino acid sequence of the Micpu-Δ6D protein (463 amino acids; 52.9 kDa)

```
MEFAQPLVAMAQEQYAAIDAVVAPAIFSATDSIGWGLKPISSATKDLPLVESPTPLILSL
LAYFAIVGSGLVYRKVFPRTVKGQDPFLLKALMLAHNVFLIGLSLYMCLKLVYEAYVNKY
SFWGNAYNPAQTEMAKVIWIFYVSKIYEFMDTFIMLLKGNVNQVSFLHVYHHGSISGIWW
MITYAAPGGDAYFSAALNSWVHVCMYTYYFMAAVLPKDEKTKRKYLWWGRYLTQMQMFQF
FMNLLQAVYLLYSSSPYPKFIAQLLVVYMVTLLMLFGNFYYMKHHASK
```

Figure 11: Translated amino acid sequence of the Pyrco-Δ6E protein (288 amino acids; 33.1 kDa)

```
MPPRDSYSYAAPPSAQLHEVDTPQEHDKKELVIGDRAYDVTNFVKRHPGGKIIAYQVGT
DATDAYKQFHVRSAKADKMLKSLPSRPVHKGYSPRRADLIADFQEFTKQLEAEGMFEPS
LPHVAYRLAEVIAMHVAGAALIWHGYTFAGIAMLGVVQGRCGWLMHEGGHYSLTGNIAF
DRAIQVACYGLGCGMSGAWWRNQHNKHHATPQKLQHDVDLDTLPLVAFHERIAAKVKSP
AMKAWLSMQAKLFAPVTTLLVALGWQLYLHPRHMLRTKHYDELAMLGIRYGLVGYLAAN
YGAGYVLACYLLYVQLGAMYIFCNFAVSHTHLPVVEPNEHATWVEYAANHTTNCSPSWW
CDWWMSYLNYQIEHHLYPSMPQFRHPKIAPRVKQLFEKHGLHYDVRGYFEAMADTFANL
DNVAHAPEKKMQ
```

Figure 12: Translated amino acid sequence (plus potential glycosylation sites – green text) of the Pavsa-Δ5D protein (425 amino acids; 48.2 kDa)

```
MASIAIPAALAGTLGYVTYNVANPDIPASEKVPAYFMQVEYWGPTIGTIGYLLFIYFGK
RIMQNRSQPFGLKNAMLVYNFYQTFFNSYCIYLFVTSHRAQGLKVWGNIPDMTANSWGI
SQVIWLHYNNKYVELLDTFFMVMRKKFDQLSFLHIYHHTLLIWSWFVVMKLEPVGDCYF
GSSVNTFVHVIMYSYYGLAALGVNCFWKKYITQIQMLQFCICASHSIYTAYVQNTAFWL
PYLQLWVMVNMFVLFANFYRKRYKSKGAKKQ
```

Figure 13: Translated amino acid sequence (plus potential glycosylation sites – green text) of the Pyrco-Δ5E protein (267 amino acids; 31.3 kDa)

```
MPPSAAKQMGASTGVHAGVTDSSAFTRKDVADRPDLTIVGDSVYDAKAFRSEHPGGAHFV
SLFGGRDATEAFMEYHRRAWPKSRMSRFHVGSLASTEEPVAADEGYLQLCARIAKMVPSV
SSGFAPASYWVKAGLILGSAIALEAYMLYAGKRLLPSIVLGWLFALIGLNIQHDANHGAL
SKSASVNLALGLCQDWIGGSMILWLQEHVVMHHLHTNDVDKDPDQKAHGALRLKPTDAWS
PMHWLQHLYLLPGETMYAFKLLFLDISELVMWRWEGEPISKLAGYLFMPSLLLKLTFWAR
FVALPLYLAPSVHTAVCIAATVMTGSFYLAFFFFISHNFEGVASVGPDGSITSMTRGASF
LKRQAETSSNVGGPLLATLNGGLNYQIEHHLFPRVHHGFYPRLAPLVKAELEARGIEYKH
YPTIWSNLASTLRHMYALGRRPRSKAE
```

Figure 14: Translated amino acid sequence of the Pavsa-Δ4D protein (447 amino acids; 49.3 kDa)

4.1.3 Protein expression in the tissues of DHA canola

> **Unpublished Study submitted:**
>
> 2016. Protein expression of DHA biosynthesis pathway enzymes in canola. Internal Nuseed/CSIRO report. Nuseed Pty Ltd Report No. 2016-015

All seven enzymes in the n-3 LC-PUFA synthesis pathway are integral membrane proteins expected to be expressed at very low level. As such, they would be classified as intractable (Bushey et al. 2014). Since functioning antibodies against the proteins could not be raised, the usual approach for quantification involving Western blot analysis could not be followed. Therefore an alternative approach using high sensitivity liquid chromatography/multiple reaction monitoring/mass spectrometry (LC/MRM/MS) was developed.

Protein quantification by MRM, using a triple quadrupole mass spectrometer, is an approach that has been successfully applied in clinical laboratory studies (Rauh 2012; Gillette and Carr 2013). The use of tandem mass spectrometry (MS/MS) provides an additional level of quality control in confirming peptide identity. Analysis of proteins by MRM is based on detection of peptides derived from proteolytic digestion (typically trypsin) of the target protein (Lange et al. 2008). For the purposes of the current analysis, a single peptide was selected from each of the seven proteins as a proxy to be used for quantification; selection was based on high signal intensity and good chromatographic properties. A precise quantification of the endogenous protein (femtomoles per 100 µg total protein) was converted to nanogram equivalent per mg total protein based on the molecular mass of each protein).

Plants from verified seed of DHA canola (B0050-027-18-20 T_3 seed – see Figure 4), together with the non-GM AV Jade parent were sampled from two locations in Horsham (Victoria, Australia) during the 2015 season. The developmental stages[18] at which harvesting was done, are given in Table 5. To avoid cross contamination, AV Jade samples were processed first, followed by the DHA canola samples in the order given in Table 5.

Quantification of the PAT protein was attempted through both LC/MRM/MS and validated Western blotting (rabbit anti-PAT antibody) of total protein isolated from each plant part.

None of the peptides were detected in any of the AV Jade samples. The seven enzymes in the n-3 LC-PUFA synthesis pathway were only detected in the developing seed and mature seed (Table 5). This is expected since the genes encoding the proteins all have seed-specific promoters. The expression levels of all seven proteins were low, with Pyrco-Δ5E having the lowest and Pavsa-Δ4D the highest. Given that the range of total protein in seeds was 5.95 – 18.0 mg/g of seed, the levels of the seven enzymes (ranging from 20 – 740 ng/mg total protein represent a negligible proportion of the total protein.

For the PAT protein, while LC/MRM/MS could detect its presence in all plant parts, it could not be quantified. Western blotting was even less sensitive.

---

[18] The canola growth stages are based on the Bayer, BASF, Ciba-Geigy and Hoechst (BBCH) Growth Stage Scale see e.g. CCC (2014)

Table 5: Mean protein levels (ng protein/mg total protein) in various tissues of DHA canola

BBCH15 = 5 leaf stage; BBCH35 = stem elongation (3 internodes): BBCH65 = 50% flowering;
BBCH79 = developing seed; BBCH90 = senescence; ND = not detected
For those plant parts where a protein was detected, the means (plus standard deviation) for both sites have been given separately.

Growth stage/plant part	Tissue sampled (x3 replicates)	Lackl-Δ12D	Picpa-ω3D	Micpu-Δ6D	Pyrco-Δ6E	Pavsa-Δ5D	Pyrco-Δ5E	Pavsa-Δ4D	PAT
BBCH15/ whole plant	3 whole plants	ND	ND	ND	ND	ND	ND	ND	<LOQ
BBCH35/ whole plant	1 whole plant	ND	ND	ND	ND	ND	ND	ND	<LOQ
BBCH65/ root	roots from 1 plant	ND	ND	ND	ND	ND	ND	ND	<LOQ
BBCH65/ raceme	all flowers from 1 plant	ND	ND	ND	ND	ND	ND	ND	<LOQ
BBCH65/ remaining parts	all leftover parts from 1 plant	ND	ND	ND	ND	ND	ND	ND	<LOQ
BBCH79/ developing seed	all siliquae[1] from 1 plant	244.2 ± 6.8	167.7 ± 24.8	87.9± 15.2	26.1 ± 1.8	63.4 ± 16.4	ND	480.5 ± 146.2	<LOQ
		222.3 ± 72.0	168.1 ± 71.0	136.1 ± 30.3	29.7 ± 6.5	65.8 ± 31.7		438.2 ± 310.2	<LOQ
BBCH90/ mature seed	all siliquae from 1 plant	212.4 ± 43.2	224.1 ± 90.3	45.3 ± 4.0	ND	62.3 ± 15.2	20.0 ± 12.1	739.5 ± 201.5	<LOQ
		265.4 ± 42.0	263.3 ± 26.3	42.8 ± 7.9		75.0 ± 20.0	28.0 ± 4.9	724.7 ± 154.7	<LOQ

[1] the botanical name for the fruit of canola is silique (plural siliquae)

4.1.4   Characterisation of the proteins produced in DHA canola

In many cases it is not possible to obtain sufficient amounts of newly-expressed proteins from the plant for safety studies. A standard practice in such cases is to produce the proteins in a microbial expression system and confirm their equivalence to the plant-produced proteins while simultaneously determining their structural and functional characteristics.

4.1.4.1   PAT

Weight of evidence from the translation of the *pat* gene present in DHA canola (Figure 7) and the protein expression analysis (Section 4.1.3) confirm the identity of the expressed PAT protein. This protein has been previously assessed for safety by FSANZ (see Section 4.1.5.1) and was therefore not further characterised for this assessment.

4.1.4.2 Fatty acid desaturases and elongases

> **Unpublished studies submitted:**
>
> 2016. Characterisation of *Lachancea kluyveri* Δ12-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-005.
> 2016. Characterisation of *Pichia pastoris* ω3-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-006.
> 2016. Characterisation of *Micromonas pusilla* Δ6-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-007.
> 2016. Characterisation of *Pyramimonas cordata* Δ6-Elongase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-008.
> 2016. Characterisation of *Pavlova salina* Δ5-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-009.
> 2016. Characterisation of *Pyramimonas cordata* Δ5-Elongase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-010.
> 2016. Characterisation of *Pavlova salina* Δ4-Desaturase. Internal Nuseed/CSIRO report, Nuseed Pty Ltd. Report Nº 2016-011.

Results from the protein expression levels (Section 4.1.3) clearly show that all seven proteins are expressed at very low levels. Attempts by the Applicant to purify or further enrich any of the proteins from DHA canola were unsuccessful.

Characterisation of such proteins therefore requires that indirect approaches be taken to provide a weight-of-evidence.

The Applicant undertook enzyme activity studies using the expression of histidine (His) tag fusion proteins[19] (designed for each protein and ± secretion peptide) in a heterologous system. All of the proteins except Pavsa-Δ4D could be expressed in *P. pastoris*[20] (PichiaPink™ strain 4). Pavsa-Δ4D was expressed in yeast S288C cells (Zhou et al. 2007) as well as showing functionality in a number of other heterologous systems - *N. benthamiana* (Petrie et al. 2010b); Arabidopsis (Petrie et al. 2012); and Camelina (Petrie et al. 2014). For the six fusion proteins expressed in *P. pastoris*, the control in each instance is *P. pastoris* without a fusion protein but it is noted that *P. pastoris* contains endogenous Δ12D and ω3D.

The enzyme studies were based on the ability of each protein to convert a particular fatty acid substrate to a specific fatty acid product within the total fatty acids present in cells of *P. pastoris* transformed with the gene encoding the enzyme of interest. Following growth in a nutrient medium (to which the substrate fatty acid had been added if required) for 3 days, the yeast cells were harvested by centrifugation and dried. Fatty acid methyl esters were prepared and analysed by gas chromatography (Zhou et al. 2006).

Results for each fusion protein + secretion peptide are presented in Table 6, which also provides references in which the functionality of each protein expressed in other heterologous systems is given. The % substrate and % product were the % of each fatty acid present inside the *P. pastoris* cells out of the total cell fatty acid, after being cultured for 3 days. The results clearly confirm that each fusion protein functions as it should.

---

[19] Fusion proteins had to be used because of the difficulty of expressing membrane proteins in either prokaryotic or eukaryotic systems. The His tag assists in purification using affinity chromatographic methods. Since it is difficult to remove the tag following purification it is essential the tag does not affect functional activity if the proteins are used in safety studies.

[20] Eukaryotic hosts such as yeast, mammalian and insect cells are more likely to be useful for obtaining higher levels of functional membrane proteins than more traditional bacterial hosts such as *E. coli* (Tate 2001; Madduri et al. 2012).

Table 6: Activity of six fusion proteins in transformed *P. pastoris* cells

Enzyme	Pichia Sample (replicates)	Substrate	Substrate (%)	Product	Product (%)	Conversion (%)**	Activity in other heterologous systems
Lackl-Δ12D	Control (10)	18:1	30.9±6.4	18:2	37.3±4.0*	58.7	Yeast (Watanabe et al. 2014); Arabidopsis (Petrie et al. 2012); Camelina (Petrie et al. 2014)
	His$_{10}$ + SP (10)		24.8±5.2		41.9±3.2*	66.6	
Picpa-ω3D	Control (3)	18:2	26.0±1.1	18:3	4.2±0.2	14.0	Yeast (Zhang et al. 2008); Arabidopsis (Petrie et al. 2012); Camelina (Petrie et al. 2014)
	His$_{10}$ + SP (10)		12.7±1.7		11.2±2.4	46.5	
Micpu-Δ6D	Control (10)	18:3	6.4±1.8	18:4	0	0	Yeast & *N benthamiana* (Petrie et al. 2010a); Arabidopsis (Petrie et al. 2012); Camelina (Petrie et al. 2014)
	His$_{10}$ + SP (10)		6.1±1.0		0.2±0.2	3.3	
Pyrco-Δ6E	Control (3)	18:4	8.1±0.0	20:4	0	0	Yeast (Petrie et al. 2010b); Arabidopsis (Petrie et al. 2012); Camelina (Petrie et al. 2014)
	His$_{10}$ + SP (7)		4.5±1.1		3.8±1.1	46.5	
Pavsa-Δ5D	Control (3)	20:4	6.2±1.1	20:5	0	0	Yeast; (Zhou et al. 2007); *N. benthamiana* (Wood et al. 2009); Arabidopsis (Petrie et al. 2012); Camelina (Petrie et al. 2014)
	His$_{10}$ + SP (8)		5.3±0.7		0.05±0.01	0.9	
Pyrco-Δ5E	Control (3)	20:5	2.8±0.0	22:5	0	0	Yeast (Petrie et al. 2010b); Arabidopsis (Petrie et al. 2012); Camelina (Petrie et al. 2014)
	His$_{10}$ + SP (9)		1.8±1.3		2.7±	62.8	

* In the case of Lackl-Δ12D, the % product included levels of 18:2 and 18:3, as the 18:2 product could be further desaturated by the endogenous *Pichia* cell ω3D.

** % conversion = 100*(% of product/(% of product + % of substrate))

Conclusion

Due to the low levels of the seven proteins in the n-3 LC-PUFA synthesis pathway, purified proteins of sufficient quantity and quality could not be extracted from the DHA canola plant to be able to directly characterise them. However the weight of evidence, provided by a) translation of the known DNA sequences of the seven introduced genes present in DHA canola, b) the fact the proteins function as predicted in the plant and c) the functionality of the seven proteins expressed in heterologous systems, is sufficient to characterise the proteins. expressed in DHA canola.

4.1.5 Safety of the introduced proteins

4.1.5.1 PAT

The PAT protein expressed in DHA canola (Section 4.1.4.1) from the *pat* gene has the same identity as the PAT protein previously assessed by FSANZ in a number of applications and considered to be safe[21]. Since there is 87% similarity in the amino acid sequence between the PAT gene expressed by *pat* and that expressed by *bar* (Wehrmann et al. 1996) a further eight approvals considered by FSANZ have also considered the safety of the PAT protein encoded by the *bar* gene[22].

In addition, the protein is accepted as having neither toxicity nor allergenicity concerns (see e.g. Hérouet et al. 2005; Delaney et al. 2008; Fard et al. 2013; Hammond et al. 2013) and updated bioinformatics analyses supplied by the Applicant do not point to any significant similarity with known protein toxins or allergens.

4.1.5.2 Fatty acid desaturases and elongases

Based on sequence similarity and functionality, the seven enzymes can be classified into three groups:
- yeast acyl Co-A-type fatty acid desaturases – **LackI-Δ12D** and **Picpa-ω3D** – that introduce a double bond at the Δ12 and Δ15 positions respectively
- algal fatty acid elongases – **Pyrco-Δ6E** and **Pyrco Δ5E** – that add two carbons to the carboxyl end of fatty acids
- algal front-end desaturases that introduce a double bond between an existing double bond and the carboxyl end of fatty acids – **Micpu-Δ6D, Pavsa-Δ5D** and **Pavsa-Δ4D**

4.1.5.2.1 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 proteins will behave like any other dietary protein.

The assessment focuses on:
- whether the novel proteins have a prior history of safe human consumption, or are 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 proteins including whether they are resistant to heat or processing.

An oral toxicity study is only deemed necessary if the results of biochemical, bioinformatic, digestibility or stability studies indicate further investigation of potential toxicity is warranted.

---

[21] The PAT protein, encoded by the *pat* gene, has now been considered in 16 FSANZ safety assessments (A372, A375, A386, A446, A481, A518, A543, A1046, A1073, A1081, A1087, A1094, A1106, A1112, A1116, and A1118).
[22] The PAT protein encoded by the bar gene has now been considered in 8 FSANZ safety assessments (A372, A380, A385, A533, A589, A1028, A1040, and A1080).

History of human consumption

This has been considered in Sections 3.2 and 4.1.2. Any organism that produces n-3 LC-PUFAs will contain enzymes with similar functionality and structure to the seven enzymes expressed in DHA canola. Humans and animals have been routinely consuming these sources (e.g., yeast, algae) as well as the elongases and desaturases that they encode.

Amino acid sequence similarity to known protein toxins

Bioinformatic analyses are useful for assessing whether introduced proteins share any amino acid sequence similarity with known protein toxins. Two approaches were taken:

- Full sequences of each of the seven proteins were compared to all protein sequences (using a Basic Local Alignment Search Tool [23]– BLAST (Altschul et al. 1990)) analysis in the NCBI Entrez[24] protein sequences database (as at December 2016), which, has integrated access to six major protein databases. Comparisons between highly homologous proteins yield *E*-values[25] approaching zero, indicating the very low probability that such matches would occur by chance. A larger *E*-value indicates a lower degree of similarity. Typically, alignments between two sequences will need to have an *E*-value of $1e^{-5}$ (1×10^{-5}) or smaller to be considered to have significant homology. However, any conclusion about the significance of a homology needs to be tempered by an understanding of its biological relevance.

- A BLAST search was used to compare the seven protein sequences against the entire NCBI Entrez protein database with a limit option provided by using the keywords 'toxin' or 'toxic'.

The results of the homology searches showed there were matches for all proteins but that these were to common homologous proteins found in highly diverse taxa i.e. no biologically relevant identities were found for any of the seven proteins with any toxic proteins from the database.

In vitro *digestibility*

See Section 4.1.5.2.2

Thermal stability

The usual method for assaying thermal stability is to heat the enzyme to various temperatures and then assay its activity. Studies looking at thermal stability of the seven proteins were not undertaken. The following points are noted:

- Sufficient purified proteins could not be obtained from DHA canola for assay.

---

[23] BLAST - https://blast.ncbi.nlm.nih.gov/Blast.cgi
[24] Entrez databases - https://www.ncbi.nlm.nih.gov/Class/MLACourse/Original8Hour/Entrez/
[25] The Statistics of Sequence Similarity Scores - https://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html

- While enzymatic activity could be measured (Section 4.1.4.2), this was done indirectly via a biological system that was totally unsuitable for quantifying activity at different temperatures.
- Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, by itself, has limitations for determining thermal denaturation of membrane proteins.
- Membrane proteins are not thermally stable and are difficult to refold once denatured (Bowie 2001). In the processing of canola seeds to produce oil, temperatures ranging from 80° – 115° C are reached and it would be unlikely any trans-membrane proteins would remain in the folded state i.e. it is unlikely any of the seven proteins would remain as an active folded protein after processing.

4.1.5.2.2   Potential allergenicity

The potential allergenicity of the 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 (discussed in Section 4.1.5.2.1); 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.

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

Unpublished studies submitted

2016. Stability of *Pichia pastora* ω3-/ Δ15-Desaturase. Internal Nuseed/CSIRO report. Nuseed Pty Ltd. Report No. 2016-012

2016. Protein Stability of *Pyramimonas cordata* Δ5-Elongase. Internal Nuseed/CSIRO report. Nuseed Pty Ltd. Report No. 2016-013

2016. Protein Stability of *Pavlova salina* Δ4-Desaturase. Internal Nuseed/CSIRO report. Nuseed Pty Ltd. Report No. 2016-014

2016. Protein characterization and safety of the proteins expressed in DHA canola (OECD ID NS-B50027-4). Internal Nuseed/CSIRO report. Nuseed Pty Ltd. Report No. 2016-016

2016. Bioinformatics analysis of the potential allergenicity and toxicity of proteins encoded by genes inserted in canola (*Brassica napus*) for production of omega 3 fatty acids. Internal Nuseed report. Nuseed Pty Ltd. Report No. 2016-017

Sources of the proteins

See Section 2.2

Similarity to known allergens

Bioinformatic analysis provides part of a weight of evidence approach for assessing potential allergenicity of novel proteins introduced to GM plants (Thomas et al. 2005; Goodman 2006). It is a method for comparing the amino acid sequence of the introduced protein with sequences of known allergens in order to indicate potential cross-reactivity between allergenic proteins and the introduced protein. As with the bioinformatic analysis that looked at similarities of the novel proteins with known protein toxins (refer to Section 4.1.5.1), the

generation of an *E* value provides an important indicator of significance of matches (Pearson 2000; Baxevanis 2005).

To evaluate the similarity to known allergens of the seven proteins a number of *in silico* strategies were carried out as follows:

- Full length sequences of each of the seven proteins were compared with all known putative allergen sequences residing in a reference allergen database, (AllergenOnline[26] version 16, released on 27 January 2016 – containing 1,956 non-redundant entries) using the Fast Alignment Search Tool – All (FASTA) version 35.04 algorithm and BLOSUM50 scoring matrix.

- An epitope search of the AllergenOnline database was done to identify any short sequences of 80 amino acids (a sliding 80 amino acid window) that might represent an isolated shared IgE binding epitope. This search is based on the recommendation of Codex (2009) and can also help identify potentially cross-reactive proteins that are not true homologues of an allergen but might have significant similarities that could provide an immunological target for IgE antibodies. A match of 35% over 80 amino acids with a known allergen suggests further testing for possible cross-reactivity.

- A word/string search routine on AllergenOnline.org was used to identify any eight contiguous amino acid sequences of each of the seven proteins that exactly match any eight amino acid segment of any of the allergens in the database. The rationale for this is an assumption that individual epitopes may be represented by peptide sequences as short as eight amino acids (Metcalfe et al. 1996).

- A BLAST search (see Section 4.1.5.2.1) was used to compare the seven protein sequences against the entire NCBI Entrez protein database with a limit option provided by using the keywords 'allergen' or 'allergy'.

Results of the full-length FASTA searches of the seven proteins did not identify any significant alignment with any allergen. No biologically relevant alignment for any of the seven proteins met or exceeded the Codex Alimentarius (Codex 2009) FASTA alignment threshold for potential allergenicity (35% identity over 80 amino acids), and no alignments of eight or more consecutive identical amino acids were found between any of the proteins and known allergens in the database. No biologically relevant matches were obtained using the keyword limited search of all proteins in NCBI Entrez.

The overall conclusion is that none of the seven proteins have significant sequence identity matches with any known allergens.

Glycosylation

The potential for glycosylation was investigated for each protein. Glycosylation essentially involves the covalent attachment of a carbohydrate to the target protein. The carbohydrate component may represent from <1% to >80% of the total molecular weight of glycoprotein. N-glycosylation can affect protein stability and/or activity and has also been implicated in contributing to possible allergenicity (Huby et al. 2000) since it may affect the susceptibility of a protein to processing and proteolysis and may introduce glycan peptides which are known to be highly cross-reactive epitopes (Altmann 2007).

Glycosylation that occurs on side chains of asparagine residues is termed N-glycosylation. It is commonly associated with an asparagine-X-serine/threonine sequence (N-X~(P)-[S/T),

---

[26] AllergenOnline – http://www.allergenonline.org/

where X~(P) indicates any amino acid except proline (Orlando and Yang 1998). Although rare, the sequence asparagine-X-Cysteine (N-X-C) can also be an N-glycosylation site (Miletich and Broze Jr. 1990).

Insufficient quantities of protein were obtained for glycostaining. Therefore only **potential** N-glycosylation sites could be identified through an *in silico* search of the plant-expressed protein sequences through GlycoEP[27] (Chauhan et al. 2013). The result of the search for each protein has been included in Figures 8 – 14. In summary, these results show:

- no potential glycosylation sites in Micpu-Δ6D, Pyrco-Δ6E, Pavsa-Δ4D
- 1 potential glycosylation site in Lackl-Δ12D, Pyrco-Δ5E
- 2 potential glycosylation sites in Picpa-ω3D, Pavsa-Δ5D

While a negative result provides some surety the protein, as expressed in the plant, is unlikely to be glycosylated, a positive result does not, conversely, suggest a protein is likely to be glycosylated. In particular it is noted that a linear sequence search does not account for the fact that glycan attachment occurs before, and influences, protein folding and it is the folding itself, which imparts biophysical characteristics to a protein (Shental-Bechor and Levy 2008; Chuang et al. 2012). Also, while N-glycosylation is the most common stable modification to impact the physicochemical properties of a protein, its occurrence is highly variable between species, environments and cells and its value for safety assessment has been questioned (Bushey et al. 2014).

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 some correlation exists between resistance to digestion by pepsin and potential allergenicity although it does not necessarily follow that resistance to digestion is always an indicator of an allergenic protein (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 a protein is allergenic. The existence of significantly sized protease-resistant fragments, generally accepted as larger than approximately 3000– 3500MW (Bannon et al. 2002) after simulated gastric digestion suggests that further studies might be appropriate to assess whether a protein has the potential to be allergenic.

An indication of potential protein digestibility can be made by *in silico* analysis of a protein. Protease cleavage sites were investigated by FSANZ using the amino acid sequence of each protein and the PeptideCutter[28] tool in the ExPASy Proteomics Site. The results are shown in Table 7 and indicate all seven proteins are potentially as susceptible to digestion as the vast majority of dietary proteins.

---

[27] GlycoEP - http://www.imtech.res.in/raghava/glycoep/index.html
[28] PeptideCutter - http://web.expasy.org/peptide_cutter/

Table 7: Predicted protease cleavage sites in each protein

Protein (no. of amino acids)	No. of predicted cleavage sites in each protein					
	Pepsin pH1.3	Pepsin pH>2	Trypsin	Chymotrypsin – high specificity	Chymotrypsin – low specificity	Endopeptidases
Lackl-Δ12D (416)	88	141	35	68	126	71
Picpa-ω3D (415)	89	135	35	64	115	73
Micpu-Δ6D (463)	84	136	48	60	123	91
Pyrco-Δ6E (288)	77	116	20	48	102	29
Pavsa-Δ5D (425)	77	124	39	49	119	78
Pyrco-Δ5E (267)	71	111	21	53	93	18
Pavsa-Δ4D (447)	102	132	36	45	125	76

In vitro digestibility studies were carried out on one representative protein (Picpa-ω3D, Pyrco Δ5E, and Pavsa-Δ4D) from each of the three classifications given at the beginning of Section 4.1.5.2 using fusion proteins expressed in, and purified from, a heterologous system (see Table 8). All fusion proteins contained a His tag at the N-terminus; one fusion protein also contained green fluorescent protein (GFP)[29]. The apparent molecular weight of each fusion protein was determined by SDS-PAGE and Western blotting (using an anti-His antibody) (Table 8).

Two test systems – simulated gastric fluid (SGF) containing pepsin, and a combined SGF-trypsin digestion - were used independently to test the stability of each fusion protein. For SGF digestion, the test proteins were evaluated following incubation at 37° C for 0, 5, 10, 15, 30 and 60 min. For the combined SGF –trypsin digestion, each SGF time point was followed by 16 h digestion with trypsin. Stability was ascertained through analysis (of five replicates) by LC/MRM/MS (see Section 4.1.3) using selected peptides (numbers given in Table 8) as proxy for the protein that a) spanned the length of the intact protein, b) could be identified with 95% confidence and c) yielded intense signals. For SGF, the increased abundance of targeted peptic peptides was used as indicator of protein digestibility while for the combined digestion, the relative abundance of tryptic peptides compared to the abundance of the same peptides at zero time SGF digestion followed by trypsin digestion was used as the indicator of digestibility.

The reason for using LC/MRM/MS analysis rather than the usual Western blotting is that, functional antibodies specific to each protein could not be raised because the proteins are integrally bound to membranes. Antibodies could be raised to the affinity His tag of the fusion proteins thus allowing their immune-purification but this His tag is specific to the fusion region rather than the whole protein, and could be cleaved off during digestion.

---

[29]The His tag aids in purification. GFP allows tracking of protein expression by monitoring its fluorescence and is widely used as a fusion partner for soluble expression. Neither peptide is considered likely to affect proteolysis or the detection of digestion products.

Table 8: Information on the fusion proteins used in the digestibility studies

Fusion protein	Predicted mw	Apparent mw	Expression system	No. of peptides monitored by LC/MRM/MS	
				Pepsin	Trypsin
His$_8$::GFP::Picpa-ω3D	76.4 kDa	50 & 60 kDa	*Escherichia coli* strain C41	8	8
His$_{10}$::Pyrco-Δ5E	33.7 kDa	27 kDa	Insect cell lines (*Sf9*) infected with *baculovirus* pFastBac vector[1]	8	1[2]
His$_{10}$::Pavsa-Δ4D	51 kDa	50 kDa	Insect cell lines (*Sf9*) infected with *baculovirus* pFastBac vector[1]	8	8

[1]The insect cell/baculovirus system is widely used for expression of membrane proteins but the yield of the expressed protein is many times less than that of 'traditional' expression systems such as *E. coli.*

[2]Only a single tryptic product could be used because the distribution of trypsin sites within the protein sequence resulted in few peptides that were amenable to LC-MS.

The characteristics of the bands obtained for the SDS-PAGE analysis of the purified fusion proteins were not as predicted. This may be because a lower than predicted apparent molecular weight on SDS-PAGE is a common and well-documented phenomenon for membrane proteins and is caused by the presence and binding of detergent to the hydrophobic regions (Rath et al. 2009). For the His$_8$::GFP::Picpa-ω3D protein, the presence of two separate bands may be due to a population where the GFP in the fusion remains (partially) folded, causing it to migrate faster (lower band) than the population where the GFP is completely denatured (upper band) (Geertsma et al. 2008).

The results of the digests of the three proteins can be summarised as follows:

- All peptic peptides monitored for His$_8$::GFP::Picpa-ω3D were produced rapidly (<15 min). The tryptic peptides monitored after the pepsin digest showed a rapid decline in the first 5 min and then a further decline over the remainder of the experiment (60 min duration). It is estimated that >97% of the protein was cleaved after 60 min on the basis of the disappearance of the eight tryptic peptides.
- There was rapid degradation of the His10::Pyrco-Δ5E protein in SGF with >75% cleavage of the N-terminal region achieved in < 5 min and a suite of products spanning the entire Pyrco-Δ5E sequence produced within 60 min. There was a decline of the single tryptic peptide after trypsin digestion.
- For His$_{10}$::Pavsa-Δ4D, six of the eight peptic peptides reached a peak at 5 min while the remaining two peptides had reached 70% of maximum response by 5 min and peaked by 30 min. The tryptic peptides monitored after the pepsin digest showed a rapid decline in the first 5-10 min and then a further decline over the remainder of the 60 min duration experiment.

Overall, the results suggest that all seven integral membrane bound proteins would be readily digestible in pepsin and/or trypsin.

4.1.6 Bioinformatic analyses of potential ORFs created by the transformation procedure

Study submitted:

2017. Bioinformatics Analysis of the Potential Fusion Proteins at DNA Junctions in Canola (*Brassica napus*) for Omega 3 Fatty Acids: Identity Comparison to Allergens and Toxins. Internal Nuseed report, Nuseed Pty Ltd. Report N° 2016-004

Bioinformatics analyses were performed to assess the similarity to known allergens and toxins of the putative polypeptides encoded by the ORFs (see Section 3.4.2) identified in a) the two inserts (excluding those ORFs intentionally introduced and already considered in the bioinformatics searches in Sections 4.1.5.2.1 and 4.1.5.2.2) and b) the junction region/joined insert segments of the DHA canola transgenic loci.

The bioinformatics analyses were carried out by comparing each ORF sequences with sequences present in the same databases as described for the toxin and allergen searches in Sections 4.1.5.1 and 4.1.5.2 respectively except that, for the two insert sequences, version 17 (released on 18 January 2017) of the AllergenOnline database was used and the searches were done in May and June of 2017.

Given the large number of potential ORFs for the two insert sequences (see discussion in Section 3.4.2) there needed to be a logical rationale for which ORFs to consider. For possible toxicity, the keyword limits used were sufficient. For possible allergenicity, sequences that were longer than 80 aa, with >35% identity were taken as positive findings. Sequences of longer alignments such as 100 aa, with lower identity matches (e.g. 32%) were then manually entered in the public online version of AllergenOnline.org to test whether the highest scoring 80mer had an identity >35%.

No biologically relevant identities were found between the query sequences and any toxic proteins from the NCBI Entrez database.

For the allergen searches, none of the full-length FASTA alignments were significant. The identities were below the 50% level that is likely to indicate cross-reactivity (Aalberse 2000) and below the 35% suggested by Codex (2009). There were no matches with either the sliding 80-amino acid window or with eight contiguous amino acid identity matches. No biologically relevant matches were obtained using the keyword limited search of all proteins in NCBI Entrez.

It is concluded that, in the event an unexpected translation product were to be derived from any of the hypothetical ORFs detected in the DHA canola inserts or junction regions, these putative polypeptides are not expected to possess functional cross-reactivity with known allergens or protein toxins.

4.1.7 Conclusion

DHA canola expresses eight novel proteins, Lackl-Δ12D, Picpa-ω3D, Pyrco-Δ6E, Pyrco Δ5E, Micpu-Δ6D, Pavsa-Δ5D, Pavsa-Δ4D and PAT.

The seven proteins in the n-3 LC-PUFA synthesis pathway were only detected in the developing and mature seed of DHA canola. This is expected since the genes encoding the proteins all have seed-specific promoters. The levels for all seven proteins were low with Pyrco-Δ5E having the lowest and Pavsa-Δ4D the highest expression. Given that the range of total protein in seeds was 5.95 – 18.0 mg/g of seed, the levels of the seven proteins (ranging from 20 – 740 ng/mg total protein) represent a negligible proportion of the total protein. The PAT protein levels were too low to be quantified in any of the plant parts analysed. Thus,

dietary exposure to any of the proteins would be very low.

Previous safety assessments of PAT indicate the protein would be rapidly degraded in the digestive system following ingestion and would be inactivated by heating. Additionally, updated bioinformatic studies considered in this assessment confirm the lack of any significant amino acid sequence similarity to known protein toxins or allergens.

The low levels of the seven proteins in the n-3 LC-PUFA pathway that were expressed in the DHA canola precluded their direct characterisation. An indirect approach was therefore used which relied upon: a) translation of the known DNA sequences of the seven introduced genes present in DHA canola; b) the fact that the proteins function as predicted in the plant; and c) the functionality of the seven proteins expressed in heterologous systems.

For these seven proteins, bioinformatic studies both confirmed the lack of any significant amino acid sequence similarity to know protein toxins or allergens and showed the proteins share significant homologies with proteins already consumed in food from other species. *In vitro* digestibility studies carried out on one representative protein (Picpa-ω3D, Pyrco Δ5E, and Pavsa-Δ4D) from each class of protein, using His fusion proteins expressed in, and purified from, a heterologous system suggest the proteins would be rapidly degraded in the gastro-intestinal tract following ingestion. Thermolability studies for the seven proteins could not be done but evidence from the literature suggests integral membrane proteins are not thermally stable and are functionally inactivated following heating.

Taken together, the evidence indicates that should Lackl-Δ12D, Picpa-ω3D, Pyrco-Δ6E, Pyrco Δ5E, Micpu-Δ6D, Pavsa-Δ5D, Pavsa-Δ4D or PAT be present in the diet they are unlikely to be toxic or allergenic in humans.

4.2.   n-3 LC PUFAs

As n-3 LC PUFAs are produced in significant quantities in DHA canola as a consequence of the genetic modification, an evaluation of their safety, particularly DHA, in the context of the Australian/New Zealand diet, has been undertaken and is provided in the SD2 report.

4.3   Herbicide metabolites

The presence of the PAT protein in DHA canola was intended for use in the early selection stages of product development and is not intended to confer agronomically useful tolerance to glufosinate ammonium. The trait was not selected for during product development and, as indicated in Section 4.1.3, the PAT protein is hardly detectable in plant parts. Nonetheless, for GM foods derived from crops containing a herbicide-tolerance gene, irrespective of the effectiveness of that gene, there are two issues that require consideration. The first is dealt with in this safety assessment and involves assessment of any novel metabolites that may be produced, should the herbicide be applied, to determine whether these are present in the final food and whether their presence raises any toxicological concerns. In particular, the assessment considers whether appropriate health-based guidance values (i.e. Acceptable Daily Intake [ADI] or Acute Reference Dose [ARfD]) need to be established.

The second consideration, which is separate from the GM food approval process and therefore not included as part of this safety assessment, relates to the presence of herbicide residues on the food. Any food products (whether derived from GM or non-GM sources) sold in both Australia and New Zealand must not have residue levels greater than the relevant maximum residue limit[30] (MRL). Where necessary, an MRL may have to be set.
In the case of PAT, the metabolic profiles resulting from the novel protein x herbicide

---

[30] http://www.foodstandards.gov.au/consumer/chemicals/maxresidue/Pages/default.aspx

interaction have been established through a significant history of use. The glufosinate-tolerance trait is present in lines from 24 previous applications submitted to FSANZ. The enzyme activity of PAT results in the acetylation of the free amino group of glufosinate to produce the non-herbicidal N-acetyl glufosinate. This is a well-known metabolite in glufosinate-tolerant plants and was previously considered in detail by FSANZ in cotton line LL25 (FSANZ 2006). There are no concerns that the spraying of DHA canola with glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicide and already used in the food supply.

5   Compositional analyses

The main purpose of compositional analyses is to determine if, as a result of the genetic modification, any unexpected changes have occurred to the food. These changes could take the form of alterations in the composition of the plant and its tissues and thus its nutritional adequacy. Compositional analyses can also be important for evaluating the intended effect where there has been a deliberate change to the composition of the food.

The classic approach to the compositional analyses of GM food is a targeted one. Rather than analysing every possible 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 such 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 (e.g. glycoalkaloids in potatoes).

5.1   Key seed components

Canola oil is the primary food product used for human consumption and is extracted from the seed by any one of a number of processes of which solvent extraction is the most efficient and leaves about 2% - 4% residual oil in the meal (NSW DPI 2014). The key seed components to be analysed for a comparison between transgenic and conventional canola are proximates, amino acids, fatty acids, vitamins E and K, glucosinolates, tannins, sinapine and phytic acid (OECD 2011).

5.2   Study design and conduct for key components

The Applicant submitted compositional data for both seed and meal (crude and hexane-extracted) of DHA canola. As meal is still used primarily for animal feed, the data were noted by FSANZ but are not reported in this assessment; the results for the meal were unremarkable.

Verified seed of DHA canola (B0050-027-18-20 T_3 seed – see Figure 4) and non-GM AV Jade were used for plantings at eight sites[31] across major canola growing regions of western Victoria (Australia) in 2015. Planting and crop maintenance were done according to local agronomic practices at each site. Additionally, a further seven non-GM cultivars[32] were also grown as reference lines at each site, in order to generate ranges for each analyte and hence to aid in the determination of the normal variation found in canola analyte levels. These reference lines represent an agronomically diverse range of well-adapted and widely grown open-pollinated cultivars in Australia.

Each site was planted as a randomised complete block consisting of five replicates. Seed samples of 350 – 400 g were harvested at physiological maturity from the middle two rows of each plot. A point to note is that although normal reproduction in canola occurs primarily through self-pollination, cross pollination can and does occur. In having all of the lines growing in such close proximity at each site it is inevitable that the 'purity' of seed samples will be compromised. Therefore, for example, in the fatty acid analysis (Table 10) minor levels of DHA were recorded for both AV Jade and the reference lines, where the level should be zero because conventional canola does not have the necessary enzymes to drive DHA synthesis. The explanation is that there has been some out-crossing of these non-GM lines with the DHA canola lines.

Methods of composition 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.

5.3 Statistical analyses of key components in seed

The compositional components of DHA canola and AV Jade were statistically analysed using a linear mixed model with genotype as a fixed factor and site as a random factor. Data were transformed into Statistical Analysis Software[33] (SAS) data sets and analysed using SAS® software (SAS, version 9.4). Descriptive statistics (mean and standard deviation (SD) were generated and are presented in Tables 9 – 15, which represent results from combined-site analyses. In assessing the significance of any difference between means, a P-value of 0.05 was used (i.e. a P-value of ≥0.05 was not significant).

Any statistically significant differences between DHA canola and the AV Jade control have been compared to the range compiled from the results of the seven commercial reference lines combined across all sites, to assess whether the differences are likely to be biologically meaningful. Additionally, the results for DHA canola and AV Jade have been compared to a combined literature range (where available) for each analyte, compiled from published literature for commercially available canola[34]. It is noted, however, that information in the published literature is limited and is unlikely to provide a broad reflection of the natural diversity that occurs within canola (Delourme et al. 2013). Therefore, even if means fall

---

[31] Nurrabiel (2 sites), Douglas, Green Lake, Toolondo, Gymbowen, Kaniva, Ararat.
[32] AV Zircon, AV Garnet, ATR Bonito, ATR Gem, ATR Wahoo, ATR Stingray, Monola 515TT
[33] SAS website - http://www.sas.com/technologies/analytics/statistics/stat/index.html
[34] Published literature for canola includes Wang et al (1999), Pritchard et al (2000), Szmigielska et al (2000), Marwede et al (2004), Barthet & Daun (2005), Brand et al (2007), Seberry et al (2007), Spragg & Mailer (2007), OECD (2011), Dairy One Cooperative (2011).

outside the published range, this is unlikely to raise a concern.

Seed samples were analysed for proximates, fibre, fatty acids, amino acids, minerals, vitamins, tocopherols/sterols and anti-nutrients (glucosinolates, phytic acid, and phenolics). In total, 102 individual analyte levels were considered and some groupings were also analysed. A total of 16 individual analytes had more than 33% of the observations below the assay limit of quantitation (LOQ) and were excluded from the statistical analysis. The data for 86 individual analytes were therefore analysed.

5.3.1 Proximates and fibre

The results are given in Table 9 and show there was no significant difference between the means for DHA canola and those for the control for protein, acid detergent fibre (ADF), neutral detergent fibre (NDF) and crude fibre. The level of crude fat was significantly lower in DHA canola than in the control and the levels of ash and carbohydrate were significantly higher. However, the means for all entries, for all proximates and fibre, were within the reference range and the range found in the literature (where available).

Table 9: Mean percentage ±SD of proximates and fibre in seed of DHA canola and the AV Jade control collected from eight locations

Analyte (% dw)	AV Jade	DHA canola	p-value[2]	Reference range	Combined literature range
Protein	30.1±1.2	30.4±1.2	0.3797	23.5–32.1	17.4–44.3
Crude Fat	33.2±2.9	30.5±2.7[3]	**<0.0001**	25.5–42.1	24.0–49.5
Ash	3.7±0.5	3.8±0.4[3]	**<0.0001**	2.7–4.5	3.36–6.02
Carbohydrate[1]	33.0±2.3	35.4±2.0[3]	**<0.0001**	27.3–42.3	-
ADF	11.2±1.4	11.4±1.4	0.4709	8.6–16.5	11.6–26.7
NDF	15.6±1.6	15.6±1.1	0.9676	13.6–18.1	16.49–34.72
Crude fibre	14.9±1.9	14.7±2.0	0.5936	10.9–22.6	N/A

[1] Carbohydrate determined by calculation
[2] p-values indicating significant differences are bolded and underlined.
[3] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.3.2 Fatty acids

The levels of 35 individual fatty acids and certain groupings were measured. Of these, eight individual fatty acids (16:3 n-3; 20:2 n-9; 20:3 n-6; 20:3 n-9; 20;4 n-6; 22:1 n-9, 22:2 n-6; 22:4 n-6) and two groupings (16:1 *trans* and 22:1 Total) were omitted from analysis because all or most results were below the LOQ for both DHA canola and AV Jade.

Results for the remaining 27 individual fatty acids and 13 groupings are listed for AV Jade in Table 10; no results have been given for DHA canola as these are CCI at the time of preparing this safety assessment. Of these, there were nine individual fatty acids (including EPA) and six groupings in which >33% of the results were less than the LOQ for AV Jade (as

would be expected in conventional canola) while there were quantifiable levels in all or most of the DHA canola samples. For most of these, a statistical analysis was therefore not feasible (indicated as N/A in column 3 of Table 10)

The fact that seed harvested from the conventional canola lines (i.e. AV Jade and the reference lines) contained a significant level of DHA and minor levels of a number of other long chain fatty acids none of which would be expected to be present at all is, as discussed previously (Section 5.2), indicative of the inevitable cross-pollination that occurred between these conventional lines and the DHA canola lines. It is also noted that this factor slightly skews the reference range.

Given that seven fatty acid pathway enzymes have been expressed in DHA canola, it is expected that there will be significant differences between DHA canola and the AV Jade control for a number of fatty acids and that these will also fall outside the normal range for canola. Means for 10 of the individual fatty acids, including DHA, that were statistically analysed were significantly higher in DHA canola than in the control. Means of three fatty acids, including OA and LA, that could be statistically analysed were significantly lower in DHA canola than in the control (and LA and OA in DHA canola was also lower than the reference range) as consistent with the fatty acid synthesis pathway being pushed towards production of EPA and DHA. For four of the individual fatty acids (myristic, palmitoleic, stearic, and lignoceric) levels in DHA canola and AV Jade were not significantly different.

As expected from the results for individual fatty acids, the groupings also showed significant differences between DHA canola and the control, where a statistical analysis was feasible.

In terms of *trans* fatty acids, DHA canola seed had a significantly higher total level than AV Jade and this was higher than the reference range. It is noted, however, that the relative level of *trans* fatty acids in DHA canola seed is low (<1% of total fatty acids). The major contributors to the *trans* fatty acids level were two *trans*-/*cis*- C18:3 isomers (i.e. isomers containing both *cis*- and *trans*- bonds) namely 18:3n3t,6c,9c and 18:3n3t,6t,9c that accounted for approximately 80% of the total *trans* fatty acid level and also occur at similar levels in non-GM retail vegetable oils such as canola, soybean, sunflower and corn (see e.g. Wolff 1992; Martin et al. 2008; Hou et al. 2012; Li et al. 2012; Feng et al. 2014).

The results in Table 10 represent those for crude oil. Most crude vegetable and marine oils undergo processing to produce what is known as a refined, bleached, and deodorised (RBD) oil. During deodorisation, the purpose of which is to remove volatile odour substances essentially by steam distillation under low pressure and high temperature, there is formation of further *trans* isomers (see e.g. Ruiz-Mendez and Dobarganes, 2017). Therefore, *trans* fatty acid levels in an RBD oil are usually higher than those in the crude oil. The extent of *trans* isomer formation depends on the specific processing conditions used and these can vary considerably. When not conducted carefully, deodorising crude DHA canola oil, rich in polyunsaturated fatty acids, could negatively affect other quality parameters in the final RBD product, leading to a commercially-unacceptable product. Advanced industrial technologies and good manufacturing practices associated with processing crude edible oils usually limit the formation of *trans* fatty acids and any compromising organoleptic parameters in the final product. At the time of preparing this safety assessment, the Applicant had not undertaken any commercial scale RBD production of DHA canola. However, in a limited laboratory scale (n = 2) test that produced commercially acceptable oil, the total *trans* fatty acid content in the DHA canola RBD oil was approximately 1% of total fatty acids.

The level of total fatty acids concurred with crude fat results presented in Table 9 and indicates the mean fat level in DHA canola seed is significantly lower than that in AV Jade but is not outside the reference range.

Table 10: Individual fatty acids and fatty acid groupings (mean ± SD) evaluated in seed from DHA canola and the AV Jade control

Fatty acid (% total fatty acids)	AV Jade	p-value for the comparison between AV Jade and DHA canola[1]	Reference range	Combined literature range or expected[2]
Myristic (14:0)	0.075±0.005	0.2289	0.055-0.096	Possible trace
Palmitic (C16:0)	4.310±0.110	**<0.0001**	3.625-4.818	2.5 – 7.0
Palmitoleic (16:1, n-7)	0.193±0.007	0.6390	0.150-0.250	ND – 0.6
16:1, n-9	0.053±0.003	**<0.0001**	0.041-0.079	Possible trace
Margaric (Heptadecanoic) (17:0	0.047±0.005	**0.0248**	0.040-0.060	ND-0.3
Heptadecenoic (17:1)	0.059±0.004	**<0.0001**	0.045-0.079	ND-0.3
Stearic (18:0)	2.212±0.082	0.0868	1.410-2.260	0.8–3.0
Vaccenic (18;1, n-7)	2.678±0.079	**<0.0001**	2.701-3.079	Expect around 4.0[35]
Oleic (18:1)	57.069±1.477	**<0.0001**	49.157-72.679	51.0-70.0
Linoleic (18:2, n-6)	19.341±0.827	**<0.0001**	11.590-23.260	15.0–30.0
18:2, n-9	0.01±0.02	**N/A**	0.00 – 0.09	Expect zero
α-Linolenic (18:3, n-3)	11.185±0.743	**<0.0001**	3.905-12.080	5.0–14.0
γ-Linolenic (18:3, n-6)	0.01±0.02	**N/A**	0.00–0.43	Expect zero
Stearidonic (18:4, n-3)	0.1±0.1	**N/A**	0.00-2.30	Expect zero
Arachidic (20:0)	0.475±0.012	**<0.0001**	0.422-0.730	0.2-1.2
Eicosenoic (20:1)	0.951±0.025	**<0.0001**	0.878-1.590	3.0–15.0
Eicosadienoic (20:2)	0.060±0.005	**<0.0001**	0.050-0.190	ND-0.1
Eicosatrienoic (20:3, n-3)	0.01±0.03	**N/A**	0.00-0.26	Expect zero
Eicosatetraenoic 20:4, n-3)	0.04-0.05	**N/A**	0.00-1.13	Expect zero
Eicosapentaenoic (20:5, n-3)	0.01±0.02	**N/A**	0.00-0.40	Expect zero
Behenic (22:0)	0.189±0.007	**<0.0001**	0.180-0.387	ND–0.6
Docosatetraenoic (22:4, n-3)	0.0±0.01	**N/A**	0.00-0.23	Expect zero

---

[35] Very limited data available – value of 4.0% obtained from Vingering et al (2010)

Fatty acid (% total fatty acids)	AV Jade	p-value for the comparison between AV Jade and DHA canola[1]	Reference range	Combined literature range or expected[2]
Docosapentiaenoic (22:5, n-3)	0.03±0.05	**N/A**	0.00-0.09	Expect zero
Docosapentaenoic (22:5, n-6)	0.00	**N/A**	0.00-0.09	Expect zero
Docosahexaenoic (22:6, n-3)	0.239±0.284	**<0.0001**	0.00-0.338	Expect zero
Lignoceric (24:0)	0.100±0.006	0.1906	0.096-0.210	ND–0.3
Nervonic (24:1)	0.096±0.007	**<0.0001**	0.00-0.170	ND–0.4
16:1 Total	0.267±0.012	**<0.0001**	0.214-0.335	Expect around 0.2
18:1 Total	59.823±1.441	**<0.0001**	51.934-74.358	Expect around 60
18:2 Total	19.441±0.825	**<0.0001**	11.706-23.361	Expect around 20
18:3 Total	11.278±0.767	**<0.0001**	3.930-12.190	5.0–14.0
18:4 Total	0.10±0.10	**N/A**	0.00-2.30	Expect zero
20:3 Total	0.01-0.03	**N/A**	0.00-0.62	Expect zero
20:4 Total	0.05±0.05	**N/A**	0.00-1.15	Expect zero
22:5 Total	0.03±0.05	**N/A**	0.00-1.13	Expect zero
24:1 Total	0.096±0.007	**<0.0001**	0.00-0.170	ND–0.4
18:1 *trans*	0.01±0.02	**N/A**	0.00-0.08	Expect <0.1
18:2 *trans*	0.00	**N/A**	0.00-0.06	Expect <0.1
18:3 *trans*[3]	0.077	**Not done**	0.00-0.161	Expect <0.2
Total *trans* fatty acids	0.108±0.033	**<0.0001[4]**	0.046-0.232	Expect <0.5
Total fatty acids (%dw)	29.91±1.95	**<0.0001**	23.94-35.78	30-48

[1] N/A indicates that while there were quantifiable levels of the fatty acid present in DHA canola, > 33% of AV Jade values were less than the LOQ and therefore a statistical analysis could not meaningfully be done. Mauve shading represents DHA canola means that were significantly lower than the AV Jade means; orange shading represents DHA canola means that were significantly higher

[2] The literature range reflects values for conventional unmodified canola but it is noted that there are conventionally-bred, high oleic acid canola lines that have different fatty acid profiles.
ND = non detectable

[3] Levels for 18:3 *trans* were obtained from preliminary analyses and were not included as part of the compositional data in the unpublished studies supplied in the original A1143 dossier. No statistical analysis was done but the Applicant indicated the DHA canola mean was higher than the AV Jade mean. The 18:3 *trans* fatty acids in DHA canola accounted for approximately 80% of the total *trans* fatty acids.

[4] Total *trans* fatty acids in DHA canola were <1% of total fatty acids

5.3.3   Amino acids

Levels of 18 amino acids were measured. Since asparagine and glutamine are converted to aspartate and glutamate respectively during the analysis, levels for aspartate include both aspartate and asparagine, while glutamate levels include both glutamate and glutamine. Results of the analysis are given in Table 11.

Mean levels of alanine, aspartate, glycine, lysine, methionine, threonine and tyrosine in DHA canola were significantly higher than those in the control while the mean level of proline was significantly lower. However, the means for all entries, for all amino acids, were within the reference range. Interestingly, the means for the majority of amino acids in both AV Jade and DHA canola were higher than reported in the literature. There is no obvious explanation for this.

Table 11: Mean ±SD amino acid composition (% dw) in seed from DHA canola and the AV Jade control collected from eight locations

Amino acid % dw	AV Jade	DHA canola	p-value[1]	Reference range	Combined literature range
Alanine	1.239±0.049	1.268±0.046[2]	**0.0053**	0.999-1.340	0.71–1.38
Arginine	1.923±0.092	1.919±0.087	0.8397	1.480-2.090	0.93–2.46
Aspartate	2.164±0.106	2.282±0.097[2]	**<0.0001**	1.680-2.420	1.20–2.03
Cystine	0.754±0.037	0.743±0.038	0.1520	0.580-0.820	0.32–0.52
Glutamate	5.681±0.258	5.599±0.269	0.1426	4.360-6.170	3.23–4.71
Glycine	1.519±0.062	1.584±0.061[2]	**<0.0001**	1.240-1.660	0.82–2.22
Histidine	0.843±0.032	0.845±0.036	0.8145	0.677-0.922	0.41–0.82
Isoleucine	1.218±0.052	1.218±0.048	0.9692	0.931-1.310	0.62–1.02
Leucine	2.129±0.092	2.120±0.086	0.6732	1.660-2.300	1.07–1.99
Lysine	1.890±0.107	1.948±0.129[2]	**0.0171**	1.490-2.140	0.96–1.85
Methionine	0.611±0.023	0.623±0.027[2]	**0.0197**	0.490-0.670	0.27–0.52
Phenylalanine	1.217±0.054	1.202±0.046	0.1830	0.949-1.310	0.64–1.07
Proline	1.925±0.086	1.865±0.091[2]	**0.0010**	1.460-2.050	0.85–3.74
Serine	1.279±0.050	1.292±0.051	0.2342	1.020-1.380	0.69–1.55
Threonine	1.280±0.044	1.318±0.045[2]	**<0.0001**	1.040-1.360	0.74–1.30
Tryptophan	0.456±0.020	0.453±0.021	04308	0.340-0.490	0.20–0.37

Amino acid % dw	AV Jade	DHA canola	p-value[1]	Reference range	Combined literature range
Tyrosine	0.789±0.035	0.817±0.029[2]	**<0.0001**	0.644-0.839	0.51–0.92
Valine	1.562±0.063	1.566±0.068	0.7330	1.160-1.650	0.8–1.55

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.3.4   Minerals

Levels of ten minerals were measured and the means are given in Table 12. Levels of iron, potassium and zinc in DHA canola were significantly higher than those in the control while calcium was significantly lower. However, the means were within both the reference range and the combined literature range (where available).

Table 12: Mean ±SD mineral composition (% dw) in seed from DHA canola and the AV Jade control collected from eight locations

Mineral % dw	AV Jade	DHA canola	p-value[1]	Reference range	Combined literature range
Calcium	0.356±0.061	0.312±0.048[2]	**<0.0001**	0.204-0.488	0.36–0.72
Copper	0.0002±0.0001	0.0003±0.0003	0.1285	0.0002-0.002	0.0001–0.0005
Iron	0.005±0.001	0.007±0.001[2]	**<0.0001**	0.004-0.008	ND–0.09
Magnesium	0.308±0.021	0.308±0.021	0.9910	0.262-0.370	0.27–0.42
Manganese	0.003±0.001	0.003±0.001	0.7774	0.002-0.004	0.0033–0.006
Phosphorus	0.655±0.122	0.669±0.123	0.0639	0.365-0.869	0.54–0.89
Potassium	0.666±0.093	0.782±0.082[2]	**<0.0001**	0.532-0.915	0.70–1.02
Sodium	0.002±0.001	0.003±0.002	0.9508	<LOQ-0.010	N/A
Sulphur	0.512±0.031	0.510±0.033	0.7002	0.380-0.650	N/A
Zinc	0.004±0.001	0.005±0.001[2]	**0.0026**	0.003-0.006	ND–0.01

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.3.5   Vitamins

Levels of nine vitamins were measured (Table 13). For many of these, the values were close to the LOQ. There was no significant difference between the mean for DHA canola and the control for folic acid. For the remaining vitamins, means for DHA canola were significantly higher than those for AV Jade. This is not unremarkable given that certain vitamins,

especially the B vitamins, have anti-oxidant activity, and hence may increase (along with tocopherols and sterols – see below) with increased levels of unsaturated fatty acids. However all means for all vitamins, were within the reference range.

Table 13: Mean ±SD vitamin composition (% dw) in seed from DHA canola and the AV Jade control collected from eight locations

Vitamin (% dw)	AV Jade	DHA canola	p-value[1]	Reference range	Combined literature range
Biotin	0.05±0.00	0.07±0.00[2]	**<0.0001**	ND-0.1	N/A
Choline	262.7±21.6	276.0±23.3[2]	**0.0249**	195.37-381.31	N/A
Folic acid	0.12±0.04	0.12±0.03	0.9802	ND-0.6	N/A
Niacin (B3)	9.66±0.96	15.14±1.91[2]	**<0.0001**	8.4-16.8	N/A
Pantothenic acid (B5)	0.46±0.10	0.56±0.11[2]	**<0.0001**	0.2-0.8	N/A
Pyridoxine (B6)	0.54±0.06	0.85±0.10[2]	**<0.0001**	0.4-1.0	N/A
Riboflavin (B2)	0.32±0.06	0.35±0.03[2]	**0.0241**	0.2-0.6	N/A
Thiamin (B1)	1.29±0.20	1.48±0.23[2]	**0.0023**	0.2-2.3	N/A
Vitamin K1	0.05±0.01	0.05±0.01[2]	**0.0205**	0.0-0.1	N/A

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.3.6   Tocopherols and sterols

Tocopherols and sterols are the major non-saponifiable components of vegetable oils. The level of tocopherols (which are anti-oxidants) is governed by the level of unsaturated fatty acids; an increase in unsaturation will result in the formation of higher levels of anti-oxidants to protect the oil. This also applies, to a lesser extent, with some sterols. An increase in both would therefore be expected in DHA canola. During the deodorisation stage of oil refining both tocopherols and sterols are largely removed (NSW DPI 2014).

The levels of four individual tocopherols and 12 individual phytosterols were tested. For three of the phytosterols (δ_7_aveansterol, sitostanol and δ_7_stigmastenol) >33% of the values were below the LOQ and they were excluded from the statistical analysis. It is also noted that for a further five of the phytosterols (cholesterol, clerosterol, delta-5-avenasterol, 24-methylene cholesterol, and delta-5 24-stigmastadienol) the values were very low and close to the LOQ.

Results in Table 14 show that the total tocopherol levels were significantly higher in DHA canola than in the control and that this was largely attributable to a significantly higher level of α-tocopherol. Nonetheless, levels were within the reference range. While all sterol levels were low and most were close to the LOQ, campesterol and sitosterol made up the greatest components of the total sterols. Most sterol means were significantly higher in DHA canola than in AV Jade but mostly fell within the reference range (the exception was

δ_5_avenasterol where both the AV Jade and DHA canola means were higher than the range).

Table 14: Mean ±SD tocopherol and sterol composition in seed from DHA canola and the AV Jade control collected from eight locations

Analyte	AV Jade	DHA canola	p-value[1]	Reference range	Combined literature range
α-tocopherol (% dw)	11.9±6.6	15.7±5.8[2]	**<0.0001**	10.9-31.3	7.1–10.8.
β-tocopherol (% dw)	0.08±0.10	0.12±0.05	**0.0021**	ND–0.6	N/A
γ-tocopherol (% dw)	21.2±1.8	22.8±1.9	0.0648	10.2-72.2	N/A
δ-tocopherol (% dw)	0.5±0.5	0.3±0.1	0.2990	ND-13.5	N/A
Total tocopherols (% dw)	33.7±7.2	38.9±5.9[2]	**0.0001**	24.5-96.9	N/A
Brassicasterol (µg/g)	0.112±0.005	0.052±0.004[2]	**<0.0001**	0.045-0.170	N/A
Campesterol (µg/g)	0.287±0.010	0.385±0.018[2]	**<.0001**	0.226-0.397	N/A
Cholesterol (µg/g)	0.002±0.002	0.002±0.003	0.7472	ND-0.020	N/A
Clerosterol (µg/g)	0.006±0.000	0.006±0.000	**0.0009**	0.004-0.006	N/A
δ_5_avenasterol (µg/g)	0.036±0.006	0.044±0.008[2]	**<0.0001**	0.008-0.037	N/A
Sitosterol (µg/g)	0.551±0.028	0.579±0.036[2]	**<0.0001**	0.346-0.580	N/A
Stigmasterol (µg/g)	0.003±0.000	0.000±0.001[2]	**<0.0001**	ND-0.005	N/A
24-methylene cholesterol (µg/g)	0.013±0.005	0.011±0.004[2]	**0.0013**	0.003-0.020	N/A
δ_5 24-stigmastadienol (µg/g)	0.007±0.001	0.009±0.001[2]	**<0.0001**	0.003-0.009	N/A
Total phytosterols (µg/g)	1.025±0.040	1.106±0.061[2]	**<0.0001**	0.702-1.097	N/A

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.3.7 Anti-nutrients

Anti-nutrients (glucosinolates, phenolics, and phytates) in canola meal make the meal problematic for food use (Tan et al. 2011b). Canola seeds are much richer in phenolic compounds than other oilseeds (Naczk et al. 1998) but most of the phenolics remain in the meal after pressing and contribute to the dark colour, bitter taste and astringency of the meal. Phenolic acids and condensed tannins are the predominant phenolics. Sinapine is the main phenolic ester in canola seeds; ferulic acid and *p*-coumaric acids are minor phenolics. While there may be glucosinolates present in the canola seed, there are essentially no

glucosinolates present in the most widely used edible product of canola – its oil. These remain behind in the meal when the oil is extracted.

Levels of 10 glucosinolates plus phytic acid and the phenolics sinapine, ferulic acid, *p*-coumaric acid and soluble condensed tannins were measured. Of these, gluconapoleiferin, neoglucobrassicin, *p*-coumaric acid and soluble tannins condensed all had >33% of values less than the LOQ and were therefore excluded from statistical analysis..

The results (Table 15) can be summarised as follows:

- The only glucosinolate where the mean for DHA canola differed significantly from the mean for AV Jade was glucobrassicin. The DHA canola mean, did, however, fall within the reference range.
- The mean for sinapine was significantly lower in DHA canola than in the control but was within the reference range.
- For all anti-nutrients, the means for both DHA canola and AV Jade were within the reference range.

Table 15: Mean ±SD anti-nutrient levels in seed from DHA canola and the AV Jade control collected from eight locations

Anti-nutrient	AV Jade	DHA canola	p-value[1]	Reference range	Combined literature range
Epi progoitrin (µmol/g dw)	0.094±0.052	0.096±0.053	0.9182	ND-0.300	N/A
Glucoalyssin (µmol/g dw)	0.349±0.147	0.363±0.150	0.7523	ND-1.800	N/A
Glucobrassicanapin (µmol/g dw)	0.311±0.155	0.250±0.121	0.0905	ND-1.300	N/A
Glucobrassicin (µmol/g dw)	0.205±0.055	0.282±0.073[2]	**<0.0001**	0.090-0.550	N/A
Gluconapin (µmol/g dw)	2.166±0.723	1.972±0.681	0.2947	0.417-6.390	N/A
Gluconasturtin (µmol/g dw)	0.094±0.052	0.135±0.084	0.1092	ND-0.520	N/A
Progoitrin (µmol/g dw)	4.914±1.895	4.936±1.874	0.9660	0.838-17.00	N/A
4-hydroxyglucobrassicin (µmol/g dw)	3.938±0.769	3.849±0.964	0.6366	ND-5.540	N/A
Phytic acid (%dw)	1.918±0.433	1.895±0.440	0.7961	0.840-2.500	N/A
Sinapine (%dw)	1.264±0.078	1.191±0.070[2]	**0.0002**	0.876-1.463	N/A
Ferulic acid (%dw)	137.238±23.68	130.084±20.96	0.1768	88.91-217.50	

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.3.8   Summary of analyses of key components

A summary of the statistically significant differences in the mean individual analyte levels (excluding the fatty acids) found between seed of DHA canola and the AV Jade control is

provided in Table 16.

Table 16: Summary of mean individual analyte levels (excluding fatty acids) found in seed of DHA canola that are significantly (p<0.05) different from those found in seed of the control AV Jade

Analyte	AV Jade	DHA canola[1]	p-value	DHA canola within reference range?	DHA canola within literature range?
Crude Fat (% dw)	33.2±2.9	30.5±2.7	**<0.0001**	yes	yes
Ash (% dw)	3.7±0.5	3.8±0.4	**<0.0001**	yes	yes
Carbohydrate (% dw)	33.0±2.3	35.4±2.0	**<0.0001**	yes	N/A
Alanine (% dw)	1.239±0.049	1.268±0.046	**0.0053**	yes	yes
Aspartate (% dw)	2.164±0.106	2.282±0.097	**<0.0001**	yes	No – control also higher
Glycine (% dw)	1.519±0.062	1.584±0.061	**<0.0001**	yes	yes
Lysine (% dw)	1.890±0.107	1.948±0.129	**0.0171**	yes	No – control also higher
Methionine (%dw)	0.611±0.023	0.623±0.027	**0.0197**	yes	No – control also higher
Proline (% dw)	1.925±0.086	1.865±0.091	**0.0010**	yes	yes
Threonine (% dw)	1.280±0.044	1.318±0.045	**<0.0001**	yes	no
Tyrosine (% dw)	0.789±0.035	0.817±0.029	**<0.0001**	yes	yes
Calcium (% dw)	0.356±0.061	0.312±0.048	**<0.0001**	yes	No – control also lower
Iron	0.005±0.001	0.007±0.001	**<0.0001**	yes	yes
Potassium (% dw)	0.666±0.093	0.782±0.082	**<0.0001**	yes	yes
Zinc (% dw)	0.004±0.001	0.005±0.001	**0.0026**	yes	yes
Biotin	0.05±0.00	0.07±0.00	**<0.0001**	yes	N/A
Choline (% dw)	262.7±21.6	276.0±23.3	**0.0249**	yes	N/A
Niacin (% dw)	9.7±1.0	15.1±1.9	**<0.0001**	yes	N/A
Pantothenic acid (% dw)	0.5±0.1	0.6±0.1	**<0.0001**	yes	N/A
Pyridoxine (% dw)	0.5±0.1	0.9±0.1	**<0.0001**	yes	N/A
Riboflavin (% dw)	0.32±0.06	0.35±0.03	**0.0241**	yes	N/A

Analyte	AV Jade	DHA canola[1]	p-value	DHA canola within reference range?	DHA canola within literature range?
Thiamin (% dw)	1.3±0.2	1.5±0.2	**0.0023**	yes	N/A
Vitamin K1 (% dw)	0.0±0.0	0.1±0.0	**0.0205**	yes	N/A
α-tocopherol (% dw)	11.9±6.6	15.7±5.8[2]	**<0.0001**	yes	No – control also higher
β-tocopherol (% dw)	0.08±0.10	0.12±0.05	**0.0021**	yes	N/A
Brassicasterol (µg/g)	0.112±0.005	0.052±0.004	**<0.0001**	yes	N/A
Campesterol (µg/g)	0.287±0.010	0.385±0.018	**<0.0001**	yes	N/A
δ_5_avenasterol (µg/g)	0.036±0.006	0.044±0.008	**<0.0001**	no	N/A
Sitosterol (µg/g)	0.551±0.028	0.579±0.036	**<0.0001**	yes	N/A
Stigmasterol (µg/g)	0.003±0.000	0.000±0.001	**<0.0001**	yes	N/A
24-methylene cholesterol (µg/g)	0.013±0.005	0.011±0.004	**0.0013**	yes	N/A
δ_5 24-stigmastadienol (µg/g)	0.007±0.001	0.009±0.001	**<0.0001**	yes	N/A
Glucobrassicin (µmol/g dw)	0.205±0.055	0.282±0.073	**<0.0001**	yes	N/A
Sinapine (%dw)	1.264±0.078	1.191±0.070	**0.0002**	yes	N/A

[1] mauve shading represents DHA canola means that are significantly lower than the AV Jade means while orange shading represents DHA canola means that are significantly higher.

5.4 Conclusions of the compositional analyses

Detailed compositional analyses were conducted on seed from DHA canola. Seed samples harvested from DHA canola grown in eight different locations were analysed for proximates, fibre, fatty acids, amino acids, minerals, vitamins, tocopherols/sterols and anti-nutrients (glucosinolates, phytic acid, and phenolics). The levels of each analyte in DHA canola were compared to levels in: a) the non-GM parental line, AV Jade; b) eight non-GM commercial reference lines grown at the same locations; and c) levels recorded in the literature.

In total, 106 individual analytes and some groupings were analysed. A total of 15 individual analytes (including eight fatty acids) had more than 33% of the observations below the assay limit of quantitation (LOQ) for both DHA canola and the AV Jade control and were excluded from the statistical analysis. For a further nine individual fatty acids, statistical analysis could not be performed because levels were at, or close to, zero in AV Jade but were detectable in DHA canola. The data for 91 individual analytes (including these nine fatty acids) were therefore considered.

Of these 91 individual analytes, a total of 48 (including 14 of the statistically analysed individual fatty acids) differed significantly between DHA canola and AV Jade. The changes in fatty acid profile were consistent with those expected as a result of the introduction of seven enzymes from the n-3 LC-PUFA synthesis pathway leading to the production of DHA

from OA.

For the 34 non-fatty acid analytes that showed significant differences, the levels of all but δ_5_avenasterol were within the reference range generated from the non-GM varieties. Comparative information in the literature was not available for many of the analytes but, where it was available, it is noted in five cases where the mean level of the DHA canola analyte fell outside the literature range it also fell outside for AV Jade. This highlights the limitations of the literature in capturing the range of variability found in canola.

The mean fatty acid profile of DHA canola seed in Table 10 is expected to be the same for crude oil extracted from the seed. On this basis, the mean *trans* fat content of crude DHA canola oil would be within the range analysed in retail canola oil. Deodorisation of DHA canola oil (as for all RBD oils from all sources) may increase the *trans* fat content but the extent will depend on processing conditions appropriate for a highly polyunsaturated food. Given the likely care that would be taken by processors to limit the formation of *trans* fatty acids across all RBD oils, the *trans* fat content of DHA canola oil is not expected to vary significantly from other retail oils.

Apart from the intended change to fatty acid profile and a small increase in *trans* fatty acids, seed from DHA canola is otherwise compositionally equivalent to seed from conventional canola varieties.

6   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 in the levels of certain nutrients in the GM food, additional nutritional assessment may assist 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. Evidence indicates that feeding studies using target livestock species will add little to the safety assessment (see e.g. OECD 2003; Bartholomaeus et al. 2013; Herman and Ekmay 2014).

In the case of DHA canola, there are significant changes in the fatty acid profile of the seed particularly the production of DHA. Preformed DHA is mainly available in the human diet through seafood and marine oils.

FSANZ has conducted a separate nutrition risk assessment of DHA canola to determine the potential risks associated with the intake of the contained n-3 LC PUFAs by the general population and whether this may cause any adverse health effects (see Supporting Document 2). However, the public health risk of *trans* fatty acids from DHA canola is given in the next subsection.

6.1   *Trans* fatty acids

To put the *trans* fatty acid level in DHA canola in perspective, the *trans* fat content of retail canola oil in Australia and New Zealand, measured in a 2013 survey, was 0.6–1.7% *trans* fatty acids (FSANZ 2014). This is less than the prescribed maximum of 2% *trans* fatty acids for the four algal sources of DHA listed in Schedule 3 – Identity and Purity in the Code.

In relation to public health risk, the mean *trans* fat content of vegetable oils used for the 2009 estimate of *trans* fatty acid intake (NSW Food Authority 2009) and the summary results of the 2011–12 National Nutrition and Physical Activity Survey component of the Australian Health Survey (FSANZ 2014) declined over the period and both were lower than 2% *trans* fatty acids. The more recent estimate of dietary intake showed that *trans* fatty acid intake was low, with the mean intake of all Australians aged two years and above (1.4 g/day) representing 0.6% of dietary energy. Also ruminant sources (milk, butter, meat) were the major dietary source of *trans* fat. Compared with international recommendations, a large majority of Australians had intakes below the WHO recommendation of less than 1% dietary energy. This was also the case for Australia and New Zealand in 2009.

Noting the *trans* fatty acid levels in DHA canola seed and oil discussed in section 5.3.2, previous consideration of the consumption data of *trans* fatty acids in the Australian and New Zealand diets indicates that consumption of food products derived from DHA canola would not pose a public health concern due to intake of *trans* fatty acids.

7   References[36]

Aalberse RC (2000) Structural biology of allergens. J Allergy Clin Immunol 106(2):228–238.

Abedi E, Sahari MA (2014) Long-chain polyunsaturated fatty acid sources and evaluation of their nutritional and functional properties. Food Sci Nutr 2(5):443–463

Ahmad M, Hirz M, Pichler H, Schwab H (2014) Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol 98:5301–5317

Altmann F (2007) The role of protein glycosylation in allergy. Int Arch Allergy Immunol 142:99–115.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

AOF (2016) Australian oilseeds. Australian Oilseeds Federation. http://www.australianoilseeds.com/

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

Barakoni R, Awal S, Christie A (2015) Growth performance of the marine microalgae *Pavlova salina* and *Dunaliella tertiolecta* using different commercially available fertilizers in natural seawater and inland saline ground water. Journal of Algal Biomass Utilization 6(1):15–25

Barthet VJ, Daun JK (2005) Effect of sprouting on the quality and composition of canola seed and oil. J Am Oil Chem Soc 82(7): 511-517

Bartholomaeus A, Parrott W, Bondy G, Walker K (2013) The use of whole food animal studies in the safety assessment of genetically modified crops: Limitations and recommendations. Crit Rev Toxicol 43(S2):1–24

Baxevanis AD (2005) Assessing Pairwise Sequence Similarity: BLAST and FASTA. Ch 11 In: Baxevanis AD, Ouellette BFF (eds) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. John Wiley & Sons, Inc., p. 295–324

Belide S, Zhou X-R, Kennedy Y, Lester G, Shrestha P, Petrie JR, Singh S (2013) Rapid expression and validation of seed-specific constructs in transgenic LEC_2 induced somatic embryos of *Brassica napus*. Plant Cell Tissue Organ Cult 113(3):543–553

---

[36] All website references were current as of 19 May 2017

Bevan M, Barnes WM, Chilton M-D (1983) Structure and transposition of the nopaline synthase gene region of T-DNA. Nucleic Acids Res 11:369–385

Bhalla PL, Singh MB (2008) Agrobacterium-mediated transformation of *Brassica napus* and *Brassica oleracea*. Nat Protoc 3(2):181–189

Bogdanov S (2016) Pollen: production, nutrition and health: a review. Bee Product Science. www.bee-hexagon.net/files/file/fileE/Health/PollenBook2Review.pdf

Bonnardeaux J (2007) Uses for canola meal. Department of Agriculture and Food, Government of Western Australia,

Bos C, Airinei G, Mariotti F, Benamouzig R, Berot S, Evrard J, Fenart E, Tome D, Gaudichon C (2007) The poor digestibility of rapeseed protein is balanced by its very high metabolic utilization in humans. J Nutr 137:594–600

Bowie JU (2001) Stabilizing membrane proteins. Curr Opin Struct Biol 11(4):397–402

Bradbury JF (1986) Guide to Plant Pathogenic Bacteria. Cambridge News Ltd, Aberystwyth

Brand TS, Smith N, Hoffman LC (2007) Anti-nutritional factors in canola produced in the Western and Southern Cape areas of South Africa. S Afr J Anim Sci 37(1):45-50.

Bushey DF, Bannon GA, Delaney BF, Graser G, Hefford M, Jiang X, Lee TC, Madduri KM, Pariza M, Privalle LS, Ranjan R, Saab-Rincon G, Schafer BW, Thelen JJ, Zhang JXQ, Harper MS (2014) Characteristics and safety assessment of intractable proteins in genetically modified crops. Regul Toxicol Pharmacol 69:154–170

Campbell L, Rempel CB, Wanasundara JPD (2016) Canola/rapeseed protein: future opportunities and directions - workshop proceedings of IRC 2015. Plants 5(2):7 pp

CCC (2012) Steps in oil and meal processing. Canola Council of Canada. http://www.canolacouncil.org/oil-and-meal/what-is-canola/how-canola-is-processed/steps-in-oil-and-meal-processing/

CCC (2014) Growth stages of the canola plant. Canola Council of Canada. http://www.canolacouncil.org/canola-encyclopedia/crop-development/growth-stages/

CERA (2011) A review of the environmental safety of the PAT protein. Environ Biosafety Res 10:73–101.

Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Correa M, Da Silva C, Just J, Falentin C, and others (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science 345(6199):950–953.

Chaudhary, S., Van Rooijen, G., Moloney, M., Singh, S. (2001) Flax seed promoters. Patent WO 2001/016340 A1. https://www.lens.org/lens/patent/WO_2001_016340_A1/fulltext

Chauhan JS, Rao A, Raghava GPS (2013) *In silico* platform for prediction of N-, O-, and C- glycosites in eukaryotic protein sequences. PLoS ONE 8(6):e67008.

Cho M-J, Widholm JM, Vodkin LO (1995) Casettes for seed-specific expression tested in transformed embryogenic cultures of soybean. Plant Mol Biol Report 13(3):255–269

Chuang G-Y, Boyington JC, Joyce MG, Zhu J, Nabel GJ, Kwong PD, Georgiev I (2012) Computational prediction of N-linked glycosylation incorporating structural proteins and patterns. Bioinformatics 28(17):2249–2255.

Codex (2009) Foods derived from modern biotechnology, 2nd edition. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme, Food and Agriculture Organization of the United Nations, Rome. http://www.codexalimentarius.org/standards/thematic-publications/

Considine DM, Considine GD (1982) Foods and food production encyclopedia. Van Nostrand Reinhold Company, New York

Coutu C, Brandle J, Brown D, Brown K, Miki B, Simmonds J, Hegedus DD (2007) pORE: a modular

binary vector series suited for both monocot and dicot plant transformation. Transgenic Res 16(6):771–781

Dairy One Cooperative Inc. (2011) Canola seed accumulated crop years: 05/01/2000 through 04/30/2011. Dairy One Cooperative Inc. Feed Composition Library. http://www.dairyone.com/Forage/FeedComp/mainlibrary.asp

Deblaere R, Reynaerts A, Höfte H, Hernalsteens J-P, Leemans J, van Montagu M (1987) Vectors for cloning in plant cells. Methods Enzymol 153:277–292

Delaney B, Astwood JD, Cunny H, Eichen Conn R, Herouet-Guicheney C, MacIntosh S, Meyer LS, Privalle LS, Gao Y, Mattsson J, Levine M, ILSI (2008) Evaluation of protein safety in the context of agricultural biotechnology. Food Chem Toxicol 46:S71–S97

Delourme R, Falentin C, Fomeju BF, Boillot M, Lassalle G, André I, Duarte J, Gauthier V, Lucante N, Marty A, Pauchon M, Pichon J-P, Ribière N, Trotoux G, Balnchard P, Rivière N, Martinant J-P, Pauquet, J (2013) High-density SNP-based genetic map development and linkage disequilibrium assessment in Brassica napus L. BMC Genomics 14:e120

Dhaese P, de Greve H, Gielen J, Seurinck J, van Montagu M, Schell J (1983) Identification of sequences involved in the polyadenylation of higher plant nuclear transcripts using *Agrobacterium* T-DNA genes as models. EMBO J 2(3):419–426

Dovuro Seeds (2008) Tarcoola and AV-Jade seeds from Dovuro seeds. http://www.fatcow.com.au/c/dovuro-seeds/tarcoola-and-av-jade-seeds-from-dovuro-seeds-n810708

Dröge W, Broer I, Pühler A (1992) Transgenic plants containing the phosphinothricin-N-acetyltransferase gene metabolize the herbicide L-phosphinothricin (glufosinate) differently from untransformed plants. Planta 187(1):142–151

Durrett TD, Benning C, Ohlrogge J (2008) Plant triacylgycerols as feedstocks for the production of biofuels. Plant J 54:593–607

FAOSTAT3 (2015) Online database of the Food and Agriculture Organization of the United Nations. http://faostat3.fao.org/browse/T/TP/E

Fard NA, Minuchehr Z, Mousavi A (2013) Allergenicity study of genetically modified herbicide resistant crops (bioinformatics assessment). Bulletin of Environment, Pharmacology and Life Sciences 2(3):24–32

Feng H, Sui X, Chang Y, Qi B, Zhang Y, Li Y, Jiang L (2014) Effect of extruding full-fat soy flakes on trans fat content. The Scientific World Journal 2014:Article ID 427423, 6 pages.

Fleet GH (2006) The commercial and community significance of yeasts in food and beverage production. Ch 1 In: Querol A, Fleet GH (eds) Yeasts in food and beverages. Springer-Verlag, Berlin, p. 1–12

FSANZ (2006) Application A533 - Food Derived from Glufosinate-Ammonium-Tolerant Cotton Line LL25. http://www.foodstandards.gov.au/code/applications/pages/applicationa533foodd2440.aspx

FSANZ (2014) Technical evaluation for labelling review recommendation 13: mandatory declaration of trans fatty acids. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/consumer/labelling/review/Pages/Technical-evaluation-for-recommendation-13-%28trans-fatty-acids%29-.aspx

Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA (1987) The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucleic Acids Res 15(8):3257–3273

Geertsma ER, Groenveld M, Slotboom D-J, Poolman B (2008) Quality control of overexpressed membrane proteins. Proc Natl Acad Sci USA 105(15):5722–5727.

Gillette MA, Carr SA (2013) Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat Methods 10:28–34.

Goodman RE (2006) Practical and predictive bioinformatics methods for the identification of potentially

cross-reactive protein matches. Mol Nutr Food Res 50:655–660

Gororo N, Burton W (2006) AV Jade. Plant Varieties Journal 19(4):145–147

Greenwell HC, Laurens LML, Shields RJ, Lovitt RW, Flynn KJ (2010) Placing microalgae on the biofuels priority list: a review of the technological challenges. J R Soc Interface 7:703–726

Guschina IA, Harwood JL (2006) Lipids and lipid metabolism in eukaryotic algae. Prog Lipid Res 45:160–186

Hall Jr G, Allen GC, Loer DS, Thompson WF, Spiker S (1991) Nuclear scaffolds and scaffold-attachment regions in higher plants. Proc Natl Acad Sci USA 88(20):9320–9324

Halweg C, Thompson WF, Spiker S (2005) The Rb7 matrix attachment region increases the likelihood and magnitude of transgene expression in tobacco cells: a flow cytometric study. Plant Cell 17:418–429

Hammond B, Kough J, Herouet-Guicheney C, Jez JM (2013) Toxicological evaluation of proteins introduced into food crops. Crit Rev Toxicol 43(S2):25–42

Harwood JL, Guschina IA (2009) The versatility of algae and their lipid metabolism. Biochimie 91:679–684

Health Canada (2016) Mustard; a priority food allergen. Government of Canada. https://www.canada.ca/en/health-canada/services/food-nutrition/reports-publications/food-safety/mustard-priority-food-allergen.html

Herman RA, Ekmay R (2014) Do whole-food animal feeding studies have any value in the safety assessment of GM crops? Regul Toxicol Pharmacol 68:171–174

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 fluidInt J Food Sci Nutr 58:125–141

Hérouet C, Esdaile DJ, Mallyon BA, Debruyne E, Schulz A, Currier T, Hendrickx K, van der Klis R-J, Rouan D (2005) Safety evaluation of the phosphinothricin acetyltransferase proteins encoded by the *pat* and *bar* sequences that confer tolerance to glufosinate-ammonium herbicide in transgenic plants. Regul Toxicol Pharmacol 41:134–149

Hou J-C, Wang F, Wang Y-T, Xu J, Zhang C-W (2012) Assessment of *trans* fatty acids in edible oils in China. Food Control 25:211–215.

Huby RDJ, Dearman RJ, Kimber I (2000) Why are some proteins allergens? Toxicol Sci 55:235–246.

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236(4806):1299–1302

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

Kützner HJ (1981) The Family Streptomycetaceae. In: Starr MP, Stolp H, Trüper HG, Ballows A, and Schlegel HG (eds) The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria. Springer Verlag, Berlin, p. 2028–2090

Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 44:222.

Lazo GR, Stein PA, Ludwig RA (1991) A DNA transformation-competent Arabidopsis genomic library in Agrobacterium. Nat Biotechnol 9(10):963–967

Li A, Ha Y, Wang F, Li W, Li Q (2012) Determination of thermally-induced trans-fatty acids in soybean oil by attenuated total reflectance Fourier transform infrared spectroscopy and gas chromatography analysis. J Agric Food Chem 60:10709–10713.

Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IAP, Zhao M, Ma J, Yu J, Huang S (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. Nat Commun

5:3930.

Madduri KM, Schafer BW, Hasler JM, Lin G, Foster ML, Embrey SK, Sastry-Dent L, Song P, Larrinua IM, Gachotte DJ, Herman R (2012) Preliminary safety assessment of a membrane-bound delta 9 desaturase candidate protein for transgenic oilseed crops. Food Chem Toxicol 50:3776–3784.

Martin CA, Visentainer JV, de Oliveira AN, de Oliviera CC, Matsushita M, de Souza NE (2008) Fatty acid contents of Brazilian soybean oils with emphasis on *trans* fatty acids. J Braz Chem Soc 19(1):117–122

Marwede V, Schierholt A, Möllers C, Becker HC (2004) Genotype x environment interactions and heritability of tocopherol contents in canola. Crop Sci 44(3):728-731

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. Crit Rev Food Sci Nutr 36 Suppl:S165–S186

Miletich JP, Broze Jr. GJ (1990) b Protein C is not glycosylated at asparagine 329. J Biol Chem 265:11397–11404

Murakami T, Anzai H, Imai S, Satoh A, Nagaoka K, Thompson CJ (1986) The bialaphos biosynthetic genes of *Streptomyces hygroscopicus*: Molecular cloning and characterization of the gene cluster. Mol Gen Genet 205:42–50

Naczk M, Amarowicz R, Sullivan A, Shahidi F (1998) Current research developments on polyphenolics of rapeseed/canola: a review. Food Chem 62(4):489–502

NSW DPI (2014) Variability of quality traits in canola seed, oil and meal - a review. NSW Department of Primary Industries. www.dpi.nsw.gov.au/agriculture/broadacre-crops/winter-crops/canola-and-safflower/canola-seeds

NSW Food Authority (2009) Report on 2009 *trans* fatty acid survey - analytical results. Report prepared by the New South Wales Food Authority.

http://www.foodstandards.gov.au/publications/Pages/transfattyacidsrepor4560.aspx

OECD (1999) Consensus document on general information concerning the genes and their enzymes that confer tolerance to phosphinothricin herbicide. ENV/JM/MONO(99)13. https://www.oecd.org/env/ehs/biotrack/46815628.pdf

OECD (2001) Consensus document on key nutrients and key toxicants in low erucic acid rapeseed (canola). ENV/JM/MONO(2001)13. Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/science/biotrack/consensus-document-for-work-on-safety-novel-and-foods-feeds-plants.htm

OECD (2003) Considerations for the safety assessment of animal feedstuffs derived from genetically modified plants. ENV/JM/MONO(2003)10. Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=ENV/JM/MONO(2003)10&doclanguage=en

OECD (2011) Revised consensus document on compositional considerations for new varieties of low erucic acid rapeseed (canola): key food and feed nutrients, anti-nutrients and toxicants. ENV/JM/MONO(2011)55. Organization for Economic Co-operation and Development, Paris. http://www.oecd.org/document/22/0,3746,en_2649_34385_46808854_1_1_1_1,00.html

Orlando R, Yang Y (1998) Analysis of Glycoproteins. Ch 9 In: Larsen BS, McEwen CN (eds) Mass Spectrometry of Biological Materials. 2nd ed, Marcel Dekker, p. 216–246

Pearson WR (2000) Flexible Sequence Similarity Searching with the FASTA3 Program Package. Ch 10 In: Misener S, Krawetz SA (eds) Methods in Molecular Biology, Volume 132: Bioinformatics Methods and Protocols. Human Press Inc., Totowa, NJ, p. 185–219

Petrie JR, Singh SP (2011) Expanding the docosahexaenoic acid food web for sustainable production: engineering lower plant pathways into higher plants. AoB Plants 2011:pir011

Petrie JR, Shrestha P, Mansour MP, Nichols PD, Liu Q, Singh SP (2010a) Metabolic engineering of

omega-3 long-chain polyunsaturated fatty acids in plants using an acyl-CoA D6-desaturase with w-3 preference from the marine microalga *Micromonas pusilla*. Metab Eng 12:233–240

Petrie JR, Shrestha P, Zhou X-R, Mansour MP, Liu Q, Belide S, Nichols PD, Singh S (2012) Metabolic engineering plant seeds with fish oil-like levels of DHA. PLoS ONE 7(11):e49165

Petrie JR, Liu Q, Mackenzie AM, Shrestha P, Mansour MP, Robert SS, Frampton DF, Blackburn SI, Nichols PD, Singh SP (2010b) Isolation and characterisation of a high-efficiency desaturase and elongases from microalgae for transgenic LC-PUFA production. Mar Biotechnol 12(4):430–438

Petrie JR, Shrestha P, Belide S, Kennedy Y, Lester G, Liu Q, Divi UK, Mulder RJ, Mansour MP, Nichols PD, Singh S (2014) Metabolic engineering *Camelina sativa* with fish oil-like levels of DHA. PLoS ONE 9(1):e85061.

Pritchard FM, Eagles HA, Norton RM, Salisbury PA, Nicolas M (2000) Environmental effects on seed composition of Victorian canola. Australian Journal of Experimental Agriculture. 40(5):679-85.

Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc Natl Acad Sci USA 106(6):1760–1765.

Rauh M (2012) LC-MS/MS for protein and peptide quantification in clinical chemistry. J Chromatogr B Analyt Technol Biomed Life Sci 883 - 884:59–67.

Rossak M, Smith M, Kunst L (2001) Expression of the FAE_1 gene and FAE_1 promoter activity in developing seeds of Arabidopsis thaliana. Plant Mol Biol 46:717–725

Ruiz-Mendez MV; Dobarganes, MC (2017) Oil refining. American Oil Chemists Society Lipid Library http://lipidlibrary.aocs.org/OilsFats/content.cfm?ItemNumber=39212

Seberry DE, Mailer RJ, Parker PA (2007) Quality of Australian canola 2007/2008. Australian Oilseeds Federation, NSW Dept of Primary Industries. Vol. 14.

Shental-Bechor D, Levy Y (2008) Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. Proc Natl Acad Sci USA 105(24):8256–8261

Spohner SC, Schaum V, Quitmann H, Czermak P (2016) *Kluyveromyces lactis*: an emerging tool in biotechnology. J Biotechnol 222:104–116

Spragg JC, Mailer RJ (2007) Canola meal value chain quality improvement: A final report prepared for AOF and CRC. Project code: 1B-103-0506.

Stålberg K, Ellerström M, Josefsson L-G, Rask L (1993) Deletion analysis of a 2S seed storage protein promoter of *Brassica napus* in transgenic tobacco. Plant Mol Biol 23(4):671–683

Szmigielska AM, Schoenau JJ, Levers V (2000) Determination of glucosinolates in canola seeds using anion exchange membrane extraction combined with the high-pressure liquid chromatography detection. J Agric Food Chem 48(10):4487-4491

Tan SH, Mailer RJ, Blanchard CL, Agboola SO (2011a) Extraction and residual antinutritional components in protein fractions of *Sinapis alba* and *Brassica napus* oil-free meals. www.australianoilseeds.com/__data/assets/pdf_file/0013/8320/S5-P3-Tan.pdf.

Tan SH, Mailer RJ, Blanchard CL, Agboola SO (2011b) Canola proteins for human consumption: extraction, profile, and functional properties. J Food Sci 76(1):R16–R28.

Tate GC (2001) Overexpression of mammalian integral membrane proteins for structural studies. FEBS Letters 504:94–98

Thomas K, Bannon G, Hefle S, Herouet C, Holsapple M, Ladics G, MacIntosh S, Privalle L (2005) In silico methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report February 23 - 24, 2005. Toxicol Sci 88:307–310

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 Chem Toxicol 47:1041–1050

Thomas K, Aalbers M, Bannon GA, Bartels M, Dearman RJ, Esdaile DJ, Fu T-J, 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. Regul Toxicol Pharmacol 39:87–98

Thompson CJ, Movva NR, Tizard R, Crameri R, Davies JE, Lauwereys M, Botterman J (1987) Characterization of the herbicide-resistance gene *bar* from *Streptomyces hygroscopicus*. EMBO J 6:2519–2523

Tzanetakis N, Hatzikamari M, Litopoulou-Tzanetaki E (1998) Yeasts of the surface microflora of feta cheese. Food and Agriculture Organization of the United Nations. http://agris.fao.org/agris-search/search.do?recordID=BE1999000440

Vanegas-Calerón M, Sayanova O, Napier JA (2010) An alternative to fish oils: metabolic engineering of oil-seed crops to produce omega-3 long chain polyunsaturated fatty acids. Prog Lipid Res 49(2):108–119

Vingering N, Oseredczuk M, Du Chaffaut L, Ireland J, Ledoux M (2010) Fatty acid composition of commercial vegetable oils from the French market analysed using a long highly polar column. OCL 17(3):185–192.

Vodkin LO, Rhodes PR, Goldberg RB (1983) cA lectin gene insertion has the structural features of a transposable element. Cell 34(3):1023–1031

Volkman JK, Dunstan GA, Jeffrey SW, Kearney PS (1991) Fatty acids from microalgae of the genus Pavlova. Phytochemistry 30:1855–1859

Wang Y, McAllister TA, Pickard MD, Xu Z, Rode LM, Cheng KJ (1999) Effect of micronizing full fat canola seed on amino acid disappearance in the gastrointestinal tract of dairy cows. J Dairy Sci 82(3):537-544

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun J-H, Bancroft I, Cheng F, et al. (2011) The genome of the mesoplyploid crop species *Brassica rapa*. Nat Genet 43(10):1035–1039.

Watanabe K, Oura T, Sakai H, Kajiwara S (2014) Yeast D12 fatty acid desaturase: gene cloning, expression, and function. Biosci Biotechnol Biochem 68(3):721–727

Wehrmann A, Van Vliet A, Opsomer C, Botterman J, Schulz A (1996) The similarities of *bar* and *pat* gene products make them equally applicable for plant engineers. Nature Biotechnology 14:1274–1278

Wojtatowicz M, Chrzanowska J, Juszczyk P, Skiba A, Gdula A (2001) Identification and biochemical characteristics of yeast microflora of Rokpol cheese. Int J Food Microbiol 69(1-2):135–140

Wolff RL (1992) *trans*-polyunsaturated fatty acids in French edible rapeseed and soybean oils. J Am Oil Chem Soc 69(2):106–110

Wood CC, Petrie JR, Shrestha P, Mansour MP, Nichols PD, Green AG, Singh S (2009) A leaf-based assay using interchangeable design principles to rapidly assemble multistep recombinant pathways. Plant Biotechnol J 7:914–924.

Yoshie-Stark Y, Wada Y, Wäsche A (2008) Chemical composition, functional properties, and bioactivities of rapeseed protein isolates. Food Chem 107(1):32–39.

Young GJ, Zhang S, Mirsky HP, Cressman RF, Cong B, Ladics GS (2012) Assessment of possible allergenicity of hypothetical ORFs in common food crops using current bioinformatic guidelines and its implications for the safety assessment of GM crops. Food Chem Toxicol 50:3741–3751

Zhang X, Li M, Wei D, Xing L (2008) Identification and characterization of a novel yeast w3-fatty acid desaturase acting on long-chain n-6 fatty acid substrates from *Pichia pastoris*. Yeast 25(1):21–27

Zhou X-R, Robert S, Singh S, Green A (2006) Heterologous production of GLA and SDA by expression of an *Echium plantagineum* D6-desaturase gene. Plant Sci 170:665–673.

Zhou X-R, Robert SS, Petrie JR, Frampton DMF, Mansour MP, Blackburn SI, Nichols PD, Green AG, Singh SP (2007) Isolation and characterization of genes from the marine microalga *Pavlova salina*

encoding three front-end desaturases involved in docosahexaenoic acid biosynthesis. Phytochemistry 68(6):785–796

Appendix 1

Results of homology searches for each of the seven n-3 LCPUFA proteins introduced into DHA canola

In silico analyses using Vector NTI software were conducted, where the sequence of each of the seven proteins in the n-3 LC-PUFA pathway was used to search for homologous sequences present in organisms used in food, food production or in animal feeds. The purpose of this search was to identify the similarity (sequence identity) of each protein to other proteins present in consumed foods or used in food production or animal feeds. The results are presented in Tables 1.1 – 1.7.

Table 1.1: Amino acid similarity (sequence identity) between Lackl-Δ12D expressed in DHA canola & other desaturase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity												
				1	2	3	4	5	6	7	8	9	10	11	12	13
1	NS-B50027-4 Lackl-Δ12D			100	64.4	58.1	67.6	69.0	66.4	37.5	36.1	34.6	36.1	33.9	35.5	36.7
2	Picpa-Δ12D	AAX20125	Yeast		100	62.3	59.2	51.5	52.5	38.0	33.9	33.8	33.6	34.1	33.1	35.1
3	Pican-Δ12D	BAN63793	Yeast			100	57.2	53.7	54.7	38.0	37.4	35.9	35.6	35.1	35.9	36.0
4	Kluma-Δ12D	BAO38850	Fungus				100	64.7	65.1	37.8	37.1	36.4	35.3	35.0	34.1	36.0
5	Zygba-Δ12D	CDH15170	Fungus					100	78.0	36.8	35.3	35.8	35.4	34.3	33.7	34.3
6	Zygro-Δ12D	XP_002495271	Fungus						100	38.3	35.2	36.3	35.2	36.6	34.4	36.1
7	Moral-Δ12D	BAAB1754	Fungus							100	39.5	40.1	40.5	41.0	40.7	42.5
8	Brana-Δ12D	AFJ19031	Canola								100	69.3	78.9	77.6	72.7	75.8
9	Orysa-Δ12D	ACN87220	Rice									100	71.6	68.5	66.8	71.9
10	Glyma-Δ12D	ABF84063	Soybean										100	79.4	74.0	79.4
11	Linus-Δ12D	ACF49508	Flax											100	73.7	78.3
12	Helan-Δ12D	AF251844	Sunflower												100	75.0
13	Sesin-Δ12D	AAX11454	Sesame													100

Δ12D, Δ12-desaturase; Brana, *Brassica napus*; Glyma, *Glycine max*; Helan, *Helianthus annus*; Kluma, *Kluyveromyces marxianus*; Lackl, *Lachancea kluyveri*; Linus, *Linum usitatissmum*; Moral, *Mortierella alpina*; Orysa, *Oryza sativa*; Pican, *Pichia angusta*; Picpa, *P. pastoris*; Sesin, *Sesamum indicum*; Zygba, *Zygosaccharomyces bailii*; Zygro, *Z. rouxii*.

Table 1.2: Amino acid similarity (sequence identity) between Picpa-ω3D expressed in DHA canola & other desaturase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity												
				1	2	3	4	5	6	7	8	9	10	11	12	13
1	NS-B50027-4 Picpa-ω3D			100	63.1	63.4	59.8	58.9	48.8	35.5	50.2	55.2	27.6	25.6	25.6	25.9
2	Wicci-ω3D	XP_011275445	Fungus		100	65.2	72.1	64.0	49.2	33.9	51.1	58.0	24.8	25.6	23.5	25.6
3	Lackl-ω3D	BAD11952	Fungus			100	57.4	61.5	51.5	33.3	52.2	54.8	24.5	23.4	23.9	23.6
4	Cybjan-ω3D	BAJ78984	Fungus				100	62.1	44.9	30.9	46.8	53.4	24.9	23.8	22.6	24.0
5	Lipko-Δ12D/Δ15D	ACF36357	Fungus					100	46.7	31.8	49.8	56.6	24.4	25.9	25.6	23.7
6	Yarli-Δ12D	CAG82952	Yeast						100	39.3	47.3	52.5	28.1	28.1	27.4	26.7
7	Moral-ω3D	BAD91495	Fungus							100	32.6	32.6	29.3	29.5	30.5	26.4
8	Zygba-Δ12D	CDH15170	Fungus								100	64.7	26.7	26.6	26.2	25.2
9	Kluma-Δ12D	BAO38850	Fungus									100	26.5	26.7	25.9	26.9
10	Brana-ω3D	P48624	Canola										100	68.7	66.3	65.5
11	Glyma-ω3D	NP_001236943	Soybean											100	67.9	62.2
12	Linus-Δ15D	ABA02172	Flax												100	65.0
13	Sesin-ω3D	XP_011080789	Sesame													100

Δ12D, Δ12-desaturase; Δ15D, Δ15-desaturase; ω3D, ω3-desaturase; Brana, *Brassica napus*; Cybji, *Cyberlindernera jadinii*; Glyma, *Glycine max*; Kluma, *Kluyveromyces marxianus*; Lackl, *Lachancea kluyveri*; Linus, *Linum usitatissmum*; Lipko, *Lipomyces kononenkoae*; Moral, *Mortierella alpine*; Picpa, *Picia pastoris*; Sesin, *Sesamum indicum*; Wicci, *Wickerhamomyces ciferrii* ; Yarli, *Yarrowia lipolytica*; Zygba, *Zygosaccharomyces bailii*.

Table 1.3: Amino acid similarity (sequence identity) between Micpu-Δ6D expressed in DHA canola & other desaturase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity 1	2	3	4	5	6
1	NS-B50027-4 Micpu-Δ6D			100	20.5	21.1	19.1	17.5	19.1
2	Moral-Δ6D	BAC82359	Fungus		100	37.3	21.9	26.6	26.6
3	Thaps-Δ6D	AAX14505	Diatom			100	23.5	25.5	26.5
4	Salsa-Δ6D	AAR21624	Salmon				100	23.7	23.2
5	Onebi-Δ6D	ACB47482	Evening primrose					100	61.2
6	Brana-Δ8D	NP_001302507	Canola						100

Δ6D, Δ6-desaturase; Δ8D, Δ8-desaturase; Brana, *Brassica napus* (canola); Moral, *Mortierella alpine* (fungus); Onebi, *Oenothera biennis* (evening primrose); Salsa, *Salmo salar* (salmon); Thaps, *Thalassiosira pseudonana* (alga).

Table 1.4: Amino acid similarity (sequence identity) between Pyrco-Δ6E expressed in DHA canola & other elongase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity 1	2	3	4	5	6	7
1	NS-B50027-4 Pyrco-Δ6E			100	35.2	25.4	27.1	26.8	25.9	21.9
2	Moral-Δ6E	AAF70417	Fungus		100	22.5	24.6	23.7	24.6	21.5
3	Thaps-Δ6E	AAV67799	Alga			100	63.8	21.3	24.2	22.5
4	Phatr-Δ6E	CAM55851	Diatom				100	23.2	25.6	22.5
5	Homsa-Elo1	XP_002040	Human					100	31.7	22.1
6	Salsa-Elo	AAO13175	Salmon						100	19.2
7	Glyma-Elo	XP_003531583	Soybean							100

Δ6E, Δ6-elogase; Elo, PUFA elongase; Glyma, *Glycine max*; Homsa, *Homo sapiens*; Moral, *Mortierella alpine*; Phatr, *Phaeodactylum tricornutum*; Pyrco, *Pyramimonas cordata*; Salsa, *Salmo salar*; Thaps, *Thalassiosira pseudonana*.

Table 1.5: Amino acid similarity (sequence identity) between Pavsa-Δ5D expressed in DHA canola & other desaturase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity 1	2	3	4	5	6	7	8	9
1	NS-B50027-4 Pavsa-Δ5D			100	52.7	52.6	46.8	24.9	22.9	22.9	18.0	20.8
2	Pavlu-Δ5D*	ALE15225	Alga		100	86.6	43.3	26.8	24.7	26.6	18.7	20.7
3	Isoga-Δ5D	AIA24277	Alga			100	42.5	25.2	24.1	24.6	17.4	19.4
4	Thrau-Δ5D	BAK08911	Protist				100	23.8	24.2	23.0	16.1	21.4
5	Octvu-Δ5D	AEK20864	Octopus					100	49.7	51.0	19.9	24.7
6	Homsa-Δ5D	AAF70457	Human						100	57.2	21.8	23.7
7	Salsa-Δ5D	AAL82631	Salmon							100	20.0	23.0
8	Moral-Δ5D	AAC72755	Fungus								100	20.7
9	Onebi-Δ6D	ACB47482	Evening primrose									100

Δ5D, Δ5-desaturase; Δ6D, Δ6-desaturase. Homsa, *Homo sapiens* (human); Isoga, *Isochrysis galbana* (alga); Moral, *Mortierella alpina* (fungus); Octvu, *Octopus vulgaris* (octopus); Onebi, *Oenothera biennis* (evening primrose); Pavlu, *Pavlova lutheri* (alga); Pavsa, *P. salina* (alga); Salsa, *Salmo salar* (salmon); Thrau, *Thraustochytrium aureum* (protist). *Pavlu-Δ5D is partial sequence. The sequence identity was based on the overlap region only.

Table .1.6: Amino acid similarity (sequence identity) between Pyrco-Δ5E expressed in DHA canola & other elongase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity							
				1	2	3	4	5	6	7	8
1	NS-B50027-4 Pyrco-Δ5E			100	37.5	25.1	27.6	27.5	24.9	25.1	22.9
2	Pavlu-Δ5E	AAV33630	Alga		100	24.4	24.2	25.5	22.0	31.0	22.4
3	Moral-Δ6E	AAF70417	Fungus			100	24.0	23.7	24.6	18.1	21.5
4	Homsa-Elo1	XP_002040	Human				100	31.7	23.2	32.2	22.1
5	Salsa-Elo	AAO13175	Salmon					100	25.6	23.8	19.2
6	Phatr-Δ6E	CAM55851	Diatom						100	19.0	22.5
7	Thaps-Δ5E	AAV67800	Alga							100	18.2
8	Glyma-Elo	XP_003531583	Soybean								100

Δ5E, Δ5-elongase; Δ6E, Δ6-elogase; Elo, PUFA elongase; Glyma, *Glycine max*; Homsa, *Homo sapiens*; Moral, *Mortierella alpine*; Pavlu, *pavola lutheri*; Phatr, *Phaeodactylum tricornutum*; Pyrco, *Pyramimonas cordata*; Salsa, *Salmo salar*; Thaps, *Thalassiosira pseudonana*.

Table .1.7: Amino acid similarity (sequence identity) between Pavsa-Δ4D expressed in DHA canola & other desaturase proteins found in food/feed sources

No.	Protein	Accession	Common Name	Sequence identity											
				1	2	3	4	5	6	7	8	9	10	11	12
1	NS-B50027-4 Pavsa-Δ4D			100	65.5	23.3	26.6	19.8	24.9	20.0	18.9	19.8	17.7	15.6	17.4
2	Pavlu-Δ4D	AQ98793	Alga		100	23.9	25.8	19.2	23.2	20.1	20.9	18.5	19.1	18.5	16.2
3	Euggr-Δ4D	AAQ19605	Alga			100	22.8	34.4	24.2	27.9	16.0	16.7	16.7	15.7	18.3
4	Phatr-Δ5D	AAL92562	Alga				100	20.6	21.7	21.9	21.0	17.9	18.9	20.1	19.6
5	Thaps-Δ4D	AAX14506	Alga					100	23.8	16.4	16.8	17.2	15.8	15.5	15.3
6	Moral-Δ5D	O74212	Fungus						100	20.9	18.1	22.5	20.7	20.9	21.5
7	Moral-Δ6D	AAF08685	Fungus							100	21.0	15.5	27.9	27.1	27.9
8	Isoga-Δ4D	AAV33631	Alga								100	12.0	20.2	20.5	19.3
9	Artpl-Δ6D	CAA60573	Cyanobacteria									100	16.7	13.9	14.5
10	Echpl-Δ6D	AAZ08559	Paterson's curse										100	85.3	62.5
11	Borof-Δ6D	O04353	Borage											100	64.3
12	Oenbi-Δ6D	ACB47482	Evening primrose												100

Artpl-Δ6D, *Arthrospira platensis* Δ6-desaturase; Borof-Δ6D, *Borago officinalis* Δ6-desaturase; Echpl-Δ6D, *Echium plantagineum* Δ6-desaturase; Euggr-Δ4D, *Euglena gracilis* Δ4-desaturase; Isoga-Δ4D, *Isochrysis galbana* Δ4-desaturase; Moral-Δ5D, *Mortierella alpina* Δ5-desaturase; Moral-Δ6D, *M. alpine* Δ6-desaturase; Oenbi-Δ6D, *Oenothera biennis* Δ6-desaturase; Pavlu-Δ4D, *Pavola luthri* Δ4-desaturase; Phatr-Δ5D, *Phaeodactylum tricornutum* Δ4-desaturase; Oenbi-Δ6D, *Oenothera biennis* Δ6-desaturase; Thaps-Δ4D, *Thalassiosira pseudonana* Δ4-desaturase.