Author/s (year) **Title (version Nº)** **Owner** **Date**	[redacted] **Characterisation of *Pavlova salina* Δ4-Desaturase** **Nuseed Pty Ltd** **December 11, 2016**
Report Nº	**2016-011**
Testing Facility	**CSIRO** **Black Mountain Laboratories** **Canberra, ACT** **Australia**
Test method	**Not Applicable**
Good Laboratory Practice	**No GLP**
Confidentiality	**None**

© 2016 Nuseed Pty Ltd. All Rights Reserved.
This document is protected under copyright law.

TITLE:
CHARACTERIZATION OF *PAVLOVA SALINA* Δ4-DESATURASE

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ALA	α-Linoleic acid, $18:3^{\Delta 9,12,15}$ (ω3)
CoA	Coenzyme A
DHA	Docosahexaenoic acid $22:6^{\Delta 4,7,10,13,16,19}$ (ω3)
DHA canola	Genetically modified canola, event NS-B50027-4
DPA	Docosapentaenoic acid, $22:5^{\Delta 7,10,13,16,19}$ (ω3)
EPA	Eicosapentaenoic acid, $20:5^{\Delta 5,8,11,14,17}$ (ω3)
ETA	Eicosatetraenoic acid, $20:4^{\Delta 8,11,14,17}$ (ω3)
FAME	Fatty acid methyl ester
GC	Gas chromatography
GRS	Generally Regarded as Safe
LA	Linoleic acid, $18:2^{\Delta 9,12}$ (ω6)
Lackl-Δ12D	*Lachancea kluyveri* Δ12-desaturase
Micpu-Δ6D	*Micromonas pusilla* Δ6-desaturase
MUFA	Monounsaturated fatty acid
OA	Oleic acid, $18:1^{\Delta 9}$
ω3 LC-PUFA	Omega-3 long-chain (≥C20) polyunsaturated fatty acids
MMT	Million metric tons
ORF	Open reading frame
Pavsa-Δ4D	*Pavlova salina* Δ4-desaturase
Pavsa-Δ5D	*Pavlova salina* Δ5-desaturase
pI	Theoretical isoelectric point
Picpa-ω3D	*Pichia pastoris* Δ15-/ω3-desaturase
Pyrco-Δ5E	*Pyramimonas cordata* Δ5-elongase
Pyrco-Δ6E	*Pyramimonas cordata* Δ6-elongase
SDA	Stearidonic acid, $18:4^{\Delta 6,9,12,15}$ (ω3)
X:Y	A fatty acid containing X carbons with Y double bonds

EXECUTIVE SUMMARY

The purpose of this report was to characterise the algal *Pavlova salina* Δ4-desaturase (Pavsa-Δ4D) protein, its amino acid sequence and homology to other proteins, and its enzymatic activity in different expression systems.

The results of the study demonstrated that Pavsa-Δ4D protein was a functional enzyme for accumulation of omega-3 long-chain (≥C20) polyunsaturated fatty acids (ω3 LC-PUFA) in different cells. Pavsa-Δ4D protein contains 447 amino acid residues and shares high homology to other Δ4-desaturases that have been consumed as food, used in food production or in animal feeds. The molecular weight of Pavsa-Δ4D is predicted to be 49.3 kDa, with an estimated isoelectric point (pI) of 8.66.

I. INTRODUCTION

The omega-3 long-chain (≥C20) polyunsaturated fatty acids (ω3 LC-PUFA) eicosapentaenoic acid (EPA, 20:5ω3), docosapentaenoic acid (DPA, 22:5ω3) and docosahexaenoic acid (DHA, 22:6ω3) are widely recognised for their beneficial roles in human health, particularly those related to cardiovascular and inflammatory health. EPA, DPA and DHA are primarily sourced from wild-caught fish oils and algal oils, with algae being the primary producer in the marine food web. These sources are under pressure due to increasing demand for ω3 LC-PUFA by aquaculture, nutraceutical and pharmaceutical applications. Additional sources of these fatty acids can be produced by engineering land-based oilseed crops to convert native fatty acids to marine-type ω3 LC-PUFA which are then accumulated in seed oil. Canola is a commonly grown oilseed with 67 million metric tons (MMT) of rapeseed produced globally in 2015/16[1].

In collaboration with the Commonwealth Scientific and Industrial Research Organization (CSIRO), Nuseed Pty Ltd has developed genetically modified canola event NS-B50027-4 (DHA canola), which accumulates significant amounts of DHA in the seed oil.

In this DHA canola, seven fatty acid desaturases and elongases were introduced to convert oleic acid (OA) to DHA in a single pathway expression vector (Figure 1). The pathway includes the following enzymes: *Lachancea kluyveri* Δ12-desaturase (Lackl-Δ12D, Watanabe et al., 2004), *Pichia pastoris* ω3-/Δ15-desaturase (Picpa-ω3D, Zhang et al. 2008),

[1] http://www.ers.usda.gov/data-products/oil-crops-yearbook/oil-crops-yearbook/#World Supply and Use of Oilseeds and Oilseed Products

Micromonas pusilla Δ6-desaturase (Micpu-Δ6D, Petrie et al., 2010c), *Pyramimonas cordata* Δ6-elongase (Pyrco-Δ6E, Petrie et al., 2010a), *Pavlova salina* Δ5-desaturase (Pavsa-Δ5D, Zhou at al., 2007), *P. cordata* Δ5-elongase (Pyrco-Δ5E, Petrie et al., 2010a) and *P. salina* Δ4-desaturase (Pavsa-Δ4D, Zhou at al., 2007). The functionalities and activities of these enzymes have been demonstrated in different heterologous expression systems (see Report Nºs 2016-005, 2016-006, 2016-007, 2016-008, 2016-009, 2016-010) and in transgenic Arabidopsis or Camelina seeds (Petrie et al., 2012, 2014). Based on the sequence similarity and functionality, these seven proteins can be classified into three groups: (1) yeast acyl-CoA-type fatty acid desaturases including Lackl-Δ12D and Picpa-ω3D (Figure 1, blue) that introduces a double bond at the Δ12 and Δ15 positions, respectively; (2) algae fatty acid elongases including Pyrco-Δ6E and Pyrco-Δ5E (Figure 1, purple) that add two carbons to the carboxyl end of fatty acids; and (3) algae front-end fatty acid desaturases that introduce a double bond between an existing double bond and the carboxyl end of fatty acids (Zhou at al., 2007) including Micpu-Δ6D, Pavsa-Δ5D and Pavsa-Δ4D (Figure 1, green).

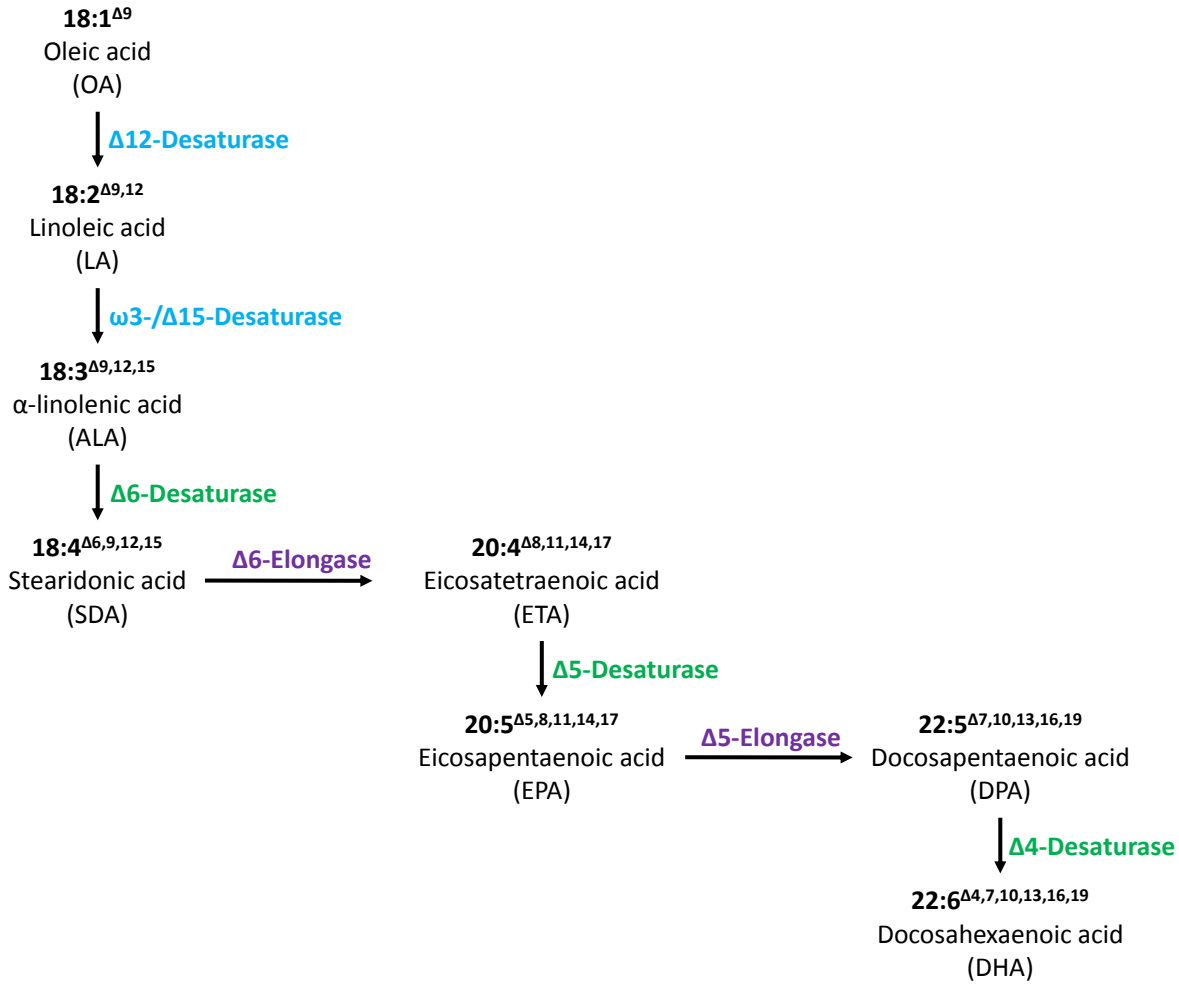

Figure 1. DHA biosynthesis pathway engineered into DHA canola event NS-B50027-4. Seven enzymes introduced in canola to convert oleic acid to final product docosahexaenoic acid were grouped into three classes: two fatty acid desaturases from yeast in blue, two elongases from microalgae in purple, and three front-end desaturases from microalgae in green (see text for detail).

II. PURPOSE

The purpose of this study was to characterise the fatty acid biosynthesis enzymes used in the engineering of DHA canola. The investigations included the amino acid sequences, their homology to other proteins with similar function or presented in consumed food or used in food production, and their enzymatic activities in heterologous expression systems. This

particular report is focused on the *Pavlova salina* Δ4-desaturase (Pavsa-Δ4D) protein that catalyses the Δ4-desaturation of DPA into DHA ($22:5^{\Delta7,10,13,16,19} \rightarrow 22:6^{\Delta4,7,10,13,16,19}$).

III. MATERIALS

A. TARGET PROTEIN

The Δ4-desaturase gene used in the DHA canola event was previously cloned from the alga *P. salina* (Zhou et al., 2007). The Pavsa-Δ4D protein was expressed as a native sequence in yeast S288C cells (Zhou et al., 2007), Arabidopsis seed (Robert et al., 2005; Petrie et al., 2012), Camelina seed (Petrie et al., 2014) and *Nicotiana benthamiana* leaf (Petrie et al., 2010b), as well as a His-tag fusion in insect cell lines (*Sf9*) infected with *baculovirus* pFastBac vector (Life Science Technologies, Germany). The His-tag fusion vector contains a coding sequence that encods a His-tag (His_{10}) and a PreScission protease cleavage site (SLEVLFQ$^{\downarrow}$GP) fused to the codon optimized *Pavsa-Δ4D* gene.

B. OTHER MATERIALS

P. salina strain CS-49 was obtained from the CSIRO Collection of Living Microalgae[2].

For *baculovirus* expression, the codon-optimized *Pavsa-Δ4D* gene was synthesized at GeneArt (Life Science Technologies, Germany) as a His-tag fusion, and cloned into the pFastBac vector.

IV. METHODS

A. SEQUENCE COMPARISON

The *Pavsa-Δ4D* gene was previously cloned from the alga *P. salina* (Zhou et al., 2007). The translated amino acid sequence homology was compared to other published Δ4-desaturases or related front-end desaturases presented in food or used in food production, with Vector NTI software (Version 11, Invitrogen, Germany).

[2] https://www.csiro.au/en/Research/Collections/ANACC

V. RESULTS AND DISCUSSION

A. GENE SOURCE AND DONOR ORGANISM

The Δ4-desaturase gene used in DHA canola was previously cloned from the microalga *Pavlova salina* (Class Prymnesiophyceae; Zhou et al., 2007). The open reading frame (ORF) of the *Pavsa-Δ4D* gene is shown in Figure 2.

ATGCCTCCGAGCGCGGCGAAGCAGATGGGCGCGAGCACGGGCGTGCATGCGGGCGTCACA
GATTCGTCGGCCTTCACGCGCAAGGATGTCGCCGACAGGCCGGACCTCACGATCGTGGGT
GACAGCGTGTACGATGCGAAGGCGTTCCGCTCCGAGCATCCGGGTGGCGCGCACTTTGTG
TCGCTGTTCGGCGGGCGCGATGCCACGGAGGCGTTCATGGAGTACCACCGGCGCGCCTGG
CCCAAGTCGCGCATGTCGCGCTTCCACGTCGGCTCTCTGGCATCGACCGAGGAGCCCGTC
GCCGCCGATGAGGGCTACCTCCAGCTGTGCGCTCGCATCGCCAAGATGGTGCCGTCGGTC
AGCAGCGGGTTCGCGCCGGCGTCGTACTGGGTGAAGGCCGGGCTGATCCTCGGCTCCGCG
ATCGCGCTCGAGGCGTACATGCTGTACGCGGGCAAGCGCCTGCTCCCGTCGATCGTGCTC
GGGTGGCTGTTTGCGCTGATTGGCCTGAACATCCAGCACGATGCCAACCACGGCGCGCTC
TCCAAGTCGGCCTCGGTCAACCTGGCGCTCGGGTTGTGCCAGGACTGGATCGGCGGGAGC
ATGATCCTCTGGCTGCAGGAGCACGTTGTCATGCACCACTTGCACACCAACGACGTTGAC
AAGGACCCGGACCAGAAGGCGCACGGCGCCCTGCGGCTCAAGCCGACCGACGCGTGGAGC
CCGATGCACTGGCTGCAGCACCTCTACCTGCTGCCTGGGGAGACGATGTACGCCTTCAAG
CTGCTGTTTCTCGACATCAGCGAGCTGGTGATGTGGCGGTGGGAGGGCGAGCCCATCAGC
AAGCTGGCCGGGTACCTCTTCATGCCCTCGCTGCTCCTCAAGCTCACCTTCTGGGCGCGC
TTTGTCGCGCTGCCGCTGTACCTCGCGCCCAGCGTGCACACGGCGGTGTGCATCGCGGCG
ACGGTAATGACGGGGAGCTTCTACCTCGCCTTCTTCTTCTTCATCTCGCACAACTTCGAG
GGCGTGGCGAGCGTCGGACCGGACGGCAGCATCACCAGCATGACGCGCGGCGCATCCTTC
CTCAAGCGGCAGGCCGAGACCTCGTCCAACGTGGGCGGCCCGCTGCTCGCCACGCTCAAC
GGCGGCCTCAACTACCAAATCGAGCACCACCTCTTCCCCAGGGTGCACCACGGCTTCTAC
CCTCGCCTCGCGCCGTTGGTCAAGGCGGAGCTCGAGGCGCGCGGCATTGAGTACAAGCAC
TACCCCACCATATGGAGCAACCTGGCATCCACGCTGAGGCACATGTACGCGCTCGGCCGC
AGGCCGCGCAGCAAGGCGGAG**TGA**

Figure 2. Nucleotide sequence of native Pavsa-Δ4D gene.
Start codon (ATG) and stop codon (TGA) are in bold.

B. PROTEIN SEQUENCE

The translated *P. salina* Δ4-desaturase (Pavsa-Δ4D) contains 447 amino acid residues (Figure 3). The amino acid sequence of Pavsa-Δ4D shares high homology to other Δ4-desaturases (Figure 3). The molecular weight of Pavsa-Δ4D is predicted to be 49.3 kDa, with an estimated isoelectric point (pI) of 8.66.

MPPSAAKQMGASTGVHAGVTDSSAFTRKDVADRPDLTIVGDSVYDAKAFRSEHPGGAHFV
SLFGGRDATEAFMEYHRRAWPKSRMSRFHVGSLASTEEPVAADEGYLQLCARIAKMVPSV
SSGFAPASYWVKAGLILGSAIALEAYMLYAGKRLLPSIVLGWLFALIGLNIQHDANHGAL
SKSASVNLALGLCQDWIGGSMILWLQEHVVMHHLHTNDVDKDPDQKAHGALRLKPTDAWS
PMHWLQHLYLLPGETMYAFKLLFLDISELVMWRWEGEPISKLAGYLFMPSLLLKLTFWAR
FVALPLYLAPSVHTAVCIAATVMTGSFYLAFFFFISHNFEGVASVGPDGSITSMTRGASF
LKRQAETSSNVGGPLLATLNGGLNYQIEHHLFPRVHHGFYPRLAPLVKAELEARGIEYKH
YPTIWSNLASTLRHMYALGRRPRSKAE

Figure 3. Amino acid sequence of Pavsa-Δ4D.

The Pavsa-Δ4D protein shares high homology to Δ4-desaturase proteins isolated from other organisms. Figure 4 shows the phylogenetic tree of amino acid sequence comparison.

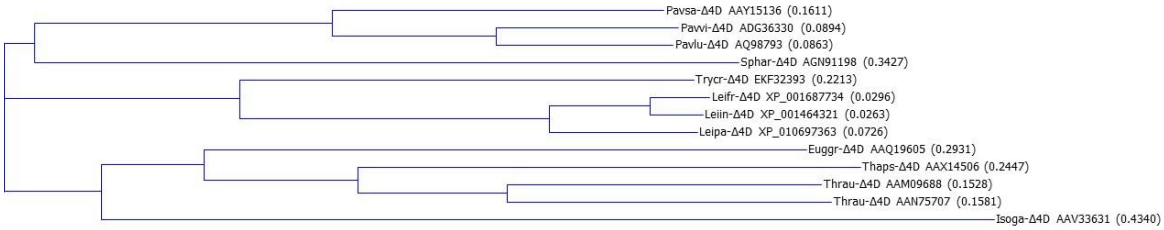

Figure 4. Phylogenetic tree for sequence comparison of Pavsa-Δ4D.
The phylogenetic tree was generated with Vector NTI software (Invitrogen, Carlsbad, USA). The protein sequences were named as 5-letter initial of species name, followed by NCBI accession numbers. Euggr, *Euglena gracilis* (Eukaryote algae); Isoga, *Isochrysis galbana* (microalga); Leifr, *Leishmania* major strain Friedlin (trypanosome); Leiin, *L. infantum* JPCM5 (trypanosome); Leipa, *L. panamensis* (trypanosome); Pavlu, *P. luthri* (microalga); Pavsa, *P. salina* (microalga); Pavvi, *P. viridis* (microalga); Sphar, *Sphaeroforma arctica* (protozoan); Thaps, *Thalassiosira pseudonana* (marine phytoplankton); Thrau, *Thraustochytrium aureum* (moncentric fungus); Trycr, *Trypanosoma cruzi* (protozoan).

C. Similarity of Pavsa-Δ4D to Other Proteins in Consumed Foods, Used in Food Production or in Animal Feeds

The Pavsa-Δ4D protein shares similarity to desaturase proteins presented in food that is consumed, used in food production or in animal feeds (Table 1). *P. salina* itself is one of microalgae used in mariculture (Brown 1991). *P. lutheri* is used for oyster larvae and clam

larvae feeds (Brown et al. 1997). Pavsa-Δ4D shares 66% sequence identity to *P. lutheri* Δ4-desaturase (AQ98793).

The Pavsa-Δ4D protein belongs to the subfamily of front-end desaturases that introduce a double bond between an existing double bond and the carboxyl end of fatty acids. The front-end desaturases include Δ4-, Δ5-, Δ6- and Δ8-desaturases, which exist in a wide range of organisms including algae, diatom, fungi, moss and bacteria. Microalgae such as *Spirulina*, *Chlorella*, *Dunaliella*, *Haematococcus* and *Schizochytrium* are classified as food sources falling under the Generally Regarded as Safe (GRS) category according to the U.S. Food and Drug Administration (Chacón-Lee and González-Marino, 2010). *Spirulina* and *Chlorella* are major microalgal genera cultivated in China for health food (Liang et al., 2004). Algal biomass is supplemented to noodles, breads, biscuits, candies, ice cream, bean curd and other common foods to enhance their nutritive and health values. The so-called blue-green algae *Spirulina* is actually a cyanobacterium. Cyanobacteria have been part of the human diet for centuries (Gantar and Svircev, 2008). Some cyanobacteria like *Arthrospira platensis* and *Aphanizomenon flos-aquae* are sold as food (Spolaore et al. 2006). Food products containing the cyanobacteria Spirulina (*A. platensis*) have been sold worldwide as supplements (Belay et al., 1993). FDA had approved Spirulina's extracted blue color for use as a natural food coloring in gum and candy.[3]. The Pavsa-Δ4D protein shows 20% sequence identity to cyanobacterium *A. platensis* Δ6-desaturase (CAA60573). The Pavsa-Δ4D protein shares 23% sequence identity to single-celled flagellate *Euglena gracilis* Δ4-desaturase (AAQ19605). From 2005, Tokyo-based Euglena Company has started marketing *Euglena*-based food and beverage products, based on their provision of both plant- and animal-based nutrients. The Pavsa-Δ4D protein shows 20% or 25% sequence identity to fungus *Mortierella alpina* Δ6-desaturase (AAF08605) or Δ5-desaturase (O74212). *M. alpina* is currently used for the commercial production of arachidonic acid for fortification of baby food. The Pavsa-Δ4D protein shares 26% sequence identity to *Phaeodactylum tricornutum* Δ5-desaturase (AAL92562). *P. tricornutum* is used to produce pigments or antioxidant for food (Chacón-Lee and González-Marino, 2010). *P. tricornutum* and *Thalassiosira pseudonana* are foods of oyster (Epifanio et al., 1981). The Pavsa-Δ4D protein shares 19% amino acid sequence identity to microalga *Isochrysis galbana* (AAV33631). *I. galbana* is an outstanding food for various bivalve larvae, and is now widely cultured in aquaculture[4]. *Pavlova lutheri* is used in the aquaculture of bivalves, carp and shrimp where it is fed either directly or indirectly to cultured larval organisms.

[3] Federal Register; FDA Approves Natural Blue Color Additive Extracted from Spirulina". 13 August 2013

[4] The hatchery culture of bivalves: a practical manual. FAO Fisheries and Aquaculture Department. 2004. p. 201. ISBN 9251052247.

The Pavsa-Δ4D protein shares ~17% amino acid sequence identity with plant Δ6D proteins from *Echium plantagineum* (echium, AAZ08559), *Borago officinalis* (borage, O04353) and *Oenothera biennis* (evening primrose, ACB47482). These species have been used to produce oils that are relatively high in GLA and/or SDA for human consumption. The oils produced by these species have been studied extensively for their anti-inflammatory effects on leukotriene and prostaglandin biosynthesis (Fan and Chapkin, 1998), and are sold as cold-pressed oils for use as dietary supplements. Evening primrose oil is commonly sold in Australian health food shops. Additionally, the flowers of *Echium* sp. have been consumed as medicinal plants in countries such as Iran (Heidari et al., 2006). Evening primrose plants have been used as ornamentals, food sources, and as medicinal herbs for more than 50 years.

Table 1. Comparison of amino acid sequence identity of Pavsa-Δ4D in DHA canola (event NS-B50027-4) to other desaturase proteins present in foods that are consumed, used in food production or in animal feeds.

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.

D. HETEROLOGOUS EXPRESSION

The enzyme functionality of Pavsa-Δ4D has been confirmed in different heterologous expression systems, including yeast (Zhou et al., 2007), Arabidopsis seed (Robert et al., 2005; Petrie et al., 2012), Camelina seed (Petrie et al., 2014) and *Nicotiana benthamiana* leaf (Petrie et al., 2010a). Table 2 shows the enzyme activity of Pavsa-Δ4D expressed in yeast S288C cells when exogenously supplied with ω3 and ω6 substrate (Zhou et al., 2007).

Table 2. Fatty acid composition of yeast S288C cells expressing Pavsa-Δ4D showing the activity of Δ4-desaturase activity

Fatty acid	Exogenous fatty acid fed in growth medium	
	$22{:}4^{\Delta7,10,13,16}$ (ω6)	$22{:}5^{\Delta7,10,13,16,19}$ (ω3)
$22{:}4^{\Delta7,10,13,16}$ (ω6, DTA)	0.97	0
$22{:}5^{\Delta4,7,10,13,16}$ (ω6, DPA)	0.03	0
$22{:}5^{\Delta7,10,13,16,19}$ (ω3, DPA)	0	1.66
$22{:}6^{\Delta4,7,10,13,16,19}$ (ω3, DHA)	0	0.04
Conversion	3.0%	2.4%

Pavsa-Δ4D is able to desaturate both ω6 $22{:}4^{\Delta7,10,13,16}$ (ω6 DTA) or ω3 $22{:}5^{\Delta7,10,13,16,19}$ (ω3 DPA) substrates at Δ4 positions, making ω6 $22{:}5^{\Delta4,7,10,13,16}$ (ω6 DPA) or ω3 $22{:}6^{\Delta4,7,10,13,16,19}$ (ω3 DHA). The conversion rates with the yeast cell expressed the Pavsa-Δ4D protein were 3.0% or 2.4% respectively.

E. EXPRESSION OF FUSION PROTEIN

In order to purify the target protein, the target enzyme was expressed as a His-tag fusion protein in *baculovirus.* The functionality in baculovirus was not tested. Nevertheless, the Pavsa-Δ4D has shown activity in different expression systems. The DHA canola expressing DHA synthesis pathway including Pavsa-Δ4D has resulted in approximately 10% of DHA in seed oil.

For *baculovirus* expression, the codon optimized *Pavsa-Δ4D* gene was fused to the coding sequence for a His-tag (His_{10}) and a PreScission protease cleavage site (SLEVLFQ$^{\downarrow}$GP) at its 5′-end, as shown in Figure 5. The expressed fusion protein from insect cells was used for protein stability assay (Report Nº 2016-014).

ATGCATCATCATCATCATCACCATCACCACCACTCCCTGGAAGTGCTGTTCCAGGGTCCC
ATGCCCCCATCCGCTGCTAAGCAGATGGGCGCTTCCACCGGTGTCCACGCTGGTGTCACC
GACTCCTCCGCTTTCACCCGCAAGGACGTGGCCGACCGTCCCGACCTGACCATCGTGGGC
GACTCCGTGTACGACGCTAAGGCTTTCCGTTCCGAGCACCCTGGTGGTGCTCACTTCGTG
TCCCTGTTCGGTGGTCGTGACGCTACCGAGGCTTTCATGGAATACCACCGTCGTGCTTGG
CCCAAGTCCCGTATGTCCCGTTTCCACGTGGGCTCCCTGGCTTCCACCGAGGAACCCGTG
GCTGCTGACGAGGGTTACCTGCAGCTGTGCGCTCGTATCGCTAAGATGGTGCCCTCCGTG
TCCTCCGGTTTCGCTCCCGCTTCCTACTGGGTCAAGGCTGGCCTGATCCTGGGTTCCGCT
ATCGCTCTGGAAGCTTACATGCTGTACGCTGGCAAGCGTCTGCTGCCCTCCATCGTGCTG
GGCTGGCTGTTCGCTCTGATCGGCCTGAACATCCAGCACGACGCTAACCACGGTGCTCTG
TCCAAGTCCGCTTCCGTGAACCTGGCTCTGGGCCTGTGCCAGGACTGGATCGGTGGTTCC
ATGATCCTGTGGCTGCAAGAGCACGTGGTCATGCACCACCTCCACACCAACGACGTGGAC
AAGGACCCCGACCAGAAGGCTCACGGCGCTCTGCGTCTGAAGCCCACCGACGCTTGGTCC
CCCATGCACTGGCTGCAGCACCTGTACCTGCTGCCCGGCGAGACTATGTACGCTTTCAAG
CTGCTGTTCCTGGACATCTCCGAGCTGGTCATGTGGCGTTGGGAGGGCGAGCCCATCTCC
AAGCTGGCTGGTTACCTGTTCATGCCCTCCCTGCTGCTGAAGCTGACCTTCTGGGCTCGT
TTCGTGGCTCTGCCCCTGTACCTGGCTCCCTCCGTGCACACCGCTGTGTGTATCGCTGCT
ACCGTGATGACCGGTTCCTTCTACCTGGCTTTCTTCTTCTTCATCTCCCACAACTTCGAG
GGTGTCGCTTCCGTGGGTCCCGACGGTTCCATCACCTCCATGACCCGTGGTGCTAGCTTC
CTGAAGCGTCAGGCTGAGACTTCCTCCAACGTGGGCGGTCCCCTGCTGGCTACCCTGAAC
GGTGGCCTGAACTACCAGATCGAGCACCACCTGTTCCCCCGTGTGCACCACGGTTTCTAC
CCCCGTCTGGCTCCCCTGGTCAAGGCCGAGCTGGAAGCTCGTGGTATCGAGTACAAGCAC
TACCCCACCATCTGGTCCAACCTGGCCTCCACCCTGCGTCACATGTACGCTCTGGGTCGT
CGTCCCCGTTCCAAGGCTGAG**TAA**

Figure 5. Nucleotide sequence of codon optimized Pavsa-Δ4D gene for baculovirus expression.
The start codon (ATG) and stop codon (TGA) are shown in bold. The start codon (ATG) for the His-tag fusion coding sequence is underlined.

The total protein extracted from insect cells expressing His-tag::Pavsa-Δ4D was confirmed by Western blot against anti His-tag antibody (see detail in Report Nº 2016-014). The expected molecular weight (MW) of His-tag::Pavsa-Δ4D is approximately 51 kDa. A specific protein band close to 50 kDa was detected in the protein pellet (Figure 6).

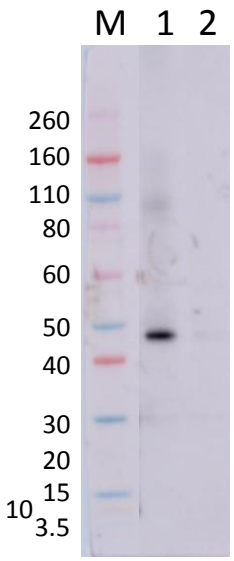

Figure 6. Western blot analysis of His-tag::Pavsa-Δ4D protein expressed in baculovirus-infected insect cells.
M, protein markers with molecular weight (in kDa) indicated to the left; lane 1, total protein in pellet; lane 2, total protein in supernatant.

F. GLYCOSYLATION ANALYSIS

Several classes of glycans exist, which are widely distributed in nature, including *N*-linked glycans glycolipids, *O*-GlcNac, and glycosaminoglycans. *N*-linked glycosylation occurs when glycans are attached to asparagine residues on the protein. *O*-linked glycans are most commonly attached to serine or threonine residues through the *N*-Acetylgalactosamine residue. N-linked glycans are the most common in plants, and typically, can only be found as a linkage to an asparagine residue (N) where it is flanked on the C-terminal side by X-S or X-T. For the Pavsa-Δ4D protein, there is no predicted *N*-glycosylation site within this amino acid sequence derived from the nucleotide sequence of the inserted DNA.

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

Figure 7. Absence of theoretical N-glycosylation sites (NXT/NXS) in Pavsa-Δ4D.

G. SEQUENCE CONFIRMATION IN TRANSGENIC CANOLA

The genome sequence including the T-DNA insertions in DHA canola was analysed. The translated amino acid sequence of Pavsa-Δ4D in the insert was confirmed to be identical to the original sequence (Figure 8).

```
                        1                                                  50
Pavsa-Δ4D_vec     (1)   MPPSAAKQMGASTGVHAGVTDSSAFTRKDVADRPDLTIVGDSVYDAKAFR
NS-B50027-4       (1)   MPPSAAKQMGASTGVHAGVTDSSAFTRKDVADRPDLTIVGDSVYDAKAFR
                        51                                                 100
Pavsa-Δ4D_vec    (51)   SEHPGGAHFVSLFGGRDATEAFMEYHRRAWPKSRMSRFHVGSLASTEEPV
NS-B50027-4      (51)   SEHPGGAHFVSLFGGRDATEAFMEYHRRAWPKSRMSRFHVGSLASTEEPV
                        101                                                150
Pavsa-Δ4D_vec   (101)   AADEGYLQLCARIAKMVPSVSSGFAPASYWVKAGLILGSAIALEAYMLYA
NS-B50027-4     (101)   AADEGYLQLCARIAKMVPSVSSGFAPASYWVKAGLILGSAIALEAYMLYA
                        151                                                200
Pavsa-Δ4D_vec   (151)   GKRLLPSIVLGWLFALIGLNIQHDANHGALSKSASVNLALGLCQDWIGGS
NS-B50027-4     (151)   GKRLLPSIVLGWLFALIGLNIQHDANHGALSKSASVNLALGLCQDWIGGS
                        201                                                250
Pavsa-Δ4D_vec   (201)   MILWLQEHVVMHHLHTNDVDKDPDQKAHGALRLKPTDAWSPMHWLQHLYL
NS-B50027-4     (201)   MILWLQEHVVMHHLHTNDVDKDPDQKAHGALRLKPTDAWSPMHWLQHLYL
                        251                                                300
Pavsa-Δ4D_vec   (251)   LPGETMYAFKLLFLDISELVMWRWEGEPISKLAGYLFMPSLLLKLTFWAR
NS-B50027-4     (251)   LPGETMYAFKLLFLDISELVMWRWEGEPISKLAGYLFMPSLLLKLTFWAR
                        301                                                350
Pavsa-Δ4D_vec   (301)   FVALPLYLAPSVHTAVCIAATVMTGSFYLAFFFFISHNFEGVASVGPDGS
NS-B50027-4     (301)   FVALPLYLAPSVHTAVCIAATVMTGSFYLAFFFFISHNFEGVASVGPDGS
                        351                                                400
Pavsa-Δ4D_vec   (351)   ITSMTRGASFLKRQAETSSNVGGPLLATLNGGLNYQIEHHLFPRVHHGFY
NS-B50027-4     (351)   ITSMTRGASFLKRQAETSSNVGGPLLATLNGGLNYQIEHHLFPRVHHGFY
                        401                                              448
Pavsa-Δ4D_vec   (401)   PRLAPLVKAELEARGIEYKHYPTIWSNLASTLRHMYALGRRPRSKAE
NS-B50027-4     (401)   PRLAPLVKAELEARGIEYKHYPTIWSNLASTLRHMYALGRRPRSKAE
```

Figure 8. Alignment of protein sequences of Pasva-Δ4D.
Δ4D sequence translated from sequenced T-DNA insert in DHA canola NS-B50027-4 event was identical to the original Δ4D sequence from *P. salina* in the binary vector (Pavsa-Δ4D_vec).

VI. CONCLUSIONS

The results of this study demonstrated that the cloned alga Pavsa-Δ4D protein has activity in heterologous expression systems, including in DHA canola event NS-B50027-4, Arabidopsis seed, Camelina seeds and yeast cells. Data for Pavsa-Δ4D expressed in yeast S288C cells when exogenously supplied with ω3 and ω6 substrate confirms this functionality. The Pavsa-Δ4D protein shares similarity to desaturase proteins present in food that is consumed, used in food production or in animal feeds.

Pavsa-Δ4D protein contains 447 amino acid residues. The molecular weight of Pavsa-Δ4D is predicted to be 49.3 kDa, with an estimated pI of 8.66. For the Pavsa-Δ4D protein, there are no potential *N*-glycosylation sites within this amino acid sequence derived from the nucleotide sequence of the inserted DNA. The study also demonstrates that canola event NS-B50027-4 contains T-DNA insertions that are translationally identical to the original Pavsa-Δ4D protein sequence.

VII. REFERENCES

Belay, A, Ota, Y, Miyaka, K, Shimamatsu, H. 1993. Current knowledge on potential health benefits of *Spirulina*. *J Appl Phycol* 5:235-241.

Brown, MR. 1991. The amino-acid and sugar composition of 16 species of microalgae used in mariculture. *J Exp Mar Biol Ecol* 145: 79-99.

Brown, MR, Jeffrey, SW, Volkman, JK, Dunstan, GA. 1997. Nutritional properties of microalgae for mariculture. *Aquaculture* 151: 315-331.

Chaćon-Lee, TL and Gonzalez-Marino, GE. 2010. Microalgae for "Healthy" Foods-Possibilities and Challenges. *Compre Rev Food Sci Food Safety* 9:655-675.

Epifanio, CE, Valenti, CC, Turk, CL. 1981. A comparison of *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* as foods for the oyster, *Crassostrea virginica*. *Aquaculture* 23:347-353.

Fan, Y-Y and Chapkin, RS. 1998. Importance of dietary gamma-linolenic acid in human health and nutrition? *J Nutr* 128:1411-1414.

Gantar, M and Svirčev, Z. 2008. Microalgae and cyanobacteria: Food for thought. *J Phycology* 44:260-268.

Heidari, MR, Azad, EM, Mehrabani, M. 2006. Evaluation of the analgesic effect of *Echium amoenum* Fisch & C.A. Mey. extract in mice: Possible mechanism involved. *J Ethnopharm* 103:345-349.

Liang, S, Liu, X, Chen, F, Chen, Z. 2004. Current microalgal health food R & D activities in China. *Hydrobiol* 512:45-48.

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

Petrie, JR, Shrestha, P, Liu, Q, Mansour, MP, Wood, CC, Zhou, X-R, S, Nichols, PD, Green, AG, Singh, SP. 2010b. Rapid expression of transgenes driven by seed-specific constructs in leaf tissue: DHA production. *Plant Methods* 6:6.

Petrie, JR, Shrestha, P, Mansour, MP, Nichols, PD, Liu, Q, Singh, SP. 2010c. Metabolic engineering of omega-3 long-chain polyunsaturated fatty acids in plants using an acyl-CoA Δ6-desaturase with omega 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, SP. 2012. Metabolic engineering plant seeds with fish oil-like levels of DHA. *PloS One* 7:e49165.

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

Robert, SS, Singh, SP, Zhou, X-R, Blackburn, SI, Mansour, PM, Nichols, PD, Liu, Q, Green, AG. 2005. Metabolic engineering of Arabidopsis to produce nutritionally important DHA in seed oil. *Func Plant Biol* 32:473-479.

Spolaore, P, Joannis-Cassan, C, Duran, E, Isambert, A. 2006. Commercial applications of microalgae. *J Biosci Bioeng* 101:87-96.

Watanabe, K, Oura, T, Sakai, H, Kajiwara, S. 2004. Yeast Δ12 fatty acid desaturase: Gene cloning, expression, and function. *Biosci Biotechnol Biochem* 68:721-727.

Zhang, X, Li, M, Wei, D, Xing, L. 2008. Identification and characterization of a novel yeast ω3-fatty acid desaturase acting on long-chain n-6 fatty acid substrates from *Pichia pastoris. Yeast* 25:21-27

Zhou, X-R, Robert, SS, Petrie, JR, Frampton, DM, 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. *Phytochem* 68:785-796.

VIII. UNPUBLISHED REFERENCES

Report Nº 2016-005. [redacted]
Nuseed Pty Ltd.

Report Nº 2016-006. [redacted]
Nuseed Pty Ltd.

Report Nº 2016-007. [redacted]
Nuseed Pty Ltd.

Report Nº 2016-008. [redacted]
Nuseed Pty Ltd.

Report Nº 2016-009. [redacted]
Nuseed Pty Ltd.

Report Nº 2016-010. [redacted].
Nuseed Pty Ltd.

Report Nº 2016-011. [redacted]
Nuseed Pty Ltd.