Food and Chemical Toxicology 126 (2019) 313–321

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

Quantitation of seven transmembrane proteins from the DHA biosynthesis pathway in genetically engineered canola by targeted mass spectrometry

Michelle L. Colgrave[a,*], Keren Byrne[a], Sapna Vibhakaran Pillai[b], Bei Dong[b], Antonio Leonforte[c], Joanne Caine[d], Lukasz Kowalczyk[d], Judith A. Scoble[d], James R. Petrie[b], Surinder Singh[b], Xue-Rong Zhou[b]

[a] CSIRO Agriculture and Food, 306 Carmody Rd, St Lucia, QLD 4067, Australia
[b] CSIRO Agriculture and Food, GPO Box 1600, Canberra, ACT 2601, Australia
[c] Nuseed, 5 Ballinger St, Horsham, VIC, 3400, Australia
[d] CSIRO Manufacturing, 343 Royal Parade, Parkville, VIC, 3052, Australia

ARTICLE INFO

Keywords:
Protein quantitation
Genetic modification
Omega-3 long-chain fatty acids
Docosahexaenoic acid (DHA)
Canola

ABSTRACT

Examining tissue-specific expression and the measurement of protein abundance are important steps when assessing the performance of genetically engineered crops. Liquid chromatography-mass spectrometry offers many advantages over traditional methods for protein quantitation, especially when dealing with transmembrane proteins that are often difficult to express or generate antibodies against. In this study, discovery proteomics was used to detect the seven transgenic membrane-bound enzymes from the docosahexaenoic acid (DHA) biosynthetic pathway that had been engineered into canola. Subsequently, a targeted LC-MS/MS method for absolute quantitation was developed and applied to the simultaneous measurement of the seven DHA biosynthetic pathway enzymes in genetically modified canola grown across three sites. The results of this study demonstrated that the enzymatic proteins that drive the production of DHA using seed-specific promoters were detected only in mature and developing seed of DHA canola. None of the DHA biosynthesis pathway proteins were detected in wild-type canola planted in the same site or in the non-seed tissues of the transgenic canola, irrespective of the sampling time or the tissues tested. This study describes a streamlined approach to simultaneously measure multiple membrane-bound proteins *in planta*.

1. Introduction

The omega-3 long-chain (≥C20) polyunsaturated fatty acids (ω3 LC-PUFA) eicosapentaenoic acid (EPA, 20:5ω3), docosapentaenoic acid (DPA, 22:5ω3) and DHA (22:6ω3) are widely recognised for their beneficial roles in human health, particularly those related to cardiovascular and inflammatory health (Galli and Calder, 2009). EPA, DPA and DHA are sourced primarily 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, as well as nutraceutical and pharmaceutical applications (Tocher, 2009). 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 74 million metric tons of seed produced globally in 2017/18 (Foreign Agricultural Service, 2018).

The development of genetically modified canola that accumulates significant amounts of DHA in the seed oil (termed DHA canola) has been described (Devine et al., 2018). In this DHA canola, seven genes encoding fatty acid desaturases and elongases that sequentially convert oleic acid (OA) to DHA, under seed-specific promoters, were introduced in a single pathway expression vector, along with a constitutive 35S promoter-driven phosphinothricin-N-acetyltransferase (PAT) gene as an initial selection marker.

Food safety assessment of genetically modified crops is required in all jurisdictions, wherein the potential allergenicity is assessed using a weight-of-evidence approach incorporating a range of techniques (Ladics, 2008). These include: bioinformatics analysis to reveal allergenic domains in the introduced proteins; gastrointestinal stability, processing stability (e.g. thermal); and protein quantitation to estimate the maximum level of exposure based on the abundance in the crop. Quantitation is often accomplished using antibodies via Western blotting or enzyme-linked immunosorbent assays (ELISA). ELISA represents

* Corresponding author.
E-mail address: michelle.colgrave@csiro.au (M.L. Colgrave).

https://doi.org/10.1016/j.fct.2019.02.035
Received 14 November 2018; Received in revised form 4 February 2019; Accepted 25 February 2019
Available online 01 March 2019
0278-6915/

a rapid, sensitive and specific method for protein detection and quantitation, but the development of suitable antibodies can be challenging especially when dealing with membrane proteins. A first challenge lies in achieving a suitable level of recombinant protein expression and purity that is required for successful antibody generation (Hamakubo et al., 2014). Antibodies may also suffer from cross-reactivity, especially when there are multiple homologous proteins as is the case for the DHA biosynthesis pathway, making it difficult to accurately quantify the individual transgenic proteins. Furthermore, membrane proteins are often buried *in situ*, presenting a limited number of epitopes for antibody recognition and the detergents required to efficiently extract such proteins may be incompatible with the immunoassay (e.g., may cause irreversible denaturation of native epitopes).

Liquid chromatography-mass spectrometry (LC-MS) offers an alternative to traditional techniques and offers a number of advantages when detecting and quantifying proteins, especially membrane bound enzymes. Sample preparation strategies used with LC-MS include the use of filter-assisted sample preparation (FASP) (Wisniewski et al., 2009) which relies on protein digestion on a molecular weight cut-off filter. This workflow enables any detergent or buffer composition to be used, because the buffer is exchanged prior to digestion and is thus compatible with downstream LC-MS analysis. There is no requirement for extensive purification of the target protein with partially purified or crude mixtures suitable for method development. In fact, peptides from the target protein may be defined *in silico* although the use of a protein standard is advantageous in selecting peptides with optimal performance. The LC-MS assay may be multiplexed such that many proteins may be investigated simultaneously. The LC-MS assay may also be highly selective and sensitive, especially when making use of targeted scan types as is the case for multiple reaction monitoring (MRM) MS assays (Aebersold et al., 2013; Anderson and Hunter, 2006). Absolute quantitation may be achieved in the absence of protein standards by making use of stable-isotope labels (SIL) (Kirkpatrick et al., 2005; Zhang et al., 2011). The utility of LC-MS for the quantitation of transgenic proteins has been reported (Schacherer et al., 2017) with application in soybean (Hill et al., 2015), maize (Hill et al., 2017; Hu and Owens, 2011). In this study we report the seed-specific expression and absolute quantitation of the levels of the seven transgenic proteins across a range of canola plant tissues from DHA canola and segregants.

2. Materials and methods

2.1. Materials

2.1.1. Plant material

The DHA canola event NS-B50027-4 has two T-DNA insertions in A05 and A02 chromosomes. The A05 T-DNA insert has two complete gene sets that form a palindromic structure with RB-LB:LB-RB orientation. The A02 T-DNA insert has a partial of the T-DNA cassette, lacking *Lackl-Δ12D, Pyrco-Δ6E, Pavsa-Δ4D* and selection marker *PAT* genes. The segregation of these two loci was achieved by back-crossing, and confirmed by loci-specific PCR analysis (Devine et al., 2018). The segregants 14E-0368-04-03 and 14E-0377-04-05 of NS-B50027-4 contained the A05 T-DNA insert only, or the A02 T-DNA insert only, respectively. Wild type (WT, AV Jade), T_5 (B0050-027-18-20-12-19) plants of DHA canola event and the segregants were planted at three field trial sites (Site 1616_NAR, 1617_DOU, and 1619_TOO) in 2016 near Horsham (Victoria, Australia). The sampling times represent specific growth stages of canola allowing for various tissue types, including leaves, roots, pods and reproductive tissues. Samples were harvested at different developmental stages (Table S1) as described by Lancashire et al. (1991). The tissues harvested were maintained on wet ice during transit, and then transferred into a −80 °C freezer until processing.

2.1.2. Reagents

All reagents were purchased from Sigma-Aldrich (Sydney, Australia) unless otherwise specified. The peptides selected for protein quantitation were synthesized at JPT Peptide Technologies GmbH (Berlin, Germany). Two types of synthetic JPT Q-tag peptides were used in this study for absolute quantitation. The first type were purified unlabeled "light" peptides (SpikeTides™ TQ). The second type were "heavy" versions (SpikeTides™ TQL) that were isotopically labelled with stable isotopes at the C-terminal arginine (R) ($^{13}C_6$; $^{15}N_4$) or lysine (K) ($^{13}C_6$; $^{15}N_2$), resulting in a mass shift of +10 or + 8 Da, respectively. The heavy peptides were used as the internal standard for each light peptide. Each peptide was provided freeze dried in 96 micro tube format with an absolute quantity (10 × 1 nmol of net peptide) for each peptide. The amount and purity of each synthesized peptide was determined by the supplier (Protocol S1).

2.2. Sample preparation

2.2.1. Characterisation and digestion of expressed proteins

Protein extracts from a variety of sources including total protein extracts from canola and recombinant proteins expressed in either yeast, bacterial or baculovirus expression systems were used (Protocol S2, Protocols S3-S9). The proteins were prepared either in-solution or as excised gel slices. The solutions were subjected to filter-assisted sample preparation (FASP) as described previously (Colgrave et al., 2015) with minor modifications. The total protein extracts were diluted in 8 M urea, 100 mM Tris-HCl and reduced with 50 mM DTT and incubated at room temperature (22 °C, RT) for 30 min prior to being applied to a 10 kDa molecular weight cutoff filter (Millipore, Australia). The filters were centrifuged (20,800*g*, 10 min) and washed with two 200 μL volumes of 8 M urea, 100 mM Tris-HCl. For cysteine alkylation, 50 mM iodoacetamide in 8 M urea, 100 mM Tris-HCl was added (100 μL) and the solution was incubated at RT in the dark for 45 min. The filters were centrifuged (20,800*g*, 10 min) to remove excess iodoacetamide and washed with two 200 μL volumes of 8 M urea, 100 mM Tris-HCl. The buffer was exchanged using 100 mM ammonium bicarbonate (pH 8.0) by two consecutive wash/centrifugation steps. Trypsin (Promega, Alexandria, Australia) (200 μL, 0.01 μg/μL in 100 mM ammonium bicarbonate, 1 mM $CaCl_2$) was added, and the mixture was incubated overnight at 37 °C. The 10 kDa filters were transferred to fresh centrifuge tubes, and the filtrates (digested peptides) were collected following centrifugation (20,800*g*, 10 min). The filters were washed with 100 μL of 100 mM ammonium bicarbonate, and the filtrates were combined and lyophilised. The tryptic peptides were resuspended in 50 μL of 1% formic acid and stored at 4 °C until LC-MS/MS analysis.

Gel slices were washed and destained according to standard protocols. In brief, the gel slices were vortexed and sonicated in 200 μL of 50:50 100 mM NH_4HCO_3/acetonitrile for 5 min. After removal of the supernatant this wash-step was repeated once. The supernatant was removed and the gel slices were dehydrated twice in 200 μL acetonitrile (100%) for 5 min at RT. Protein disulfides were reduced with 50 μL of 10 mM dithiothreitol (60 min, RT), the supernatant was removed and the sulfhydryls formed were alkylated with 50 μL of 50 mM iodoacetamide (60 min, RT). After removal of the supernatant the gel slices were dehydrated twice in 200 μL acetonitrile (100%) for 5 min at RT. The supernatant was removed and gel slices were air dried for 10 min. Dried gel slices were covered with 50 μL of sequencing grade porcine trypsin (Promega, Alexandria, Australia) at 40 ng/μL in 100 mM NH_4HCO_3. Digestion proceeded overnight at 37 °C in a shaking heating block. Resultant peptides were eluted from the gel pieces in three steps. Firstly, the excess digest solution was removed and combined with two successive 5 min sonication washes; firstly 50 μL of 1% formic acid in 50% acetonitrile, followed by 50 μL of 1% formic acid in 70% acetonitrile. Peptide eluates were evaporated and reconstituted in 1% formic acid prior to mass spectrometry.

2.2.2. Protein extraction and digestion from canola

The frozen canola samples were ground with mortar and pestle into

a fine powder with liquid nitrogen. All samples were maintained frozen on dry ice during the process. To avoid cross contamination, WT samples were processed first, then transgenic samples, in the order of BBCH15, BBCH35, BBCH65 root, BBCH65 other parts, BBCH65 flower, BBCH79, BBCH90. Total protein was extracted following an established protocol for plant membrane protein extraction (Wang et al., 2006) from 1.8 g of samples in 15 mL plastic tubes in order to obtain more than 1 mg of total protein. The tubes with samples were filled with 12 mL of 10% TCA in acetone and vortexed and sonicated at frequency of 25% amplitude for 20 s using a Branson digital probe sonicator (St Louis, MO, USA). Samples were centrifuged at 16,000 × *g* for 3 min at 4 °C. The supernatant was removed by careful decanting. The pellet was re-suspended in 12 mL of 0.1 M ammonium acetate in 80% methanol, mixed by vortex mixing and centrifuged at 16,000 × *g* for 3 min at 4 °C. The supernatant was discarded by careful decanting. The pellet was re-suspended in 12 mL of 80% acetone, vortexed until the pellet was fully dispersed, and centrifuged at 16,000 × *g* for 3 min at 4 °C. The supernatant was discarded, and the pellet was air dried to remove the residual acetone. The air-dried pellet was re-suspended in 3.6 mL of UltraPure buffer-saturated phenol (Invitrogen, catalogue #15513-039, Carlsbad CA, USA) and 3.6 mL freshly prepared sodium dodecyl-sulfate (SDS) polyacrylamide buffer (30% sucrose, 2% SDS, 0.1 M Tris-HCl pH 8.8, 0.1 M dithiothreitol), mixed thoroughly and incubated for 5 min at room temperature. The samples were centrifuged at 16,000 × *g* for 5 min at room temperature. The upper phenol phase was transferred to a new 15 mL tube, and 12 mL of 0.1 M ammonium acetate in 80% methanol was added. The proteins were precipitated at −20 °C overnight. Samples were centrifuged at 16,000 × *g* for 5 min at 4 °C. The supernatant was carefully discarded, and the pellet was first washed with 100% methanol then with 80% acetone. The proteins were pelleted by centrifuging at 16,000 × *g* for 5 min at 4 °C. The final protein pellet was left to air dry.

The extracted proteins from different tissues (n = 252) were dissolved in 8 M urea, 0.1 M Tris-HCl, pH 8.5. Protein estimations were performed using the Bio-Rad microtiter Bradford protein assay (California, USA). The instructions provided by the reagent manufacturer (version: Lit 33 Rev C) were followed. Samples were diluted in water over three dilutions in duplicate and measurements were made at 595 nm using a SpectraMax Plus. Bovine serum albumin (BSA) standard was used in the linear range 0.05 mg/mL to 0.5 mg/mL. The BSA standard concentration was determined by high sensitivity AAA at Australian Proteomics Analysis Facility (Sydney, Australia). Blank-corrected standard curves were run in duplicate. Linear regression was used to fit the standard curve. Protein samples were stored at −80 °C prior to processing.

Protein (250 μg) was subjected to filter-assisted sample preparation (FASP) workflow (as described in 2.2.1). Prior to digestion with trypsin, the pooled stock solution containing the heavy Q-tag peptides was added (0.025 pmol/μL in 100 μL of 50 mM ammonium bicarbonate) as an internal standard (IS). Sequencing grade porcine trypsin (Promega, Alexandria, Australia) was then added (5 μg, 0.5 μg/μL in 100 μL of 50 mM ammonium bicarbonate, 2 mM $CaCl_2$) to the protein on the 10 kDa filters and incubated for 16 h at 37 °C in a wet chamber. The filters were transferred to fresh centrifuge tubes and the filtrate (digested peptides) was collected following centrifugation (20,800 × *g*, 10 min). The filters were washed with 200 μL of 100 mM ammonium bicarbonate and the filtrates were combined and lyophilised. The resultant peptides were re-suspended in 62.5 μL of 1% formic acid containing 0.04 pmol/μL of the IS peptides and 25 μL (equivalent to ~100 μg of total protein and 1 pmol of IS) was analysed by LC-MS/MS.

2.3. LC-MS/MS analysis

2.3.1. Protein characterisation

Trypsin digested protein were analysed as described previously (Colgrave et al., 2014) with chromatographic separation (2%/min linear gradient from 2 to 40% acetonitrile) using a nano HPLC (high performance liquid chromatography) system (Shimadzu Scientific, Rydalmere, Australia) directly coupled to a TripleTOF 5600 MS (AB SCIEX, Redwood City, CA, USA). ProteinPilot™ 4.0 software (AB SCIEX) with the Paragon Algorithm (Shilov et al., 2007) was used for protein identification. Tandem mass spectrometry data was searched against *in silico* tryptic digests of a custom-built database comprising the transgenic proteins. The search parameters were defined as: iodoacetamide modified for cysteine alkylation and trypsin as the digestion enzyme.

2.3.2. MRM method development

After database searching where possible a minimum of five tryptic peptides were selected for each of the seven target proteins (Lackl-Δ12D, Picpa-ω3D, Micpu-Δ6D, Pavsa-Δ5D, Pavsa-Δ4D, Pyrco-Δ6E and Pyrco-Δ5E) and the control protein phosphinothricin-N-acetyltransferase (PAT). The selection of peptides was based on the criteria: MS response (high intensity), absence of amino acids within the peptide sequence that are likely to be modified (for example, oxidation of methionine) or containing missed cleavages (due to the presence of dibasic residues at either terminus), specific/unique to the target protein and of a size amenable to LC-MS (~6–20 amino acids in length). The MS/MS data and subsequent database search results were used to select the highest intensity precursor ion (Q1 m/z) and product ions (Q3 m/z, n = 5) for each peptide which were subsequently defined as the MRM method.

2.3.3. Targeted proteomics

Data were acquired using Analyst 1.6.3 software (AB SCIEX) on a 6500 QTRAP LC-MS/MS system (AB SCIEX). Peptide extracts after trypsin digestion were chromatographically separated on a Shimadzu Nexera UHPLC (ultra-high performance liquid chromatography) and analysed on a 6500 QTRAP mass spectrometer (AB SCIEX) as described previously (Colgrave et al., 2014). Quantitation was achieved using scheduled MRM scanning experiments using a 120 s detection window for each MRM transition and a 0.3 s cycle time. Peaks were integrated using MultiQuant v3.0 (AB SCIEX) wherein all three MRM transitions were required to co-elute at the same RT with a signal-to-noise (S/N) > 3 for detection and a S/N > 5 for quantitation.

The two best performing peptides per protein (where possible) were selected as the signature peptides for each enzyme based on criteria such as chromatographic performance (good peak shape), intensity in MS, free from interference (as assessed in sample matrix). For each selected peptide, both the endogenous light peptides (SpikeTides™ TQ) and ^{15}N and ^{13}C heavy peptides (SpikeTides™ TQL) were synthesized (Table S2) and the MRM transitions are defined in Table S3. Using the synthetic peptides, the linear range of quantitation was determined using spiked WT-canola extracts. An aliquot (25 μL) of spiked canola peptide extract was analysed as described above.

2.3.4. Assessment of technical variation

The technical variation from protein extraction through to data acquisition was assessed by preparing five tryptic digests of three biological replicates of DHA canola seed from a single batch (grown under identical conditions). The twelve peptides selected were assessed for the co-efficient of variation (CV) with all values < 26% CV and 31/36 values were < 15% CV. No single peptide yielded CVs for each of the biological replicates > 15% (Table S4). The biological variation was noted to be < 25% for all peptides.

2.3.5. Generation of calibration curves

Calibration curves were generated (Protocol S1) to determine the linear range of quantitation suitable for determining the absolute peptide concentration. The analyte concentration (light peptide) was varied and a defined amount (equivalent to 1 pmol on-column) of IS (heavy peptide) was spiked into WT canola matrix. The ratio (summed analyte peak area)/(summed IS peak area) was plotted against the

known analyte concentration using Graphpad Prism v6. Absolute quantitation of the target proteins was achieved by interpolation from the calibration curve. The amount of each target peptide (as femtomoles per 100 µg total protein) was determined by interpolation from the calibration curve using the best model to fit the data (non-linear regression, third order polynomial cubic, Fig. S1). The adjusted R^2 values were improved using this interpolation method (e.g. LAPLVK $R^2 = 0.9984$, linear; vs. 0.9993, cubic). The amount of protein detected in these samples was then calculated based on the protein molecular mass, by conversion to a nanogram equivalent per mg total protein. Two calibration curves were run three weeks apart to assess reproducibility of the analytical method. The analytical parameters for quantitation of canola peptides wherein limit of detection (LOD), lower limit of quantification (LLOQ) and upper limit of quantification (ULOQ) are given in femtomoles for the non-linear regression model with the adjusted ULOQ, the standard deviation of the residuals (Sy.x), and the sum-of-the-squares (SS) are listed in Table S5.

2.3.6. Protein quantitation

For absolute quantitation, a reference peptide master mix containing all of the heavy IS peptides was prepared and spiked into 250 µg of each canola total protein extract and measured in biological triplicates. The heavy peptides served as internal standards (IS) for determining the correct RT, MRM transition order, and for absolute quantitation. The spiking was performed prior to protein trypsin digestion to ensure cleavage of the proprietary JPT Q-tag. The extracted and digested protein from the seven different tissue types (representing five growth stages) from three biological replicates taken across three sites, comprising wild type (WT), DHA canola (NS-B50027-4), and the segregants (14E-0368-04-03 and 14E-0377-04-05) (a total of 252 samples) were analysed by LC-MRM-MS.

3. Results and discussion

3.1. Expressed transgenic proteins in DHA canola

The pathway consists of the *Lachancea kluyveri* Δ12-desaturase (Lackl-Δ12D (Watanabe et al., 2004)), *Pichia pastoris* Δ15-/ω3-desaturase (Picpa-ω3D (Zhang et al., 2008)), *Micromonas pusilla* Δ6-desaturase (Micpu-Δ6D (Petrie et al., 2010b)), *Pyramimonas cordata* Δ6-elongase (Pyrco-Δ6E (Petrie et al., 2010a)), *Pavlova salina* Δ5-desaturase (Pavsa-Δ5D (Zhou et al., 2007)), *P. cordata* Δ5-elongase (Pyrco-Δ5E (Petrie et al., 2010a)) and *P. salina* Δ4-desaturase (Pavsa-Δ4D (Zhou et al., 2007)). 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 (Fig. 1, blue) that introduce a double bond at the Δ12 and Δ15 positions, respectively; (2) algal fatty acid elongases including Pyrco-Δ6E and Pyrco-Δ5E (Fig. 1, purple) that elongate the carbon chain by adding two carbons to the carboxyl end of fatty acids; and (3) algal front-end fatty acid desaturases that introduce a double bond between an existing double bond and the carboxyl end of fatty acids (Zhou et al., 2007) including Micpu-Δ6D, Pavsa-Δ5D and Pavsa-Δ4D (Fig. 1, green). The elite event of DHA canola contains two T-DNA inserts, one on chromosome A02 and the other on chromosome A05. A02 chromosome has a partial T-DNA insert containing complete genes Micpu-Δ6D, Pyrco-Δ5E, Pavsa-Δ5D and Picpa-ω3D, while A05 chromosome has two complete T-DNA copies that each contain the eight-gene set linked by a 156 bp palindromic left border sequence in a palindromic structure with an RB-LB:LB-RB orientation (Devine et al., 2018). The segregants of DHA canola containing either only the A02 insert or only the A05 insert were also used for protein expression analysis.

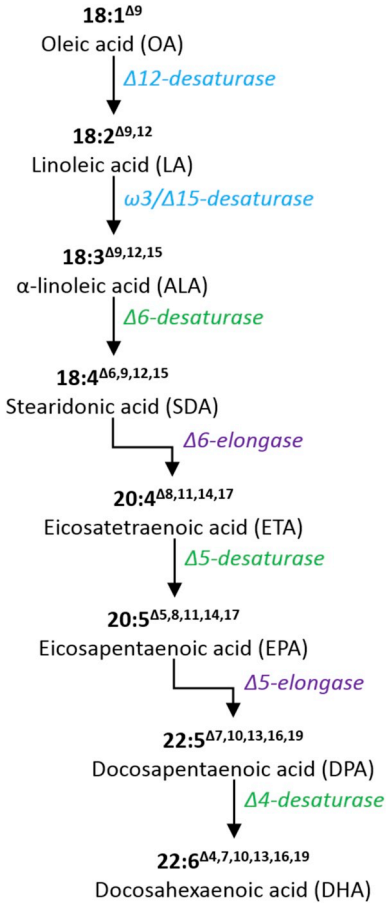

Fig. 1. DHA biosynthesis pathway engineered into canola. The seven enzymes introduced in canola to convert oleic acid to final product docosahexaenoic acid were grouped into three classes: two fatty acid desaturases from yeast in blue; two elongases from microalgae in purple; and three front-end desaturases from microalgae in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.2. Attempts at producing individual antibodies against DHA pathway enzymes

For each enzyme in the DHA biosynthesis pathway, three synthetic peptides were selected for immunisation into rabbits, but no antibodies were produced that could bind specifically to any protein of the expected molecular weight in a Western blot (data not shown). Domains predicted to be outside the transmembrane regions of the enzymes were expressed in *E. coli*, either alone or as a fusion with green fluorescent protein (GFP). Only the domains from ω3, Δ6 and Δ4 desaturases were able to be purified and used as inoculants, and of these only ω3 and Δ6 proteins produced antibodies that were able to detect the full length protein in Western blots (data not shown). Attempts were made to express each of the full length enzymes either as GFP fusions in *E. coli*, or without a fusion partner in *P. pastoris*, but insufficient protein was purified for use as an inoculant. All of the enzymes were eventually expressed in *Sf9* insect cells, however by the time these proteins were prepared, the LC-MS assay described in this paper had been developed and the purified protein was directed towards LC-MS assay optimisation rather than towards time consuming antibody production.

3.3. Development and validation of a multiplexed LC-MS assay

In the absence of functioning antibodies against these membrane integral proteins required for protein quantitation by traditional Western blot or ELISA, an alternative approach using high sensitivity

LC-MRM-MS quantitation was developed to quantify the target proteins.

Using protein extracts from a variety of sources, including total protein extracts from canola and recombinant proteins expressed in either yeast, bacterial, or baculovirus expression systems (Protocol S2), the peptides liberated after tryptic digestion were characterised (Fig. S2). Using these data, peptide markers for each of the transgenic proteins were selected and a quantitative method employing multiple reaction monitoring (MRM) mass spectrometry was developed and optimized using up to five peptides from each target protein. The two peptides with optimal performance characteristics (high signal intensity, good chromatographic properties; Fig. S3) for each protein were selected as the signature peptides for protein quantitation and were synthesized as light and heavy versions (Table S2). The exceptions were the elongases (Δ6, Δ5), Δ5-desaturase and PAT, wherein only one peptide gave performance deemed suitable for absolute quantitation. The isotopically labelled heavy peptides were used as internal standard (IS) partner peptides for quantitation, as well as for determining the correct RT and MRM transition order.

3.4. Calibration curves and linearity

Calibration curves (Fig. S4) were generated by spiking light and heavy labelled internal standard peptides into WT canola to account for possible matrix effects, which included retention time deviation. Peptide markers representing seven of the eight proteins gave a linear response over the concentration range of 31.25 to 1,000 femtomoles (Fig. S4). This was suitable for quantitating all proteins in transgenic canola with the exception of the Pavsa-Δ4D peptide LAPLVK, which yielded an MS response outside of the linear range. In order to quantitate the more abundant Pavsa-Δ4D the calibration curve range was extended to 7.8-4,000 femtomoles to accommodate the full dynamic range of the seven pathway enzymes. Four target proteins yielded two peptide markers that could be used for detection and the peptides showed a highly similar pattern across all samples tested (Fig. S5). In the case where two peptides were available, the peptide marker demonstrating the lowest co-efficient of variation (CV), a lower limit of detection (LOD) and/or lower limit of quantitation (LLOQ) and/or higher R^2 value were employed for the final quantitative method. The calibration standards were analysed in duplicate three weeks apart and the variation spanning the linear range was shown to be < 5% for all peptides except the Δ5E peptide marker which showed variation of ~7%.

3.5. Protein quantitation in transgenic canola

Total protein was extracted from seed yielding values in the range of 1.6–2.2 mg/g from developing seed or 2.8–4.4 mg/g from mature seed. LC-MRM-MS quantitation confirmed that none of the targeted peptides were detected in the total protein extracts from WT canola, including all seven tissues sampled at five growth stages (Table S1), and from three field trial sites in Victoria, Australia grown over the period May to November, 2016. The seed-specific expression of the DHA biosynthesis pathway enzymes was confirmed as the transgenic desaturase and elongase proteins were detected only in developing and mature seeds. There was no detection of DHA biosynthesis pathway enzymes in the non-seed tissues of DHA canola at any growth stage. These tissues included whole plant at BBCH15 (stage at 5 true leaves), whole plant at BBCH35 (stage at 3 visibly extended internodes), flowers, roots and remaining plant material at BBCH65 (stage at 50% full flowering). Four of the seven enzymes (*i.e.* Picpa-ω3D, Micpu-Δ6D, Pavsa-Δ5D and Pyrco-Δ5E) were detected in the partial A02 T-DNA insert segregant that contained only these four genes. Fig. 2 shows an example of an LC-MS chromatogram revealing detection of the target proteins in mature DHA canola seed (Fig. 2A) with reference to detection of the isotopically labelled (heavy) internal standards in WT-canola (Fig. 2B). Evidence for detection of all eight target proteins (including phosphinothricin-N-acetyltransferase, PAT) across the four lines is provided as Figs. S6–S13.

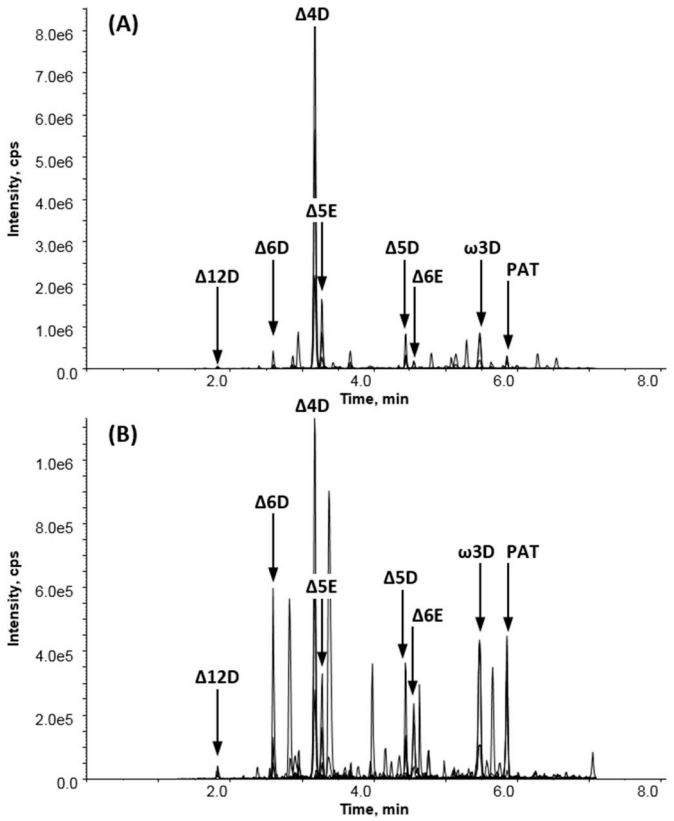

Fig. 2. LC-MS detection of light signature peptides representing seven target proteins and the control protein PAT in developing seed of DHA canola (A); and of heavy internal standard peptides in AV-Jade WT canola (B).

Based on the molecular mass of each protein, the level of each transgenic protein was determined (on a per mg protein basis, Tables 1–3, Fig. 3). The data show that the biological variation between batches was higher (average of ~40%) than the technical variation (typically < 15%). Examining the protein levels in the DHA canola on a site-by-site basis, the variation was notably lower in the mature seed (average across all proteins of 11%) compared to the developing seed (~31%) indicating that the final protein level in the mature seed was more uniform than in the dynamic environment in the developing seed. The developing seed harvested at BBCH79 was defined as "nearly all fruits have reached final size" (Lancashire et al., 1991), while mature seeds were fully developed. In practice, a proportion of the harvested developing seeds would span a range of developmental stages, which would lead to higher protein level variation than in mature seeds. In addition, the environmental variation between different sites is also expected to contribute to variation in the levels of expressed protein. For example, in this study, the levels of several transgenic proteins in seeds harvested from site 3 were higher than that observed from the two other sites (Fig. 3). Environmental variation has previously been noted to exceed genetic variation as a contributor to alterations in seed protein levels (Pritchard et al., 2000).

The Pavsa-Δ4D protein was the most abundant among the seven transgene products detected in developing seed or mature seed with up to 1,500 ng per mg of total protein (0.15%) in mature seed and 5,600 ng per mg of total protein (0.56%, range of 3,200-5,600 ng/mg total protein) in developing seed. Being the most abundant protein and hence easiest to detect, we searched for, but found no evidence of the Pavsa-Δ4D peptide marker LAPLVK in whole plant taken at the first growth stage (5 true leaves, BBCH15), whole plant taken at the second growth stage (3 visibly extended internodes, BBCH35), or in root, flower or the

Table 1
Absolute quantitation of transgenic proteins in developing and mature seed of DHA canola (NS-B50027-4).

Protein	Mature Seed (BBCH90)[a]			Developing Seed (BBCH79)[a]		
	Site 1	Site 2	Site 3	Site 1	Site 2	Site 3[b]
Lackl-Δ12D	412 ± 95	476 ± 60	406 ± 81	670 ± 610	1,050 ± 170	850 ± 120
Picpa-ω3D	483 ± 94	430 ± 30	351 ± 92	650 ± 530	1,057 ± 28	635.1 ± 9.1
Micpu-Δ6D	190 ± 91	220 ± 140	210 ± 180	470 ± 370	870 ± 300	794 ± 95
Pyrco-Δ6E[c]	D	D	D	230 ± 120	430 ± 220	480 ± 50
Pavsa-Δ5D	742 ± 92	930 ± 290	760 ± 480	940 ± 750	2,030 ± 270	1,900 ± 200
Pyrco-Δ5E[c]	409 ± 78	500 ± 270	440 ± 320	670 ± 490	1,310 ± 390	1,290 ± 150
Pavsa-Δ4D	1,550 ± 130	1,430 ± 300	1,340 ± 520	3,200 ± 2,700	5,600 ± 1,300	5,560 ± 920
PAT	23.1 ± 4.5	33 ± 11	29.2 ± 5.5	330 ± 140	390 ± 210	605 76

[a] Units are ng of transgene protein per mg total protein extracted reported as mean ± SD, n = 3, except.
[b] Where n = 2. D, detected but below LOQ.
[c] The elongases showed limited resistance to trypsin proteolysis and it is estimated that the content of Pyrco-Δ6E is 5.2-fold higher and Pyrco-Δ5E is 3.2-fold higher.

Table 2
Absolute quantitation of transgenic proteins in developing and mature seed of A05 segregant (14E-0368-04-03).

Protein	Mature Seed (BBCH90)[a]			Developing Seed (BBCH79)[a]		
	Site 1	Site 2	Site 3	Site 1	Site 2	Site 3
Lackl-Δ12D	640 ± 270	720 ± 220	894 ± 180	700 ± 500	510 ± 460	224 ± 45
Picpa-ω3D	350 ± 160	390 ± 94	351 ± 53	210 ± 130	220 ± 190	126 ± 36
Micpu-Δ6D	130 ± 110	130 ± 50	261 ± 81	140 ± 130	180 ± 170	83 ± 22
Pyrco-Δ6E[b]	D	D	D	330 ± 170	390 ± 160	207 ± 50
Pavsa-Δ5D	600 ± 380	550 ± 260	1,200 ± 200	440 ± 370	440 ± 480	287 ± 77
Pyrco-Δ5E[b]	320 ± 200	334 ± 81	379 ± 87	350 ± 260	320 ± 350	222 ± 32
Pavsa-Δ4D	1,190 ± 590	1,350 ± 360	5,250 ± 880	2,000 ± 1,800	1,900 ± 2,100	1,140 ± 170
PAT	53 ± 25	63 ± 27	147 ± 64	350 ± 220	340 ± 180	248 ± 52

[a] Units are ng of transgene protein per mg total protein extracted reported as mean ± SD, n = 3.
[b] The elongases showed limited resistance to trypsin proteolysis and it is estimated that the content of Pyrco-Δ6E is 5.2-fold higher and Pyrco-Δ5E is 3.2-fold higher.

Table 3
Absolute quantitation of transgenic proteins in developing and mature seed of A02 segregant (14E-0377-04-05).

Protein	Mature Seed (BBCH90)[a]			Developing Seed (BBCH79)[a]		
	Site 1	Site 2	Site 3	Site 1	Site 2[b]	Site 3
Lackl-Δ12D	ND	ND	ND	ND	ND	ND
Picpa-ω3D	170 ± 70	230 ± 100	200 ± 130	450 ± 140	388 ± 68	520 ± 77
Micpu-Δ6D	53 ± 32	56 ± 21	49 ± 16	107 ± 33	117 ± 11	128 ± 42
Pyrco-Δ6E[c]	ND	ND	ND	ND	ND	ND
Pavsa-Δ5D	270 ± 160	353 ± 140	311 ± 75	390 ± 140	330 ± 45	371 ± 54
Pyrco-Δ5E[c]	290 ± 150	370 ± 130	320 ± 87	540 ± 140	450 ± 150	540 ± 170
Pavsa-Δ4D	ND	ND	ND	ND	ND	ND
PAT	ND	ND	ND	ND	ND	ND

[a] Units are ng of transgene protein per mg total protein extracted reported as mean ± SD, n = 3, except.
[b] Where n = 2. D, detected but below LOQ. ND = not detected.
[c] The elongases showed limited resistance to trypsin proteolysis and it is estimated that the content of Pyrco-Δ6E is 5.2-fold higher and Pyrco-Δ5E is 3.2-fold higher.

other tissues sampled when 50% of plants were flowering (BBCH65). This confirmed the seed-specificity of the introduced proteins.

The Pyrco-Δ6E protein was the protein with lowest abundance in DHA canola. It was detected, but was below the limit of quantitation (< LOQ) in mature seeds in both the DHA canola and A05 segregant. The Pyrco-Δ6E protein was quantified in the immature developing seeds, ranging from 230 to 480 ng per mg of total protein (0.023–0.048%) in the DHA canola, and 210–390 ng per mg of total protein (0.021–0.039%) in the A05 segregant. It should be noted that these estimates of the Δ6E and Δ5E amounts in seed are likely to be an under-estimation. The elongases are smaller than the desaturases and comprise 5–6 transmembrane domains such that most of their structure is buried within the membrane restricting access to the tryptic sites. In a related study, the *in vitro* digestibility of the same suite of proteins was investigated (Colgrave et al., 2019). The desaturases yielded maximum levels of the tryptic peptide markers indicating that they were efficiently liberated from the intact protein and thus suitable as proxies for protein quantitation. However, the elongases showed limited resistance to trypsin proteolysis. When the elongases were subjected to a brief period of pepsin digestion (5 min) and then subsequently digested with trypsin (overnight), maximal levels of the tryptic markers were observed representing a 5-fold increase in peptide liberation from Pyrco-Δ6E and 3-fold increase for Pyrco-Δ5E compared to digestion with trypsin alone. Digestion with the pepsin disrupts the tertiary structure exposing the tryptic sites and allowing maximal peptide liberation. From this observation, the adjusted abundance of Pyrco-Δ6E is

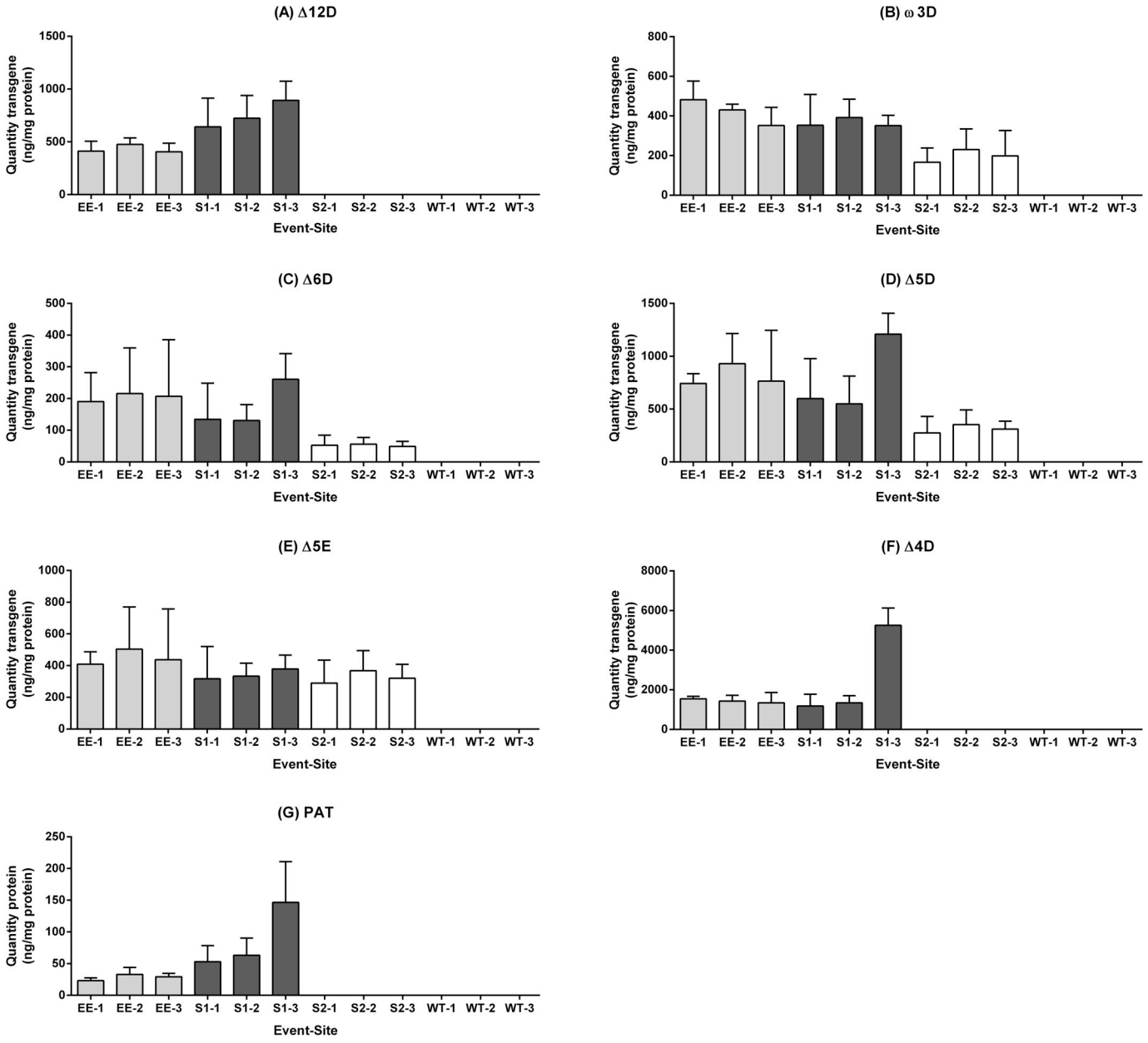

Fig. 3. Absolute quantitation (mean + SD, n = 3 replicates), presented as ng transgenic protein per mg total protein, of six transgenic proteins and the control protein PAT in mature seed of DHA canola: EE represents the DHA canola (elite event NS-B50027-4); S1 represents the A05 segregant (14E-0368-04-03); S2 represents the A02 segregant (14E-0377-04-05); and WT represents canola AV-Jade. Canola was sampled across three field sites. The transgenic protein Δ6E was not detected in mature seed and is not graphed.

conservatively estimated to be ~1.2–2.5 μg of per mg total protein (0.12–0.25%) in developing seed. Likewise, the abundance of Pyrco-Δ5E is estimated to be 1.3–1.6 μg per mg total protein (0.13–0.16%) in mature seed and 2.2–4.2 μg per mg total protein (0.22–0.42%) in developing seed.

3.6. Quantitation of PAT protein in transgenic canola

In contrast to desaturase and elongase genes, expression of the selection marker *Streptomyces viridochromogenes pat* gene was driven by a constitutive 35S promoter. The PAT protein was expected to be in all tissues of DHA canola. Indeed, the PAT protein was detected in all tissue types from the non-seed DHA canola and A05 segregant that contained the PAT selection marker, at levels of 46–310 and 63–390 ng per mg total protein respectively (Table 4). The highest signal intensity was detected for the flower tissue at stage BBCH65 in the DHA canola and the A05 segregant (the full, palindromic insert). In the seed, the PAT protein was also detected in DHA canola and the A05 segregant, with higher expression levels in the developing seed (330–605 and 248–350 ng per mg total protein, respectively) compared to the mature seed (23–33 and 53–147 ng per mg total protein, respectively) (Tables 1 and 2). However, no PAT was detected (< LOD) in the A02 segregant (partial insert, Table 3) or the WT (AV-Jade) in any of the tissues as expected from the molecular characterisation data.

4. Conclusions

The presence and quantity of seven DHA biosynthetic pathway enzymes and the selection marker PAT were assessed in DHA canola. The results of this study demonstrate that the enzymatic proteins that drive the production of DHA using seed-specific promoters were only detected in mature seed and developing seed of DHA canola at levels

Table 4
Detection of PAT in canola non-seed plant parts.

Canola	Whole plant BBCH15 Leaf	Whole plant BBCH35 Node	Root BBCH65	Flower BBCH65	Other BBCH65
DHA canola (NS-B50027-4)	46 ± 15	107 ± 66	173 ± 33	310 ± 180	130 ± 92
A05 segregant (14E-0368-04-03)	147 ± 38[a]	84 ± 63[b]	63 ± 95[b]	390 ± 240[a]	110 ± 67[a]
A02 segregant (14E-0377-04-05)	ND	ND	ND	ND	ND
Control AV Jade	ND	ND	ND	ND	ND

The amount of the PAT peptide SVVAVIGLPNDPSVR detected is reported in units of ng per mg total protein, as mean ± SD, n = 3. ND, not detected. The data represents the average of three biological replicates from three sites (n = 9), except where indicated.
[a] 8/9.
[b] 7/9 replicates.

ranging from 190 ng/mg total protein for the least abundant transgenic protein (Δ6D) to 5,600 ng/mg total protein for the most abundant (Δ4D), while none of the DHA pathway enzymes were detected in other tissues of transgenic canola at any sampling time. No DHA pathway enzymes were detected in any tissues tested in WT canola. The selection marker includes a constitutive promoter, PAT, which was detected in all tissues during the growing season. The LC-MS methodology employed in this study enabled the simultaneous detection of seven transmembrane proteins enabling an estimate of the potential exposure level based on the transgenic protein abundance in the crop. The approach described herein is well suited to the analysis of membrane proteins that are notoriously difficult to express and generate antibodies against, and could see wide application in performance testing in genetically modified crops.

Author contributions

Conceptualization: MC, JP, SS, XZ. Investigation: KB, JC, LK, SVP, BD, AL. Methodology: MC. Formal analysis and data curation: JS, XZ, KB, MC. Writing: KB, MC, XZ.

Data availability statement

All relevant data are within the paper and its Supporting information files.

Conflict of interest

The authors report grants from Nuseed Pty Ltd, during the conduct of the study. In addition, the authors have a patent US 7807849 B2 "Synthesis Of Long-chain Polyunsaturated Fatty Acids By Recombinant Cells" issued.

Acknowledgments

The authors would like to thank the Molecular and Cellular Proteomics Mass Spectrometry Facility at the University of Queensland for access to some of the mass spectrometers used in this study. The authors also thank members of the Plant Oil Engineering Group in CSIRO Agriculture and Food for assistance processing the samples for protein extraction. The study was in part funded by Nuseed Pty Ltd.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fct.2019.02.035.

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.fct.2019.02.035.

References

Aebersold, R., Burlingame, A.L., Bradshaw, R.A., 2013. Western blots versus selected reaction monitoring assays: time to turn the tables? Mol. Cell. Proteomics 12, 2381–2382.
Anderson, L., Hunter, C.L., 2006. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol. Cell. Proteomics 5, 573–588.
Colgrave, M.L., Byrne, K., Caine, J., Kowalczyk, L., Vibhakaran Pillai, S., Dong, B., Lovrecz, G., MacIntosh, S., Scoble, J.A., Petrie, J.R., Singh, S.P., Zhou, X.R., 2019. Assessing the in vitro protein digestibility of seven transmembrane enzymes from the DHA biosynthesis pathway using LC-MS. Food Chem. Toxicol (data not shown).
Colgrave, M.L., Goswami, H., Blundell, M., Howitt, C.A., Tanner, G.J., 2014. Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. J. Chromatogr. A 1370, 105–114.
Colgrave, M.L., Goswami, H., Byrne, K., Blundell, M., Howitt, C.A., Tanner, G.J., 2015. Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination. J. Proteome Res. 14, 2659–2668.
Devine, M., Leonforte, A., Gororo, N., Buzza, G., Tang, S., Gao, W., Petrie, J.R., Singh, S.P., 2018. Elite Event Canola NS-B50027-4. Nuseed Pty Ltd., pp. 1–103.
Foreign Agricultural Service, 2018. Table 01: Major Oilseeds: World Supply and Distribution (Commodity View) United States Department of Agriculture.
Galli, C., Calder, P.C., 2009. Effects of fat and fatty acid intake on inflammatory and immune responses: a critical review. Ann. Nutr. Metab. 55, 123–139.
Hamakubo, T., Kusano-Arai, O., Iwanari, H., 2014. Generation of antibodies against membrane proteins. Biochim. Biophys. Acta 1844, 1920–1924.
Hill, R.C., Oman, T.J., Shan, G., Schafer, B., Eble, J., Chen, C., 2015. Development and validation of a multiplexed protein quantitation assay for the determination of three recombinant proteins in soybean tissues by liquid chromatography with tandem mass spectrometry. J. Agric. Food Chem. 63, 7450–7461.
Hill, R.C., Wang, X., Schafer, B.W., Gampala, S.S., Herman, R.A., 2017. Measurement of lipid transfer proteins in genetically engineered maize using liquid chromatography with tandem mass spectrometry (LC-MS/MS). GM Crops Food 8, 239–252.
Hu, X.T., Owens, M.A., 2011. Multiplexed protein quantification in maize leaves by liquid chromatography coupled with tandem mass spectrometry: an alternative tool to immunoassays for target protein analysis in genetically engineered crops. J. Agric. Food Chem. 59, 3551–3558.
Kirkpatrick, D.S., Gerber, S.A., Gygi, S.P., 2005. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. Methods 35, 265–273.
Ladics, G.S., 2008. Current codex guidelines for assessment of potential protein allergenicity. Food Chem. Toxicol. 46 (Suppl. 10), S20–S23.
Lancashire, P.D., Bleiholder, H., Vandenboom, T., Langeluddeke, P., Stauss, R., Weber, E., Witzenberger, A., 1991. A uniform decimal code for growth-stages of crops and weeds. Ann. Appl. Biol. 119, 561–601.
Petrie, J.R., Liu, Q., Mackenzie, A.M., Shrestha, P., Mansour, M.P., Robert, S.S., Frampton, D.F., Blackburn, S.I., Nichols, P.D., Singh, S.P., 2010a. Isolation and characterisation of a high-efficiency desaturase and elongases from microalgae for transgenic LC-PUFA production. Mar. Biotechnol. 12, 430–438.
Petrie, J.R., Shrestha, P., Mansour, M.P., Nichols, P.D., Liu, Q., Singh, S.P., 2010b. Metabolic engineering of omega-3 long-chain polyunsaturated fatty acids in plants using an acyl-CoA Delta6-desaturase with omega3-preference from the marine microalga Micromonas pusilla. Metab. Eng. 12, 233–240.
Pritchard, F.M., Eagles, H.A., Norton, R.M., Salisbury, P.A., Nicolas, M., 2000. Environmental effects on seed composition of Victorian canola. Aust. J. Exp. Agric. 40, 679–685.
Schacherer, L.J., Owens, M.A., Hu, T.X., 2017. Quantification of intractable membrane proteins in genetically engineered crops by liquid chromatography coupled with tandem mass spectrometry. Anal. Meth.-Uk 9, 2821–2829.
Shilov, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., Hunter, C.L., Nuwaysir, L.M., Schaeffer, D.A., 2007. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. Mol. Cell. Proteomics 6, 1638–1655.
Tocher, D., 2009. Issues surrounding fish as a source of omega-3 long-chain polyunsaturated fatty acids. Lipid Technol. 21, 13–16.
Wang, W., Vignani, R., Scali, M., Cresti, M., 2006. A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. Electrophoresis 27, 2782–2786.
Watanabe, K., Oura, T., Sakai, H., Kajiwara, S., 2004. Yeast Delta 12 fatty acid desaturase:

gene cloning, expression, and function. Biosci. Biotechnol. Biochem. 68, 721–727.
Wisniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., 2009. Universal sample preparation method for proteome analysis. Nat. Methods 6, 359–362.
Zhang, H., Liu, Q., Zimmerman, L.J., Ham, A.J., Slebos, R.J., Rahman, J., Kikuchi, T., Massion, P.P., Carbone, D.P., Billheimer, D., Liebler, D.C., 2011. Methods for peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. Mol. Cell. Proteomics 10 M110 006593.
Zhang, X., Li, M., Wei, D., Xing, L., 2008. Identification and characterization of a novel yeast omega3-fatty acid desaturase acting on long-chain n-6 fatty acid substrates from Pichia pastoris. Yeast 25, 21–27.
Zhou, X.R., Robert, S.S., Petrie, J.R., Frampton, D.M., Mansour, M.P., Blackburn, S.I., Nichols, P.D., Green, A.G., Singh, S.P., 2007. Isolation and characterization of genes from the marine microalga Pavlova salina encoding three front-end desaturases involved in docosahexaenoic acid biosynthesis. Phytochemistry 68, 785–796.