Plant Biotechnology Journal

Plant Biotechnology Journal (2014), pp. 1–11 doi: 10.1111/pbi.12286

Transgenic soya bean seeds accumulating β-carotene exhibit the collateral enhancements of oleate and protein content traits

Monica A. Schmidt[1,*], Wayne A. Parrott[2], David F. Hildebrand[3], R. Howard Berg[4], Amanda Cooksey[5], Ken Pendarvis[5], Yonghua He[1], Fiona McCarthy[5] and Eliot M. Herman[1]

[1]*School of Plant Sciences/BIO5 Institute, University of Arizona, Tucson, AZ, USA*
[2]*Crop and Soil Sciences, University of Georgia, Athens, GA, USA*
[3]*Plant and Soil Sciences, University of Kentucky, Lexington, KY, USA*
[4]*Danforth Plant Science Center, St.Louis, MO, USA*
[5]*School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, AZ, USA*

Received 30 July 2013;
revised 19 September 2014;
accepted 24 Septmber 2014
*Correspondence (Tel (520) 626 1643;
fax 520 626 4824;
email monicaschmidt@email.arizona.edu)

Summary

Transgenic soya bean (*Glycine max*) plants overexpressing a seed-specific bacterial phytoene synthase gene from *Pantoea ananatis* modified to target to plastids accumulated 845 μg β carotene g^{-1} dry seed weight with a desirable 12:1 ratio of β to α. The β carotene accumulating seeds exhibited a shift in oil composition increasing oleic acid with a concomitant decrease in linoleic acid and an increase in seed protein content by at least 4% (w/w). Elevated β-carotene accumulating soya bean cotyledons contain 40% the amount of abscisic acid compared to nontransgenic cotyledons. Proteomic and nontargeted metabolomic analysis of the mid-maturation β-carotene cotyledons compared to the nontransgenic did not reveal any significant differences that would account for the altered phenotypes of both elevated oleate and protein content. Transcriptomic analysis, confirmed by RT-PCR, revealed a number of significant differences in ABA-responsive transcripton factor gene expression in the *crtB* transgenics compared to nontransgenic cotyledons of the same maturation stage. The altered seed composition traits seem to be attributed to altered ABA hormone levels varying transcription factor expression. The elevated β-carotene, oleic acid and protein traits in the β-carotene soya beans confer a substantial additive nutritional quality to soya beans.

Keywords: soybean, carotenoid, oleic acid, elevated protein, seed.

Introduction

Soya bean (*Glycine max*) is a global source of edible vegetable oil and high-quality protein, accounting for 25% of the world's edible oils and 75% of the world's protein meal (Soystats, 2012). Breeding efforts to enhance protein content in soya bean have shown a negative correlation with both oil concentration (Li and Burton, 2002; Piper and Boote, 1999) and yield (Cober and Voldeng, 2000; for review Wilson, 2004). Approximately 25% of the calories consumed in the United States are derived from plant oils (Broun *et al.*, 1999). The fatty acid composition of edible oils is an influencing factor of some major human diseases, such as rheumatoid arthritis, inflammatory bowel (Calder, 2001), atherosclerosis, coronary heart disease (Mozaffarian *et al.*, 2006; Shekelle *et al.*, 1981) and certain cancers (Chan *et al.*, 2005). Coronary heart disease is responsible for 1 in every 4 deaths in the United States, making it the leading cause of death for Americans (Michimi *et al.*, 2013). Consumption of trans fats increases the risk of coronary heart disease by raising levels of 'bad' LDL cholesterol and lowering the levels of 'good' HDL cholesterol. In response, food manufacturers have reduced, or eliminated, trans fats from their products (Korver and Katan, 2006). Foods manufactured with a healthier soya bean oil can contribute to the reduction of the incidence of coronary heart disease in the United States.

Like unsaturated fats, the lipophilic carotenoids are known to have a number of benefits to human health. In the human diet, carotenoids have been shown to have antioxidant activity that may help to prevent certain kinds of cancers, arthritis and atherosclerosis (Stahl and Sies, 2003), and there is evidence that carotenoids can improve gut and immune health including allergic reactions (Chew, 1993). Carotenoids are potent antioxidants and might help prevent aliments due to oxidative damage. Although there are over 700 carotenoids identified in nature (Delgado-Vargas *et al.*, 2000), the consumption of only six of them, α-carotene, β-carotene, lutein, lycopene, zeaxanthin and astaxanthin, has been documented to have positive health benefits (Johnson, 2002). Among these beneficial carotenoids, β-carotene has a key role in vision health as pro-vitamin A that is processed in vertebrates on demand into two molecules of active vitamin A (retinol) in the intestine (Olson, 1989). β-carotene has been shown to alleviate deficiencies leading to night blindness

Please cite this article as: Schmidt, M.A., Parrott, W.A., Hildebrand, D.F., Berg, R.H., Cooksey, A., Pendarvis, K., He, Y., McCarthy, F. and Herman, E.M. (2014) Transgenic soya bean seeds accumulating β-carotene exhibit the collateral enhancements of oleate and protein content traits. *Plant Biotechnol. J.*, doi: 10.1111/pbi.12286

and other related nutritional insufficiencies (Haskell *et al.*, 2005), and the enhancement of β-carotene has been a primary goal of biofortification projects of staple crops developed for consumption in less developed nations.

Carotenoid biochemistry has been extensively studied, with most of its pathway enzymes characterized and the relevant encoding genes cloned from numerous organisms. The synthesis of carotenoids begins with the conversion of geranylgeranyl diphosphate to phytoene by phytoene synthase. Subsequently, through the action of four eznymes, phytoene desaturase, ζ-carotene desaturase, ζ-carotene isomerase and carotenoid isomerase, phytoene is converted to lycopene. With additional enzymes, the carotenoid flux progresses onward to α and β-carotene that are then used as substrates to form additional carotenoids such as lutein and zeaxanthin (review Cazzonelli and Pogson, 2010). Because of its key role in human health and frequent incidence of vitamin A deficiency in developing nations, many different crops have been modified to attempt to enhance β-carotene levels in the edible parts. Suppression strategies used in *Brassica* seeds (Yu *et al.*, 2007) and potato tubers (Diretto *et al.*, 2007) targeted lycopene εcyclase to direct the carotenoid flux to produce β-carotene from lycopene and to impede the accumulation of lutein. A spontaneous orange mutation (*Or*) that resulted in an enhanced β-carotene accumulation in cauliflower (*Brassica oleracea*) was identified (Li *et al.*, 2003) and later characterized to have a mutation in a plastid-associated protein; its function is reasoned to be in the regulation of the cellular differentiation of plastids to chromoplasts (Lu *et al.*, 2006). When the *Or* allele was expressed in transgenic potato tubers, chromoplast structures containing carotenoids were seen in the transgenic tissue, but such structures were not observed in potato cultivars accumulating comparable levels of carotenoids (Lopez *et al.*, 2008). Welsch *et al.* (2010) studied a naturally occurring mutant responsible for β-carotene accumulation in cassava (*Manihot esculenta*) roots and discovered the genetic basis to be a variant allele of the phytoene synthase gene that has a single amino acid mutation, increasing its catalytic activity. Its function was confirmed by elevated carotenoid accumulation in transgenic yeast and *E. coli* cells (Welsch *et al.*, 2010). The widely reported success in fortifying rice endosperm tissue with β-carotene producing the original 'Golden Rice' and Golden Rice 2 were achieved by the insertion of two enzymes of the carotenoid pathway, phytoene synthase, phytoene desaturase and lycopene β-cyclase (Paine *et al.*, 2005; Ye *et al.*, 2000). These various reports show the importance of the early carotenoid enzymatic reactions in directing the flux of metabolites into carotenoid biosynthesis to achieve elevated accumulation of β-carotene.

We report here the seed-specific overexpression of a bacterial phytoene synthase gene modified to target to plastids into soya bean results in the accumulation of a high level β-carotene. The β-carotene accumulating seeds also exhibit two additional collateral traits that confer elevated desaturated oil and increased protein that substantially enhances its nutritional value with potential end uses in food and feed.

Materials and methods

Soya bean transformation

A plant expression cassette containing the phytoene synthase gene (*crtB*) from the bacteria *Pantoea ananatis* [formally *Erwinia uredovora* (GenBank accession D90087)] was constructed by placing the 931-bp open reading frame behind the 170-bp chloroplast signal from pea (*Pisum sativum*) rib 1,5-bisphosphate carboxylase (GenBank X04334.1). The chimeric open reading frame was then placed flanking the seed-specific regulatory elements of Le 1 (GenBank K00821.1), specifically a 970-bp lectin promoter and a 324-bp polyadenylation signal. This 2.4 kb seed-specific chloroplast targeted cassette of *crtB* was then placed into a pUC-derived vector engineered to contain constitutive expression of hygromycin resistance gene previously constructed (Schmidt and Herman, 2008a,b; Schmidt *et al.*, 2011). Sequencing of the chimeric open reading was performed using a primer lying near the end of the lectin promoter (5′GATTCCCAGTGAAGAAGGTC3′) to ensure both correct orientation behind the lectin promoter and the in-frame placement of *crtB* gene and the chloroplast signal. The resultant construct was referred to as pcrtB and was used to transform cultivar 'Jack' of soya bean via biolistics as previously described (Schmidt *et al.*, 2005, 2011). Soya bean plants, both *crtB* transgenic and nontransgenic, were grown under greenhouse conditions of 25 °C under 16 h daylight with ~1000 μm^{-2}/s and took approximately 120 days to set seed. Transgenic plants were confirmed to contain both the selectable marker and the *crtB* cassette by genomic PCR (Appendix S1).

β-carotene quantification by high-performance liquid chromatography

Extracts for HPLC analysis were performed on *crtB* and nontransgenic seeds. Methods were as Ravanello *et al.* (2003) with the following modifications of using a Waters Breeze 2 HPLC system with a reverse phase C18 column (Gemini 3 u 110A 150 × 2 mm 3 micron; Phenomenex) and 82% acetonitrile/10% dioxane/8% methanol/0.1% 150 mM ammonium acetate/triethylamine (v/v) as the mobile phase for separation. β-carotene was detected using a photodiode array detector at 450 nm. A standard curve of a dilution series of commercially available β-carotene (Sigma-Aldrich; St.Louis MO USA) was used to determine the amount of β-carotene present in the soya bean seed samples.

Light microscopy

Fresh mid-mature (~150 mg) cotyledons from both nontransgenic and *crtB* transgenic soya bean plants were analysed by light microscopy by placing a thin hand cut section in water on a glass slide. Brightfield images were made on a Nikon Eclipse 800, using a 60×/1.2 NA water immersion objective.

Transmission electron microscopy (TEM)

Mid-mature cotyledons from both *crtB* transgenic and nontransgenic plants were prepared for TEM by cryofixing with a Balzer's high pressure freezer. Samples were freeze substituted for 5 days at −80 °C in 2% osmium tetraoxide in acetone, thawed to room temperature, rinsed in acetone and embedded in Spurr's resin. TEM thin sections were stained with uranyl acetate (70 mg/mL) and lead citrate (27 mg/mL). All TEM was performed using a LEO 912AB microscope with imagery captured using a 2 × 2 k CCD camera operated in montage mode.

Mass spectroscopy proteomic analysis

Three biological replicates of mid-mature cotyledons from both nontransgenic and *crtB* transgenic soya beans were each separated into five protein fractions using differential detergent fractionation as described by McCarthy *et al.* (2005), subsequently digested with trypsin and identified by mass spectroscopy. (detailed proteomic methods, Appendix S1).

Lipid analysis

Oil extracts from mature seeds were analysed by gas chromatography (GC). Fatty acid methyl esters retention were matched with authentic standards and verified by GC-MS (details Appendix S1).

Determination of protein and oil levels

The protein and oil levels of *crtB* transgenic and nontransgenic seeds were determined in bulk by near infrared (NIR) spectroscopy using a Perten (Springfield, IL) DA7200. This NIR seed analyser was calibrated with > one hundred soya bean samples and the calibration samples determined by combustion for protein and Soxhlet for oil (AOAC, 1995; De Castro and Priego-Capote, 2010; Rotundo *et al.*, 2011; Soxhlet, 1879). Protein levels were calculated as total nitrogen X 6.25. Every set of NIR determinations was validated by running 13–20 calibration standards of known values and adjusting the bias settings if needed such that the protein and oil readings are in ±1% of the wet chemistry values. These core set of standards were also analysed for protein and oil by Kjeldahl, acid hydrolysis (Mojonnier flask method) and NMR (Ashraf-Khorassani *et al.*, 2002; Hakoda *et al.*, 2011; Ullah *et al.*, 2011). An ANOVA and *post hoc* comparisons using Tukey HSD test were conducted to compare the mean values obtained in all composition factors of *crtB* transgenics to nontransgenics. Gravimetric determination of moisture levels of seed samples involved drying samples in a convection oven at 103 °C for 36–72 h or until the weights stopped changing. Crude protein (Kjeldahl) was also determined using 8 g of homozygous seeds from each of the 3 *crtB* transgenic lines by Chemical Laboratories, University of Missouri-Columbia (www.aescl.missouri.edu).

Lipidomics analysis

Seeds from *crtB* transgenic and nontransgenic soya beans, five replicates per sample, were extracted for lipids, dried under nitrogen gas and the pellet shipped to Kansas Lipidomics Research Center (http://www.k-state.edu/lipid/lipidomics/) under nitrogen gas for oil composition analysis (details Appendix S1).

Nuclear magnetic resonance spectrometry (NMR)

Nondestructive NMR was performed on both *crtB* soya bean lines and nontransgenic Jack seeds and subsequently compared with regard to total protein oil and moisture content. Seed samples were sent to Seed Laboratory at Oregon State University http://seedlab.oregonstate.edu/node/158 for NMR analysis.

Phytohorome quantification

Liquid chromatography/mass spectroscopy analysis was performed on 80% methanol/1% acetic acid (v/v) extracts from immature soya bean cotyledons from two *crtB* lines and a nontransgenic cv Jack plant at Colarado State University Proteomic and Metabolomics Facility (www.pmf.colostate.edu) for quantification of the phytohormones abscisic acid and phaseic acid. (details Appendix S1).

Nontargeted metabolite analysis

Mid-mature (~150 mg) cotyledonary tissue from both *crtB* transgenic and nontransgenic seeds was analysed for 300 nontargeted metabolite analysis by Metabolomics Research Laboratory (Iowa State University). Two *crtB* lines were used with two independently grown Jack plants as two biological controls for each, with each sample then having three technical replications.

Expression analysis

Immature cotyledons from two *crtB* lines and a nontransgenic had RNA isolated and subsequent transcript sequencing by The University of Arizona Genetics Core (www.uagc.arl.arizona.edu). All data were submitted to the sequence read archive (SRA) and can be found under the BioProject accession SRP041621. For analysis, all reads were quality trimmed and adapters were removed using Trimmomatic 0.25 (Lohse *et al.*, 2012). Paired-end mRNA reads were mapped to the *Glycine max* genome (NCBI Gma_ref_V1.1) and transcriptome (NCBI ref_v1.1_top_level.gff3) using Tophat (Trapnell *et al.*, 2009) with the Bowtie2 algorithm (Langmead and Salzberg, 2012). Transcripts were assembled and their abundances estimated using Cufflinks (Trapnell *et al.*, 2010) (details Appendix S1). Quantitative reverse-transcription PCR was performed to confirm expression levels of genes of interest (Appendix S1).

Results

Production of seed-specific β-carotene enhanced soya bean seeds

Transgenic *crtB* soya bean events were produced that express a seed-specific phytoene synthase gene from *Pantoea ananatis* (Genbank D90087) (Figure 1). The T1 *crtB* cotyledons are a deep orange colour shown in Figure 2a. That β-carotene accumulation is restricted to the cotyledonary tissue, as demonstrated by the absence of colour in the maturing hypocotyl (Figure 2a). The accumulation of the carotenoid in cotyledonary tissue is stable and remains as an accumulated product in the mature dry seeds (Figure 2b) and throughout seed germination. The three transgenic *crtB* lines that were produced were grown to homozygosity and have been grown over six generations. One event has been field-grown, and β-carotene maintained its initial level, demonstrating that β-carotene accumulation is a stable trait in soya bean.

The intracellular site of β-carotene accumulation was evaluated by light and electron microscopy. Some carotenoids accumulate as intraplastid lipid droplets termed plastoglobules that form in close association with the chloroplast thylakoids as the plastid transitions to a chromoplast in senescing or ripening cells (Harris and Spurr, 1969). The presence of β-carotene in the *crtB* seed

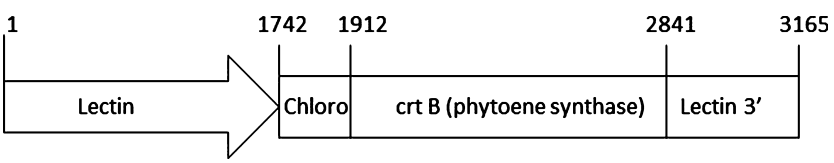

Figure 1 Graphical representation of construct used to direct seed-specific expression of a chloroplast targeted phytoene synthase (*crtB*) gene from *Pantoea* in soya bean. Numbers denote number of nucleotides.

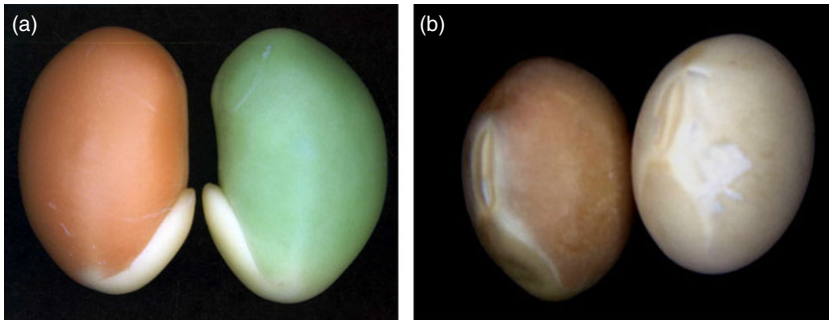

Figure 2 Overall seed phenotype of seed-specific overexpressing *crtB* transgenic soya bean seeds. (left) *crtB* transgenic samples (right) nontransgenic. (a) immature cotyledons (~200 mg) with seed coat removed showing orange colour in the cotyledon tissue only in the *crtB* transgenic; (b) dry soya bean seeds showing the enhanced β-carotene colour seen through the dry seed coat.

plastids was confirmed by direct visualization of the orange coloured organelles in the cotyledon cells. Figure 3a,b show light micrographs of the *crtB* cotyledon cells compared with nontransgenic tissue, demonstrating that orange coloured discrete structures are present in the *crtB* tissue but absent in the nontransformed cotyledon cells. This observation was amplified by transmission electron microscopy using cryofixed samples that showed that the *crtB* plastids accumulated lipid-dense osmophilic structures (Figure 3c) that are not observed in the corresponding nontransgenic cotyledon plastids (Figure 3d). These results show that the cotyledon plastids accumulate β-carotene as intraplastid lipid moieties comparable to the plastoglobuli observed in the organs of many species accumulate β-carotene.

To ascertain the identity of the carotenoid accumulated in the *crtB* transgenic soya bean seeds, the lipid soluble portion of the seeds was extracted and subjected to both TLC and HPLC analysis. The TLC results (Figure S2) indicated the accumulated carotenoid in the transgenic seeds was β-carotene, as demonstrated by co-migration of the cotyledon-derived sample with a β-carotene standard. Quantification the *crtB* transgenic samples was performed by HPLC using a commercial β-carotene standard to calculate the total abundance of β-carotene. The HPLC protocol was designed to resolve both α- and β-carotene as well as the phytoene precursor. The crtB seeds accumulate both α- and β-carotene. Homozygous seeds of the three *crtB* events produced an average β-carotene content of 845 μg β-carotene g^{-1} (dry weight), individually line 1, 2, 3 being 882, 850 and 808 μg/g, respectively. This represents an increase of β-carotene accumulation of over 1500-fold of nontransgenic levels of β-carotene in the soya bean seeds. β-carotene accumulation

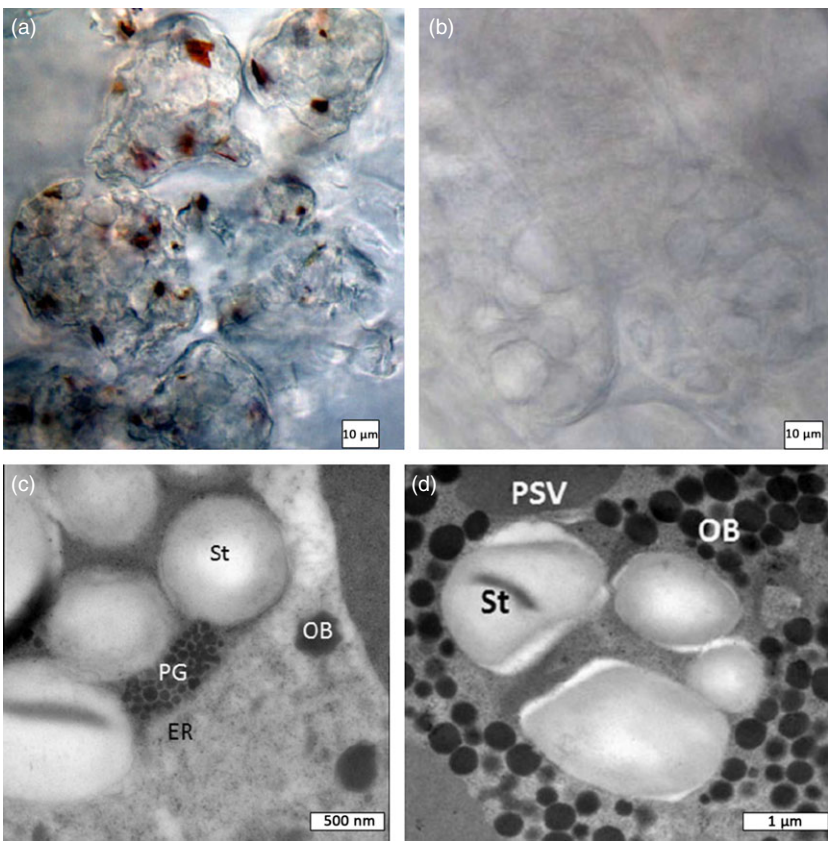

Figure 3 Cellular morphology of *crtB* transgenics compared to nontransgenic. (a) Light micrograph of immature cotyledon tissue from *crtB* transgenic displaying discrete orange structures indicating β-carotene is accumulated in a cellular organelle. (b) Light micrograph of immature cotyledon tissue from nontransgenic soya bean showing no discernible coloured bodies. (c) Cryo-fixed conventional electron micrograph (EM) of cotyledonary tissue from a *crtB* transgenic showing lipid-dense structures, plastoglobuli, present in the chloroplast (d) cryo-fixed conventional EM of cotyledon tissue from nontransgenic soya bean showing a close up of a chloroplast. No plastoglobuli are present in nontransgenic chloroplasts. St: starch gran; PG: plastoglobuli; ER: endoplasmic reticulum; OB: oil body; PSV: protein storage vacuole.

was strongly favoured over α-carotene accumulation, being present in a 12 : 1 β-carotene : α-carotene ratio. The *crtB* seeds also exhibited detectable levels of phytoene in the transgenics, not observed in the nontransgenics (Figure 4), which likely represents a precursor pool of phytoene present in the *crtB* transgenic seeds.

Compositional analysis of enhanced β-carotene seeds show an increase in overall protein content without specific proteome modifications

The *crtB* transgenic soya bean seeds were determined to have an overall increase in protein seed content. Protein content of *crtB* soya bean seeds were assessed by three methods: nuclear magnetic resonance (NMR) analysis (Table 1), near infrared spectroscopy (NIR) (Figure 5) and total Kjeldahl nitrogen. Crude protein determined by Kjeldahl method was 42.5%, 43.7% and 41.7% (w/w) for *crtB* lines 1, 2 and 3, respectively, compared to the nontransgenic counterpart 39.2%. An ANOVA, followed by *post hoc* Tukey HSD test, indicated a significant difference in NIR protein in *crtB* transgenics at the $P < 0.0001$ level. The *crtB* seeds have at least a 4% increase by NIR results, or up to 12% increase in overall protein by NMR analysis with a parallel 3.6–5% overall decrease in oil content (Figure 5). To assess whether the increase in protein content conferred a change in protein composition, the total soluble proteins of mature seeds was analysed by two-dimensional IEF/SDS PAGE. The side-by-side comparison of the polypeptide distribution in broad range IEF and SDS PAGE gels showed that the increased protein content in the *crtB* seeds did not result in significant alteration of the seed proteome (Figure S3). This is further amplified by the comparison of total amino acid content of the *crtB* and nontransgenic seeds that shows a similar relative distribution for both (Figure S1). Mass spectroscopy (MS) analysis of differential detergent fractions of cotyledonary tissue was used to identify proteins in the samples. In the nontransgenic and *crtB* cotyledonary tissue, 2126 and 2235 proteins, respectively, were identified; combined the total proteins identified by differential detergent fractionation and MS analysis was 2642 proteins, representing about 6% of the predicted soya bean proteome. Ultimately, 428 proteins with *P* value less than 0.01 were considered to be differentially expressed when the seed proteome of *crtB* transgenics were compared with nontransgenic. Functional group clustering analysis, performed by REVIGO software (Supek *et al.*, 2011), revealed that differentially expressed proteins involved in photosynthesis were down-regulated, while oxidation-reduction processes were up-regulated in *crtB* compared to nontransgenic tissue (Figure S4). These results show that the increase of protein in the *crtB* seeds does not significantly alter the distribution of the types and composition of the soya bean seed proteins but rather appears to increase the total content of the distributed components of the seed proteome.

Table 1 Nuclear magnetic resonance (NMR) analysis of *crtB* transgenic soya bean compared to nontransgenic. Each reading is a measurement of at least 100 pooled homozygous seeds

Seed sample	Protein (%)	Oil (%)	Moisture (%)
Nontransgenic	28.20	19.83	7.01
crtB-1	45.38	11.94	7.20
crtB-2	47.07	11.58	7.10
crtB-3	40.72	14.48	7.11

Transcript profiling and RT-PCR expression analysis of the β-carotene enhanced seeds indicates a number of significant alterations in the expression of ABA-responsive transcription factors, minor changes in the carotenoid biosynthetic encoding transcripts and the fatty acid desaturase transcripts

The relative differential abundance of transcripts detected in two independent *crtB* transgenic lines compared to nontransgenic transcripts is shown in Figure S7. In total, 21 984 transcripts were identified in the nontransgenic sample and 36 203 and 35 341 transcripts in *crtB* transgenic lines 1 and 2, respectively. Pairwise comparisons using a *P* value less than 0.05 revealed a total of 157 differentially expressed transcripts between the nontransgenic and *crtB* line 1, with 134 of those transcripts up-regulated and 23 down-regulated in the *crtB* sample. Likewise, comparisons of nontransgenic to *crtB* line 2 revealed 118 differentially expressed transcripts with 108 being up-regulated and 10 down-regulated in the *crtB* transgenic line 2. Overall, transcript profiling revealed minor differences in transcripts encoding for carotenoid biosynthetic pathway enzymes, including the phytohormone ABA. Of note is the up-regulation of lycopene β-cyclase and the

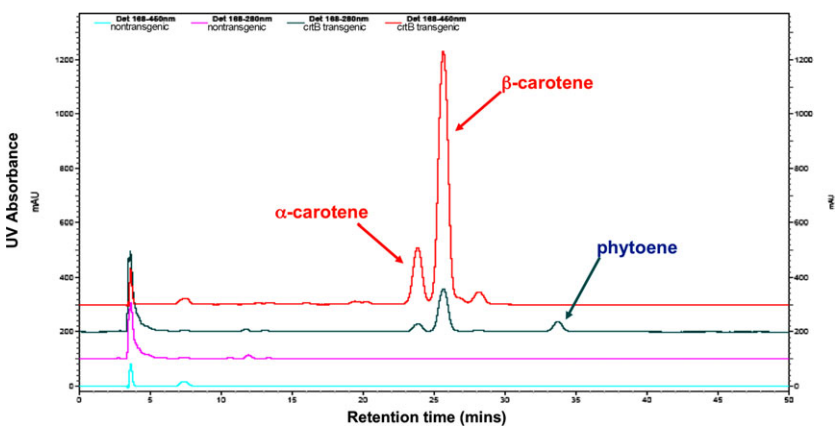

Figure 4 High pressure liquid chromatography (HPLC) of lipid extracts from *crtB* transgenics compared to nontransgenics. Transgenic samples are the top two (the first detection up to 460 nm and the second 280 nm) and the bottom two samples are nontransgenic (third line, 460 nm detection and the bottom line 280 nm detection). Both α and β-carotene are detected at 450 nm and their peaks are discernible in the transgenic but not the nontransgenic. Phytoene is detected at 280 nm and its peak is seen only in the *crtB* sample. *crtB* transgenics have 12 : 1 β-carotene : α carotene.

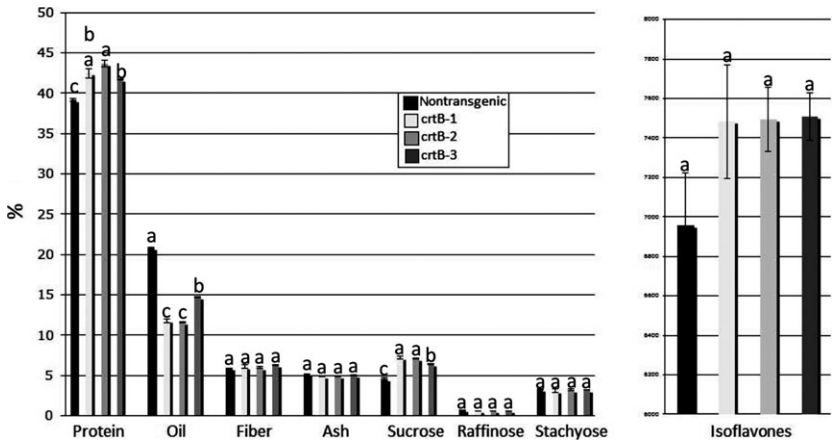

Figure 5 Composition of 3 lines of seed-specific β-carotene enhanced *crtB* transgenic soya bean seeds. Mature dry seeds were analysed by near infrared spectroscopy to determine overall composition of *crtB* transgenics. The *crtB* transgenic seed significantly differ in protein, oil and sucrose levels compared to nontransgenic seeds (ANOVA significant $P < 0.0001$). Values are represented as an average of 3 replicates ± standard error.

down-regulation of both lycopene ε cyclase and ABA-aldehyde oxidase in the *crtB* transgenics. Together, this result accounts for the 12 : 1 ratio of β-carotene : α carotene in the *crtB* transgenic seeds as the simultaneous up-regulation of lycopene β-cyclase and down-regulation of lycopene ε cylase would preferentially produce β-carotene, the favoured precursor to bioactive vitamin A (Figure 6). Figure 7 depicts the average log fold difference detected in transcript abundance in two independent *crtB* transgenic lines compared to nontransgenic. Of the FAD2 genes and FAD3 genes, all but FAD3a was detected to be down-regulated in the *crtB* transgenics. RT-PCR was performed to verify gene expression of FAD genes and confirmed the down-regulation of FAD2-1 in *crtB* transgenics (Figure S8 and Table S3). The suppression of FAD2 genes likely accounts of the elevated amount of oleic acid in the *crtB* seeds.

Of the differentially expressed transcripts detected in the two independent *crtB* lines compared to nontransgenic, a number of ABA-responsive transcription factors were detected to have significantly increased transcript abundance, namely ABI3 (LOC100810343) 5.8-fold, AP2 (LOC100527371) 4.7-fold, ABI4 (LOC100779779) fivefold, YABBY-4 (LOC100798441) 2.2-fold, YABBY-5 (LOC100804731) 3.5-fold and NAC2 3.9-fold. RT-PCR was performed on a mixture of *crtB* immature cotyledons and compared to a mixture of nontransgenic tissue. Of the transcription factors tested, most verified the transcript data set in that the genes were overexpressed in *crtB* tissue (Figure S8). YABBY (not

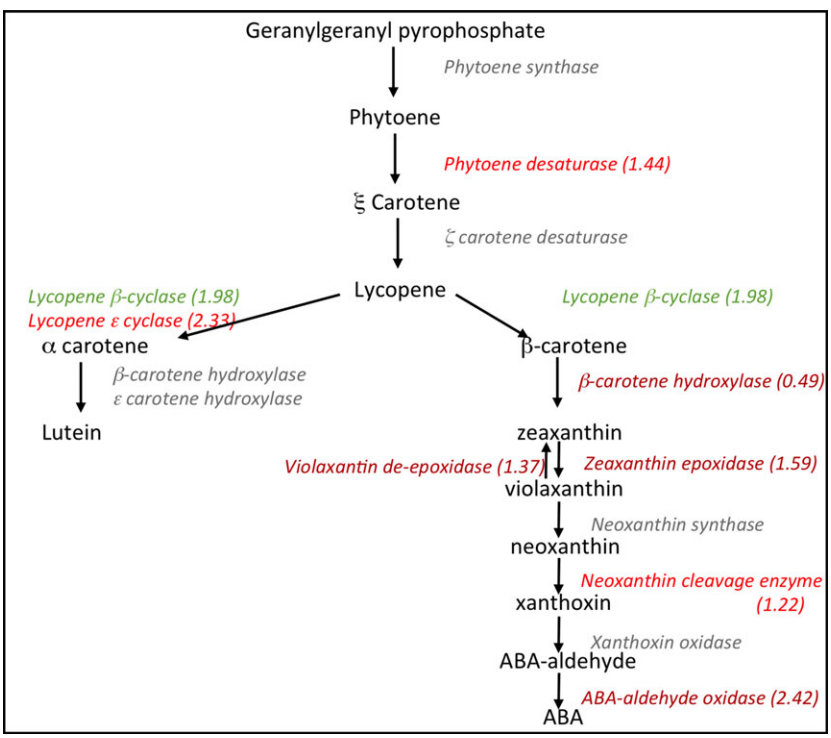

Figure 6 Carotenoid and ABA biosynthetic pathway with detected transcript abundance. Red: down-regulated transcript, green up-regulated transcript in *crtB* transgenics compared to nontransgenic. Numbers in () indicate the average log2 fold change in transcript abundance.

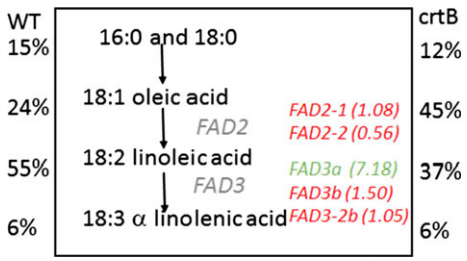

Figure 7 Simplified triglyceride saturation pathway with corresponding transcript abundance. % of oil moieties shown for both WT and *crtB* transgenics. Transcripts for all FAD2 and FAD3 were detected and indicated in red if underexpressed in crtB transgenics and in green if overexpressed in crtB transgenics. Number in () indicates the average log2 fold change in transcript abundance compared to WT.

shown) is a gene family with close homology it was not technically feasible to amplify individual genes.

Oil compositional analysis of enhanced β-carotene seeds show an overall decrease in oil content, with an elevated amount of oleic acid

The *CrtB* transgenic seeds had increased protein content, decreased oil content, increase in sucrose, with other components unchanged (Figure 5). The fatty acid composition of the β-carotene seeds showed a shift favouring unsaturated fatty acids, with the *CrtB* seeds containing 12% saturated fats, 45% oleic acid, 37% linoleic acid and 6% linolenic acid. In contrast, nontransgenic seeds grown side-by-side in the greenhouse had a fatty acid composition of 15% saturated fats, 24% oleic acid, 55% linoleic acid and 6% linolenic acid 6% (Figure 8). The overall fatty acid composition of the *CrtB* soya bean is characterized by a collateral change to a 'mid-oleic' composition similar to that achieved by breeders selecting for enhanced oleic acid content. To further characterize the oil composition, seeds were subjected to nontargeted lipidomics analysis (Kansas Lipodomics Research Center). The most noteworthy changes in lipid composition

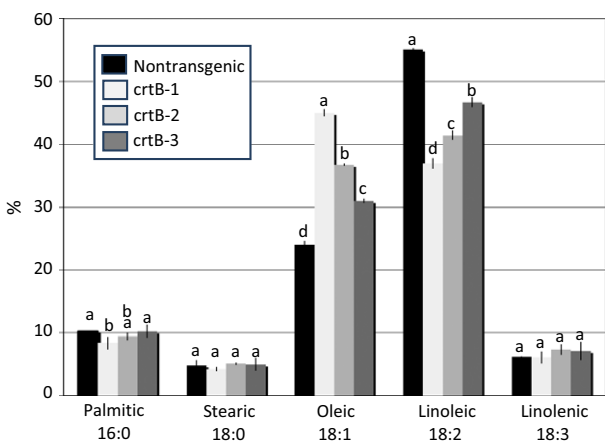

Figure 8 Oil compositional analysis of *crtB* transgenic seeds. Lipid extracts from 3 lines of *crtB* transgenics and nontransgenic were analysed by gas chromatography. The *crtB* transgenic seeds significantly differ in 18 : 1 and 18 : 2 oil composition compared to nontransgenic seeds (ANOVA significant $P < 0.0001$). Values are represented as an average of 3 replicates ± standard error.

Table 2 Phytohormone levels in enhanced β-carotene (*crtB*) mid-mature cotyledons compared to nontransgenic samples. Three technical replicates were performed and results shown are an average ± standard error

Phytohormone (ng/mL)	crtB line 1	crtB line 2	Nontransgenic
Abscisic acid	3.243 ± 0.163	3.216 ± 0.210	5.444 ± 0.339
Phaseic acid	0.377 ± 0.004	0.325 ± 0.008	0.530 ± 0.006

observed were the plastid lipids, MGDG and PG, with the former increasing and the later decreasing in content in the *CrtB* seeds (Figure S5; Table S1).

Enhanced β-carotene seeds have a 40% reduction in the phytohormone absciscic acid

Two lines of *crtB* transgenic soya bean cotyledons were used to quantitate the phytohormones abscisic acid and phaseic acid. The *crtB* seeds were determined to contain 3.243 ± 0.163 and 3.216 ± 0.210 ng/mL ABA compared to nontransgenic levels 5.444 ± 0.339 ng/mL, indicating a reduction of 40% in the transgenic seed tissue. Similarly, the transgenic tissue was determined to have a 30% reduction (*crtB* 0.377 ± 0.004 ng/mL and 0.325 ± 0.008 ng/mL, nontransgenic 0.530 ± 0.006 ng/mL) of phaseic acid (Table 2). Further metabolic profiling, revealed minor differs in *crtB* seeds within the 300 distinct molecules detected and quantified (Table S2). Although phytosterols and tocopherols have a common precursor metabolite, there was no significant difference in phytosterol type and quantity between the transgenic and nontransgenic seed (Figure S6a). Of the four forms of tocopherols, there was only minor changes in both γ-tocopherol and δ-tocopherol in the transgenic seeds (Figure S6).

Discussion

Vitamin A deficiency is the most prevalent nutritional insufficiency in developing countries. This situation has prompted efforts to biofortify crops with β-carotene as a novel delivery system for vitamin A for consumption in developing countries. The most prominent example of this approach, 'Golden Rice' (Enserink, 2008; Paine *et al.*, 2005) pioneered this strategy. Numerous other crops have been engineered for increased levels of β-carotene accumulation, predominantly by the overexpression of phytoene synthase (e.g. *CrtB*). Examples of β-carotene enhanced crops include potato (*Solanum tuberosum*, Ducreux *et al.*, 2005), tomato (*Solanum lycopersicon*; Romer *et al.*, 2000; Fraser *et al.*, 2002), canola (*Brassica napus*; Shewmaker *et al.*, 1999), flaxseed (*Linum usitatissimum;* Fujisawa *et al.*, 2008), soya bean (*Glycine max*; Kim *et al.*, 2012) and maize (*Zea mays*; Aluru *et al.*, 2008). The daily required intake for an adult male is 900 μg/day vitamin A (Penniston and Tanumihardjo, 2003). However, the conversion of β-carotene to vitamin A is not very efficient, so the resulting dietary requirement is nearly 5 mg/day (Von Lintig, 2010; ; Haskell, 2012). These results show that soya beans will produce high levels of carotene with an optimal 12 : 1 β to α ratio, so that nearly all of the β-carotene produced will be potentially biologically active. β-carotene enhanced soya beans can produce the adult male RDA levels with as few as 5 g of soya beans. In contrast, the other examples of β-carotene production in seeds report 1 : 1 or 3 : 1 β : α carotene ratio (Kim *et al.*, 2012; Ravanello *et al.*, 2003). Currently, a child would need to consume

100–150 g of Golden Rice to meet the daily requirements (Tang *et al.*, 2012). Similarly, red palm oil has approximately 500–700 ppm carotene with half of that being β-carotene (Chou *et al.*, 1992; Solomons, 2009), while transgenic *crtB* soya bean seeds have almost 425 ppm with β carotene favoured ratio over α carotene of 12 : 1.

Our average β-carotene amount is 845 μg/g β-carotene dry weight in soya bean seeds. Unlike our single-gene engineering strategy, the production of similarly high levels of β-carotene in canola (Ravanello *et al.*, 2003) was achieved with the phytoene synthase gene plus 2 additional genes (*crtI*, lycopene synthase; *crtY* lycopene β-cyclase). Kim *et al.* (2012) reported seed-specific expression of *CrtB* phytoene synthase from *Capsicum* fused to *crtI* lycopene desaturase from *Pantoea* in soya bean and reported 112 μg/g as their highest level of β-carotene in soya bean seeds. The difference in β-carotene amount achieved demonstrates the importance of gene expression cassettes and choice of the gene of interest, as enzymes in the carotenoid pathway are able to act on numerous substrates with varying affinities (Schmidt-Dannert *et al.*, 2000). Developing nontransgenic soya bean seeds contain 0.5 μg/g β-carotene (Simonne *et al.*, 2000). With addition of the phytoene synthase gene, the accumulation of β-carotene was enhanced 1500-fold over nontransgenic soya bean seeds. The 12 : 1 ratio of β- : α-carotene in the *crtB* seeds is the result of the up-regulataon of lycopene β-cyclase with a simultaneous repression of lycopene ε cyclase.

Geranylgeranyl diphosphate (GGPP) is a key metabolic branch point, as it is the substrate for chlorophyll, tocopherols, gibberellins and carotenoids. Shunting more GGPP to carotenoid production via phytoene synthase would be expected to impact the accumulation of other GGPP-derived molecules. There was a slight decrease in the amount of phytosterols accumulated, especially β-sitosterol, but unexpectedly the *CrtB* soya beans produced increases amounts of δ-tocopherols. This result is in contrast to the finding that canola seeds engineered to express β-carotene had decreased tocopherol levels (Shewmaker *et al.*, 1999), but similar to previous results in other soya bean seed research, which had no effect on tocopherol levels and a slight (1.5-fold) increase in the phytosterol β-sitosterol (Kim *et al.*, 2012). Together, these results suggest a complex regulation of the isoprenoid pathway to balance the production of all of its important end-product molecules and that these might vary among various plant species, tissues and amounts produced. The shift in fatty acid composition in the *CrtB* soya beans is noteworthy from both a biological and biotechnology perspective. The *CrtB* transgenic soya bean seeds were analysed for lipid composition and these results showed a significant increase in oleic acid (18 : 1), with a similar decrease in linoleic acid (18 : 2). The change in lipid composition results in nearly a twofold increase in 18 : 1 over the controls. The β-carotene- producing canola seeds also exhibit a shift of fatty acid composition to oleic acid (Shewmaker *et al.*, 1999). Increasing oleic acid content is a key compositional goal of the soya bean industry and has long been a major target of breeding and biotechnology programs. High oleic acid (~85%) is industrially significant, as oleic acid is stable to oxidation without the off-flavours and odours created by oils high in linoleic (18 : 2) or linolenic (18 : 3) acid content (Liu and White, 1992; Mounts *et al.*, 1994; Neff *et al.*, 1992). Oleic acid content of oil in recent years has been the focus of much health concerns with rising rates of obesity and cardiac health (Korver and Katan, 2006). The processed food industry, under pressure to support healthier diets, has reformulated much of their production, there has been an increased demand for sources higher in oleic acid. In the case of soya bean, these can come from the use of novel mutants to breed superior lines (Pham *et al.*, 2010, 2012) and most recently by the development of transgenic soya beans that produce high levels of oleic acid. Each of these strategies targeted the FAD2 locus (encodes the enzyme that catalyses oleic to linoleic acid) (Okuley *et al.*, 1994). (Kinney and Knowlton, 1998; Pham *et al.*, 2012). The excess accumulation of oleic acid in the *CrtB* soya beans is noteworthy because this favourable shift in composition was derived from a novel path that did not involve direct manipulating the FAD2 expression. The mid-oleic acid levels (45%) in *CrtB* soya bean seeds coupled with enhanced production of β-carotene together significantly enhance the soya bean's composition and performance in the context of food health. Transcript abundance for both FAD2 genes was down-regulated in the *crtB* transgenics, accounting for the elevated oleic acid phenotype. The modification of oil saturation levels by the accumulation of carotenoids in the plastid indicates a yet discovered coordination between plastid production of fatty acids and ER-processing, possibly mediated by signals from the phytohoromone ABA and its alteration of targeted transcription factor gene expression.

Soya bean protein is the dominant global vegetable protein and is the foundation of global domestic animal production. Increasing soya bean protein content has been a major producer goal for many decades, primarily approached through selective breeding. The experience of breeders has shown that the ratio of protein to oil in soya bean is inversely correlated, and selection for higher protein results in a re-allocation to carbon producing less oil with a decrease in oil content at about two mass units for every one mass unit of protein increase (Li and Burton, 2002; Piper and Boote, 1999; Wilcox and Shibles, 2001). The accumulation of β-carotene has a novel, and potentially valuable, collateral consequence, of increased protein content. The *CrtB* soya bean seeds exhibit 4–12% higher protein content than those of its parental line grown side-by-side in the greenhouse and retained their elevated protein content when grown in the field. Typically, high protein soya bean lines produced through breeding are not widely deployed due to yield limitations (Concibido *et al.*, 2003; Helms and Orf, 1998). Increasing soya bean seed's protein content could have immense value in addressing the burgeoning need to more than double the production of animal feed by 2050. The β-carotene enhanced soya bean seeds elevated protein content is likely the result of modified phytohoromone levels ABA with a consequent alteration in the transcriptome involving altered expression of pivotal seed ABA-responsive transcription factors.

ABA is a central hormone in plant development and abiotic stress responses. It is biosynthetically derived from isoprenoid carotenoids. Free ABA, through its subsequent signalling, has been correlated with protein accumulation in maturing soya bean cotyledons (Ackerson, 1984; Higgins, 1984; Kagaya *et al.*, 2005). Increased ABA levels during embryogenesis promotes accumulation of protein storage reserves demonstrated with embryo/seed *in vivo* cultures, mutants and transgenics (Finkelstein *et al.*, 2002; Rock and Quatrano, 1995). Promoter sequences for storage proteins genes contain elements shown to be responsive to ABA (Rock and Quatrano, 1995). ABA-mediated reserve protein accumulation during seed maturation is controlled by a coordinated regulation of a number of transcription factors, including ABi3, ABi4, ABi5, Lec1, Lec2 and FUS3 (Finkelstein *et al.*, 2002). ABI3/VP1 (abscisic acid insensitive/vivipary) are a large group of

transcription factors that have been shown to be involved in ABA regulation and ABA-inducible gene expression during seed development (Nakamura *et al.*, 2001). Maize vp mutants are viviparous, exhibiting wilted leaves and reduced levels of carotenoids (McCarty, 1995). Additionally, *abi* Arabidopsis mutant seeds do not lose their chlorophyll and remain green, have reduced storage proteins, decreased anthocyanin production and have little desiccation tolerance (Nambara and Marion-Poll, 2003; Parcy *et al.*, 1997). ABI3 overexpression has been associated with an increase in storage protein expression (Lara *et al.*, 2003). Similarly, *crtB* transgenic soya bean seeds have a 55-fold significantly increased expression of the ABI3 transcript as observed in our RNA-seq data comparing transcript profiles from two independent soya bean *crtB* lines to nontransgenic cotyledons. The prior results that *abi* mutants have reduced storage proteins and that ABI3 overexpression lines have an increase in protein storage expression suggests a mechanism by which enhanced β-caroene lines have elevated protein content in their seeds due to the increased expression of ABI3. CrtB trangenics also exhibit enhanced expression of members of the YABBY transcription factor family. Prior expression data on soya bean seed development has shown that YABBY-regulated genes include AP2 transcription factors and fatty acid desaturases (Shamimuzzaman and Vodkin, 2013). Elevated expression of YABBY transcription factors likely induces both the elevated protein and modified fatty acid desaturation in the β-carotene enhanced transgenic soya bean seeds.

In this report, we have shown with the addition of a single-gene insertion overexpressing phytoene synthase produced soya beans with three distinct enhanced value traits: record levels of β-carotene, mid-oleic acid content and elevated protein. These results indicate that the production and accumulation of β-carotene has a significant systems biology impact of the seed inducing it to modify its composition. The actual carbon flow into β-carotene is proportionally small (<0.1% mass) for the effect it has on protein and oil content/composition. This indicates it is not the re-allocation of carbon into β-carotene that causes these collateral effects, but rather β-carotene production/accumulation induces a re-allocation of resource through the phytohormone ABA signal. Understanding the carotenoid feedback on both protein and oil accumulation in soya bean seeds will assist in deriving new engineering approaches to enhance this globally important crop.

References

Ackerson, R. (1984) Regulation of soybean embryogenesis by abscisic acid. *J. Exp. Bot.* **35**, 403–413.

Aluru, M., Xu, Y., Guo, R., Wang, Z., Li, S., White, W., Wang, K. and Rodermel, S. (2008) Generation of transgenic maize with enhanced provitamin A content. *J. Exp. Bot.* **59**, 3551–3562.

AOAC. (1995) Oils and Fats. In *Official methods of analysis of AOAC international* (AOAC, ed.), pp.1–33. Gaithersburg: AOAC International.

Ashraf-Khorassani, M., Ude, M., Doane-Weideman, T., Tomczak, J. and Taylor, L.T. (2002) Comparison of gravimetry and hydrolysis/derivatization/gas chromatography–mass spectrometry for quantitative analysis of fat from standard reference infant formula powder using supercritical fluid extraction. *J. Agric. Food Chem.* **50**, 1822–1826.

Broun, P., Gettner, S. and Somerville, C. (1999) Genetic engineering of plant lipids. *Annu. Rev. Nutr.* **19**, 197–216.

Calder, P.C. (2001) Polyunsaturated fatty acids, inflammation and immunity. *Lipids*, **36**, 1007–1024.

Cazzonelli, C. and Pogson, B.J. (2010) Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci.* **15**, 266–274.

Chan, J.M., Gann, P.H. and Giovannucci, E.L. (2005) Role of diet in prostate cancer development and progression. *J. Clin. Oncol.* **23**, 8152–8160.

Chew, B.P. (1993) Role of carotenoids in the immune response. *J. Dairy Sci.* **76**, 2804–2811.

Chou, Y.M., Yap, S.C., Ooi, C.K., Ong, A.S.H. and Goh, S.H. (1992) Production of palm oil carotenoid concentrate and its potential application in nutrition. In *Lipid-Soluble Antioxidants: Biochemistry and Clinical Applications* (Ash, A.S.H. and Packer, L., eds), pp. 243–254. Basel: Birkhauser.

Cober, E.R. and Voldeng, H.D. (2000) Developing high-protein, high-yield soybean populations and lines. *Crop Sci.* **40**, 39–42.

Concibido, V.C., LaVellee, B., Mclaird, P., Pineda, N., Meyer, J., Hummel, L., Yang, J., Wu, K. and Delannay, X. (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor. Appl. Genet.* **106**, 575–582.

De Castro, L.M.D. and Priego-Capote, F. (2010) Soxhlet extraction: past and present panacea. *J. Chromatogr. A*, **1217**, 2383–2389.

Delgado-Vargas, F., Jimenez, A.R. and Paredes-Lopez, O. (2000) Natural pigments: carotenoids, anthocyanins, and betalains – characteristics, biosynthesis, processing and stability. *Crit. Rev. Food Sci. Nutr.*, **40**, 173–289.

Diretto, G., Welsch, R., Tavazza, R., Mourgues, F., Pizzichini, D., Beyer, P. and Giuliano, G. (2007) Silencing of beta-carotene hydroxylase increases total carotenoid and beta-carotene levels in potato tubers. *BMC Plant Biol.* **7**, 11–16.

Ducreux, L.J.M., Morris, W.L., Hedley, P.E., Shephard, T., Davies, H.V., Milam, S. and Taylor, M.A. (2005) Metabolic engineering of high carotenoid potato tubers containing enhanced levels of β-carotene and lutein. *J. Exp. Bot.* **56**, 81–89.

Enserink, M. (2008) Tough lessons from golden rice. *Science*, **320**, 468–471.

Finkelstein, R.R., Gampala, S.S.L. and Rock, C.D. (2002) Abscisic acid signaling in seeds. *Plant Cell*, **14**, S15–S45.

Fraser, P.D., Romer, S., Shipton, C.A., Mills, P.B., Kiano, J.W., Misawa, N., Drake, R.G., Schuck, W. and Bramley, P.M. (2002) Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proc. Natl Acad. Sci. USA*, **99**, 1092–1097.

Fujisawa, M., Watanabe, M., Choi, S., Termanoto, M., Ohyama, K. and Misawa, N. (2008) Enrichment of carotenoids in flaxseed (*Linum usitatissimum*) by metabolic engineering with introduction of bacterial phytoene synthase gene crtB. *J. Biosci. Bioeng.* **105**, 636–641.

Hakoda, A., Ii, Y., Naito, S., Suzuki, T. and Yasui, A. (2011) Determination of crude protein in macaroni products by the combustion method and comparison with the Kjeldahl method: interlaboratory study. *Food Sci. Technol. Res.* **17**, 227–232.

Harris, W.M. and Spurr, A.R. (1969) I. Ultrastructure of low-pigment and high beta mutants. Carotene analysis. *Am. J. Bot.* **56**, 369–379.

Haskell, M.J. (2012) The challenge to reach nutritional adequency for vitamin A: β-carotene bioavailability and conversion – evidence in humans. *Am. J. Clin. Nutr.* **96**(suppl), 1193–1203.

Haskell, M.J., Pandey, P., Graham, J.M., Peerson, J.M., Shrestha, R.K. and Brown, K.H. (2005) Recovery from impaired dark adaptation in nightblind pregnant Napali women who receive small daily does of vitamin A as amaranth leaves, carrots, goat liver, vitaminA-fortified rice, or retinyl palmitate. *Am. J. Clin. Nutr.*, **81**, 461–471.

Helms, T.C. and Orf, J.H. (1998) Protein, oil and yield of soybean lines selected for increased protein. *Crop Sci.* **38**, 707–711.

Higgins, T.J.V. (1984) Synthesis and regulation of major proteins in seeds. *Annu. Rev. Plant Biol.* **35**, 191–221.

Johnson, E.J. (2002) The role of carotenoids in human health. *Nutr. Clin. Care*, **5**, 56–65.

Kagaya, Y., Okuda, R., Ban, A., Toyoshima, R., Tsutsumida, K., Usui, H., Yamamoto, A. and Hattori, T. (2005) Indirect ABA-dependent regulation of seed storage protein genes by FUSCA3 transcription factor in Arabidopsis. *Plant Cell Physiol.* **46**, 300–311.

Kim, M., Kim, J.K., Kim, H.J., Pak, J.H., Lee, J., Kim, D., Choi, H.K., Jung, H.W., Lee, J., Chung, Y. and Ha, S. (2012) Genetic modification of the soybean to enhance the β-carotene content through seed-specific expression. *PLoS ONE*, **7**, e48287.

Kinney, A.J. and Knowlton, S. (1998) Designer oils: the high oleic acid soybean. In *Genetic Modification in the Food Industry* (Roller, S. and Harlander, S., eds), pp. 193–213. London, Blackie: Springer Publisher.

Korver, O. and Katan, M.B. (2006) The elimination of trans fats from spreads: how science helped turn an industry around. *Nutr. Rev.* **64**, 275–279.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.

Lara, P., Onate-Sanchez, L., Abraham, Z., Ferrandiz, C., Diaz, I., Carbonero, P. and Vincente-Carbajosa, J. (2003) Synergistic activation of seed storage protein gene expression in Arabidopsis by ABI3 and two bZIPs related to OPAQUE2. *J. Biol. Chem.* **278**, 1003–1011.

Li, H. and Burton, J.W. (2002) Selecting increased seed density to increase indirectly soybean seed protein concentration. *Crop Sci.* **42**, 393–398.

Li, L., Lu, S., O'Halloran, D.M., Garvin, D.F. and Vrebalov, J. (2003) High resolution genetic and physical mapping of the cauliflower high β-carotene gene Or (orange). *Mol. Genet. Genomics*, **270**, 132–138.

Liu, H.R. and White, P.J. (1992) Oxidative stability of soybean oils with altered fatty acid composition. *J. Am. Oil Chem. Soc.* **69**, 528–532.

Lohse, M., Bolger, A.M., Negel, A., Fernie, A.R., Lunn, J.E., Stitt, M. and Usadel, B. (2012) RobiNA: a user friendly integrated software solution for RNA-Seq based transcriptomics. *Nucleic Acids Res.* **40**, W622–W627.

Lopez, A.B., Van Eck, J., Conlin, B.J., Paolillo, D.J., O'Neill, J. and Li, L. (2008) Effect of cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.* **59**, 213–223.

Lu, S., Van Eck, J., Zhou, X., Lopez, A.B., O'Halloran, D.M., Cosman, K.M., Conlin, B.J., Paolillo, D.J., Garvin, D.F., Vrebalov, J., Kochian, L.V., Kupper, H., Earle, E.D., Cao, J. and Li, L. (2006) The cauliflower Or gene encodes a DNAJ cysteine-rich domain containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*, **18**, 3594–3605.

McCarthy, F.M., Burgess, S.C., van den Berg, B.H., Koter, M.D. and Pharr, G.T. (2005) Differential detergent fractionation for non-electrophoretic eukaryote cell proteomics. *J. Proteome Res.* **4**, 316–324.

McCarty, D.R. (1995) Genetic control and integration of maturation and germination pathways in seed development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 71–93.

Michimi, A., Ellis-Griffith, G., Nagy, C. and Peterson, T. (2013) Coronary heart disease prevalence and occupational structure in US metropolitan areas: a multilevel analysis. *Health Place*, **21**, 192–204.

Mounts, T.L., Warner, K., List, G.R., Neff, W.E. and Wilson, R.F. (1994) Low linolenic acid soybean oils – alternatives to frying oils. *J. Am. Oil Chem. Soc.* **5**, 495–499.

Mozaffarian, D., Katan, M.B., Ascherio, A., Stampfer, M.J. and Willett, W.C. (2006) Trans fatty acids and cardiovascular disease. *N. Engl. J. Med.* **354**, 1601–1613.

Nakamura, S., Lynch, T.J. and Finkelstein, R.R. (2001) Physical interactions between ABA response loci of Arabidopsis. *Plant J.* **6**, 627–635.

Nambara, E. and Marion-Poll, A. (2003) ABA action and interactions in seeds. *Trends Plant Sci.* **8**, 213–217.

Neff, W.E., Selke, E., Mounts, T.L., Rinsch, W., Frankel, E.N. and Zietoun, M.M. (1992) Effect of triacylglycerol composition and structures on oxidative stability of oils from selected soybean germplasm. *J. Am. Oil Chem. Soc.* **69**, 111–118.

Okuley, J., Lighter, J., Feldmann, K., Yadav, N., Lark, E. and Browse, J. (1994) Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell*, **6**, 147–158.

Olson, J.A. (1989) Provitam A function of carotenoids: the conversion of beta-carotene into vitamin A. *J. Nutr.* **119**, 105–108.

Paine, J.A., Shipton, C.A., Sunandha, C., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hinchliffe, E., Adams, J.L., Silverstone, A.L. and Drake, R. (2005) Improving the nutritional value of golden rice through increased pro-vitamin A content. *Nat. Biotechnol.* **23**, 482–487.

Parcy, F., Valon, C., Kohara, A., Misera, S. and Giraudat, J. (1997) The abscisic acid insensitive 3, fusca 3 and leafy cotyledon 1 loci act in concert to control multiple aspects of Arabidopsis seed development. *Plant Cell*, **9**, 1265–1277.

Penniston, K.L. and Tanumihardjo, S.A. (2003) Vitamin A dietary supplements and fortified foods: too much of a good thing? *J. Am. Diet. Assoc.* **103**, 1185–1187.

Pham, A., Lee, J., Shannon, J.G. and Bilyeu, K.D. (2010) Mutant alleles of FAD2-1 and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. *BMC Plant Biol.* **10**, 195.

Pham, A., Shannon, J.G. and Bilyeu, K.D. (2012) Combinations of mutant FAD2 and FAD3 genes to produce high oleic acid and low linolenic acid soybean oil. *Theor. Appl. Genet.* **125**, 503–515.

Piper, E.L. and Boote, K.J. (1999) Temperature and cultivar effects on soybean seed oil and protein concentrations. *J. Am. Oil Chem. Soc.* **76**, 1233–1241.

Ravanello, M.P., Ke, D., Alvarez, J., Huang, B. and Shewmaker, C.K. (2003) Coordinated expression of multiple bacterial carotenoid genes in canola leading to altered carotenoid production. *Metab. Eng.* **5**, 255–263.

Rock, C. and Quatrano, R. (1995) The role of hormones during seed development. In *Plant Hormones: Physiology, Biochemistry and Molecular Biology* (Davies, P. ed.), pp. 671–697. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Romer, S., Fraser, P.D., Kiano, J.W., Shipton, C.A., Misawa, N., Schuch, W. and Bramley, P.M. (2000) Elevation of the provitamin A content of transgenic tomato plants. *Nat. Biotechnol.* **18**, 666–669.

Rotundo, J.L., Borras, L. and Westgate, M.E. (2011) Linking assimilate supply and seed developmental processes that determine soybean seed composition. *Eur. J. Agron.* **35**, 184–191.

Schmidt, M.A. and Herman, E.M. (2008a) Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. *Plant Biotechnol. J.* **6**, 832–842.

Schmidt, M.A. and Herman, E.M. (2008b) Suppression of soybean oleosin produces micro-oil bodies that aggregate into oil body/ER complexes. *Mol. Plant*, **1**, 910–924.

Schmidt, M.A., Tucker, D.M., Cahoon, E.B. and Parrott, W.A. (2005) Towards normalization of soybean somatic embryo maturation. *Plant Cell Rep.* **24**, 383–391.

Schmidt, M.A., Barbazuk, W.B., Stanford, M., May, G., Song, Z., Hong, W., Nikolau, B.J. and Herman, E.M. (2011) Silencing of soybean seed storage proteins results in a rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. *Plant Physiol.* **156**, 330–345.

Schmidt-Dannert, C., Umeno, D. and Arnold, F.H. (2000) Molecular breeding of carotenoid biosynthetic pathways. *Nat. Biotechnol.* **18**, 750–753.

Shamimuzzaman, M. and Vodkin, L. (2013) Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by ChIP-Seq and RNA-Seq. *BMC Genomics*, **14**, 477–494.

Shekelle, R.B., Shryock, A.M., Paul, O., Lepper, M., Jeremiah, S., Liu, S. and Raynor, W.J. (1981) Diet, serum cholesterol and death from coronary heart disease – the western electric study. *N. Engl. J. Med.* **304**, 65–70.

Shewmaker, C.K., Sheehy, J.A., Daley, M., Colburn, S. and Ke, D.Y. (1999) Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. *Plant J.* **20**, 401–412.

Simonne, A.H., Smith, M., Weaver, D.B., Vail, T., Barnes, S. and Key, C.I. (2000) 'Retention and changes of soy isoflavones and carotenoids in immature soybean seeds (Edamame) during processing'. *J. Agric. Food Chem.*, **48**, 6061–6069.

Solomons, N.W. (2009) Plant sources of vitamin A and human nutrition: red palm oil does the job. *Nutr. Rev.* **56**, 309–311.

Soxhlet, F. (1879) Die gewichtsanalytische bestimmung des milchfettes. *Polytechnisches J. (Dingler's)*, **232**, 461.

Soystats (2012). Available at: www.soystats.com/2012.

Stahl, W. and Sies, H. (2003) Antioxidant activity of carotenoids. *Mol. Aspects Med.* **24**, 345–351.

Supek, F., Bosnjak, M., Skunca, N. and Smuc, T. (2011) REVIGO summaries and visualizes long lists of gene ontology terms. *PLoS ONE*, **6**, e21800.

Tang, G., Hu, Y., Yin, S., Wang, Y., Dallal, G.E., Grusak, M.A. and Russell, R.M. (2012) β-carotene in golden rice is as good as β-carotene in oil at providing vitamin A to children. *Am. J. Clin. Nutr.* **96**, 658–664.

Trapnell, C., Pachter, L. and Salzberg, S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515.

Ullah, S.M.R., Murphy, B., Dorich, B., Richter, B. and Srinivasan, K. (2011) Fat extraction from acid and base-hydrolyzed food samples using accelerated solvent extraction. *J. Agric. Food Chem.* **59**, 2169–2174.

Von Lintig, J. (2010) Colors with Functions: elucidating the biochemical and molecular basis of carotenoid metabolism. *Annu. Rev. Nutr.* **20**, 35–56.

Welsch, R., Arango, J., Bar, C., Salazar, B., Al-Babili, S., Beltran, J., Chavarriaga, P., Ceballos, H., Tohme, J. and Beyer, P. (2010) Provitamin A accumulation in cassava (*Manihot escaulenta*) roots driven by a single nucleotide polymophism in a phytoene synthase gene. *Plant Cell*, **22**, 3348–3356.

Wilcox, J. and Shibles, R.M. (2001) Interrelationships among seed quality attributes in soybean. *Crop Sci.* **41**, 11–14.

Wilson, R.F. (2004) Seed composition. In *Soybeans: Improvement, Production and Uses*, 3rd edn (Boerma, H.R. and Specht, J.E., eds), pp. 621–679. Madison, WI: American Society of Agonomy.

Ye, X., Al-Babili, S., Kloti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000) Engineering the provitam A (β-carotene) biosynthetic pathway into (carotenoid free) rice endosperm. *Science*, **287**, 303–305.

Yu, B., Lydiate, D.J., Young, L.W., Schafer, U.A. and Hannoufa, A. (2007) Enhancing the carotenoid content of *Brassica napus* seeds by downregulating lycopene epsilon cyclase. *Transgenic Res.* **17**, 573–585.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Total amino acid composition analysis of *crtB* transgenic seeds.
Figure S2 Thin-layer chromatography (TLC) analysis of *crtB* transgenics.
Figure S3 Proteomic analysis of *crtB* transgenic seeds.
Figure S4 Proteomic changes in *crtB* transgenics compared to nontransgenic.
Figure S5 Lipidomics analysis of *crtB* transgenic soya bean seeds.
Figure S6 Metabolite analysis of *crtB* transgenic soya bean seeds.
Figure S7 Pairwise comparisons of transcript expression (FPKM).
Figure S8 Expression analysis of both fatty acid desaturases (FAD) and transcription factors genes in crtB transgenic seeds.
Table S1 Lipidomics profile of *crtB* transgenic seeds.
Table S2 Metabolomics profile of *crtB* transgenic seeds.
Table S3 Primer sequences used in quantitative reverse-transcription polymerase chain reaction to confirm transcriptome data set.
Appendix S1 Materials and Methods.

