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

Monica A. Schmidt1,*, Wayne A. Parrott2, David F. Hildebrand3, R. Howard Berg4, Amanda Cooksey5, Ken Pendarvis5, Yonghua He1, Fiona McCarthy5 and Eliot M. Herman1

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

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 absicic 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 transcription 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 ailments 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.
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 enzymes, 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 ξ 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 (DiRieto et al., 2007) targeted lycopene cyclase to direct the carotenoid flux to produce β-carotene from lycopene and to impede the accumulation of lutein. A spontaneous orange mutation (O) 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 chloroplasts (Lu et al., 2006). When the O allele was expressed in transgenic potato tubers, chloroplast 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 collat- eral 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 [formerly 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 chloroplastic 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'-GAT-TCCCATGGAAGAGGTC3') 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 pcrB 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 nontransgen- ic 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 tetroxide 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 sepa- rated into five protein fractions using differential detergent fractionation as described by McCarthy et al. (2005), subse- quently 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 Pertem (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 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 Colorado 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 SRR041621. 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 D900877) (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 line has an initial expression pattern similar to Pantoea and a wide range of beta-carotene detection in the mature dry seeds (Figure 2a).
plastids was confirmed by direct visualization of the orange coloured organelles in the cotyledon cells. Figure 3a,b show light micrographs of the \textit{crtB} cotyledon cells compared with nontransgenic tissue, demonstrating that orange coloured discrete structures are present in the \textit{crtB} tissue but absent in the nontransformed cotyledon cells. This observation was amplified by transmission electron microscopy using cryofixed samples that showed that the \textit{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 \(\beta\)-carotene as intraplastid lipid moieties comparable to the plastoglobuli observed in the organs of many species accumulate \(\beta\)-carotene.

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

---

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

---

**Figure 3** Cellular morphology of \textit{crtB} transgenics compared to nontransgenic. (a) Light micrograph of immature cotyledon tissue from \textit{crtB} transgenic displaying discrete orange structures indicating \(\beta\)-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 \textit{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 $\alpha$-carotene accumulation, being present in a $12:1$ $\beta$-carotene : $\alpha$-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 $\beta$-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.

**Transcript profiling and RT-PCR expression analysis of the $\beta$-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 $\beta$-cyclase and the

![Figure 4](image)

*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 $\alpha$ and $\beta$-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 $12:1$ $\beta$-carotene : $\alpha$ carotene.

<table>
<thead>
<tr>
<th>Seed sample</th>
<th>Protein (%)</th>
<th>Oil (%)</th>
<th>Moisture (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nontransgenic</td>
<td>28.20</td>
<td>19.83</td>
<td>7.01</td>
</tr>
<tr>
<td>$crtB$-1</td>
<td>45.38</td>
<td>11.94</td>
<td>7.20</td>
</tr>
<tr>
<td>$crtB$-2</td>
<td>47.07</td>
<td>11.58</td>
<td>7.10</td>
</tr>
<tr>
<td>$crtB$-3</td>
<td>40.72</td>
<td>14.48</td>
<td>7.11</td>
</tr>
</tbody>
</table>

**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
down-regulation of both lycopene \( \varepsilon \)-cyclase and ABA-aldehyde oxidase in the \( \text{crtB} \) transgenics. Together, this result accounts for the 12:1 ratio of \( \beta \)-carotene : \( \alpha \)-carotene in the \( \text{crtB} \) transgenic seeds as the simultaneous up-regulation of lycopene \( \beta \)-cyclase and down-regulation of lycopene \( \varepsilon \)-cyclase would preferentially produce \( \beta \)-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 \( \text{crtB} \) transgenic lines compared to nontransgenic. Of the \( \text{FAD2} \) genes and \( \text{FAD3} \) genes, all but \( \text{FAD3a} \) was detected to be down-regulated in the \( \text{crtB} \) transgenics. RT-PCR was performed to verify gene expression of \( \text{FAD} \) genes and confirmed the down-regulation of \( \text{FAD2-1} \) in \( \text{crtB} \) transgenics (Figure S8 and Table S3). The suppression of \( \text{FAD2} \) genes likely accounts for the elevated amount of oleic acid in the \( \text{crtB} \) seeds.

Of the differentially expressed transcripts detected in the two independent \( \text{crtB} \) 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 \( \text{crtB} \) 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 \( \text{crtB} \) tissue (Figure S8). YABBY (not

![Figure 5](image-url) Composition of 3 lines of seed-specific \( \beta \)-carotene enhanced \( \text{crtB} \) transgenic soya bean seeds. Mature dry seeds were analysed by near infrared spectroscopy to determine overall composition of \( \text{crtB} \) transgenics. The \( \text{crtB} \) 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.

![Figure 6](image-url) Carotenoid and ABA biosynthetic pathway with detected transcript abundance. Red: down-regulated transcript, green up-regulated transcript in \( \text{crtB} \) transgenics compared to nontransgenic. Numbers in () indicate the average log2 fold change in transcript abundance.
ANOVA significant 18 : 1 and 18 : 2 oil composition compared to nontransgenic seeds by gas chromatography. The extracts from 3 lines of fatty acid composition of 15% saturated fats, 24% oleic acid, nontransgenic seeds grown side-by-side in the greenhouse had a acids, with the components unchanged (Figure 5). The fatty acid composition of the decreased oil content, increase in sucrose, with other components unchanged (Figure S5; Table S1).

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 soybean is characterized by a collateral change to a ‘mid-oleic’ composition similar to that achieved by breeders selecting for enhanced oleic acid content.

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

Figure 8 Oil compositional analysis of ctfB transgenic seeds. Lipid extracts from 3 lines of ctfB transgenics and nontransgenic were analysed by gas chromatography. The ctfB 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>
<thead>
<tr>
<th>Phytohormone (ng/mL)</th>
<th>ctfB line 1</th>
<th>ctfB line 2</th>
<th>Nontransgenic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abscisic acid</td>
<td>3.243 ± 0.163</td>
<td>3.216 ± 0.210</td>
<td>5.444 ± 0.339</td>
</tr>
<tr>
<td>Phaseic acid</td>
<td>0.377 ± 0.004</td>
<td>0.325 ± 0.008</td>
<td>0.530 ± 0.006</td>
</tr>
</tbody>
</table>

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

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’ (Ensink, 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; Shevemaker et al., 1999), flaxseed (Linum usitatissimum; Fujisawa et al., 2008), soybean (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 (Pennistone 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 soybean plants will produce high levels of carotenoids with an optimal 12 : 1 β β to α ratio, so that nearly all of the β-carotene produced will be potentially biologically active. β-carotene enhanced soybean plants can produce the adult male RDA levels with as few as 5 g of soybeans. 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 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 abscisic acid

Two lines of ctfB transgenic soybean cotyledons were used to quantitate the phytohormones asbestic acid and phaseic acid. The ctfB 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 (ctfB 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 ctfB 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).
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 crtY lycopene desaturase from Pantoaea 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-regulation of lycopene β-cyclase with a simultaneous repression of lycopene β-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 (Kover 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 Shibie, 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 AB13, AB14, AB15, Lec1, Lec2 and FUS3 (Finkelstein et al., 2002). AB13/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 \( \beta \)-carotene lines have elevated protein content in their seeds due to the increased expression of ABI3. CrtB transgenics 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 \( \beta \)-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 \( \beta \)-carotene, mid-oleic acid content and elevated protein. These results indicate that the production and accumulation of \( \beta \)-carotene has a significant systems biology impact of the seed inducing it to modify its composition. The actual carbon flow into \( \beta \)-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 \( \beta \)-carotene that causes these collateral effects, but rather \( \beta \)-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


© 2014 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd, Plant Biotechnology Journal, 1-11


**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 soybean seeds. **Figure S6** Metabolite analysis of crtB transgenic soybean 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.