# CO-ORDINATED EXPRESSION OF *CRTB*, *At-VTE3*, AND *VTE4* TO ENHANCE PRO-VITAMIN A AND VITAMIN E IN TRANSGENIC SOYBEAN

---

A Thesis

presented to the Faculty of the Graduate School

at the University of Missouri-Columbia

---

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

By

DUNG THU PHAM

Dr. Zhanyuan J. Zhang, Thesis Supervisor

JULY 2013

The undersigned, appointed by the dean of the Graduate School,

have examined the Thesis entitled

CO-ORDINATED EXPRESSION OF *CRTB*, *At-VTE3*, AND *VTE4* TO ENHANCE PRO-VITAMIN A AND VITAMIN E IN TRANSGENIC SOYBEAN

Presented by Dung Thu Pham

A candidate for the degree of

MASTER OF SCIENCE

And hereby certify that, in their opinion, it is worthy of acceptance.

---

Dr. Zhanyuan J. Zhang

---

Dr. Kristin D. Bilyeu

---

Dr. Henry T. Nguyen

# ACKNOWLEDGEMENTS

I would like to address my special thanks to my supervisor Dr. Zhanyuan Jon Zhang for his guidance and encouragement during the entire course of this research. I would like to thank Dr. Kristin Bilyeu and Dr. Henry Nguyen for their collaboration, distribution, and contribution. Also, I would like to thank all current and past members in Dr. Zhang Zhanyuan J. Zhang's lab for their direct and indirect assistance during my studies: Xiaoyan Yin, Yinxia He, Wan Neng for helping me design construct, soybean tissue culture, and take care transformed plants in green house, SoYon Park, Ulku Baykal, Hyeyong Lee, Phat Do, Sha Lu, and Michelle Folta who shared their knowledge, lab techniques to help me during my research. My special thanks also are for Christine Cole (Dr. Kristin's lab) for helping me with tocopherol and fatty acid analysis, Tristan Lipkie (Purdue University) for carotenoid analysis. I am grateful to my parents and my brother for their love, understanding and encouragement. At last, many thanks to all friends that I was fortunate to meet in Columbia MO for help and friendship.

# TABLE OF CONTENTS

# LIST OF FIGURES

**CHAPTER 1**



**CHAPTER2**

# LIST OF TABLES

# ABSTRACT

Although soybean is an excellent general nutritional source, it is not very rich in particular vitamins. The main goal of this study is to enhance both pro-vitamin A (carotenoids) and vitamin E (tocopherols) content in soybean seeds. We have genetically engineered the carotenoid and the tocopherol biosynthetic pathways in soybean seeds by ectopically expressing three genes: *Erwinia uredovora phytoene synthase* (*crt*B) to increase carotenoid content, *Arabidopsis 2-methyl-6-phytylbenzoquinol methyl transferase (At-VTE3*) and soybean γ-*tocopherol methyl transpherase* (*VTE4*) to increase α-tocopherol using the self-cleavage activity of FMDV 2A sequence to join two adjacent proteins. This 2A-polyprotein construct was introduced to soybean via *Agrobacterium*-mediated cotyledonary node transformation method. One inheritable transgenic event displayed "golden"- colored seeds. The presence and expression of the three genes were detected in golden soybean event HYX-7-1 by qRT-PCR analyses. HPLC analysis of individual golden soybean lines revealed that the seeds accumulated as high as 128μg/g of total carotenoids, approximately 25- fold higher than wild type and 45-fold higher than empty vector control seeds. Of total carotenoids, 98% was pro-vitamin A (β-carotene and β-carotene equivalents). By contrast, golden seeds showed a decrease of tochopherol and a significant change of fatty acid profile. Transgenic golden soybean lines also displayed a delay in germination with a decreased rate (50%) and dwarf phenotype at early developmental stage, but showed normal development later.

# CHAPTER 1

# INTRODUCTION AND LITERATURE REVIEW

Soybean [*Glycine max* (L). Merrill] is one of the most important crops worldwide with high quality and inexpensive protein and oil (Hartman et al., 2011). Whether for oil, protein or other functional components, soybeans play a necessary role in human diet. Although China is the country that originally domesticated soybean, China has steadily decreased in soybean production and ranked fourth in world's production. At present the USA ranks at the top in soybean production with 83.2 metric tons in 2011. The USA, Brazil and Argentina are now among the biggest soybean-producing countries. These three countries harvested 81% of the world soybean production in 2011 (Chen et al., 2012).

## Values of soybeans

Soybean is one of the most important crops in the world with food, nutritional, industrial, and pharmaceutical uses (Yamada et al., 2012).

### *Nutritional values*

Soybean seeds contain approximately 40% protein, 35% carbohydrate, 20% edible oil and 5% ash (Pednekar et al., 2010; Hartman et al., 2011). It is one of the few plants providing complete protein (Henkel, 2000; Chen et al., 2012), and is, therefore, used as substitute for animal protein and other dairy products (Fukushima, 2001; Azadbakht et al., 2003; Hartman et al., 2011). In addition, supplement of soybean protein avoids lysine deficiency of cereals (Pednekar et al., 2010). Furthermore, soybeans with

high protein have a great potential in solving the protein malnutrition in the world especially in developing countries where consumption of protein is below the recommend level. Young (1991) indicated that soybean protein can meet the protein needs for both children and adults as a sole protein source when consumed at 0.6g/kg body wt.

*Oil* is extracted from soybean seeds, then refined and blended for different applications, and most of them used as dietary oil. Soybean oil is low in saturated fatty acids but high in polyunsaturated fatty acid (Bennett et al., 2003; Hartman et al., 2011). Because the high content of polyunsaturated fatty acids can reduce cholesterol (Kinsell et al., 1954; Grundy and Denke, 1990; Bennett et al., 2003), consumption of soybean oil dramatically increased.

***Pharmaceutical value***

There are many health benefits from intake of soybean or soy products. They can be used as dietary supplements or substitution of animal protein for diabetic patients (Azadbakht et al., 2003; Villegas et al., 2008); use in weight loss (Chen et al., 2003); reducing risk of certain types of cancer (MacDonald et al., 2005; Messina et al., 2006; Hamilton-Reeves et al., 2007; Nagata et al., 2007); reducing total cholesterol content (Kinsell et al., 1954; Grundy and Denke, 1990; Young, 1991; Fukushima, 2001; Azadbakht et al., 2003). Consumption of as little as 25g of soybean protein per day as a diet, in addition to the low intake in cholesterol and saturated fatty acid, may reduce risk of heart disease and renal disease (Azadbakht et al., 2003; Hartman et al., 2011). Moreover, the trend of using soybean oil instead of animal fats increases greatly because of reducing risk of cardiovascular when compared to using animal fats (Hartman et al., 2011). In addition, soybeans are also rich in calcium, which benefits bone health, and

isoflavones, which have received special attention and becomes interesting topics for many scientists. Finally, soybean isoflavones are believed to play a great role in the contribution to health benefits (Chen et al., 2012) such as reducing risk of cardiovascular (Zyriax and Windler, 2000); preventing different cancers and having positive effects on bone health (Chen et al., 2012).

**Soy products**

85% of the world's soybean crop is processed into soybean meal and vegetable oil (http://www.soystats.com). 98% of soybean meal is used in livestock and aquaculture feed because of its high level and good quality of protein (Hartman et al., 2011) which makes soybeans a complete protein source containing essential amino acids that the human body can't synthesize. Over the past 2000 years, soybean has been consumed in East Asia as traditional soy foods, such as nimame (cooked whole soy), edamame (green fresh soy), soymilk, tofu, kori-tofu (freeze-denatured and dry tofu), abura-age (deep-fat-fried tofu), sufu or tofu-yo (fermented tofu), soy sauce, miso, natto, tempeh, etc (Fukushima, 2001). In recent years, consumption of these foods has widely expanded to outside Asia. With improving technology and advanced processes, new products of soy protein have also been developed such as soy flour, soy protein concentrates, soy protein isolates, and their texturized products that enhance variety of foods that are familiar to consumers but contain soybean for nutritional purpose, including baked goods, snack bars, noodles, and infant formula (Hartman et al., 2011).

In addition to protein, oil is also one of the most valuable products from soybeans. And, among legumes, soybeans are also classified as an oilseed. About 95% soybean oil is used as edible oil, of which 70% of is consumed by humans in the U.S. (in Tarva and

Kim 2007). After oil extraction, the remaining material can be processed in variety ways to produce further meal for swine, poultry feed, soy concentrate, soy protein isolate, or other industrial products (Galloway et al., 2008).

## Vitamin A

Vitamin A is a lipid soluble vitamin and is dissolved in the lipid fraction of foods. It is essential for normal function of vision, growth and development, maintenance of epithelial cellular integrity, immune function, reproduction and spermatogenesis (Duerbeck and Dowling, 2012). Vitamin A is found in a wide range of foods, either as preformed vitamin A in animal products, such as meats, fish, fish oil, milk, egg yolks, liver, or as pro-vitamin A carotenoids, mainly β-carotene in plant products, such as green or yellow vegetables and fruits.

Vitamin A, in fact, is not a single compound, but it is a general name for a whole group of related nutrients. There are two major forms of vitamin A: retinoids (preformed vitamin A) and carotenoids (pro-vitamin A). While retinoids (retinol, retinal, retinoic acids and retinyl esters) are associated with foods of animal origin, carotenoids (carotenes and xanthophylls) can be found in plant foods. They play important role in contributing to good health for animals and humans.

### *Carotenoid biosynthetic pathway in plants and equivalent steps in bacteria*

Carotenoid synthesis begins with the condensation of IPP (isopentenyl diphosphate) and DMAPP (dimethylallyl diphosphate) synthesized in the plastids to generate GGPP/GGDP (genarylgenaryl diphosphate) with the catalysis of GGPP/GGDP synthase (GGPPS/GGDPS). The first committed step of carotenoid biosynthetic pathway is the condensation of two GGPP to generate phytoene, the first colorless carotenoid by

the enzyme phytoene synthase (PSY in plants and CrtB in bacteria). Then phytoene is converted into lycopene by the action of two desaturases (PDS, phytoene desaturase, and zeta –carotene desaturase, ZDS). This pathway gives rise to poly-cis compounds that are converted to poly-trans forms by the action of carotenoid isomerase CrtISO and ZISO in non-green tissue, whereas in green tissue the reaction occurs spontaneously in the presence of light and chlorophyll. However, in bacteria, function of single enzyme CrtI equals to the functions of these four plant enzymes in desaturation and isomerization reactions. Therefore, CrtI is usually used in metabolic engineering approaches.

Lycopene is the substrate of two competing cyclases downstream in the pathway. Linear molecule lycopene can be cyclized at both ends by lycopene β-cyclase alone (LCYB in plant, CrtY in bacteria) to generate β-carotene, whereas it can be cyclized at two ends by LCYE and LCYB to form α-carotene. Both these molecules can be converted to downstream products via carotene hydroxylation. In the β-carotene pathway, β-carotene converts to β-cryptoxanthin then zeaxanthin. Furthermore, zeaxanthin can further enter the xanthophyll cyclase. In the α-carotene pathway, conversion yields lutein (Figure 1.1A).

***Vitamin A deficiency in the World***

Vitamin A deficiency (VAD) is a common health problem in the world that is mostly known as the main cause of blindness. An estimated 500,000 children worldwide go blind every year from vitamin A deficiency (Ramakrishnan and Darnton-Hill, 2002). Nearly all the cases of VAD are in developing countries whose populations use single staple crop for their sustenance. The highest prevalence is found in Southeast Asia and Africa, and, the most affected are the infants, young children and pregnant women.

Vitamin A deficiency occurs in children and adults who do not consume adequate yellow and green vegetables, fruits which are rich in β-carotene, or animal-based vitamin A (retinol) products such as butter, cheese, eggs, liver. In addition, other factors may cause deficiency of vitamin A such as low fat diet, chronic exposure to oxidants such as cigarette smoke, chronic diarrhea genetic inheritance 2 (Duerbeck and Dowling, 2012); http://www.whfoods.com/genpage.php?tname=nutrient&dbid=106). In addition to eye symptoms, VAD also decreases functions of immune and inflammatory systems, leading to increased risks in infectious diseases such as measles, diarrhea and can cause high morbidity and mortality, especially among children (Zimmermann and Qaim, 2004).

***Efforts to combat VAD***

During the last decade, many strategies have been deployed to reduce VAD in developing countries. Food fortification, supplementation and dietary education programs have been applied (Zimmermann and Qaim, 2004). Dietary supplement in the form of vitamin tablets or suspension and many other foodstuffs that fortified with vitamin A (such as fats, oils, margarine and cereal) have been highly successful in high income countries (Dutra-de-Oliveira et al., 1998; Farre et al., 2010). Few other vitamin A fortification programs currently exist in lower income countries. Many approaches to enrich major staple foods with β-carotene through plant breeding or genetic engineering have been carried out for some crop species, such as maize, sweet potato, canola, tomato, and rice known as "Golden Rice" (Giuliano et al., 2000; Ye et al., 2000; Diretto et al., 2007; Wurbs et al., 2007; Aluru et al., 2008). Supplementation is the most widely used to control VAD in most high risk countries. The Recommend Dietary Allowance (RDA) for vitamin A is 2700 IU for non-pregnant and pregnant women and 5000 IU for men

(Duerbeck and Dowling, 2012). Moreover, improving the intake of vitamin A through dietary education should be used to enhance nutritional status of a population. Therefore, dietary education program is really necessary to change dietary habits, as well as providing better access to vitamin A or pro-vitamin A-rich foods.

**Vitamin E**

Similar to vitamin A, vitamin E is a lipid-soluble vitamin and synthesized by photosynthetic organisms, has antioxidant properties and plays important role in human and animal nutrition (Mayne and Parker, 1988; Hercberg et al., 1999; Bramley et al., 2000; Tavva et al., 2007). Daily intake of appropriate amount of vitamin E results in decreased risk of cardiovascular disease, certain cancers, degenerative diseases, prevent loss of memory and eye disorders or maintenance of immune system (Bramley et al., 2000; Constantinou et al., 2008; Farbstein et al., 2010).

***Source of vitamin E***

Vegetable oil is the best natural source of vitamin E, together with other natural sources including green leafy vegetables,wheat germ, some nuts, eggs, and others. Furthermore, vitamin E supplements (providing mostly α-tocopherol) are also common source for humans for prevention and therapy of vitamin E deficiency (Traber and Sies, 1996; Bramley et al., 2000).

Vitamin E is not a single compound, but a general name of a whole family of different components including tocopherols and tocotrienols (generally called tocochromanols). They are amphipathic molecules consisting of a polar chromanol head group and a lipophilic tail. Tocopherol differs from tocotrienol only in lipophilic tail which is derived from phytyl-diphosphate for tocopherol and genarylgenaryl diphosphate

with three double bonds at carbon positions 3', 7' and 11' for tocotrienol (Bramley et al., 2000; Munne-Bosch and Alegre, 2002; Cahoon et al., 2003). There are four natural forms of tocopherol and four corresponding natural forms of tocotrienol: α-, β-, γ- and δ-tocopherol/tocotrienol which are different in number and position of methyl groups on the chromanol ring that cause different "vitamin E activity/biological activity". Although all tocopherols and tocotrienols are potent antioxidants in vitro, α-tocopherol is generally regarded as having the highest nutritional value in term of "vitamin E activity" because it is the most readily absorbed and retained by human and other mammalian cells preferentially to other tocopherols and tocotrienols. This is also a reason that α-tocopherol is labeled on vitamin E supplements (Traber and Sies, 1996; Stoppani, 2004; Dörmann, 2007). Many studies have also demonstrated that the antioxidant of α-tocopherol was used in various aspects of human health, including heart disease, cancer, inflammatory responses (Winklhofer-Roob et al., 2003). Moreover, α-tocopherol is the most common form in nature compare to the others, and natural α-tocopherol is believed to have advantages than chemically synthesized α-tocopherol (Eitenmiller, 1997; Traber et al., 1998). Vitamin E activities of α-, β-, γ-, δ-tocopherols, α-, β-tocotrienol and synthetic α-tocopheryl acetate are 100%, 50%, 10%, 3%  30%, 5% and 74% equivalent that of α-tocopherol activity (Kamal-Eldin and Appelqvist, 1996; Shintani and DellaPenna, 1998; Bramley et al., 2000).

***Tocochromanols in plants***

Tocochromanols are found in most plant species, green algae and cyanobacteria (Dörmann, 2007). They play a number of important functions such as protection of chloroplast from photooxidative damage, or preventing oxidative damage to lipid

components in the period of seed storage and seed germination (Tavva et al., 2007), and may have additional role in photosynthesis (Munne-Bosch and Alegre, 2002). While tocopherols are widely found in dicot species, tocotrienols are the major form of vitamin E in seed of most monocot and rarely found in dicots (Cahoon et al., 2003; Dörmann, 2007). In plant, tocopherols occur in leaves, stems, flower petals, but the highest concentration is found in seeds. While α-tocopherol with the highest vitamin E activity (10 folds higher than γ-tocopherol) is predominant in leaves of most plants, the major tocopherol in seed is γ-tocopherol, (Bramley et al., 2000; Van Eenennaam et al., 2003; Tavva et al., 2007; Dwiyanti et al., 2011). α-tocopherol is especially rich in germ oil, safflower oil, sunflower oil, whereas soybean oil store predominantly γ-tocopherol (Traber and Sies, 1996).

***Vitamin E deficiency***

Vitamin E deficiency due to poor nutrition is rare, but still occurs, especially in developing countries (Dror and Allen, 2011). It may occur in premature infants with hemolytic anemia, in low-birth-weight babies or when vitamin E transferred from mother to fetus is not sufficient (Koletzko et al., 1995). Similar to vitamin A, digestion of vitamin E requires fat to absorb it. Thus, people having problems of malabsorption, metabolism, delivery from diets to tissue are more likely to become deficient than people without such problem (Traber and Sies, 1996). Having vitamin E deficiency, humans may meet symtoms such as blood disorder, poor reflexes, loss of balance, poor muscle coordination with shaky movements (ataxia), sight problem with difficulty seeing at night, decreased immune response (Kowdley et al., 1992; Antinoro, 2000; Monsen, 2000).

**Naturally biosynthetic vitaminE (natural vitamin E) versus chemically synthetic vitamin E (synthetic vitamin E)**

Natural vitamin E is commonly known as d-α-tocopherol or RRR-α-tocopherol and other less common forms of d-α-tocopheryl acetate, d-α-tocopheryl succinate. In contrast, synthetic forms of vitamin E are labeled with dl-α-tocopherol or all-rac-α-tocopherol. While natural α-tocopherol consists of one stereoisomer, synthetic α-tocopherol contains eight different stereoisomers in equal amounts. Only one of eight stereoisomers (about 12.5% of synthetic molecule) is identical to d-α-tocopherol (natural form), the remaining seven stereoisomers have different molecular configurations due to the manufacturing process and their biological activity is only 50-74% of that of the natural α-tocopherol (Clemente and Cahoon, 2009). Moreover, synthetic dl- α-tocopherol has bioavaiability only half that of natural α-tocopherol (Burton et al., 1998; Bramley et al., 2000), and natural source d-α-tocopherol is believed to have more advantages than chemically synthesized α-tocopherol (Eitenmiller, 1997; Traber et al., 1998). Both natural source d-α-tocopherol and synthetic dl- α-tocopherol are absorbed well in the body. However, after absorption, discrimination between natural and synthetic vitamin E occurs in the liver. A hepatic α-tocopherol transfer protein (α-TTP) in the liver recognized only the naturally occuring forms and transfer them to cell when needed. As a result, synthetic vitamin E are preferentially excreted, and natural vitamin E is retained significantly longer in body tissues (Traber et al., 1998; Niki and Traber, 2012). In addition, natural source vitamin E is thought to be transferred from pregnant women to their babies three times more efficiently than synthetic vitamin E. Many studies also presented that other components of natural vitamin E have their own specific physiological benefits. For

example, γ-tocopherol inhibits growth of cancer, prevents inflammation or improves kidney functions; and α-tocotrienol play role in neuroprotection (Jiang et al., 2001; Stoppani, 2004). In fact, most of the vitamin E on the market is synthesized by chemical reaction (Dwiyanti et al., 2011).Therefore, finding a way to improve natural tocopherol intake in human and animal diets is a current interest (Tavva et al., 2007).

***Tocopherol biosynthetic pathway in plants*** **(Figure 1.1B)**

Tocopherols are synthesized from precursors derived from two pathways: the head group derived from homogentistic acid (HGA) in Shikimate pathway and lipophilic tail derived from phytyl diphosphate in MEP (Methylerythritol phosphate) pathway to produce 2-methyl-6-phytylbenzoquinol. Then, enzyme 2-methyl-6-phytylbenzoquinol methyltransferase (VTE3) converts 2-methyl-6-phytylbenzoquinol to 2,3-dimethyl5phytylbenzoquinol, while tocopherol cyclase (VTE1) will convert it to δ-tocopherol. VTE1 also use 2,3-dimethyl-5phytylbenzoquinol as a substrate to convert to γ-tocopherols. Finally, γ-tocopherol methyltransferase (VTE4) is found to use δ-and γ-tocopherols as substrates to produce β- and α- tocopherols, respectively. Biosynthetic pathway of tocotrienols differs from that of tocopherols in using genarylgenaryldiphosphate (GGPP/GGDP) instead of phytyldiphosphate (PDP) as the side chain (Bramley et al., 2000; Van Eenennaam et al., 2003). Tocopherol biosynthesis takes place in the plastids (Li et al., 2011), therefore, it is highly correlated to other metabolic pathways in the plastid compartment including biosynthesis of carotenoids, chlorophylls, gibberellins, phylloquinone (Figure 1.1C) (DellaPenna and Pogson, 2006).

**Previous approaches of enhancing pro-vitamin A in plants**

Because of high nutritional and commercial values as well as grave effects of VAD, vitamin A becomes an interesting topic for many scientists in the world. Research on metabolic engineering of carotenoid biosynthetic pathway to enhance vitamin A for consumers has been achieved recently in many plants. The best known metabolic engineering of pro-vitamin A is "Golden rice". Ye and his coworkers transferred gene encoding daffodil phytoene synthase (*psy*), *Erwinia uredovora* phytoene desaturase (*crtI*) into rice endosperm in single transformation under the control of the endosperm – specific glutelin (Gt1) and the constitutive CaMV (cauliflower mosaic virus) 35S promoter, respectively via *Agrobacterium*-mediated transformation result in "Golden rice1". Carotenoid content of Golden rice 1 events was significantly high, up to 1.6µg carotenoids (0.8µg β-carotene) per gram of dry rice. Furthermore, that of a new version of Golden rice ("Golden rice 2") was developed up to 35µg β-carotene per gram of dry rice by using *phytoene synthase* (*psy)* from maize instead of daffodil as previous version (Ye et al., 2000; Al-Babili and Beyer, 2005; Tang et al., 2009). These transgenic crops contribute in part to enhancing pro-vitamin A intake and combat the VAD problem for people, especially in developing countries that use rice as a major food in their diet. Expression of plant *phytoene synthase* was also studied on tomato. However, this expression of tomato *phytoene synthase* reduced ripe level of transgenic tomato due to gene silencing with the endogenous genes and make dwarfism phenotype that is considered as redirection of GGDP from gibberellin pathway (Fray et al., 1995). Then, the use of *Erwinia uredovora phytoene desaturase* (*crtI*) gene on tomato was approached that resulted in normal morphological phenotype compare to the control.

Therefore, using bacterial *crtI* may have prevented gene silencing that was observed in earlier studies using the tomato phytoene synthase gene. The *crtI* transgenic tomato showed an increase in β-carotene, lutein but a decrease in lycopene, total carotenoids, γ-carotene and δ-carotene(Römer et al., 2000). Transgenic wheat has been also generated by expressing the maize *psy* driven by an endosperm-specific 1Dx5 promoter in the elite wheat, together with the *Erwinia uredovora crtI* under the control of constitutive CaMV 35S promoter. The total carotenoids content was increased up to 10.8-fold as compared with the non-transgenic cultivar (Cong et al., 2009). Bacterial carotenoids enzymes were also applied for various crops such as potato, wheat, maize, canola, soybean and even flaxseed and tobacco. Linseed flax is an industrially important oil crop containing large amounts of α-linolenic acid (18:3) and lignin in its seed oil. Transgenic flax was generated by overexpressing of *Erwinia uredovora phytoene synthase* (*crtB*), resulted in total carotenoid amounts were 65.4–156.3μg/g fresh weight, which corresponded to 7.8- to 18.6-fold increase, compared with those of untransformed controls. The seeds showed orange-colored phenotype with newly accumulation of phytoene, α-carotene, β- carotene (Fujisawa et al., 2008). In addition to flax plants, canola was also engineered to express *crtB*. Transgenic canola seeds showed orange color and carotenoid content rise up to 50-fold increase with the predominant carotenoids are α-carotene and β-carotene (Shewmaker et al., 1999). An approach of seed specific expression of endogenous *Arabidopsis psy* resulted in 43-fold increase in β-carotene and other carotenoid forms (Lindgren et al., 2003). Potato is one of the most common staple crops in the world (ranking after wheat, rice, and maize) and is a source of lutein and violaxanthin which are non-provitamin A activity. Three genes from *Erwinia uredovora: phytoene synthase*

(*crtB*), *phytoene desaturase* (*crtI*) and *lycopene beta-cyclase* (*crtY*) under tuber-specific and constitutive promoter were transferred to potato. While constitutive expression of the *crtY* and/or *crtI* interferes with the establishment of transgenesis and with the accumulation of leaf carotenoids, expression of all three genes under control of tuber specific promoter resulted in tubers with a dark yellow phenotype without any adverse leaf phenotypes. Furthermore, carotenoid amount of these tubers increase approximately 20-fold, to 114 µg/g dry weight and β-carotene 3600-fold, to 47 µg/g dry weight (Diretto et al., 2007). Recently, Kim and coworkers successfully transferred recombinant PAC (Phytoene synthase -2A-Carotene desaturase) gene in a Korean soybean. They tested either β-conglycinin or CaMV-35S promoter. While PAC gene that driven by β-conglycinin was expressed strongly in the seeds, it showed high level in leaves under the control of CaMV-35S promoter. Analysis of $T_2$ generation of β-conglycinin –PAC plants showed that seeds increase 62-fold higher than non-transgenic seeds, accumulated 146µg/g of total carotenoids, of which 112µg/g was β-carotene. In contrast, there was no significant difference in carotenoid components in leaves between transgenic and non-transgenic plants (Kim et al., 2012)

**Previous approaches of enhancing vitamin E in plants**

In 1998, Shintani and DellaPenna reported that Arabidopsis increased α-tocopherol by seed-specific overexpression of *γ-tocopherol methytransferase* (*VTE4*) that made γ-tocopherol completely convert to α-tocopherol and vitamin E increased nine-fold (Shintani and DellaPenna, 1998). Moreover, when overexpression of *Homogentisate phytyltransferase* (*VTE2*) and *VTE4* resulted in tocochromanol increase together with complete conversion of all γ-tocopherol to α-tocopherol and 12-fold increase in vitamin E

activity (Collakova and DellaPenna, 2003). Van Eenennaam and his coworkers reported that by over-expression of *Arabidopsis thaliana VTE4* (*At-VTE4*), transgenic soybean seeds showed that almost γ-tocopherol and δ-tocopherol completely converted to α-tocopherol, and β-tocopherol that make 7-fold increase and 10-fold increase in total tocopherol in soybean seeds, respectively. Moreover, they also mentioned that overexpression of *At-VTE4* together with *At- VTE3* (*2methyl-6-phytylbenzoquinol methyltransferase*) resulted in eight-fold increase in α-tocopherol and fivefold increase in vitamin E activity in transgenic compared to non-transgenic soybean seeds (Van Eenennaam et al., 2003). Tavva and coworkers reported the transgenic soybean with 10.4-fold increase of α-tocopherol and 14.9-fold increase of β-tocopherol in $T_2$ seeds by seed specific expression of *VTE4* from *Perilla frutescens* (Tavva et al., 2007). Cahoon also increased total tocotrienols plus total tocopherols in *Arabidopsis thaliana* leaves by ten to fifteen- fold and six-fold in corn seeds (Cahoon et al. 2003) via overexpression of barley HGGT (homogentesic acid genarylgenaryl transferase) which catalyzes the committed step of tocopherol/tocotrienol biosynthesis compare to previous studies of 4.4-fold and 2-fold increase in leaves (Collakova and DellaPenna, 2003) and seeds (Soll and Schultz, 1979), respectively. Collakova and Della Penna also mentioned that *Homogentisate phytyltransferase* (*HPT*) is limiting in different A*rabidopsis* tissues (leaves and seeds), overexpressing this gene resulted in ten-fold increase in HPT activity and 4.4-fold increase in total tocopherols in leaves and 40% higher in seeds compare to wild type because of higher amount of γ-tocopherol. Moreover, crossing these HPT lines with lines overexpressing γ-tocopherol methyltransferase resulted in seeds increase of 12-

fold in vitamin E activity because of increase α-tocopherol (Collakova and DellaPenna, 2003).

**Soybean oil**

Soybean tocopherols: Soybean seeds contain average 60-70% γ-tocopherol, 20-25% δ-tocopherol, and less than 10% α-tocopherol of total tocopherols (Almonor et al., 1998; Van Eenennaam et al., 2003). γ- and δ-tocopherol that occur principally in soybean seeds, have been reported to greatest degree of oxidative stability to vegetable oil for frying applications , although their vitamin E activities are only 10% and 3% of that of α-tocopherol (Huang et al., 1995; Warner et al., 2003). Recently, many studies on tocopherol enhancement in soybean seeds have focused on strategies to convert different forms to α-tocopherol. Although tocopherols are only accounted for a small percentage of the oil extracted from soybean seeds, they are critical for their oxidative stability in food process and play important role in high temperature performance of soybean oil in biodiesel and bio-based lubricants (Warner et al., 2003; Krahl et al., 2005; Warner, 2005).

In summary, various efforts have been made previously to enhance either pro-vitamin A or vitamin E, but not both. Therefore, it would be advantageous to increase both in soybean seeds to improve soybean seed nutritional values and combat vitamin A and E deficiency.

## Strategies to express multiple genes in plants

### *Conventional stacking methods/crossing/iterative processes*

Iterative process is an early introduced strategy that starts with introduction of desired genes into different cassettes and transfers each cassette into different independent plants by plant transformation. Subsequently, these different transgenic lines will be crossed to yield progeny that carry all of the genes of interest. This approach has been successfully used to generate plants that carry and express polyhydroxybutyrate biosynthetic pathway, comprising three genes (Nawrath et al., 1994) or tobacco transgenic plant that generate and assemble of functional antibodies by co-expression of four genes of a secretory immunoglobulin A (Ma et al., 1995). Another strategy is sequential transformation, which means available transgenic plants are transformed with additional transgenes (Lapierre et al., 1999; Jobling et al., 2002; Qi et al., 2004). However, these methods are labor intensive and time consuming involving multiple construction of expression cassettes, separate transformations of each gene, and several crosses with multiple verifications per cross. Moreover, the level of expression of transgenes is highly variable and influenced by many factors including the use of promoters with different strengths or gene silencing induced the use homologous promoters. Segregation of subsequent generations, if these genes are unlinked, is also a potential problem (Ma and Mitra, 2002; Naqvi et al., 2010).

### *Co-transformation*

Two or more transgenes will be introduced simultaneously. Advantages of this method are transgenic plants can carry multiple transgenes in one generation and

preventing subsequent segregation (Naqvi et al., 2010). It could be divided into two broad approaches:

*Co-transformation of linked genes***:** assembly of multiple genes on the same plasmid (Twyman et al., 2002; Karunanandaa et al., 2005).

*Co-transformation of unlinked genes*: desired genes that are carried on separate plasmids and transferred into the same cell. This method is advantageous in that selectable marker need not attach each foreign gene.

There is an equal efficiency between linked and unlinked co-transformation of two transgenes. However, when the number of transgenes increases, linked co-transformation becomes much less efficient because of vector instability, lack of unique restriction sites in cloning process, trouble in cloning multiple transgene expressions into a single construct, and the fact that larger input DNA sequences are more likely to fragments. In contrast, unlinked co-transformation shows to be much better. Particle bombardment with multiple input plasmids can achieve transgenic plants carrying all the genes with high efficiency (Ye et al., 2000; Cong et al., 2009; Jing et al., 2009; Naqvi et al., 2009). Cahoon et al. produced transgenic soybean seeds with highly increased total tocopherol and tocotrienol by co-transformation of two plasmids containing two genes of barley *HGGT* and soybean *VTE4* (Cahoon et al., 2003). Although, this method has been used with some success, it also has potential drawbacks such as gene silencing caused by different expression cassettes being driven by same promoter.

***Self-processing polyprotein***

The best strategy to overcome the above problems is the expression of multiple genes in the form of a self-processing polyproteins as *Picornavirus* and *Potyviruses*. In

these viruses, the genome often encodes a single, long open reading frame. Expression of the viral genome results in production of polyproteins which are processed into discrete proteins by proteinase encoded within the polyprotein itself (Ryan and Flint, 1997).

***Advantages of self-cleaving 2A systems***

Food-and-mouth disease virus (FMDV) contains a sequence 2A of just 20 amino acids with its unique capability to mediate cleavage at its own Carboxyl- terminus of the 2A region.

This 2A self-cleaving system can be employed to express multiple genes using a single construct in coordinated expression of transgenes with just one transformation (especially for two genes (Halpin et al., 1999)). Co-expression of multiple genes using 2A element has been used for variety of applications on animals especially on mice (Trichas et al., 2008; Carey et al., 2009; Casales et al., 2010) and plants (Halpin et al., 1999; Ma and Mitra, 2002).

Single construct with a single set of promoter and terminator sequence can avoid silencing and minimize transgene expression instability.

2A is a short amino acid sequence (just 20 amino acids), and is energetically efficient so plants don't utilize much energy to synthesize (Halpin et al., 1999).

Efficiency in targeting constituent proteins in polyproteins with their own targeting information to different subcellular has been reported (Halpin et al., 1999; Lorens et al., 2004; Provost et al., 2007).

Its self-cleaving activity is independent with any cellular factors. Thus it can be effective in all tissues and cells at all developmental stages that make it be more

advantageous in particular application where crossing is not applied for some tree species (Halpin et al., 1999).

Recognition site in its sequence is for ubiquitous ribosome-associated eukaryotic protease. However, 2A polyproteins do not cleave in *E.coli* (Donnelly et al., 1997).

Expression of multiple genes in the cassette has been shown to be at equal levels. (Halpin et al., 1999; Trichas et al., 2008).

***Mechanism of "self-cleavage" 2A system***

In fact, the "self-cleavage" mechanism is still not completely known. However, recent studies demonstrated that mechanism of "self-cleavage" of 2A sequence is "ribosomal skipping". Ribosome pauses near the end of the 2A coding sequence at the junction of Glycine (G) and Proline (P) disrupting the formation of single peptide bond and release of the nascent upstream protein, while allowing continued translation of the downstream protein (Donnelly et al., 1997; Doronina et al., 2008). It means that, after cleavage, 2A remains attached to the C-terminus of the upstream protein while single Proline residue remains attached to the N-terminus of the downstream protein (Halpin et al., 1999; Doronina et al., 2008). There is generally no problem in additional Proline of following protein (de Felipe et al., 2006), however longer extension (2A remaining) at C-termius of the upstream protein might have unpredictable effects-(Fisicaro et al., 2011).

Although FMDV 2A region functions properly within different contexts, efficiency of cleavage varies when flanking contexts change. Moreover, multiple copies of 2A in a single open reading frame function properly and independently and can be used for co-ordinate expression of multiple genes (Ma and Mitra, 2002).

The goal of this research is to ascertain the utility of the coordinated gene expression by 2A system while simultaneously increase pro-vitamin A and E contents in soybean seeds.

**Figure legend**

**Figure 1.1A:** Carotenoid biosynthesis pathway in plants and corresponding steps in bacteria. CRTB, bacterial phytoene synthase; CRTE, bacterial geranylgeranyl diphosphate synthase; CRTI, bacterial phytoene desaturase/isomerase; CRTISO, carotenoid isomerase; CRTY, bacterial lycopene cyclase; CRTZ, bacterial β-carotene hydroxylase; CYP97C, carotene ε-ring hydroxylase; DMAPP, dimethylallyl diphosphate; GGPP/GGDP, geranylgeranyl diphosphate; GGPPS/GGDPS, GGPP synthase/GGDP synthase; HYDB, β-carotene hydroxylase [non-heme di-iron hydroxylases, β-carotene hydroxylase (BCH) and heme-containing cytochrome P450 β-ring hydroxylases, CYP97A and CYP97B]; IPP, isopentenyl diphosphate; IPPI, isopentenyl diphosphate isomerase; LYCB, lycopene β-cyclase; LYCE, lycopene ε-cyclase; PDS, phytoene desaturase; PSY, phytoene synthase; VDE, violaxanthin de-epoxidase; ZDS, ζ-carotene desaturase; ZEP, zeaxanthin epoxidase; Z-ISO, ζ-carotene isomerase. (Farré et al., 2011).

**Figure 1.1B:** Tocopherol and tocotrienol structures and biosynthesis.

(a) Detailed chemical structures of tocopherol and tocotrienol. The canonical structures are shown, with R-group modifications illustrated in the gridded box at top right.

(b) Final steps of the tocopherol biosynthetic pathway. MEP: methylerythritol phosphate; VTE1: tocopherol cyclase; VTE2: homogentisic acid prenyltransferase; VTE3: 2-methyl-6-phytylbenzoquinol (MPBQ) methyltransferase; VTE4, γ-tocopherol methyltransferase.

**Figure 1.1C:** Overview of carotenoid and tocopherol biosyntheses in plants. The 2-C-methyl- d-erythritol-4-phosphate (MEP) pathway provides isopetenylpyrophosphate (IPP) for synthesis of the central intermediate geranylgeranyl diphosphate (GGDP).

GGDP can be used for synthesis of phytoene, chlorophylls, and tocotrienols or reduced to phytyl-diphosphate (PDP) used for phylloquinone, chlorophyll, and tocopherol synthesis. Phytol released from chlorophyll degradation is also used for tocopherol synthesis (not shown, see text). The pathway shown by orange arrows provides the carotenoids found in leaves of most plant species. Other carotenoids and carotenoid leavage/modification products are produced in certain species and/or particular tissues, and when known, the primary substrate for cleavage is given in parentheses. The pathway shown by green arrows is the synthesis of tocopherols from homogentisate, a product of the shikimate pathway. For clarity, only tocopherols are shown, but when GGDP is condensed with homogentisate, the corresponding tocotrienols are produced by the same pathway. ABA, abscisic acid; MPBQ, methyl-6-phytyl-1, 4-benzoquinone.

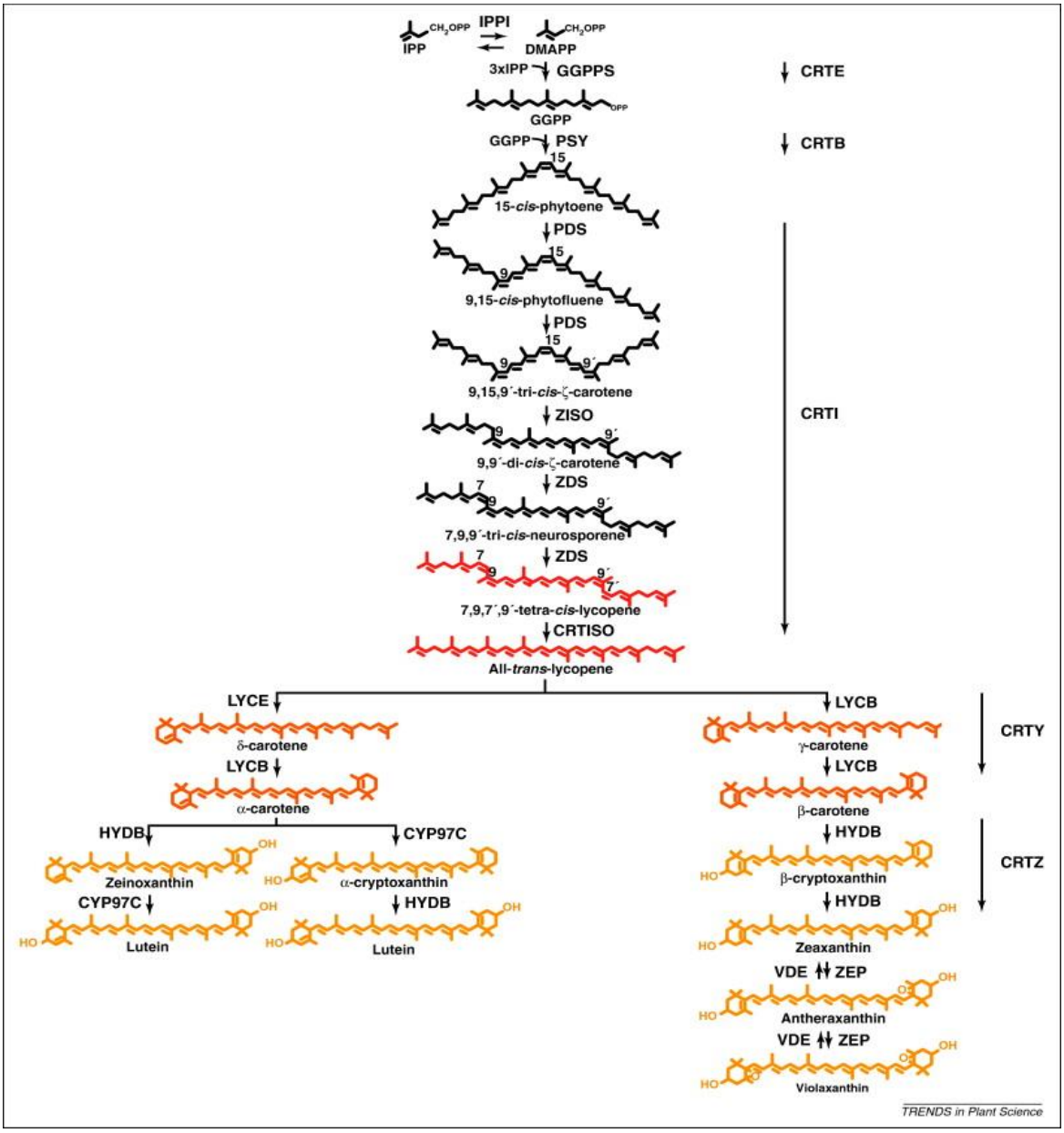
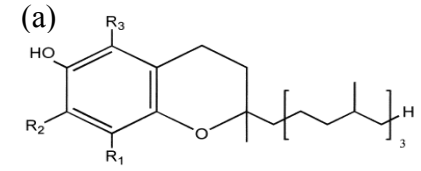
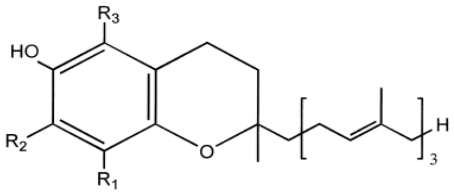

IPP
IPPI
DMAPP
3xIPP
GGPPS
CRTE
GGPP
GGPP
PSY
CRTB
15-cis-phytoene
PDS
9,15-cis-phytofluene
PDS
9,15,9´-tri-cis-ζ-carotene
ZISO
CRTI
9,9´-di-cis-ζ-carotene
ZDS
7,9,9´-tri-cis-neurosporene
ZDS
7,9,7´,9´-tetra-cis-lycopene
CRTISO
All-trans-lycopene
LYCE
δ-carotene
LYCB
α-carotene
HYDB
Zeinoxanthin
CYP97C
α-cryptoxanthin
CYP97C
Lutein
HYDB
Lutein
LYCB
γ-carotene
CRTY
LYCB
β-carotene
HYDB
β-cryptoxanthin
CRTZ
HYDB
Zeaxanthin
VDE ZEP
Antheraxanthin
VDE ZEP
Violaxanthin
TRENDS in Plant Science

**Figure 1.1A**

(a)



Tocopherols



Tocotrienols

| | $R_1$ | $R_2$ | $R_3$ |
|---|---|---|---|
| α | Me | Me | Me |
| β | Me | H | Me |
| δ | Me | H | H |
| γ | Me | Me | H |

(b)

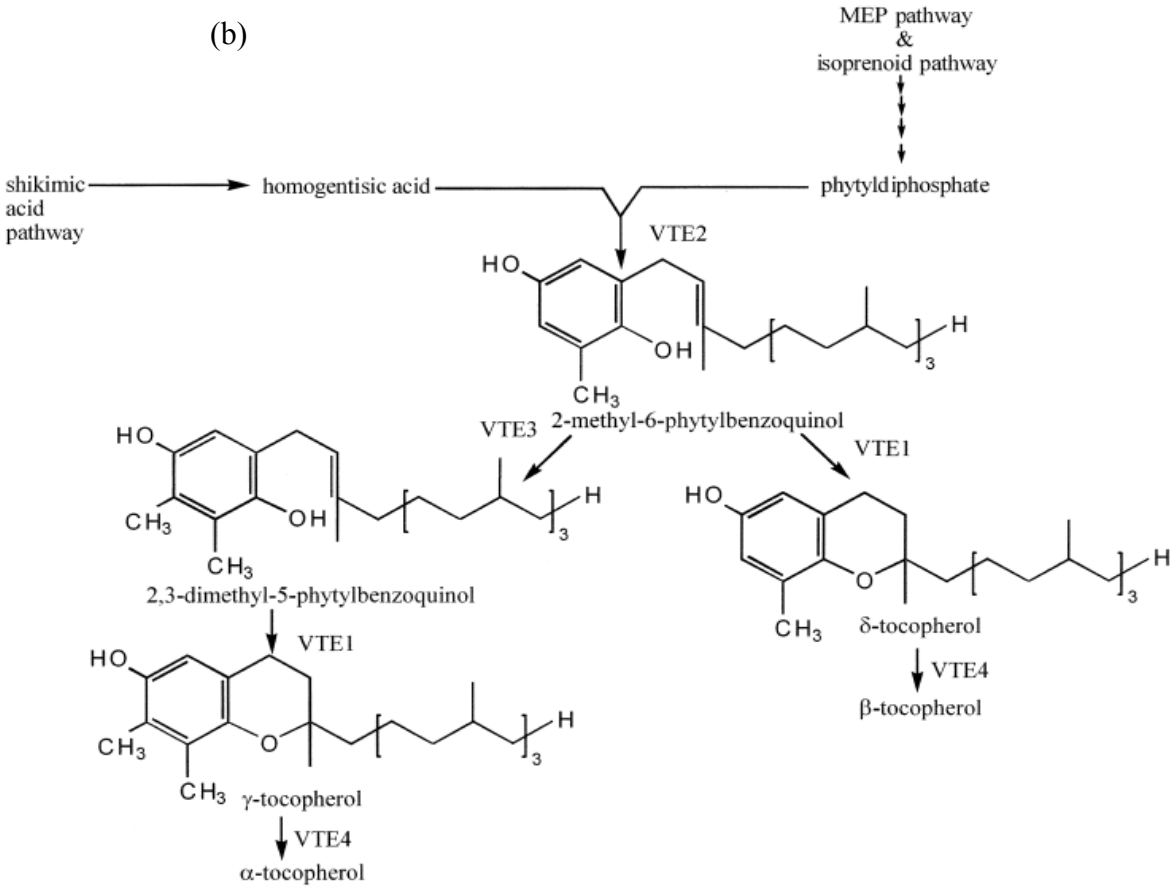



**Figure 1.1B**

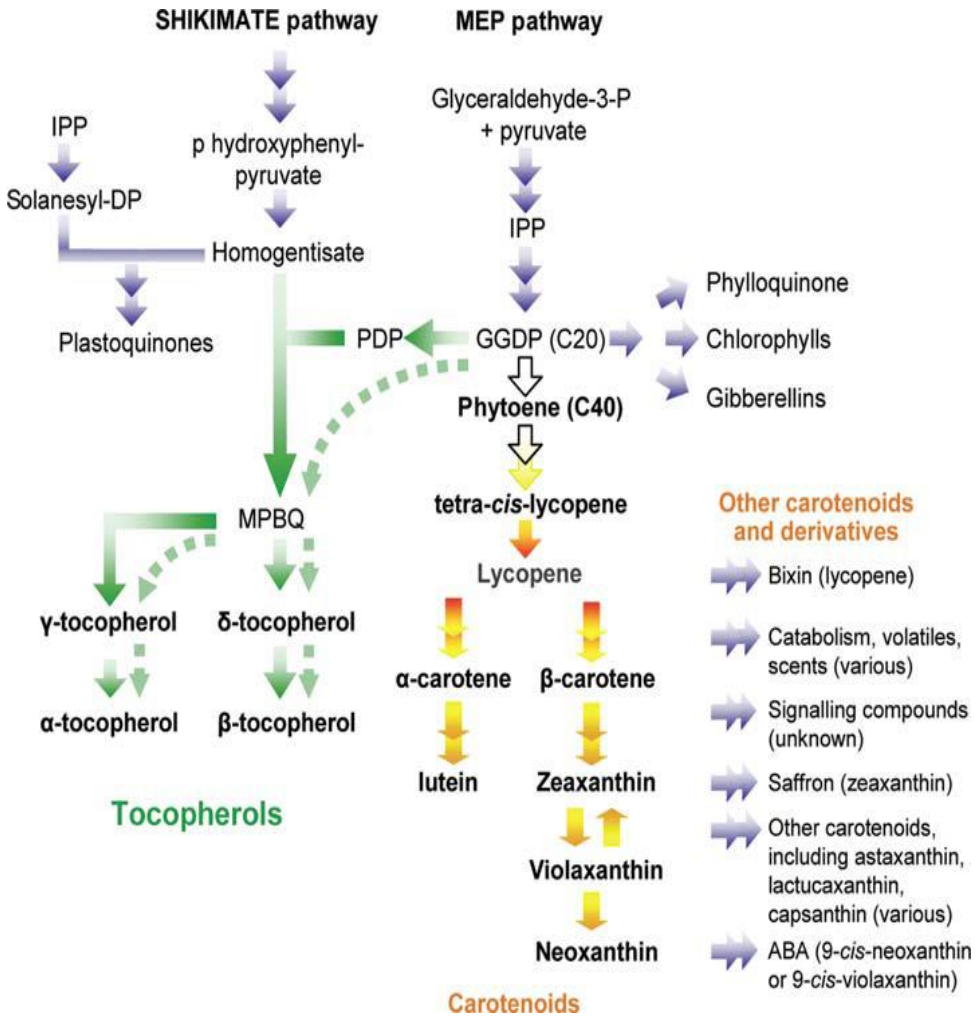

SHIKIMATE pathway
MEP pathway
IPP
p hydroxyphenyl-pyruvate
Glyceraldehyde-3-P + pyruvate
Solanesyl-DP
IPP
Homogentisate
Phylloquinone
Plastoquinones
PDP
GGDP (C20)
Chlorophylls
Gibberellins
Phytoene (C40)
tetra-cis-lycopene
Other carotenoids and derivatives
MPBQ
Lycopene
Bixin (lycopene)
γ-tocopherol
δ-tocopherol
α-carotene
β-carotene
Catabolism, volatiles, scents (various)
Signalling compounds (unknown)
α-tocopherol
β-tocopherol
lutein
Zeaxanthin
Saffron (zeaxanthin)
Other carotenoids, including astaxanthin, lactucaxanthin, capsanthin (various)
Violaxanthin
Tocopherols
Neoxanthin
ABA (9-cis-neoxanthin or 9-cis-violaxanthin)
Carotenoids

**Figure 1.1C**

## LITERATURE CITED


http://www.soystats.com/2012/Default-frames.htm

Al-Babili, S. and P. Beyer. 2005. Golden Rice – five years on the road – five years to go? Trends in Plant Science 10: 565-573.

Almonor, G.O., R.F. Wilson and G.P. Fenner. 1998. Temperature effects on tocopherol composition in soybeans with genetically improved oil quality. Journal of the American Oil Chemists' Society 75: 591-596.

Aluru, M., Y. Xu, R. Guo, Z. Wang, S. Li, W. White, et al. 2008. Generation of transgenic maize with enhanced provitamin A content. Journal of Experimental Botany 59: 3551-3562.

Antinoro, L. 2000. Has The Vitamin E Bubble Burst? EN Evaluates Whether To Supplement. (Cover story). Environmental Nutrition 23: 1.

Azadbakht, L., R. Shakerhosseini, S. Atabak, M. Jamshidian, Y. Mehrabi and A. Esmaill-Zadeh. 2003. Beneficiary effect of dietary soy protein on lowering plasma levels of lipid and improving kidney function in type II diabetes with nephropathy. European Journal of Clinical Nutrition 57: 1292.

Bennett, J.O., H.B. Krishnan, W.J. Wiebold and A.H. Krishnan. 2003. Positional effect on protein and oil content and composition of soybeans. Journal of Agricultural and Food Chemistry 51: 6882-6886.

Bramley, P.M., W. Schuch, H.E. Roxborough, K.H. Wagner, P.J.A. Sheehy, A. Kafatos, et al. 2000. Vitamin E. Journal of the Science of Food and Agriculture 80: 913-938.

Burton, G.W., H. Kayden, L. Hughes, K.U. Ingold, M.G. Traber, R.V. Acuff, et al. 1998. Human plasma and tissue alpha-tocopherol concentrations in response to supplementation with deuterated natural and synthetic vitamin E. American Journal of Clinical Nutrition 67: 669-684.

Cahoon, E.B., S.E. Hall, K.G. Ripp, S.G. Thaya, W.D. Hitz and S.J. Coughlan. 2003. Metabolic redesign of vitamin E biosynthesis in plants for tocotrienol production and increased antioxidant content. Nature Biotechnology 21: 1082.

Carey, B.W., S. Markoulaki, J. Hanna, K. Saha, Q. Gao, M. Mitalipova, et al. 2009. Reprogramming of murine and human somatic cells using a single polycistronic vector. Proceedings of the National Academy of Sciences of the United States of America 106: 157-162.

Casales, E., A. Aranda, J.I. Quetglas, M. Ruiz-Guillen, J.R. Rodriguez-Madoz, J. Prieto, et al. 2010. A novel system for the production of high levels of functional human

therapeutic proteins in stable cells with a Semliki Forest virus noncytopathic vector. New Biotechnology 27: 138-148.

Chen, K.-I., M.-H. Erh, N.-W. Su, W.-H. Liu, C.-C. Chou and K.-C. Cheng. 2012. Soyfoods and soybean products: from traditional use to modern applications. Applied Microbiology & Biotechnology 96: 9-22.

Chen, Y.-M., S.C. Ho, S.S.H. Lam, S.S.S. Ho and J.L.F. Woo. 2003. Soy isoflavones have a favorable effect on bone loss in Chinese postmenopausal women with lower bone mass: a double-blind, randomized, controlled trial. The Journal Of Clinical Endocrinology and Metabolism 88: 4740-4747.

Clemente, T.E. and E.B. Cahoon. 2009. Soybean oil: genetic approaches for modification of functionality and total content. Plant Physiology 151: 1030-1040.

Collakova, E. and D. DellaPenna. 2003. Homogentisate phytyltransferase activity is limiting for tocopherol biosynthesis in Arabidopsis. Plant Physiology 131: 632-642.

Cong, L., G. Yang, G. He, H. Liu, C. Wang and L. Chen. 2009. Expression of phytoene synthase1 and Carotene Desaturase crtI Genes Result in an Increase in the Total Carotenoids Content in Transgenic Elite Wheat (Triticum aestivum L.) [electronic resource]. Journal of Agricultural and Food Chemistry 57: 8652-8660.

Constantinou, C., A. Papas and A.I. Constantinou. 2008. Vitamin E and cancer: An insight into the anticancer activities of vitamin E isomers and analogs. International Journal Of Cancer. Journal International Du Cancer 123: 739-752.

de Felipe, P., G.A. Luke, L.E. Hughes, D. Gani, C. Halpin and M.D. Ryan. 2006. E unum pluribus: multiple proteins from a self-processing polyprotein. Trends in Biotechnology 24: 68-75.

DellaPenna, D. and B.J. Pogson. 2006. Vitamin Synthesis in Plants: Tocopherols and Carotenoids. Annual Review of Plant Biology 57: 711-738.

Diretto, G., S. Al-Babili, R. Tavazza, V. Papacchioli, P. Beyer and G. Giuliano. 2007. Metabolic engineering of potato carotenoid content through tuber-specific overexpression of a bacterial mini-pathway. PLoS ONE 2: e350-e350.

Donnelly, M.L.L., M.D. Ryan, S. Monaghan and D. Gani. 1997. The cleavage activities of aphthovirus and cardiovirus 2A proteins. Journal of General Virology 78: 13-21.

Dörmann, P. 2007. Functional diversity of tocochromanols in plants. Planta 225: 269-276.

Doronina, V.A., C. Wu, P. de Felipe, M.S. Sachs, M.D. Ryan and J.D. Brown. 2008. Site-specific release of nascent chains from ribosomes at a sense codon. Molecular And Cellular Biology 28: 4227-4239.

Dror, D.K. and L.H. Allen. 2011. Vitamin E deficiency in developing countries. Food And Nutrition Bulletin 32: 124-143.

Duerbeck, N.B. and D.D. Dowling. 2012. Vitamin A: too much of a good thing? Obstetrical & Gynecological Survey 67: 122-128.

Dutra-de-Oliveira, J.E., R.M. Fávaro, M.V. Junqueira-Franco, C.G. Carvalho, A.A. Jordão Júnior and H. Vannucchi. 1998. Effect of heat treatment on the biological value of beta-carotene added to soybean cooking oil in rats. International Journal Of Food Sciences And Nutrition 49: 205-210.

Dwiyanti, M.S., T. Yamada, M. Sato, J. Abe and K. Kitamura. 2011. Genetic variation of γ-tocopherol methyltransferase gene contributes to elevated α-tocopherol content in soybean seeds. BMC Plant Biology 11: 152-152.

Eitenmiller, R.R. 1997. Vitamin E content of fats and oils--nutritional implications. Food Technology 51: 78-81.

Farbstein, D., A. Kozak-Blickstein and A.P. Levy. 2010. Antioxidant Vitamins and Their Use in Preventing Cardiovascular Disease. Molecules 15: 8098-8110.

Farré, G., C. Bai, R.M. Twyman, T. Capell, P. Christou and C. Zhu. 2011. Nutritious crops producing multiple carotenoids – a metabolic balancing act. Trends in Plant Science 16: 532-540.

Farre, G., K. Ramessar, R.M. Twyman, T. Capell and P. Christou. 2010. The humanitarian impact of plant biotechnology: recent breakthroughs vs bottlenecks for adoption. Current Opinion In Plant Biology 13: 219-225.

Fisicaro, N., S.L. Londrigan, J.L. Brady, E. Salvaris, M.B. Nottle, P.J. O'Connell, et al. 2011. Versatile co-expression of graft-protective proteins using 2A-linked cassettes. Xenotransplantation 18: 121-130.

Fray, R.G., P. Hedden, P.M. Bramley, D. Grierson, A. Wallace, P.D. Fraser, et al. 1995. Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. Plant Journal 8: 693-701.

Fujisawa, M., M. Watanabe, S.-K. Choi, M. Teramoto, K. Ohyama and N. Misawa. 2008. Enrichment of carotenoids in flaxseed (Linum usitatissimum) by metabolic engineering with introduction of bacterial phytoene synthase gene crtB. Journal of Bioscience & Bioengineering 105: 636-641.

Fukushima, D. 2001. Recent progress in research and technology on soybeans. Kikkoman Corp., Noda, Chiba 278-0037, Japan. E-mail danjif@ka2.so-net.ne.jp. p. 8-16.

Galloway, R., P.J. White and L.A. Johnson. 2008. Soybeans : chemistry, production, processing, and utilization / editors, Lawrence A. Johnson, Pamela J. White, Richard Galloway. Urbana, IL : AOCS Press, c2008.

Giuliano, G., R. Aquilani and S. Dharmapuri. 2000. Metabolic engineering of plant carotenoids. Trends in Plant Science 5: 406-409.

Grundy, S.M. and M.A. Denke. 1990. Dietary influences on serum lipids and lipoproteins. Journal Of Lipid Research 31: 1149-1172.

Halpin, C., A. El Amrani, M.D. Ryan, S.E. Cooke and A. Barakate. 1999. Self-processing 2A-polyproteins -- a system for co-ordinate expression of multiple proteins in transgenic plants. Plant Journal 17: 453-459.

Hamilton-Reeves, J.M., S.A. Rebello, W. Thomas, J.W. Slaton and M.S. Kurzer. 2007. Soy protein isolate increases urinary estrogens and the ratio of 2:16α -hydroxyestrone in men at high risk of prostate cancer. Journal of Nutrition 137: 2258-2263.

Hartman, G.L., E.D. West and T.K. Herman. 2011. Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. Food Security 3: 5-17.

Henkel, J. 2000. Soy: health claims for soy protein, questions about other components. FDA Consumer 34: 7p.

Hercberg, S., P. Galan and P. Preziosi. 1999. Antioxidant Vitamins and Cardiovascular Disease: Dr Jekyll or Mr Hyde? American Journal of Public Health 89: 289-291.

Huang, S.W., J.B. German and E.N. Frankel. 1995. Effects of individual tocopherols and tocopherol mixtures on the oxidative stability of corn oil triglycerides. Journal of Agricultural and Food Chemistry 43: 2345-2350.

Jiang, Q., S. Christen, M.K. Shigenaga and B.N. Ames. 2001. Gamma-tocopherol, the major form of vitamin E in the US diet, deserves more attention. American Journal of Clinical Nutrition 74: 714-722.

Jing, J., W. Gang, W. Jiehua and W. Ping. 2009. Functional analysis of multiple carotenogenic genes from Lycium barbarum and Gentiana lutea L. for their effects on β-carotene production in transgenic tobacco. Biotechnology Letters 31: 305-312.

Jobling, S.A., R.J. Westcott, A. Tayal, R. Jeffcoat and G.P. Schwall. 2002. Production of a freeze-thaw-stable potato starch by antisense inhibition of three starch synthase genes. Nature Biotechnology 20: 295-299.

Kamal-Eldin, A. and L.A. Appelqvist. 1996. The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31: 671-701.

Karunanandaa, B., Q. Qi, M. Hao, S.R. Baszis, P.K. Jensen, Y.-H.H. Wong, et al. 2005. Metabolically engineered oilseed crops with enhanced seed tocopherol. Metabolic Engineering 7: 384-400.

Kim, M.-J., Y.-S. Chung, S.-H. Ha, J.K. Kim, H.J. Kim, J.H. Pak, et al. 2012. Genetic modification of the soybean to enhance the β-carotene content through seed-specific expression. PloS ONE 7: e48287.

Kinsell, L.W., G.D. Michaels, G.C. Cochrane, J.W. Partridge, J.P. Jahn and H.E. Balch. 1954. Effect of vegetable fat on hypercholesterolemia and hyperphospholipidemia: observations on diabetic and nondiabetic subjects given diets high in vegetable fat and protein. Diabetes 3: 113-119.

Koletzko, B., T. Decsi and G. Sawatzki. 1995. Vitamin E status of low birthweight infants fed formula enriched with long-chain polyunsaturated fatty acids. International Journal for Vitamin and Nutrition Research 65: 101-104.

Kowdley, K.V., J.B. Mason, S.N. Meydani, S. Cornwall and R.J. Grand. 1992. Vitamin E deficiency and impaired cellular immunity related to intestinal fat malabsorption. Gastroenterology 102: 2139-2142.

Krahl, J., J.H. Van Gerpen and G. Knothe. 2005. The biodiesel handbook / editors, Gerhard Knothe, Jon Van Gerpen, Jurgen Krahl. Champaign, Ill. : AOCS Press, c2005.

Lapierre, C., W. Boerjan, J.C. Leple, V.d. Nadai, V. Ferret, G. Pilate, et al. 1999. Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid O-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. Plant Physiology 119: 153-163.

Li, Y., G. Wang, R. Hou, Y. Zhou, R. Gong, X. Sun, et al. 2011. Engineering tocopherol biosynthetic pathway in lettuce. Biologia Plantarum 55: 453-460.

Lindgren, L.O., K.G. Stålberg and A.-S. Höglund. 2003. Seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll, and abscisic acid. Plant Physiology 132: 779-785.

Lorens, J.B., D.M. Pearsall, S.E. Swift, B. Peelle, R. Armstrong, S.D. Demo, et al. 2004. Stable, stoichiometric delivery of diverse protein functions. Journal of Biochemical & Biophysical Methods 58: 101.

Ma, C.L. and A. Mitra. 2002. Expressing multiple genes in a single open reading frame with the 2A region of foot-and-mouth disease virus as a linker. Molecular Breeding 9: 191-199.

Ma, J.K., A. Hiatt, M. Hein, N.D. Vine, F. Wang, P. Stabila, et al. 1995. Generation and assembly of secretory antibodies in plants. Science (New York, N.Y.) 268: 716-719.

MacDonald, R.S., G. JuYuan, J. Copeland, J.D. Browning Jr, D. Sleper, G.E. Rottinghaus, et al. 2005. Environmental Influences on Isoflavones and Saponins in Soybeans and Their Role in Colon Cancer. Journal of Nutrition 135: 1239-1242.

Mayne, S.T. and R.S. Parker. 1988. HPLC analysis of vitamin E by conversion to alpha-tocopheryl acetate in samples containing canthaxanthin of other coeluting compounds. Journal of Agricultural and Food Chemistry 36: 483-485.

Messina, M., J.W. Lampe and O. Kucuk. 2006. An Overview of the Health Effects of Isoflavones with an Emphasis on Prostate Cancer Risk and Prostate-Specific Antigen Levels. Journal of AOAC International 89: 1121-1134.

Monsen, E.R. 2000. Commentary. Dietary reference intakes for the antioxidant nutrients: vitamin C, vitamin E, selenium, and carotenoids [corrected] [published erratum appears in J AM DIET ASSOC 2000 Sep; 100(9): 1008]. Journal of the American Dietetic Association 100: 637-640.

Munne-Bosch, S. and L. Alegre. 2002. The function of tocopherols and tocotrienols in plants. Critical Reviews in Plant Sciences 21: 31-57.

Nagata, Y., T. Sonoda, M. Mon, N. Miyanaga, K. Okumura, K. Goto, et al. 2007. Dietary Isoflavones May Protect against Prostate Cancer in Japanese Men. Journal of Nutrition 137: 1974-1979.

Naqvi, S., G. Farré, G. Sanahuja, T. Capell, C. Zhu and P. Christou. 2010. When more is better: multigene engineering in plants. Trends in Plant Science 15: 48-56.

Naqvi, S., C. Zhu, G. Farre, K. Ramessar, L. Bassie, J. Breitenbach, et al. 2009. Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. Proceedings of the National Academy of Sciences of the United States of America 106: 7762-7767.

Nawrath, C., C. Somerville and Y. Poirier. 1994. Targeting of the polyhydroxybutyrate biosynthetic pathway to the plastids of Arabidopsis thaliana results in high levels of polymer accumulation. Proceedings of the National Academy of Sciences of the United States of America 91: 12760-12764.

Niki, E. and M.G. Traber. 2012. A History of Vitamin E. Annals of Nutrition & Metabolism 61: 207-212.

Pednekar, M., A.K. Das, V. Rajalakshmi and A. Sharma. 2010. Radiation processing and functional properties of soybean (Glycine max). Radiation Physics and Chemistry 79: 490-494.

Provost, E., J. Rhee and S.D. Leach. 2007. Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos. Genesis (New York, N.Y.: 2000) 45: 625-629.

Qi, B., T. Fraser, S. Mugford, G. Dobson, O. Sayanova, J. Butler, et al. 2004. Production of very long chain polyunsaturated omega-3 and omega-6 fatty acids in plants. Nature Biotechnology 22: 739-745.

Ramakrishnan, U. and I. Darnton-Hill. 2002. Assessment and Control of Vitamin A Deficiency Disorders. Journal of Nutrition 132: 2947S.

Römer, S., P.D. Fraser, J.W. Kiano, C.A. Shipton, N. Misawa, W. Schuch, et al. 2000. Elevation of the provitamin A content of transgenic tomato plants. Nature Biotechnology 18: 666.

Ryan, M.D. and M. Flint. 1997. Virus-encoded proteinases of the picornavirus super-group. The Journal Of General Virology 78 ( Pt 4): 699-723.

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

Shintani, D. and D. DellaPenna. 1998. Elevating the vitamin E content of plants through metabolic engineering. Science 282: 2098.

Soll, J. and G. Schultz. 1979. Comparison of geranylgeranyl and phytyl substituted methylquinols in the tocopherol synthesis of spinach chloroplasts. Biochemical And Biophysical Research Communications 91: 715-720.

Stoppani, J. 2004. VITAMIN E STACKED! Joe Weider's Muscle & Fitness 65: 180.

Tang, G., J. Qin, G.G. Dolnikowski, R.M. Russell and M.A. Grusak. 2009. Golden Rice is an effective source of vitamin A. American Journal of Clinical Nutrition 89: 1776-1783.

Tavva, V.S., Y.-H. Kim, I.A. Kagan, R.D. Dinkins, K.-H. Kim and G.B. Collins. 2007. Increased alpha-tocopherol content in soybean seed overexpressing the Perilla frutescens gamma-tocopherol methyltransferase gene. Plant Cell Reports 26: 61-70.

Traber, M.G., D. Rader, R.V. Acuff, R. Ramakrishnan, H.B. Brewer and H.J. Kayden. 1998. Vitamin E dose-response studies in humans with use of deuterated RRR-alpha-tocopherol. American Journal of Clinical Nutrition 68: 847-853.

Traber, M.G. and H. Sies. 1996. Vitamin E in humans: demand and delivery. Annual Review of Nutrition 16: 321-347.

Trichas, G., J. Begbie and S. Srinivas. 2008. Use of the viral 2A peptide for bicistronic expression in transgenic mice. BMC Biology 6: 1-13.

Twyman, R.M., A. Kohli, E. Stoger and P. Christou. 2002. Foreign DNA: integration and expression in transgenic plants. Genetic Engineering 24: 107-136.

Van Eenennaam, A.L., K. Lincoln, T.P. Durrett, H.E. Valentin, C.K. Shewmaker, G.M. Thorne, et al. 2003. Engineering vitamin E content: from Arabidopsis mutant to soy oil. The Plant Cell 15: 3007-3019.

Villegas, R., Y.T. Gao, G. Yang, H.L. Li, T.A. Elasy, W. Zheng, et al. 2008. Legume and soy food intake and the incidence of type 2 diabetes in the Shanghai Women's Health Study. American Journal of Clinical Nutrition 87: 162-167.

Warner, K. 2005. Effects on the flavor and oxidative stability of stripped soybean and sunflower oils with added pure tocopherols. Journal of Agricultural and Food Chemistry 53: 9906-9910.

Warner, K., W.E. Neff and F.J. Eller. 2003. Enhancing quality and oxidative stability of aged fried food with gamma-tocopherol. Journal of Agricultural and Food Chemistry 51: 623-627.

Winklhofer-Roob, B.M., E. Rock, J. Ribalta, D.H. Shmerling and J.M. Roob. 2003. Effects of vitamin E and carotenoid status on oxidative stress in health and disease. Evidence obtained from human intervention studies. Molecular Aspects of Medicine 24: 391.

Wurbs, D., S. Ruf and R. Bock. 2007. Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. Plant Journal 49: 276-288.

Yamada, T., K. Takagi and M. Ishimoto. 2012. Recent advances in soybean transformation and their application to molecular breeding and genomic analysis. Breeding Science 61: 480-494.

Ye, X., S. Al-Babili, A. Kloti, J. Zhang, P. Lucca, P. Beyer, et al. 2000. Engineering the Provitamin A (Beta-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice E. Science 287: 303.

Young, V.R. 1991. Soy protein in relation to human protein and amino acid nutrition. Journal of the American Dietetic Association 91: 828-835.

Zimmermann, R. and M. Qaim. 2004. Potential Health Benefits of Golden Rice: A Philippine Case Study. Food Policy 29: 147-168.

Zyriax, B.C. and E. Windler. 2000. Dietary fat in the prevention of cardiovascular disease - A review. European Journal of Lipid Science and Technology 102: 355-365.

# CHAPTER 2

# CO-ORDINATED EXPRESSION OF CRTB, AT-VTE3, AND VTE4 TO ENHANCE PRO-VITAMIN A AND VITAMIN E VIA FMDV 2A SEQUENCE IN SOYBEAN SEEDS

## INTRODUCTION

Soybean [*Glycine max* (Merr.) ] is one of the most valuable crops in the world not only as a staple food in some countries, especially Asian countries, providing a good source of protein for human diet, but also as an oil seed crop, feed for livestock, aquaculture, and biofuel feedstock. To date, soybean has become an important commercial crop with growing acreage and food products (Hartman et al., 2011). Therefore, functional and nutritional improvements of the soybean seeds for food stability and human health have gained more attention (Sattler et al., 2004). Both carotenoids and tocopherols are well-known antioxidants and intake of elevated levels of carotenoids and tocopherols in human diet is believed to reduce several diseases (Shintani and DellaPenna, 1998; Winklhofer-Roob et al., 2003). However, these two health-beneficial vitamins are only synthesized in photosynthetic organisms.

Engineering carotenoid biosynthetic pathway has been applied on various crops. The "Golden rice" is the best known transgenic crop that was engineered to achieve high level of β-carotene with 0.8μg/g β-carotene in "Golden rice 1" and 35μg/g β-carotene in "Golden rice 2" of dry rice seeds (Ye et al., 2000; Al-Babili and Beyer, 2005; Tang et al.,

2009). While two distinct cassettes containing genes encoding daffodil *phytoene synthase* (*psy*) and *Erwinia uredovora phytoene desaturase* (*crtI*) were introduced to rice in "Golden rice 1" project; maize *phytoene synthase* employed to replace daffodil gene and caused a dramatic increase of β-carotene in seeds in "Golden rice 2". Recently, "New Golden rice" has been reported to increase total carotenoid which was achieved through simultaneously expressing bicistronic gene *psy*-2A-*crt*I (PAC) involving 2A sequence, and 2A expression system was demonstrated to be a hopeful tool for multi-gene expression in crop biotechnology (Ha et al., 2010). Transgenic soybean using the same PAC gene system was also reported with total carotenoid amount of 146 µg/g (62-fold higher than non-transgenic seeds) in which 77% of carotenoids was β-carotene (112µg/g) (Kim et al., 2012). On the other hand, by expression of only bacterial *Erwinia uredovora phytoene synthase* (*crtB*), canola seeds also showed distinct orange color and a total carotenoid increment up to 50-fold increase with α- carotene (400µg/g fresh weight) and β-carotene (700µg/g fresh weight) as predominant carotenoid forms (Shewmaker et al., 1999). Transgenic flax plants showed 7.8 to 18.6-fold carotenoid increase (156.3µg/g fresh weight ) as compared to non-transgenic plant (Fujisawa et al., 2008). Carotenoid enhancement was also reported on various other crops as tobacco, potato, wheat, and maize (Diretto et al., 2007; Aluru et al., 2008; Fujisawa et al., 2008; Cong et al., 2009; Jing et al., 2009).

Several groups reported metabolic engineering of tocopherol biosynthetic pathway in *Arabidopsis* by overexpressing various enzymes (Shintani and DellaPenna, 1998; Savidge et al., 2002; Collakova and DellaPenna, 2003). They showed that overexpression of HPT (*Homogentisate phytyltransferase*) and *At*-VTE4 (*γ-tocopherol*

*methyltransferase*) each increased total tocopherol and α-tocopherol contents, respectively in *Arabidopsis* seeds. While γ-tocopherol is a biosynthetic precursor of α-tocopherol and is predominant form in soybean seeds, it has a low (only 1/10) biological activity of α-tocopherol. Furthermore, α-tocopherol may be the only type of vitamin E that human blood can maintain and transfer to cells when needed (Traber and Sies, 1996). Consequently, manipulating the tocopherol biosynthetic pathway in soybean seeds to convert the less active tocopherols to α-tocopherol could have significant health benefits (Tavva et al., 2007). In 2003, Van Eenennaam et al. reported that by overexpressing *AtVTE3* (2-*methyl-6-phytylbenzoquinol methyl transferase*) or *AtVTE4* (*γ-tocopherol-methyltransferase*) alone or both altered tocopherol compositions but did not significantly alter total level of tocopherols. While ectopic *AtVTE3* increased γ- and α-tocopherol levels and reduced δ- and β-tocopherols; overexpressing *AtVTE4* alone resulted in almost complete γ- tocopherol conversion to α- tocopherol in soybean seeds. Moreover, overexpression of both *AtVTE3* and *AtVTE4* shifted α-tocopherol from only 10% to 90% in soybean seeds (Van Eenennaam et al., 2003).

Although many successes were reported in improvement of carotenoids and tocopherols in various plant species, no research has been done to enhance both carotenoids and tocopherols simultaneously in any plants. Thus, the main objective of this research is to evaluate the capability to increase and alter the carotenoid and tocopherol compositions in soybean seeds by seed specific, coordinated overexpression of key genes in carotenoid and tocopherol pathways using 2A sequence.

# MATERIALS AND METHODS

## Plant material and growth conditions

Soybean seeds of elite genotype "Maverick" (2010) was generous gifts from the Missouri Foundation Seeds, MO and used for all experiments. All soybean plants including wild type (WT) and transgenic plants were grown in greenhouse rooms inside the Sears Plant Growth Facility complex at 26/24$^0$C day/night temperatures with combined supplemental metal halide and high sodium pressure lights. They were grown in 3 gallon pods (Nursery supply nurses, # 149637) filled with mix soil and fertilizer (Osmocote®14-14-14). All soils used were professional growing mix (SunGro, USA). Plants were watered every day for 4 months and one additional month without watering for drying and harvesting the dry seeds. Especially, regenerated plants from soybean transformation were first grown in Jiffy pot (Jiffy, cat. # 141451) containing moistened mix soil at 24$^0$C under 18:6 (light: dark) photoperiod for at least one week, and watered as needed. When plantlets have at least three healthy trifoliates were scored for Liberty® resistance prior to potting in 3 gallon pods grown in green house

## Binary constructs and bacterial strain

An entire expression cassette for a cooridinated expression of polyproteins was synthesized (GenScript, USA). This single cassette included a soybean *Glycinin* gene promoter (seed-specific) (Flores et al., 2008), the full-length cDNAs of *Erwinia uredovora phytoene synthase* (*crtB*), *Arabidopsis thaliana* 2-*methyl-6-phytylbenzoquinol methyl transferase* (*At-VTE3*), *Glycine max* γ-*tocopherol-methyltransferase* (*VTE4*), and soybean vegetative storage protein gene terminator (*Tvsp*) (Rhee and Staswick, 1992) to

generate a single transcriptional unit. The nucleotide sequences encoding *crtB* as well as *At-VTE3* and *VTE4* were from GenBank NCBI (Accession No D90087, NM_116206 and NM_001249796, respectivly). Each of these genes was fused at 5' end with soybean *Rubisco* small subunit transit peptide (Acession VOO458 J01307) and the three cDNAs were separated by 2 copies of 2A sequence (Ryan and Drew, 1994; Halpin et al., 1999) to create a single open reading frame encoding the polyprotein (crtB-2A-AtVTE3-2A-VTE4) (Fig. 2A). Two *Eco*RI sites were created to flank this cassette and the resultant sequence was synthesized and cloned into the plasmid pUC57. The entire cassette flanked by the *Eco*RI sites was consequently then excised from pUC57 and cloned into binary plant transformation vector pZY101 (Zeng et al., 2004).The resulting construct was mobilized into *Agrobacterium tumefaciens* strain AGL1 and the integrity of the binary vector within the *Agrobacterium* cells was confirmed by plasmid rescue and restriction digests.

**Agrobacterium mediated transformation of soybean cotyledonary-node**

The soybean transformation procedure followed our lab protocol with minor modifications from previous ones (Zhang et al., 1999; Zeng et al., 2004; Wright et al., 2010). These modifications included the deploying "dip-wounding" during the explant preparation (Barampuram and Zhang, 2011) as well as using glufosinate concentrations at 0, 10, and 4mg/L during first and second shoot induction and shoot enlongation stages, respectively.

**Transgene integration and segregation analysis**

All regenerated plants recovered under the herbicide glufosinate selection were subjected to leaf-painting assay with 200mg/L Liberty® (Zhang et al., 1999) three times

to confirm the presence of transgene. Leaf tissue that did not exhibit necrosis at 5 days post application was scored as herbicide tolerance.

All leaf-painting tolerant $T_0$ plants were analyzed by polymerase chain reaction (PCR) for a quick screen of the presence of transgenes. To do this, genomic DNA samples from young healthy leaf were extracted using REDExtract-N-Amp$^{TM}$ Plant PCR kits (Sigma-Aldrich). Sequences and positions of primers using in PCR were described in Table 2.1A and Figure 2.1A. The primer pairs for *bar* gene were forward primer (*bar*-F) and reverse primer (*bar*-R) to amplify *bar* gene partial sequence yielding DNA fragment of 268bp. To detect gene of interest (GOI), fragment from 5'end of *crtB* and 3'end of *VTE4* (*crtB-2A-AtVTE3-2A-VTE4*), 2 primer sets BE3-F/BE3-R and E3E4-F/ E3E4-R were used. The BE3-F/BE3-R primer pair amplified the region from 5'end of *crtB* gene to 3'end of *VTE3* gene yielding 2000bp product whereas E3E4-F/ E3E4-R primer pairs amplified the region from 5' end of *VTE3* gene to 3' end of *VTE4* yielding 2000bp product. These two primer pairs were tested on putative transgenic plants for integration of GOI from 5' end of *crtB* gene and 3'end of *VTE4* that overlap *VTE3* region. The PCR reaction conditions were as the follow: for GOI: $94^0$C for 3 min hot start, 35 cycles of $94^0$C for 30 sec denature, $58^0$C for 30 sec annealing, $72^0$C for 2 min extension, and $72^0$C for 10 min final extension. For *bar* gene, all conditions were the same as for GOI except that the each cycle extension time was reduced to 30 sec. The amplified DNA products were then run on 0.8% agarose gel for size analysis. Progeny segregation analysis was also performed on $T_1$ plants from seeds of putative transgenic $T_0$ events by leaf-painting assay and PCR analysis for screening the presence of *bar* gene and GOI as $T_0$ analysis.

**Quantitative real time PCR (qRT-PCR)**

Immature seeds of a transformed plant were collected and immediately frozen in -80$^0$C freezer before use. Then, each frozen seed was ground in liquid nitrogen, and total RNA was extracted using Trizol reagent (Invitrogen) following manufacturer's instructions. DNase treated RNA samples were prepared using "DNase I, RNase-free" (Thermo Scientific, USA). First strand cDNA was synthesized using iScript$^{TM}$ Reverse Transcription Supermix for RT-PCR (Bio-Rad). qRT-PCR was performed in 96-well plates with CFX-96$^{TM}$ Real-Time system (Bio-Rad, USA). Each reaction contained 20ng of cDNA, 10µM of each primer, and 10µl SsoFast$^{TM}$ Evagreen Supermix (Bio-Rad) and total reaction volume was adjusted to 20 µl with DNase, RNase–free water (Qiagen). The PCR conditions were 95$^0$C for 3 min, 35 cycles of 95$^0$C for 10 sec denaturation, 55$^0$C for 20 sec annealing and 72$^0$C for 20 sec extension, followed 10 sec at 95$^0$C; a melting curve was generated by increasing temperature from 65$^0$C to 95$^0$C to examine amplification specificity. Three primer pairs were used to detect transcript levels of three different genes in the cassette crtB-2A-AtVTE3-2A-VTE4: *crtB* gene (primer *crtB*-F/*crtB*-R), *At-VTE3* gene (primer *VTE3*-F/*VTE3*-R), and *VTE4* gene (*VTE4*-F/*VTE4*-R). The reference gene used in this study was *cons7*, one of the genes indicated as having the most stable and consistent expression level across different soybean tissues including leaf and seed (Libault et al ., 2008). Sequences of these primers were described in Table 2.1B.

**Carotenoid extraction and HPLC (High performance liquid chromatography) analysis**

Carotenoid extraction (Kim et al., 2012) and HPLC analysis (Kean et al., 2008) were used to examine the amount of carotenoid of individual seeds from the T1

generation of regenerated plants from soybean transformation. Seeds used for analysis were: six replications from transformed seeds using the empty binary vector pZY101, three replications from the binary vector containing expression cassette, and three replications from wild type Maverick seeds. To avoid carotenoid degradation, entire carotenoid extraction as well as carotenoid analysis steps were done in specific dim light condition as far from strong direct light as possible (Kean et al., 2008). Briefly, single dried seed was ground into a fine powder by liquid nitrogen in mortar and pestle (0.1g), mixed with 3ml of ethanol containing 0.1% ascorbic acid (w/v), vortexed for 20 seconds and placed in water bath at $85^0$C for 5 min. Potassium hydroxide (120µl, 80% w/v) was added in the carotenoid extract to saponify in the $85^0$C water bath for 10min and then placed immediately on ice after 1.5ml cold deionized water was added. Carotenoids were extracted twice with 1.5ml hexane followed by centrifugation at 1,200g. Collected aliquots of the extracts were dried under a stream of nitrogen. Dried carotenoid extracted samples were then subjected to HPLC analysis. Carotenoids were re-dissolved in 50:50 (v/v) dichloromethane/methanol and separated on C30 YMC column (Waters Corp., Milford) with a Hewlett-Packard 1090A HPLC equipped with a model 79880A diode array detector. Chromatograms were generated at 450nm with standards and solvents as described by Kean et al. (Kean et al., 2008).

**Tocopherol extraction and HPLC analysis**

Tocopherol extraction and quantification of soybean seeds were performed based on the procedure previously described (Dwiyanti et al., 2011) with minor modifications (Kristin Bilyeu, unpublished). Tocopherols were analyzed from single dried mature seed. Each seed was grounded using liquid nitrogen and 50 mg of seed powder was transferred

into a glass tube (VWR #89003-568) with PTFE screw cap (VWR #89003-568) and then mixed with 1ml extraction solvent [80% EtOH with 1.5µg/ml tocol internal standard (Matreya, #1797)]. The samples were sonicated (Brasonic® ultrasonic cleaner, Sigma-Aldrich) at room temperature for 15 minutes. After the sonication, the samples were incubated at $4^0$C for 30min and then centrifuged at 2,500 rpm for 10min. 0.5µl of the supernatant was transferred into HPLC vial (Agilent #5182-0716). The analysis was performed in Agilent C18 column with a consistent temperature at $40^0$C. The assay was done in a mobile phase of Acetonitrile with 90% methanol(v/v), flow rate 1ml/min, and injection volume 25µl. Tocopherols were detected by fluorescence with excitation at 292nm and emission at 330nm (Yang et al., 2011).

**Fatty acid analysis**

Oil extraction and gas chromatography for fatty acid methyl esters followed the previous study (Beuselinck et al., 2006) with minor modifications. Single seed was bashed in an envelope by hammer. Crushed sample was extracted overnight in 1ml of chloroform-hexane-methanol (8:5:2, v/v/v). Then, 150µl extracted solvent was transferred to 1.5ml vial and mixed with 75µl methylating reagent (0.5 M methanolic sodium methoxide–petroleum ether–ethyl ether, 1:5:2, v/v/v) then diluted in hexane to 1ml. An Agilent (Palo Alto, CA) series 6890 capillary gas chromatography was used with a flame ionization detector (2758C), and AT-Silar capillary column (Alltech Associates, Deerfield, IL). Standard fatty acid mixtures (Animal and Vegetable Oil ReferenceMixture 6, AOACS) were used as controls.

# RESULTS

## Transgene integration and expression analysis of soybean events

A total of 4 putative plants were regenerated from transformation experiments using binary construct carrying *crtB-2A-AtVTE3-2A-VTE4* (GOI) and 3 independent empty vector control events (pZY101-1, pZY101-2, and pZY101-3) were developed from a separate transformation experiment (600 explants in total). Of these plants, 3 of 4 putative GOI and all pZY101 plants were confirmed to carry the transgene *bar* as indicted by three times leaf-painting tolerance. These events were further anlyzed by PCR using genomic DNA extracts and primer pairs specific to *bar* gene and GOI, separately. Our first round PCR screen was conducted on the *bar* gene with primer pair *bar*-F/*bar*-R, and second round screen was on GOI (*crtB-2A-AtVTE3-2A-VTE4)* with primer pairs BE3-F/BER-R and E3E4-F/E3E4-R. PCR results showed that 3 out of 4 events contained the *bar* gene (HYX-7-1, HYX-7-3, and HYX-7-4 (Figure 2.1B), which was consistent with leaf-painting result. The PCR further confirmed that all three events each carried a complete GOI cassette (*crtB-2A-AtVTE3-2A-VTE4*) since the two sets of primers (BE3-F/BE3-R and E3E4-F/ E3E4-R) together covering the entire cassette yielded expected bands (Figure 2.1C).

## Molecular analysis of T1 generation

Seedlings of T1 plants from three $T_0$ events (HYX-7-1, HYX-7-3, and HYX-7-4) were grown in greenhouse and were subjected to leaf-painting and PCR analysis to confirm transgene inheritance and segregation by confirming the presence of *bar* gene and GOI. Leaf-painting and PCR results agreed, showing that *bar* gene was present only in T1 plants of HYX-7-1 event whereas none of the other two events showed Liberty

herbicide tolerant or presence of the *bar* (Figure 2.1D). The presence of GOI in HYX-7-1 progeny plants but not in HYX-7-3 and HYX-7-4 progeny was also verified (Figure 2.1E). These results were summarized in Table 2.2B. Thus, only golden soybean HYX-7-1 event was further analyzed molecularly and biochemically.

**Seed phenotype of transformed soybean plants**

Of the three GOI events, one, i.e., HYX7-1, displayed "Golden soybean" seed phenotype as shown by orange color in the seed, indicating a drastic increase in carotenoid content (Figure 2.2A and B). Therefore, we expected at least *crtB* encoding *phytoene synthase* was expressed in the "Golden soybean" seeds. Because these seeds were derived from the first generation and therefore were still segregating, HYX-7-1 plant displayed a mixture of transgenic golden and null seeds (golden: null, 3:1 ratio). Observation of transgenic golden soybean seeds of HYX-7-1G showed different intensity of color from medium orange to dark orange, suggesting that there might be different carotenoid contents in individual seeds. Quite surprisingly, however, seeds of other regenerated plants (HYX-7-3; HYX-7-4) showed null phenotype that was similar to wild type (Maverick) and pZY101 (Table 2.2A). Seedlings of $T_1$ plants from these $T_0$ events (HYX-7-1, HYX-7-3, and HYX-7-4) were grown. However, the "golden seed" event HYX-7-1 showed many seedless pods that indicated a decrease in seed yield for next generation of golden seeds (Figure 2.2C)

**Quantitative real time PCR (qRT-PCR) to determine transgene expression level**

To analyze the expression levels of the three transgenes in crtB-2A-AtVTE3-2A-VTE4 construct, qRT-PCR analyses were performed using total RNAs extracted from single mid-mature $T_1$ seed of transgenic soybean event HYX-7-1 (golden seeds HYX-7-

1G) along with single mid-mature seed from wild type Maverick (WT) and empty vector pZY101 transgenic plants. We collected and analyzed soybean mid-mature seeds because the glycinin promoter driving our transgene cassette has peak activity during this seed developmental stage (Flores et al., 2008). Because HYX-7-1 transgenic event displayed distinct "golden" seed phenotype, transgenic seeds were readily chosen by visual inspection. Each individual golden seed was then ground in liquid nitrogen and its total RNA was analyzed in qRT-PCR. The result showed that HYX-7-1 expressed all three genes strongly. Conversely, as we expected, WT and pZY101 controls expressed endogenous soybean VTE4 gene homolog at 10-fold and 3-fold lower level than golden seeds, respectively (Figure 2.3A). This result also suggested that soybean seeds expressed *VTE4* even though the expression level is relatively low and there hasn't been previous study on the presence and expression of this gene.

To confirm the above results, we examined more golden as well as WT and pZY101 control seeds, showing the same results. Notably, although individual golden seeds (4 seeds) of HYX-7-1 expressed all three genes strongly, they also showed variation in expression level of these genes among individual seeds (Figure 2.3B).

## Carotenoid and tocopherol contents in transgenic soybean seeds

### *Carotenoid analysis by HPLC*

The drastic change and different intensities in seed color and much higher levels of transgene expression in golden soybean seeds than controls prompted us to conduct seed carotenoid analysis using HPLC. These analyses were performed to determine the total amount and compositions of carotenoids of individual dry seeds from each plant of HYX-7-1 (3 replications), WT (3 replications) and pZY101 controls (6 replications) with all

samples being measured as fresh weight. Several peaks in the HPLC chromatogram of transgenic plant HYX-7-1 indicated that carotene contents had changed drastically compared to WT, and most of these changes are in β-carotene components (Figure 2.4A). The different carotenoids measured by HPLC included 9 or 9'-cis-lutein, all-trans-lutein, zeaxanthin, 13- or 13'-cis-lutein, α-cryptoxanthin, β-crytpoxanthin, 15-cis-β-carotene, 13-cis-β-carotene, α-carotene, all-trans-β-carotene, 9-cis-β-carotene. Golden seeds of HYX-7-1 showed an increase in total carotenoid (sum of all carotenoid measured) as well as pro-vitamin A (sum of β-crytpoxanthin, 15-cis-β-carotene, 13-cis-β-carotene, α-carotene, all-trans-β-carotene, 9-cis-β-carotene) compared to WT and pZY101 (Figure 2.4B and Table 2.3A). HYX-7-1 event produced very high levels of total carotenoid (highest amount of 128 μg/g), which was 25-fold higher than in WT (highest amount of 5.03μg/g) and 45-fold higher than in empty vector pZY101 seeds (highest amount of 2.85 μg/g). The pro-vitamin A content of golden seeds of HYX-7-1 also significantly increased, accounting for more than 98% of total carotenoid, as compared with 2-3% in WT or 3-16% in pZY101. Hence, the use of bacterial *phytoene synthase crtB* in the 2A-containing construct driven by G*lycinin* promoter enabled accumulation of high β-carotenoids in soybean seeds. This result also suggested that the α-branch in the soybean seed was almost completely changed to the β-branch through the ectopic *crtB*. Surprisingly, there was a large variation in carotenoid amount among individual seeds of same event, ranging from 23.54μg/g -128.20μg/g fresh weight in HYX-7-1, from 1.8μg/g - 5μg/g in WT, and 0.4μg/g - 2.8μg/g in pZY101 controls (Figure 2.4B Table 2.3A). This result indicated a possible correlation between the levels of *crtB* gene and various intensities of the golden color in distinct seeds which could stem from different carotenoid contents.

Notably, among individual components of the carotenoids in WT, lutein (95% - 98% of total carotenoid) was found to be far more abundant than the other carotenoids. By contrast, the lutein together with zeaxanthin decreased in golden seeds coupled with a dramatic increase in all-trans-β carotene, 13-cis-β carotene, α-carotene, 15-cis-β carotene, 9-cis-β carotene, and β-cryptoxanthin (pro-vitamin A activity).

**Tocopherol analysis by HPLC**

Next, we wanted to know how much the coordinated expression of the two genes encoding for *2methyl-6-phytylbenzoquinol methyltransferase* (*At-VTE3*) and *γ-tocopherolmethyltransferase* (*VTE4*) affected tocopherol content in transgenic seeds. To accomplish this, dry individual seeds in six replications of each HYX-7-1 (golden soybean), pZY101, and three replications of WT were analyzed by HPLC. The result showed that there was a decrease in total tocopherol in golden soybean seeds of HYX-7-1 (ranging from 0.106μg/mg-0.173μg/mg) compared to WT (ranging from 0.176μg/mg-0.229μg/mg) and pZY101 controls (0.176 μg/mg-0.221 μg/mg). More specifically, the golden soybean seeds showed a decrease in γ-tocopherol (0.07μg/mg-0.12μg/mg) and δ-tocopherol (0.03μg/mg), but no significant change, even though some golden seeds showed increase in α-tocopherol (0.006μg/mg-0.023μg/mg) compared to WT (with γ-tocopherol, δ-tocopherol, and α-tocopherol being 0.12-0.14μg/mg, 0.05-0.06μg/mg, and 0.006-0.011μg/mg, respectively) (Figure 2.4C and Table 2.3B). Meanwhile, the golden seeds generally showed decrease in all tocopherol compositions, when compared to pZY101 controls (γ-tocopherol, δ-tocopherol, and α-tocopherol were 0.09-0.19μg/mg, 0.03-0.05μg/mg, and 0.015-0.023μg/mg, respectively). Of note, pZY101 showed

decrease in δ-tocopherol but increase in α-tocopherol, and no change in γ- tocopherol compared to WT.

**Fatty acid analysis**

It was tempting for us to find out if coordinated expression of the three proteins from the crtB-2A-AtVTE3-2A-VTE4 and the changes in tocopherol and carotenoid could cause any change in another components in the transgenic seeds. Although tocopherols are considered to contribute both the nutritional value and oxidative stability of soybean oil, how the changes in tocopherols correlate to the content and compositions of fatty acids in soybean remains unknown. To investigate this, three single seeds of each HYX-7-1 (golden seeds), pZY101, and WT were analyzed the profiles of fatty acids by gas chromatography (GC). The result showed that there was a change in fatty acid profile in golden soybean seeds with decrease in oleic acid, linoleic acids, and palmitic acid, but increase in linolenic acid (2-2.5 fold higher) compared to WT and pZY101 (Figure 2.5 and Table 2.4).

## DISCUSSION

We observed an increase in carotenoids and a decrease in tocopherol (Table 2.3A and B) in transgenic golden soybean seeds. In our study, the seed specific overexpression of *Erwinia uredovora psy* (*crtB*) led to much increased accumulation of carotenoids mainly in β-carotene and β-carotene equivalents that increased pro-vitamin A content to more than 98%. It is important to generate high pro-vitamin A soybean to combat VAD in developing countries using metabolic engineering approaches. Furthermore, this result confirmed that the source of *phytoene synthase* (*psy*) gene has significant impact on the level of carotenoid accumulation. This gene was overexpressed before in "Golden Rice

1" and "Golden Rice2" (Paine et al., 2005). Our present result further indicated that overexpressing only one bacterial gene *phytoene synthase* could increase high amount of β-carotene (pro-vitamin A). The same strategy of using only *psy* and similar high levels of pro-vitamin A accumulation were implicated in transgenic canola (Shewmaker et al., 1999) and flax plants (Fujisawa et al., 2008). On the other hand, two or more transgenes in carotenoid pathway were introduced to other plant species to increase β-carotene (Shewmaker et al., 1999; Ye et al., 2000; Diretto et al., 2007; Aluru et al., 2008; Fujisawa et al., 2008; Cong et al., 2009; Jing et al., 2009; Ha et al., 2010; Kim et al., 2012). We also noticed that $T_1$ seeds of golden soybean event HYX-7-1 showed only 50% germination rate with one week delay as compared to $T_1$ seeds of WT and pZY101 controls. Delayed germination in transgenic seeds that were genetically engineered to increase β-carotene has been previously reported in *Arabidopsis* (Lindgren et al., 2003). Because β-carotene are also precursor of abscissic acid (ABA), higher amount of ABA resulting from higher β-carotene could inhibit seed germination by inhibiting seed uptake water and affecting a transition from germination to post-germination growth (Manz et al., 2005). At the first three week growth, $T_1$ seedling plants of golden soybean (HYX-7-1) showed dwarf phenotype with small and wrinkled leaves but after three weeks, there was not any significant phenotypic difference compared to the other $T_1$ plants (data not shown). Dwarf phenotype was also described in transgenic tobacco and tomato plants expressing *phytoene synthase* (Fray et al., 1995; Jing et al., 2009). These above phenomena implied a likely competition for GGDP (geranylgeranyl diphosphate) between gibberellic acid (GA) and carotenoid pathways. The transgenic plants showed the negative correlation of carotenoid level and germination, positive correlation of

carotenoid level and ABA was also discussed by Lindgren (Lindgren et al., 2003). Therefore, overexpressing *phytoene synthase* enabled to evaluate the importance of regulating the flux through the isoprenoid pathway for normal hormone-controlled growth. It is also a probable explanation for tocopherol reduction in golden soybean seed that the source of GGDP for tocopherol as well as GA, chlorophyll, phylloquinone is limited, since this is shifted toward carotenoid pathway by *phytoene synthase.* When $T_1$ golden soybean plants (HYX7-1) developed pods, many pods were seedless leading to a significant yield loss in $T_2$ seeds (Fig 2.2C).

Although all three genes *crtB*, *At-VTE3*, *VTE4* were integrated and expressed in $T_1$ seeds from the transgenic plant (HYX-7-1), only carotenoid showed significant increased whereas tocopherol decreased. This result is consistent with the transgenic canola seeds overexpressing *crtB* that increased total carotenoids predominantly as α-, β-carotenoid but reduced tocopherols (Shewmaker et al., 1999). We hypothesize that the decrease of tocopherol in golden seeds is a competition of the same GGDP pool between two pathways of carotenoids and tocopherols. Moreover, in the previous report, Collakova and DellaPenna (2003) mentioned that *homogentisate phytyltransferase* catalyzes the committed step in tocopherol biosynthesis, and overexpressing of this gene resulted in total tocopherol increase 40% higher than wildtype in *Arabidopsis* seeds. Consistent with this, co-expression of three key plastid-localized enzyme genes, i.e., *Arabidopsis homogentisate phytyltransferase AtVTE2*, *AtVTE3,* and *AtVTE4* in tocopherol pathway along with *crtB* in carotenoid pathway could not only shift biosynthesis toward the accumulation of more nutritious α-tocopherol and β-carotene, but also increase total contents of both tocopherols and carotenoids. Moreover, the order of

these genes should be rearranged in the 2A construct so that three genes encoding for enzymes in tocopherol biosynthesis (*AtVTE2, AtVTE3,* and *AtVTE4*) would be placed upstream of *crtB*. This would enhance the chance of a higher protein level of the *AtVTE2, AtVTE3,* and *AtVTE4*, if a reduced protein translation of the promoter-distal open-reading frame(s) does occur.

Several previous studies showed a correlation between fatty acid and tocopherol contents. They indicated that changes in the fatty acids of soybean oil can alter its tocopherol content and compositions (McCord et al., 2004; Scherder et al., 2006; Baumgartner et al., 2010). In our study, golden soybean seeds HYX-7-1 showed a significant change in the fatty acid compositions including increase in stearic acid and especially, a remarkable increase in linolenic acid, but a decrease in oleic acid, linoleic acid and palmitic acid as well as all tocopherol compositions. This result was contrast to previous observations. For example, McCord and Wang illustrated that the reduced linolenic contents is correlated with decreased total tocopherol, including particularly α- and γ-tocopherol (McCord et al., 2004; Scherder et al., 2006). The disagreement between these previous reports and our study is still unclear. By contrast, reduced palmitic acid content was shown to correlate to an average increase in total tocopherol, but average decrease in seed yield. Results from our present study agreed with Mounts (Mounts et al., 1996) who genetically modified soybean with reduced palmitic acid and stearic acid as well as less total tocopherol. Finally, both Salvin (Slavin et al., 2009) and our study implied that a tocopherol content change in soybeans could potentially altered fatty acid profile and vice versa.

Lastly, we want to emphasize that we were unable to analyze the protein levels of the three transgenes as well as the cleavage-efficiency of 2A in construct (crtB-2A-AtVTE3-2A-VTE4) due to the unavailability of antibodies against these individual proteins. Clearly, it is important to estimate how the cleavage efficiency would affect tocopherol and carotenoid contents in transgenic soybean seeds and therefore deserves future study.

**Figure legend**

**Figure 2.1A:** Schematic diagram of cassette (2A-construct) for coordinated expression of open-reading frames (ORFs without start and stop codon) of three genes to increase carotenoid and tocopherol contents. The three genes are *crtB* (*phytoene synthase*), *At-VTE3* (*Arabidopsis thaliana* 2-*methyl-6-phytylbenzoquinol methyl transferase*), and *VTE4* (*Glycine max γ-tocopherol-methyltransferase*). Each ORF is fused with TP (soybean *Rubisco* small subunit transit peptide) at 5' end and seperated from adjacent ORFs by FMDV 2A sequence. The cassette was driven by *Glycinin* gene promoter and Tvsp terminator to create a single open reading frame (with one start and stop codon) encoding polyprotein (crtB-2A-AtVTE3-2A-VTE4). Restriction site of *Eco*RI was created to flank the construct which was carried within cloning plasmid pUC57. The arrows show the location of designed primer sets used in PCR analysis (1) BE3-F, (1') BE3-R, (2) E3E4-F, (2') E3E4-R.

**Figure 2.1B:** PCR analysis of *bar* gene from $T_0$ plants using primer set of *bar*-F/*bar*-R with empty vector pZY101 as positive control and Maverick as negative control. The 100bp DNA size markers were from NewEngland Biolab.

**Figure 2.1C:** PCR analysis of gene of interest (GOI) from $T_0$ plants using primer set of BE3-F/BE3-R (first six lanes from left marker) and E3E4-F/E3E4-R (last six lanes from the middle marker). The 1kb DNA size markers were from Thermo Scientific.

**Figure 2.1D:** Representative PCR analysis of *bar* gene from $T_1$ soybean lines using primer set of *bar* gene (*bar*-F/*bar*-R) with $T_1$ leave of HYX-7-1 event (HYX-7-1-1) as

positive control and Maverick (wild type) as negative control. The 100bp DNA size markers were from NewEngland Biolab.

**Figure 2.1E:** PCR analysis of gene of interest (GOI) from $T_1$ plants using primer set of BE3-F/BE3-R (first eight lanes from left marker) and E3E4-F/E3E4-R (last eight lanes from the right edge of gel) with HYX-7-1 event (HYX-7-1-1) as positive and Maverick (wild type) as negative controls. The 1kb DNA size markers were from Thermo Scientific.

**Figure 2.2:** Seed phenotypes of HYX-7-1 and Maverick plants.

**A:** Maverick (WT) and HYX-7-1 transgenic soybean seeds at mid-mature stage. (a) Seed pods. (b) Cross-sections of seeds.

**B:** WT and HYX-7-1 transgenic soybean seeds at different development stage. (a) mid-mature stage of cross-section seeds. (b) full-mature stage of whole seeds.

**C:** Pod phenotype of $T_1$ plants (a) WT, (b) transgenic HYX-7-1(HYX-7-1-11). The arrow indicates seedless-pods of transgenic HYX-7-1-11 plant.

**Figure 2.3:** Quantitative real-time PCR (qRT-PCR).

**A:** qRT-PCR of mid-mature seeds. qRT-PCR was carried out with RNA extraction from single mid-mature seeds of transgenic HYX-7-1G (golden seed), empty vector pZY101 and WT (Maverick) controls using three target genes of *crtB, At-VTE3, VTE4.* The *cons7* gene was used as the normalization control. WT and pZY101 were not tested for target genes *crtB* and *At-VTE3*, and their value was set as zero.

**B**: qRT-PCR of individual golden seeds. qRT-PCR was carried out with RNA extraction from immature seeds of transgenic HYX-7-1 (four individual golden seeds, HYX-7-1G), wild type Maverick (two individual seeds), and empty vector pZY101 (one seed) using three target genes of *crtB, At-VTE3, VTE4.* The *cons7* gene was used as the normalization control. Maverick and pZY101 have not been tested for target genes *crtB* and *At-VTE3*, and their value was set as zero.

**Figure 2.4:** Biochemical analysis of seed compositions.

**A:** Representative HPLC chromatograms of individual seeds of HYX-7-1 and wild type Maverick. Chromatograms were generated at 450nm. The peaks are a) 9 or 9'-cis-lutein; b) 9' or 9-cis-lutein; c) all-trans-lutein; d) zeaxanthin; e) 13-or 13'-cis-lutein; f) β-apo-8'-carotenal; g) α-cryptoxanthin; h) β- cryptoxanthin; i) 15-cis- β-carotene; j) 13-cis- β-carotene; k) α-carotene; l) all-trans- β-carotene; and m) 9-cis- β-carotene.

**B**: Total carotenoid and pro-vitamin A content of individual seeds of WT (Maverick) (3 replications), empty vector pZY101 (6 replications), and golden seeds of transgenic HYX-7-1 (3 replications) using HPLC analysis. Total carotenoid levels were calculated as the sum of twelve carotenoid subtype levels i.e. 9 or 9'-cis-lutein, 9' or 9-cis-lutein, all-trans-lutein, zeaxanthin, 13-or 13'-cis-lutein, α-cryptoxanthin, β- cryptoxanthin, 15-cis- β-carotene, 13-cis- β-carotene, α-carotene, all-trans- β-carotene, and 9-cis- β-carotene. The provitamin A includes β- cryptoxanthin, 15-cis- β-carotene, 13-cis- β-carotene, α-carotene, all-trans- β-carotene, and 9-cis- β-carotene.

**C**: α-, μ-, and δ-tocopherol content of individual soybean seeds of wild type Maverick (3 replications), empty vector pZY101 (6 replications), and HYX-7-1 (3 replications) using HPLC analysis.

**Figure 2.5**

Five fatty acid compositions in seeds of WT (Maverick), empty vector pZY101, and HYX-7-1. The measurements shown are from 3 replications of each plant. The values are percentage of each fatty acid components in total fatty acid content (total fatty acid is calculated as sum of five major fatty acids analyzed).

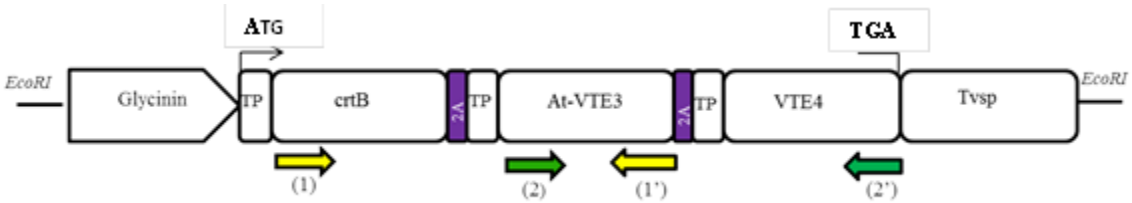

ATG
TGA
EcoRI
Glycinin
TP
crtB
2A
TP
At-VTE3
2A
TP
VTE4
Tvsp
EcoRI
(1)
(2)
(1')
(2')

**Figure 2.1A**

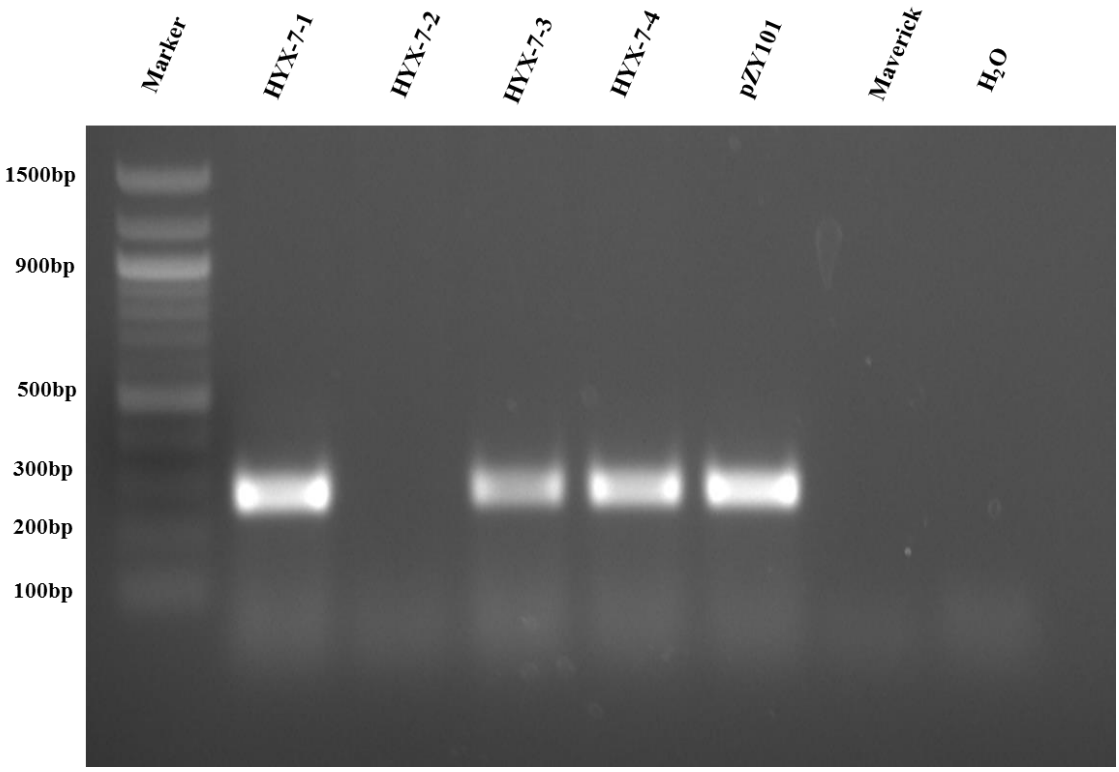

Marker
HYX-7-1
HYX-7-2
HYX-7-3
HYX-7-4
pZY101
Maverick
H2O
1500bp
900bp
500bp
300bp
200bp
100bp

**Figure 2.1B**

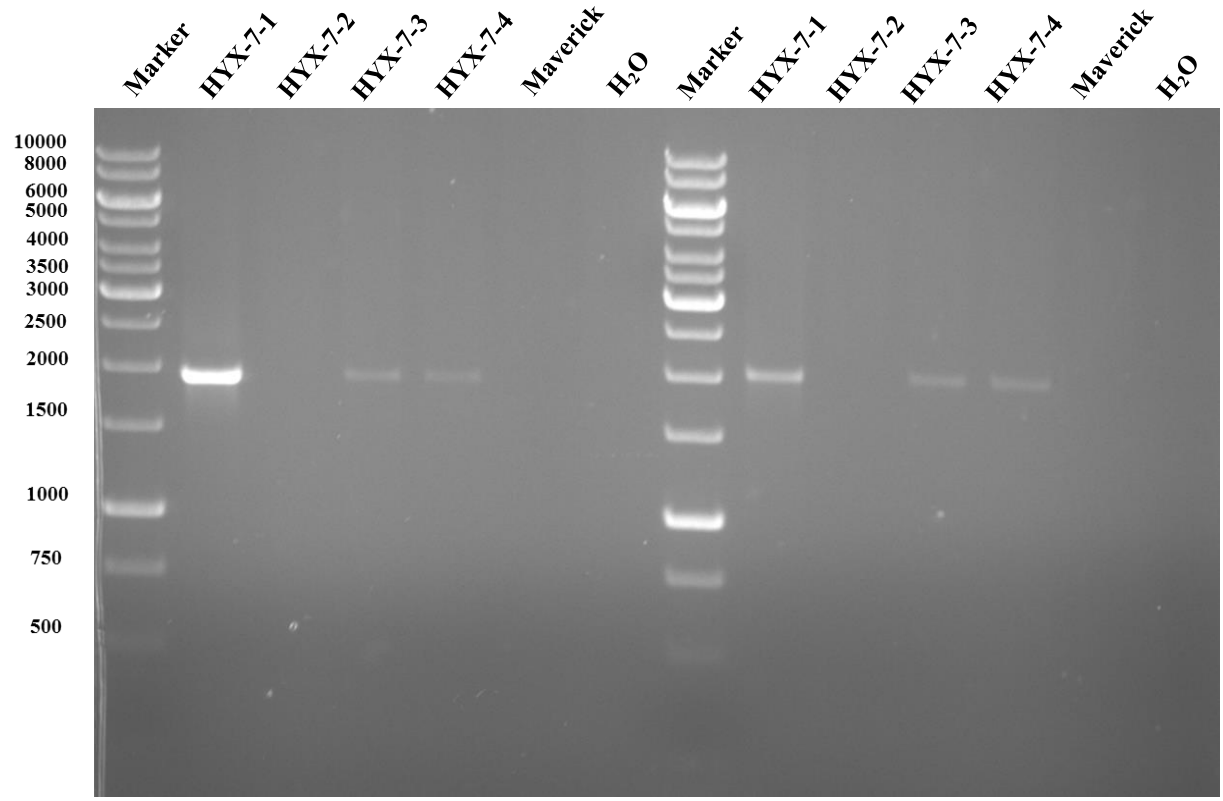

Marker
HYX-7-1
HYX-7-2
HYX-7-3
HYX-7-4
Maverick
H2O
Marker
HYX-7-1
HYX-7-2
HYX-7-3
HYX-7-4
Maverick
H2O
10000
8000
6000
5000
4000
3500
3000
2500
2000
1500
1000
750
500

**Figure 2.1C**

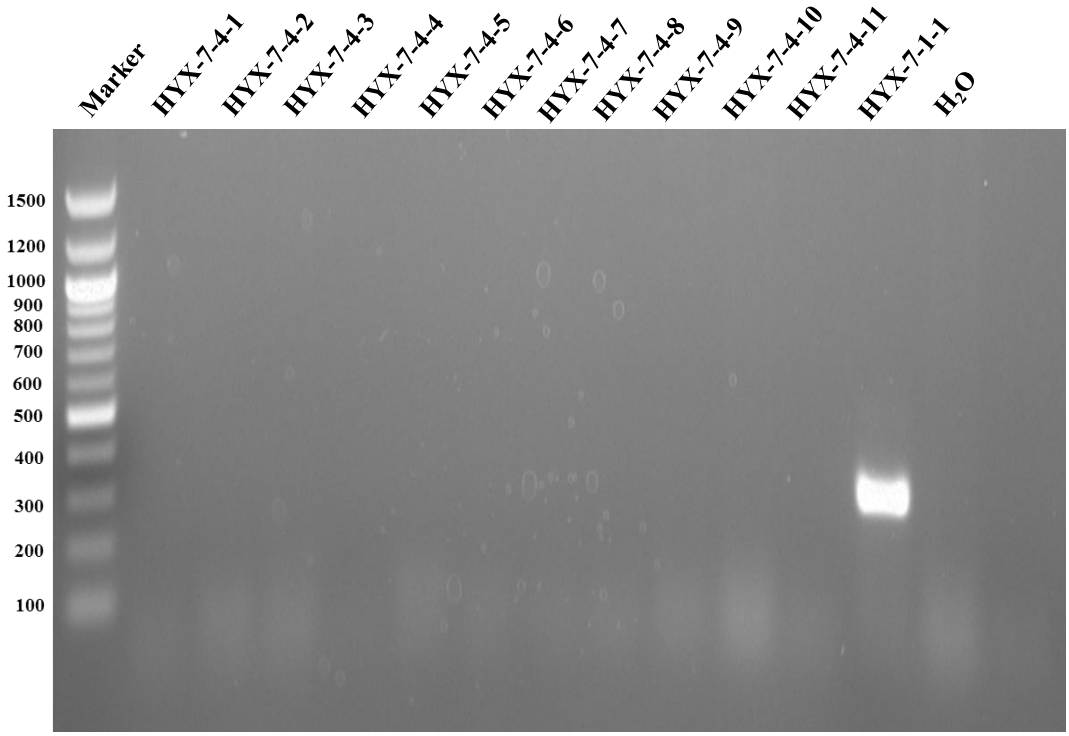

Marker
HYX-7-4-1
HYX-7-4-2
HYX-7-4-3
HYX-7-4-4
HYX-7-4-5
HYX-7-4-6
HYX-7-4-7
HYX-7-4-8
HYX-7-4-9
HYX-7-4-10
HYX-7-4-11
HYX-7-1-1
H2O
1500
1200
1000
900
800
700
600
500
400
300
200
100

**Figure 2.1D**

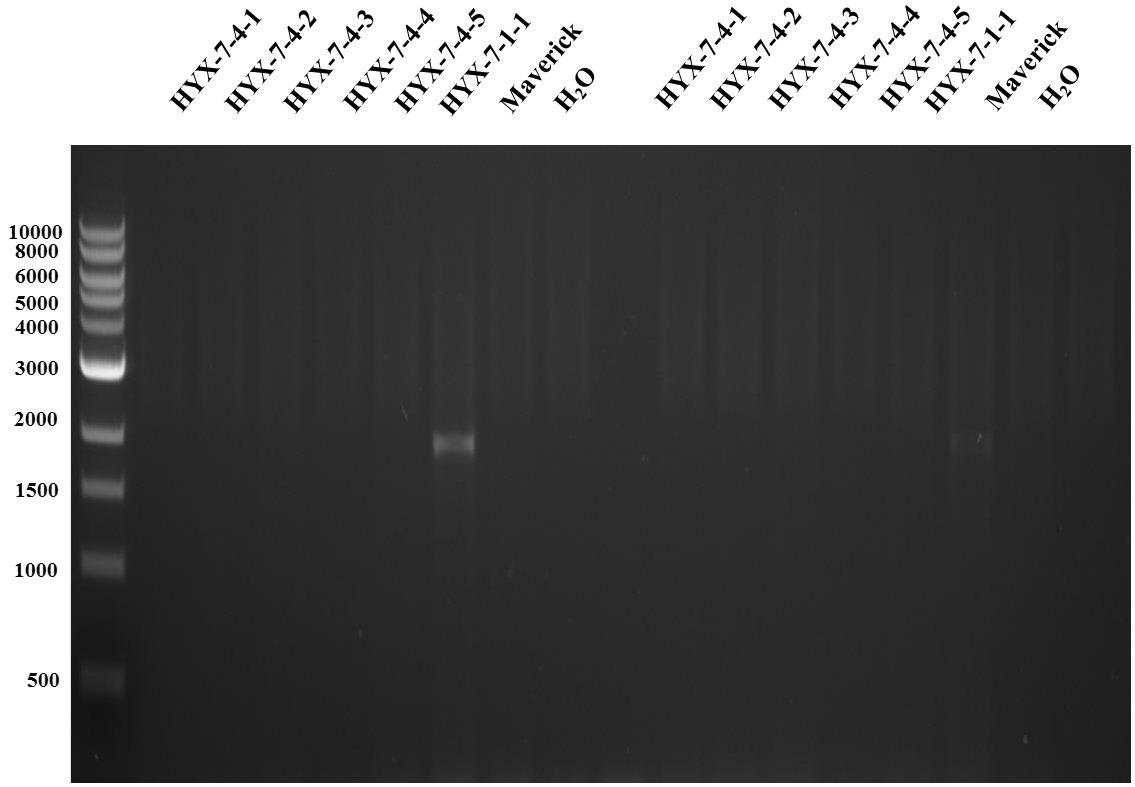

HYX-7-4-1
HYX-7-4-2
HYX-7-4-3
HYX-7-4-4
HYX-7-4-5
HYX-7-1-1
Maverick
H2O
HYX-7-4-1
HYX-7-4-2
HYX-7-4-3
HYX-7-4-4
HYX-7-4-5
HYX-7-1-1
Maverick
H2O
10000
8000
6000
5000
4000
3000
2000
1500
1000
500

**Figure 2.1E**

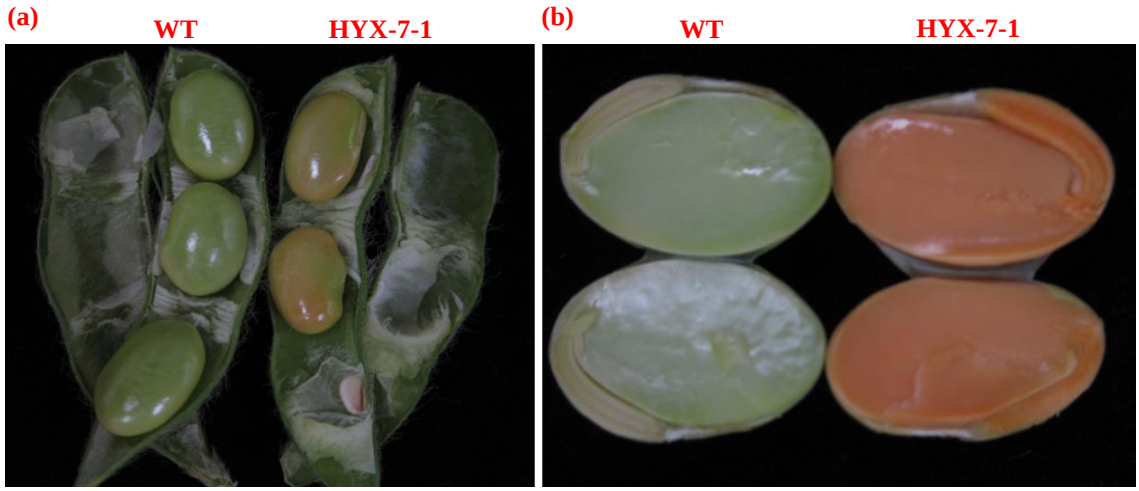



**Figure 2.2A**

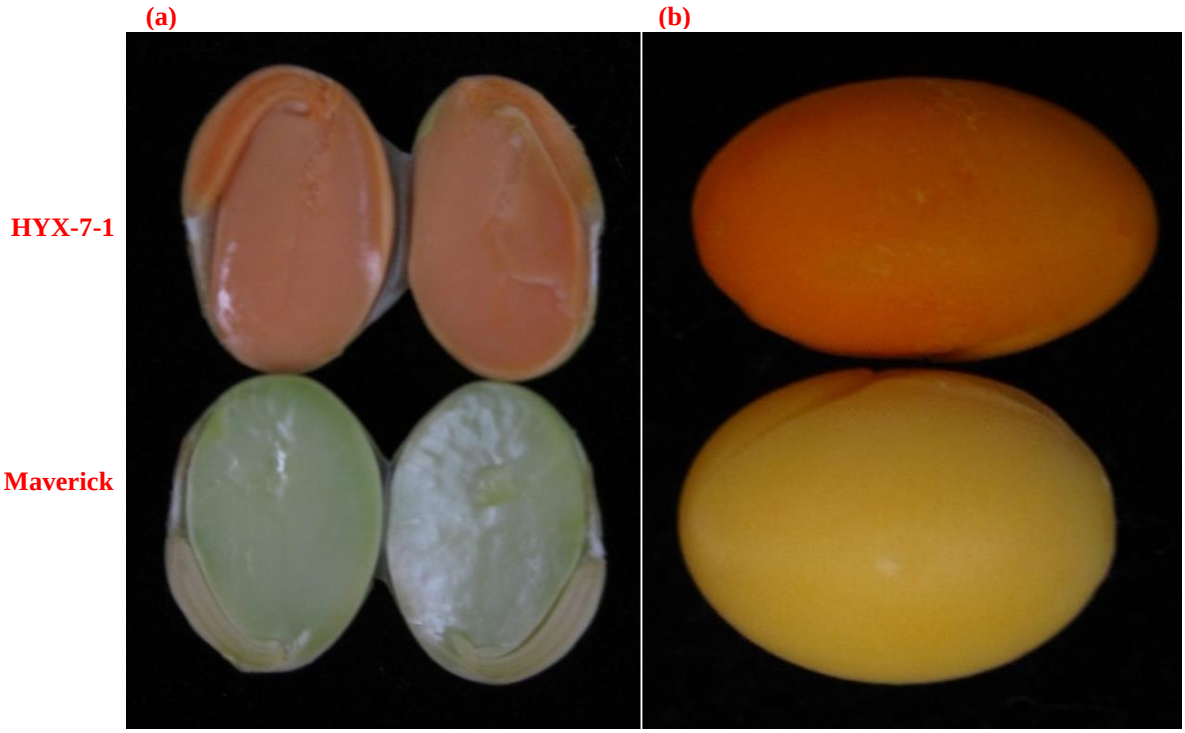



**Figure 2.2B**

**(a) Maverick** **(b) HYX-7-1-11**

**Figure 2.2C**

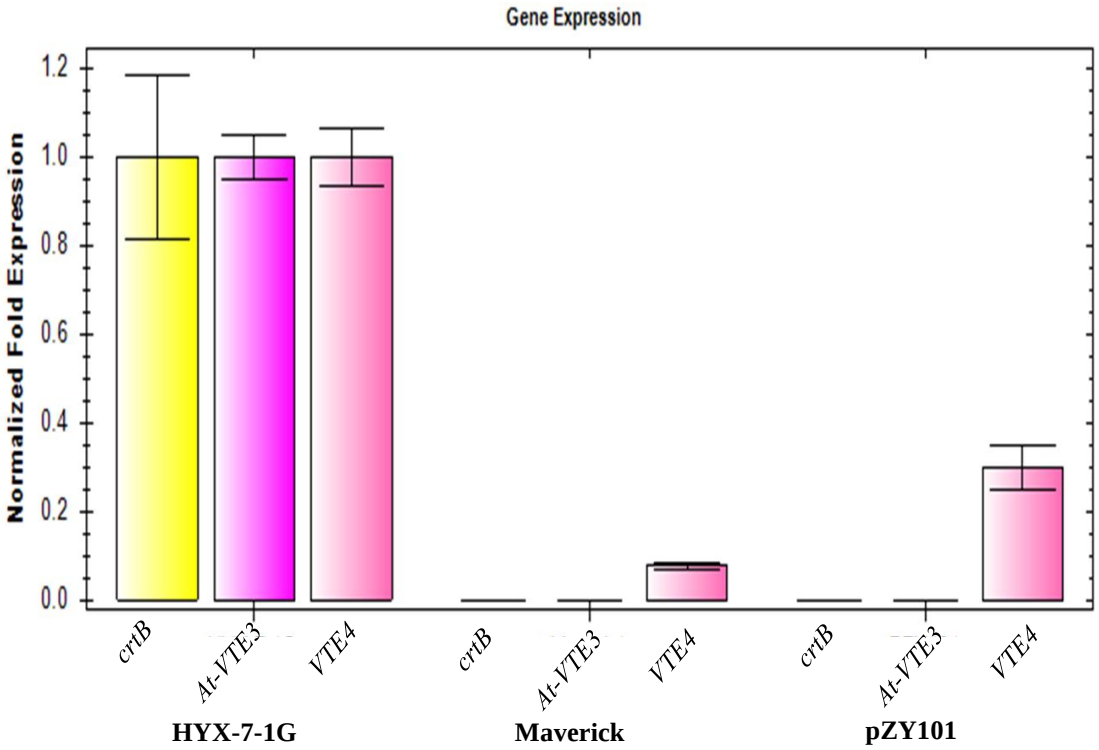

Gene Expression
Normalized Fold Expression
1.2
1.0
0.8
0.6
0.4
0.2
0.0
crtB
At-VTE3
VTE4
crtB
At-VTE3
VTE4
crtB
At-VTE3
VTE4
HYX-7-1G
Maverick
pZY101

**Figure 2.3A**

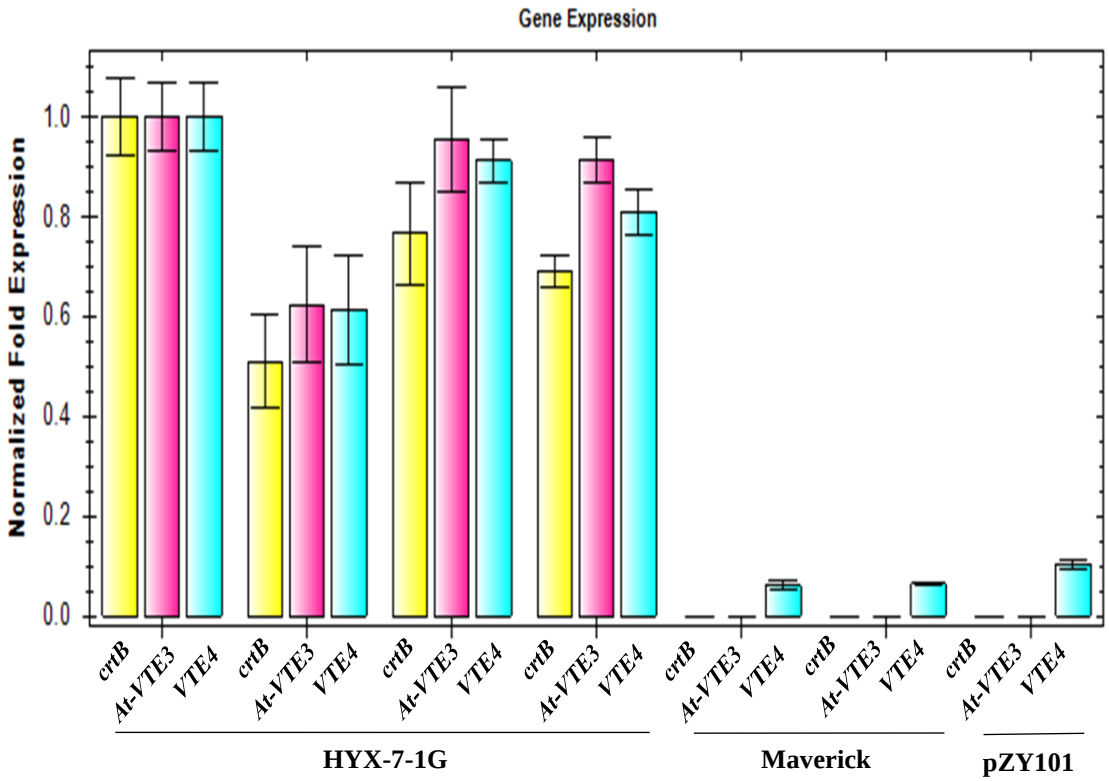

Gene Expression
Normalized Fold Expression
1.0
0.8
0.6
0.4
0.2
0.0
crtB
At-VTE3
VTE4
HYX-7-1G
Maverick
pZY101

**Figure 2.3B**

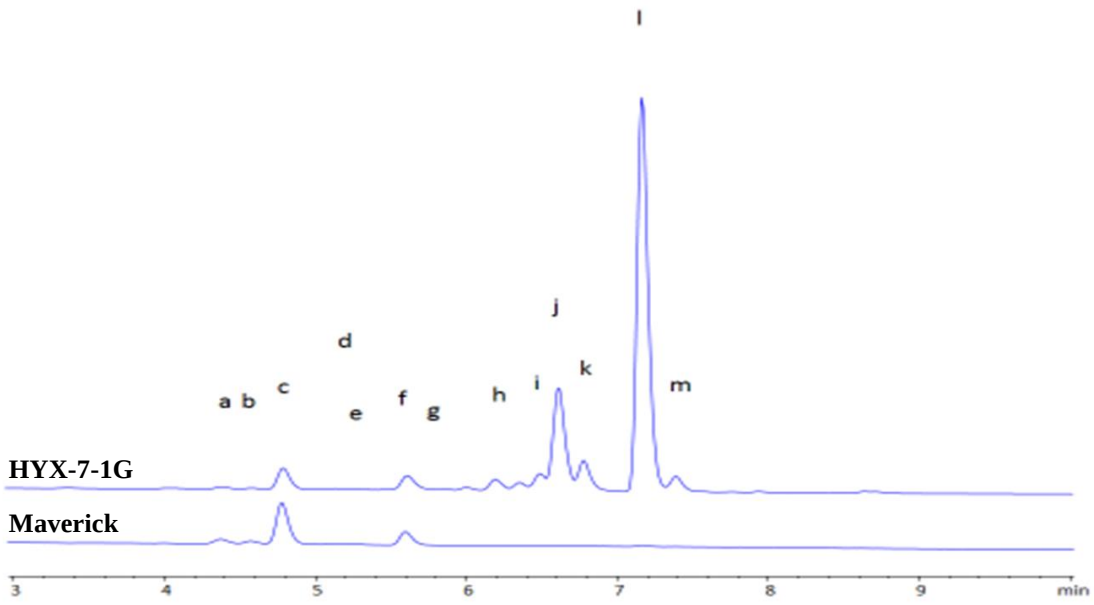



**Figure 2.4A**

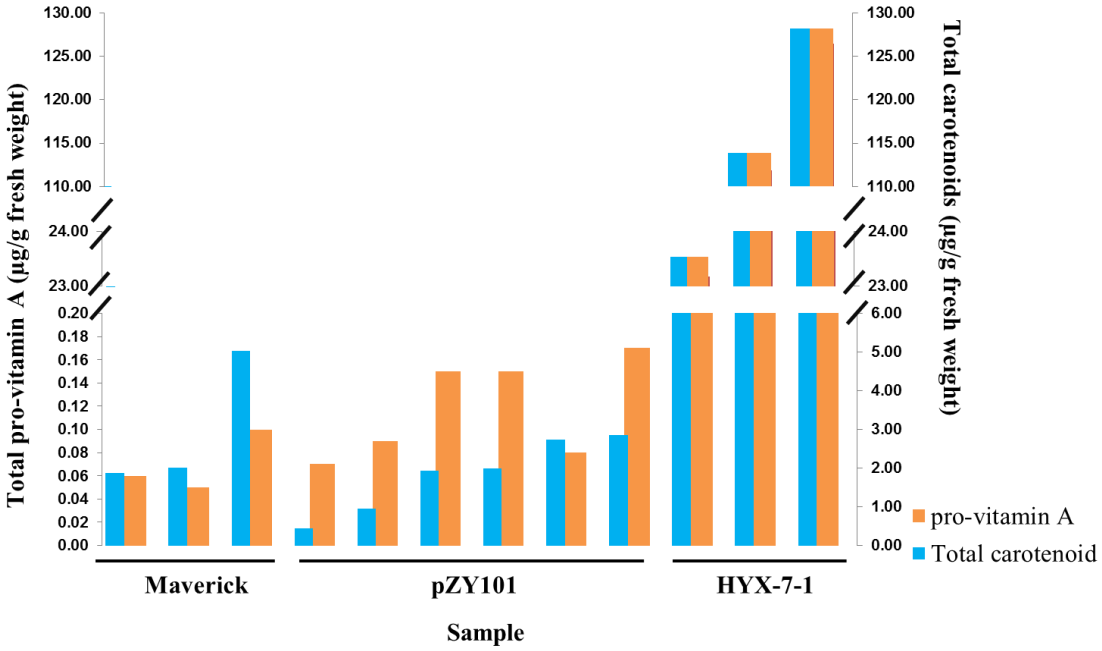

Total pro-vitamin A (µg/g fresh weight)
Total carotenoids (µg/g fresh weight)
130.00
125.00
120.00
115.00
110.00
24.00
23.00
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
6.00
5.00
4.00
3.00
2.00
1.00
pro-vitamin A
Total carotenoid
Maverick
pZY101
HYX-7-1
Sample

**Figure 2.4B**

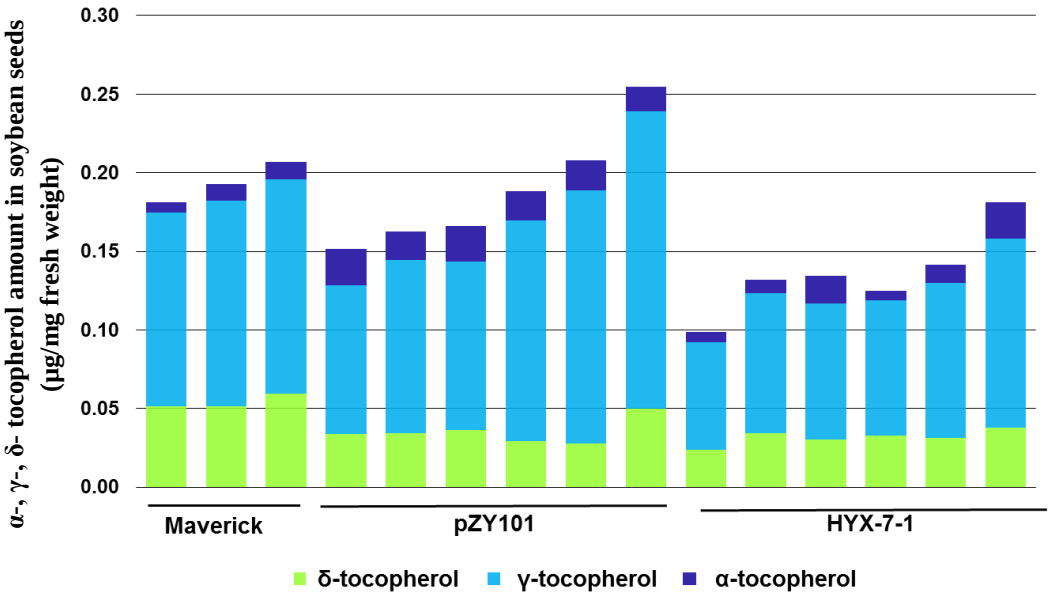

α-, γ-, δ- tocopherol amount in soybean seeds (μg/mg fresh weight)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Maverick
pZY101
HYX-7-1
δ-tocopherol
γ-tocopherol
α-tocopherol

**Figure 2.4C**

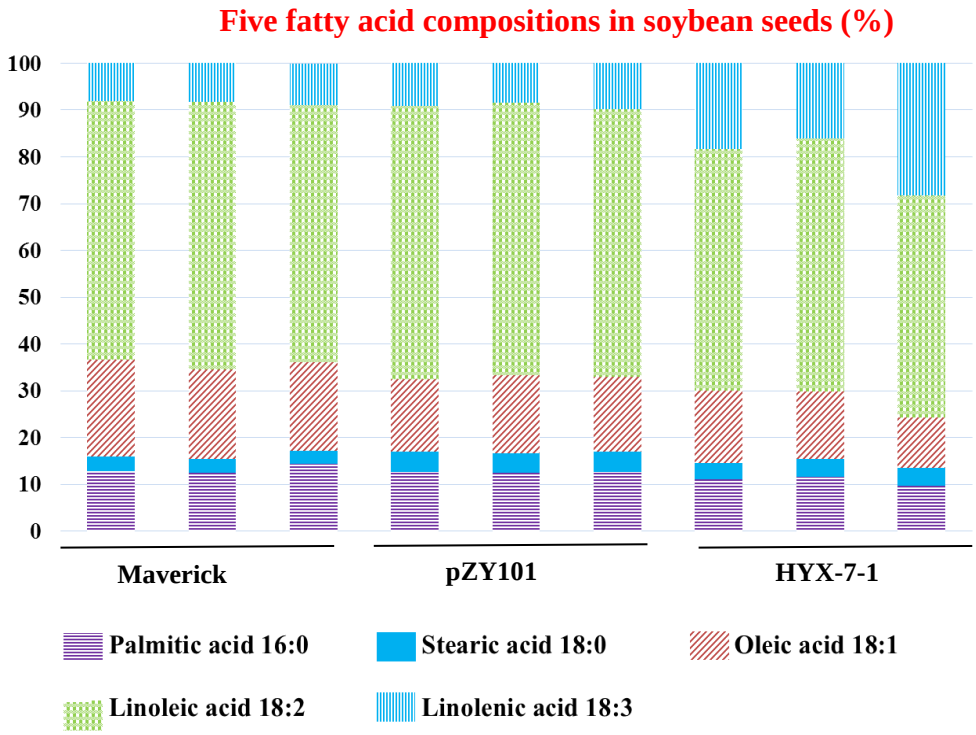

Five fatty acid compositions in soybean seeds (%)
100
90
80
70
60
50
40
30
20
10
0
Maverick
pZY101
HYX-7-1
Palmitic acid 16:0
Stearic acid 18:0
Oleic acid 18:1
Linoleic acid 18:2
Linolenic acid 18:3

**Figure 2.5**

**Table 2.1A: Primers for PCR**

| No. | Name of primer | Sequence | Amplified products (bp) |
|---|---|---|---|
| 1 | BE3-F | ATGAATAATCCGTCGTTACTCAATCATGCG | 2000 |
| | BE3-R | TTGAAACCGGCATTCTTGAACCACTCAATG | |
| 2 | E3E4-F | ATTGGACCGAGGATATGAGAGACGAC-3 | 2000 |
| | E3E4-R | TCAGGTTTTCGACATGTAATGATGGCAAAC | |
| 3 | *bar*-F | CAGCAGGTGGGTGTAGAGCGT | 268 |
| | *bar*-R | CACCATCGTCAACCACTACATCG | |

**Table 2.1B: Primers for qRT-PCR**

| No. | Name of primer for qRT PCR | Sequence (5' to 3') |
|---|---|---|
| 1 | *crtB*-F | TTATTGACGATCAGACGCTGG |
| 2 | *crtB*-R | CTTCCTGAAAAGCCGCAAAC |
| 3 | *AtVTE3*-F | GAGCCGTTGAAAGAATGCAAG |
| 4 | *AtVTE3*-R | CCCTGTACGCTTCCCTTATTC |
| 5 | *VTE4*-F | CATGGAGAGTGGAGAGCATATG |
| 6 | *VTE4*-R | TGTAAGGATTGTTCGTCAGGG |
| 7 | *cons*7-F | ATGAATGACGGTTCCCATGTA |
| 8 | *cons7*-R | GGCATTAAGGCAGCTCACTCT |

These primer sets yield amplified products of 150-200bp

**Table 2.2A: Soybean yield and segregation of T1 seeds of four regenerated plants**

| Regenerated plants | Seed yield (# seed) | Segregation Golden/Null number |
|---|---|---|
| HYX-7-1 | 248 | 189 /59 |
| HYX-7-2 | 290 | 0/290 |
| HYX-7-3 | 444 | 0/290 |
| HYX-7-4 | 268 | 0/290 |

**Table 2.2B: Transgenic integration and inheritance in T1 plants.**

| Transformed plants ($T_1$) | Total progeny | Leaf-painting | | Presence of *bar* gene | | Presence of GOI | |
|---|---|---|---|---|---|---|---|
| | | R | S | Detected | Not detected | Detected | Not detected |
| HYX-7-1(G) | 12 | 12 | 0 | 12 | 0 | 12 | 0 |
| HYX-7-3 | 12 | 0 | 12 | 0 | 12 | 0 | 12 |
| HYX-7-4 | 12 | 0 | 12 | 0 | 12 | 0 | 12 |

R and S: Resistance and Sensitive to herbicide Liberty®, respectively in leaf painting.

**Table 2.3A: Total content and compositions of carotenoids of transgenic soybean seeds of HYX-7-1 (golden seeds), empty vector pZY101transgenic and wild type Maverick plants.**

| Samples | Maverick | pZY101 | HYX-7-1(G) |
|---|---|---|---|
| 9 or 9'-cis-lutein | 0.29±0.20 | 0.19±0.11 | 0.15±0.1 |
| 9' or 9-cis-lutein | 0.17±0.12 | 0.13±0.07 | 0.12±0.09 |
| all-trans-lutein | 2.33±1.35 | 1.34±0.74 | 1.1±0.72 |
| zeaxanthin | 0.07±0.07 | 0.02±0.02 | * |
| 13- or 13'-cis-lutein | 0.05±0.03 | 0.03±0.01 | * |
| α-cryptoxanthin | * | * | * |
| β-crytpoxanthin | * | 0.01± 0.00 | 0.83±0.54 |
| 15-cis-β-carotene | * | 0.01±0.00 | 2.64±1.67 |
| 13-cis-β-carotene | 0.02±0.01 | 0.01±0.01 | 18.08±11.49 |
| α-carotene | * | 0.02±0.01 | 5.59±3.83 |
| all-trans-β-carotene | 0.04±0.02 | 0.06±0.02 | 58.12±37.29 |
| 9-cis-β-carotene | 0.01±0.01 | 0.02±0.01 | 1.90±1.19 |
| pro-vitamin A | 0.07±0.01 | 0.12±0.04 | 87.16±55.88 |
| total carotenoids | 2.98±1.78 | 1.82±0.96 | 88.53±56.74 |

Data (µg/g fresh weight) are the means and standard deviations of three replications of WT Maverick, three replications of golden seeds of HYX-7-1, and six replications of empty vector pZY101 transgenic plant. Total carotenoid levels are calculated as the sum of twelve carotenoid subtype levels i.e. 9 or 9'-cis-lutein, 9' or 9-cis-lutein, all-trans-lutein, zeaxanthin, 13-or 13'-cis-lutein, α-cryptoxanthin, β- cryptoxanthin, 15-cis- β-carotene, 13-cis- β-carotene, α-carotene, all-trans- β-carotene, and 9-cis- β-carotene. The provitamin A includes β- cryptoxanthin, 15-cis-β-carotene, 13-cis- β-carotene, α-carotene, all-trans- β-carotene, and 9-cis- β-carotene.

**Table 2.3B: Tocopherol analysis of mature seeds of HYX-7-1 (golden seed, HYX-7-1G), empty vector pZY101 and wild type Maverick plants.**

| Samples | Tocopherol content (μg/mg) | | | Total tocopherol |
|---|---|---|---|---|
| | δ-tocopherol | γ-tocopherol | α-tocopherol | |
| Maverick | 0.05±0.01 | 0.13±0.01 | 0.01±0.00 | 0.19±0.02 |
| pZY101 | 0.03±0.01 | 0.13±0.04 | 0.02±0.00 | 0.19±0.04 |
| HYX-7-1G | 0.03±0.00 | 0.09±0.02 | 0.01±0.01 | 0.14±0.02 |

Data (μg/g fresh weight) are the means and standard deviations of 6 replications of golden seeds of HYX-7-1, 6 replications of empty vector pZY101 transgenic plants, 3 replications of WT (Maverick). Total tocopherol levels are calculated as the sum of three δ-, γ-, and α-tocopherol.

**Table 2.4: Fatty acid analysis of mature seeds of HYX-7-1 (golden seeds), empty vector pZY101 and wild type Maverick plants.**

| Samples | Fatty acid content (%) | | | | Linolenic acid (18:3) |
|---|---|---|---|---|---|
| | Palmitic acid (16:0) | Stearic acid (18:0) | Oleic acid (18:1) | Linoleic acid (18:2) | |
| Maverick | 13.10±0.96 | 3.00±0.17 | 19.60±0.96 | 55.83±1.19 | 8.53±0.42 |
| pZY101 | 12.53±0.12 | 4.17±0.06 | 16.17±0.60 | 57.93±0.64 | 9.30±0.66 |
| HYX-7-1G | 10.70±0.98 | 3.70±0.26 | 13.50±2.48 | 51.10±3.44 | 21.00±6.52 |

Data (percentage of each fatty acid) are the means and standard deviations of 3 replications of golden seeds of HYX-7-1, empty vector pZY101 transgenic plants, and WT (Maverick). Total amount of five fatty acids is measured as 100%).

## LITERATURE CITED


Al-Babili, S. and P. Beyer. 2005. Golden Rice – five years on the road – five years to go? Trends in Plant Science 10: 565-573.

Aluru, M., Y. Xu, R. Guo, Z. Wang, S. Li, W. White, et al. 2008. Generation of transgenic maize with enhanced provitamin A content. Journal of Experimental Botany 59: 3551-3562.

Barampuram, S. and Z.J. Zhang. 2011. Recent advances in plant transformation. Methods In Molecular Biology (Clifton, N.J.) 701: 1-35.

Baumgartner, R.M., G. Wang, T. Wang and W.R. Fehr. 2010. Tocopherol Content of Soybean Lines with Mid-oleate and 1%-Linolenate [electronic resource]. Crop Science 50: 770-774.

Beuselinck, P.R., D.A. Sleper and K.D. Bilyeu. 2006. An Assessment of Phenotype Selection for Linolenic Acid Using Genetic Markers. Crop Science 46: 747-750.

Collakova, E. and D. DellaPenna. 2003. Homogentisate phytyltransferase activity is limiting for tocopherol biosynthesis in Arabidopsis. Plant Physiology 131: 632-642.

Cong, L., G. Yang, G. He, H. Liu, C. Wang and L. Chen. 2009. Expression of phytoene synthase1 and Carotene Desaturase crtI Genes Result in an Increase in the Total Carotenoids Content in Transgenic Elite Wheat (Triticum aestivum L.) [electronic resource]. Journal of Agricultural and Food Chemistry 57: 8652-8660.

Diretto, G., S. Al-Babili, R. Tavazza, V. Papacchioli, P. Beyer and G. Giuliano. 2007. Metabolic engineering of potato carotenoid content through tuber-specific overexpression of a bacterial mini-pathway. PLoS ONE 2: e350-e350.

Dwiyanti, M.S., T. Yamada, M. Sato, J. Abe and K. Kitamura. 2011. Genetic variation of γ-tocopherol methyltransferase gene contributes to elevated α-tocopherol content in soybean seeds. BMC Plant Biology 11: 152-152.

Flores, T., O. Karpova, X. Su, P. Zeng, K. Bilyeu, D.A. Sleper, et al. 2008. Silencing of GmFAD3 gene by siRNA leads to low [alpha]-linolenic acids (18:3) of fad3-mutant phenotype in soybean [Glycine max (Merr.)]. Transgenic Research 17: 839.

Fray, R.G., P. Hedden, P.M. Bramley, D. Grierson, A. Wallace, P.D. Fraser, et al. 1995. Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. Plant Journal 8: 693-701.

Fujisawa, M., M. Watanabe, S.-K. Choi, M. Teramoto, K. Ohyama and N. Misawa. 2008. Enrichment of carotenoids in flaxseed (Linum usitatissimum) by metabolic engineering

with introduction of bacterial phytoene synthase gene crtB. Journal of Bioscience & Bioengineering 105: 636-641.

Ha, S.H., Y.S. Liang, H. Jung, M.J. Ahn, S.C. Suh, S.J. Kweon, et al. 2010. Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. Plant Biotechnology Journal 8: 928-938.

Halpin, C., A. El Amrani, M.D. Ryan, S.E. Cooke and A. Barakate. 1999. Self-processing 2A-polyproteins -- a system for co-ordinate expression of multiple proteins in transgenic plants. Plant Journal 17: 453-459.

Hartman, G.L., E.D. West and T.K. Herman. 2011. Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. Food Security 3: 5-17.

Jing, J., W. Gang, W. Jiehua and W. Ping. 2009. Functional analysis of multiple carotenogenic genes from Lycium barbarum and Gentiana lutea L. for their effects on β-carotene production in transgenic tobacco. Biotechnology Letters 31: 305-312.

Kean, E.G., M.G. Ferruzzi and B.R. Hamaker. 2008. Carotenoid Bioaccessibility from Whole Grain and Degermed Maize Meal Products [electronic resource]. Journal of Agricultural and Food Chemistry 56: 9918-9926.

Kim, M.-J., Y.-S. Chung, S.-H. Ha, J.K. Kim, H.J. Kim, J.H. Pak, et al. 2012. Genetic modification of the soybean to enhance the β-carotene content through seed-specific expression. PloS ONE 7: e48287.

Lindgren, L.O., K.G. Stålberg and A.-S. Höglund. 2003. Seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll, and abscisic acid. Plant Physiology 132: 779-785.

Libault M., Thibivilliers S., Bilgin D.D., Radwan O., Benitez M., Clough S.J., Stacey G. 2008. Identification of Four Soybean Reference Genes for Gene Expression Normalization. Plant Gen. 1:44-54.

Manz, B., K. Müller, B. Kucera, F. Volke and G. Leubner-Metzger. 2005. Water Uptake and Distribution in Germinating Tobacco Seeds Investigated in Vivo by Nuclear Magnetic Resonance Imaging. Plant Physiology 138: 1538-1551.

McCord, K.L., S.R. Cianzio, S.R. Schnebly, G.A. Welke, W.R. Fehr and T. Wang. 2004. Tocopherol content of soybean lines with reduced linolenate in the seed oil. Crop Science 44: 772-776.

Mounts, T.L., K.A. Rennick and S.L. Abidi. 1996. Effect of genetic modification on the content and composition of bioactive constituents in soybean oil. Journal of the American Oil Chemists' Society 73: 581-586.

Paine, J.A., C.A. Shipton, S. Chaggar, R.M. Howells, M.J. Kennedy, G. Vernon, et al. 2005. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nature Biotechnology 23: 482-487.

Rhee, Y. and P.E. Staswick. 1992. Nucleotide Sequence of a Soybean Vegetative Storage Protein vspB Gene. Plant Physiology 98: 794-795.

Ryan, M.D. and J. Drew. 1994. Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. The EMBO Journal 13: 928-933.

Savidge, B., J.D. Weiss, Y.-H.H. Wong, M.W. Lassner, T.A. Mitsky, C.K. Shewmaker, et al. 2002. Isolation and characterization of homogentisate phytyltransferase genes from Synechocystis sp. PCC 6803 and Arabidopsis. Plant Physiology 129: 321-332.

Scherder, C.W., T. Wang, G.A. Welke and W.R. Fehr. 2006. Tocopherol Content and Agronomic Performance of Soybean Lines with Reduced Palmitate [electronic resource]. Crop Science 46: 1286-1290.

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

Shintani, D. and D. DellaPenna. 1998. Elevating the vitamin E content of plants through metabolic engineering. Science 282: 2098.

Slavin, M., L. Yu and W. Kenworthy. 2009. Antioxidant Properties, Phytochemical Composition, and Antiproliferative Activity of Maryland-Grown Soybeans with Colored Seed Coats [electronic resource]. Journal of Agricultural and Food Chemistry 57: 11174–11185-11174–11185.

Tang, G., J. Qin, G.G. Dolnikowski, R.M. Russell and M.A. Grusak. 2009. Golden Rice is an effective source of vitamin A. American Journal of Clinical Nutrition 89: 1776-1783.

Tavva, V.S., Y.-H. Kim, I.A. Kagan, R.D. Dinkins, K.-H. Kim and G.B. Collins. 2007. Increased alpha-tocopherol content in soybean seed overexpressing the Perilla frutescens gamma-tocopherol methyltransferase gene. Plant Cell Reports 26: 61-70.

Traber, M.G. and H. Sies. 1996. Vitamin E in humans: demand and delivery. Annual Review of Nutrition 16: 321-347.

Van Eenennaam, A.L., K. Lincoln, T.P. Durrett, H.E. Valentin, C.K. Shewmaker, G.M. Thorne, et al. 2003. Engineering vitamin E content: from Arabidopsis mutant to soy oil. The Plant Cell 15: 3007-3019.

Winklhofer-Roob, B.M., E. Rock, J. Ribalta, D.H. Shmerling and J.M. Roob. 2003. Effects of vitamin E and carotenoid status on oxidative stress in health and disease.

Evidence obtained from human intervention studies. Molecular Aspects of Medicine 24: 391.

Wright, T.R., G. Shan, T.A. Waish, J.M. Lira, C. Cui, P. Song, et al. 2010. Robust crop resistance to broadleaf and grass herbicides provided by aryloxyalkanoate dioxygenase transgenes. Proceedings of the National Academy of Sciences of the United States of America 107: 20240-20245.

Yang, W., R.E. Cahoon, S.C. Hunter, C. Zhang, J. Han, T. Borgschulte, et al. 2011. Vitamin E biosynthesis: functional characterization of the monocot homogentisate geranylgeranyl transferase. Plant Journal 65: 206-217.

Ye, X., S. Al-Babili, A. Kloti, J. Zhang, P. Lucca, P. Beyer, et al. 2000. Engineering the Provitamin A (Beta-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice E. Science 287: 303.

Zeng, P., J.C. Polacco, Z. Zhang and D.A. Vadnais. 2004. Refined glufosinate selection in Agrobacterium-mediated transformation of soybean (Glycine max (L.) Merrill). Plant Cell Reports 22: 478-482.

Zhang, Z., T.E. Clemente, P. Staswick and A. Xing. 1999. The use of glufosinate as a selective agent in Agrobacterium-mediated transformation of soybean. Plant Cell, Tissue and Organ Culture 56: 37-46.