CHARACTERIZATION OF THE CYP97 AND HYD CAROTENE HYDROXYLASE ENZYMES

by

RENA FIONA QUINLAN

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2012

This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Dr. Eleanore Wurtzel

Date

Chair of Examining Committee

Dr. Laurel Eckhardt

Date

Executive Officer

Dr. Edward Kennelly

Dr. Haiping Cheng

Dr. Akira Kawamura

Dr. David Calhoun

Dr. Loredana Quadro

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

CHARACTERIZATION OF THE CYP97 AND HYD CAROTENE HYDROXYLASE ENZYMES

by

Rena F. Quinlan

Advisor: Professor Eleanore T. Wurtzel

Vitamin A deficiency is a serious and widespread public health issue in developing countries. Provitamin A carotenoids such as β-carotene have therefore recently attracted interest as important neutraceuticals. Due to the nutritional value of carotenoids there is currently considerable interest in developing rational strategies for metabolic engineering of crops for enhanced carotenoid content. Efforts to improve the provitamin A content of cereal endosperm in staple crops such as maize will require characterization of the carotene ring hydroxylases involved in controlling the conversion of provitamin A carotenes to non-provitamin A xanthophylls.

Based on early modeling by (Cunningham and Gantt, 1998) the cytochrome P450 CYP97 and diiron HYD carotene hydroxylases were predicted to localize to chloroplast membranes to function in separate multi-enzyme complexes for the respective conversions of the provitamin A carotenes α- and β-carotene to the non-provitamin A xanthophylls lutein and zeaxanthin. To gain a better understanding of the respective roles of the CYP97 and HYD enzymes in these conversions, the

activities and localization/interaction of these enzymes were examined using both *in vitro* and *in vivo* approaches.

Escherichia coli functional complementation systems were used to assess rice P450 CYP97A4 (β-ring hydroxylase) and CYP97C2 (ε-ring hydroxylase) as well as maize diiron HYD3 and HYD4 (β-ring hydroxylases) activities and substrate specificities. Preliminary investigations examining CYP97 enzyme activity only via *E. coli* complementation showed that the CYP97A4 exhibits major activity toward β-rings to convert β-carotene to β-cryptoxanthin (pathway intermediate) as well as a low amount of zeaxanthin (pathway end-product). In addition, this enzyme exhibited minor activity toward the ε-rings of ε-ε-carotene to convert this substrate to a low amount of lactucaxanthin (pathway end-product). These studies also indicated that the CYP97C2 appeared to be exclusively active toward the ε-rings of ε-ε carotene, converting this substrate to a low amount of lactucaxanthin; no activity toward the β-rings of β-carotene was detected for this enzyme.

Subsequent complementation studies tested both individual and combined CYP97/HYD enzyme activities in *E. coli* accumulating both α- and β-carotene substrates. These studies demonstrated that the CYP97A4 and CYP97C2 enzymes function optimally when expressed together in the conversion of their preferred substrate α-carotene to produce lutein. Cells engineered to produce α- and β-carotene and which co-expressed these enzymes generated almost 30% lutein (% total carotenoids); a roughly 10-fold higher amount of lutein relative to zeaxanthin was observed. In contrast, when expressed as individual enzymes the CYP97A4 and CYP97C2 showed suboptimal

activity (ie., no lutein produced; only approx. 14%, and 1% of the intermediates zeinoxanthin and α-cryptoxanthin generated respectively) regardless of substrate choice. In these cells, the CYP97A4, when expressed alone, preferred the β-carotene substrate to the α-carotene substrate generating a low amount of zeaxanthin (the monohydroxylated intermediate β-cryptoxanthin accumulated); the CYP97A4 was only moderately active toward the α-carotene substrate as only the intermediate zeinoxanthin accumulated (no lutein was produced). When expressed as an individual enzyme, the CYP97C2 was minimally active toward both α- and β-carotene substrates to respectively generate barely detectable amounts of the intermediates α-cryptoxanthin and β-cryptoxanthin.

HYD3 + CYP97C2 and HYD4 + CYP97C2 combinations were also tested using this complementation system. Both enzyme combinations were moderately active toward α-carotene, respectively producing low amounts of lutein; although cells co-expressing the HYD3 + CYP97C2 appeared to be somewhat more active toward α-carotene than β-carotene, generating a more than two-fold higher amount of lutein relative to zeaxanthin. When expressed as an individual enzyme, the HYD3 was preferentially active toward β-carotene to convert this substrate to a low amount of zeaxanthin (end-product); this enzyme was only moderately active toward the α-carotene substrate as only the intermediate zeinoxanthin accumulated. The HYD3 exhibited suboptimal activity in our complementation systems whether expressed alone or in combination with the CYP97C2. By contrast, the HYD4 functioned optimally when expressed as an individual enzyme.

This enzyme was preferentially active toward β-carotene and efficiently converted this substrate to zeaxanthin.

This work also examined CYP97/HYD protein localization and protein interaction. *In vitro* chloroplast import and *in vivo* GFP fusion assays confirmed that these enzymes are localized to chloroplasts. In addition, import assays were used to determine the suborganellar locations of these enzymes, and *in vivo* Bimolecular Fluorescence (BiFC) assays were performed to assess protein-protein interaction. Taken together, these studies demonstrated that the CYP97A4 and CYP97C2 enzymes are peripherally-associated to chloroplast membranes where they interact to form a heterodimer complex to function in the efficient conversion of α-carotene to lutein. It was expected that these enzymes, which functioned optimally together toward the α-carotene substrate in the *E. coli* complementation system, would interact and localize to the same location in the chloroplast membrane. These data also indicated that the HYD3 and HYD4 enzymes are integrally-bound to the chloroplast membrane where they interact to function in the conversion of β-carotene to zeaxanthin. These enzymes were expected to localize to the same suborganellar location since they were both preferentially active toward the same substrate (ie., β-carotene) in the complementation system. In addition, BiFC analysis indicated that HYD4 formed a homodimer complex.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. Eleanore Wurtzel, for her guidance and support throughout my doctoral studies. I also thank my committee members, Dr. Edward Kennelly, Dr. Haiping Cheng, Dr. David Calhoun, Dr. Akira Kawamura, and Dr. Loredana Quadro for their advice and suggestions which significantly advanced the progress of my research.

I wish to acknowledge my lab mates past and present for their help and collaborative efforts: Justin, Ratnakar, Christina, Nicholas, Zhu, Yu, Chucho, Xiaoling, Oren, Abby, Regina, and Charles. I extend special thanks to Dr. Maria Shumskaya, for sharing her expertise in protein studies, and Dr. Louis Bradbury, for his help with cloning and HPLC analysis. I also thank the following members of the Kennelly lab: Drs. Chunhui Ma, Keyvan Dastmalchi, Amy Keller, Gema Flores, and Adam Kavalier for sharing their expertise in LC-MS analysis.

I thank all the members of the Biology Department staff, especially Dr. Joseph Rachlin and Dr. Dwight Kincaid for their support.

I extend a very special thanks to my friends, Jack Henning and Christina West, who have offered much kind encouragement over the years.

Finally, I am especially grateful to my family for their encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	amino acid
ABA	absisic acid
bp	base pair
BSA	bovine serum albumin
CCD	carotenoid cleavage dioxygenase
cDNA	complementary DNA
CRTISO	carotenoid isomerase
CTP	chloroplast transit peptide
CYP97	cytochrome P450 carotene hydroxylase
CYP97A	P450 β-hydroxylase
CYP97C	P450 ε-hydroxylase
DAP	days after pollination
DMAPP	dimethylallyl diphosphate
DXP	1-deoxy-D-xylulose 5-phosphate
DXS	deoxyxylulose 5-phosphate synthase
DXR	1-deoxy-D-xylulose 5-phosphate reductoisomerase
E. coli	*Escherichia coli*
FAD	flavin adenine dinucleotide
GAP	glyceraldehyde 3-phosphate
GPP	geranyl diphosphate
GGPP	geranylgeranyl pyrophosphate
GGPPS	geranylgeranyl pyrophosphate synthase
HDR	hydroxymethylbutenyl diphosphate reductase
HMBPP	hydroxymethylbutenyl diphosphate
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
HPLC	high-performance liquid chromatography
HYD	diiron carotene hydroxylase
HYDB	diiron β-hydroxylase
IPPI	isopentenyl diphosphate isomerase
IPP	isopentenyl diphosphate
IPTG	isopropylthio-β-D-galactoside
kb	kilo base pair
kDa	kilo Dalton
LB	Luria-Bertani
LC	Liquid chromatography
LCYB	lycopene β-cyclase
LCYE	lycopene ε-cyclase
LcyB	lycopene β-cyclase
LcyE	lycopene ε-cyclase
LHC	light-harvesting complex
MEP	2-C-methyl-D-erythritol 4-phosphate

MS	Mass spectrometry
MVA	mevalonic acid
NADPH	nicotinamide adenine dinucleotide phosphate
NCED	9-*cis*-epoxycarotenoid dioxygenase
NXS	neoxanthin synthase
ORF	open reading frame
PCR	polymerase chain reaction
PDS	phytoene desaturase
PSY	phytoene synthase
RT-PCR	reverse transcriptase polymerase chain reaction
VDE	violaxanthin de-epoxidase
ZDS	zeta-carotene desaturase
ZEP	zeaxanthin epoxidase
Z-ISO	15-*cis*-zeta-carotene isomerase

CHAPTER 1

Introduction

1.1 Carotenoids: background and significance

Carotenoids are a large class of isoprenoid pigments found in plants, fungi, and bacteria. More than 750 carotenoids have been characterized in nature thus far (Britton et al., 2004). In plants these compounds function as essential structural and functional components of the photosynthetic apparatus. Specifically, carotenoids serve as light-harvesting pigments in photosynthesis and protect against photooxidative damage (Liang et al., 2006). Plants also exploit the distinctive yellow, orange, and red pigments of carotenoids to attract pollinators and various agents involved in seed dispersal (Kato et al., 2004). Carotenoids also serve as precursors to plant hormones such as abscisic acid (ABA) (Cuttriss et al., 2011) and apocarotenoids which function in stress and developmental responses (Walter et al., 2010). In addition, these pigments are essential components of mammalian diets and have vital anti-oxidant and provitamin A activities (van den Berg et al., 2000; von Lintig, 2010).

Vitamin A deficiency is a serious public health issue in developing countries and places up to 140 million children at risk every year for a variety of disorders including depressed immunological responses and increased susceptibility to infectious disease, as well as xerophthalmia and blindness (Sommer and Davidson, 2002). Provitamin A carotenoids, such as β-carotene, have therefore recently attracted interest as important neutraceuticals. Various plant-derived carotenoids have also

attracted attention for their potential as anticancer agents. In addition, these compounds may play roles in preventing certain cardiovascular and eye diseases, and may serve to enhance immune system functions (Kopsell and Kopsell, 2006). Due to the nutritional value of carotenoids there is currently considerable interest in developing rational strategies for breeding and/or metabolic engineering of crops for enhanced carotenoid content (Wurtzel et al., 2012).

1.2 Carotenoid structure

Carotenoids are C_{40} terpenoids that are derived from isoprene, and are synthesized by tail-to-tail linkage of two molecules of C_{20} geranylgeranyl diphosphate. The polyene skeleton of carotenoids is the most structurally distinguishing feature, and consists of a long system of alternating conjugated double and single bonds in which the π-electrons are delocalized along the entire length of the polyene chain; this feature is responsible for the characteristic molecular shape, light-absorbing properties, and chemical reactivity of carotenoids. Geometric isomerization of the double bonds in the polyene chain of carotenoids generates either *E*-configurations (*trans*-configurations) or *Z*-configurations (*cis* configurations), depending on the arrangement of substituent groups (Britton et al., 1995).

Carotenoids are divided into two groups; the carotenes (ie., β-carotene), which are non-oxygenated carotenoids that may be linear or possess cyclic hydrocarbons at one or both ends of the molecule, and the xanthophylls (ie., lutein), which are oxygenated derivatives of carotenes. Carotenoids such as β-carotene and zeaxanthin,

serve as precursors to cleavage products including the hormone abscisic acid (ABA) and to other apocarotenoids (Walter et al., 2010; Cuttriss et al., 2011).

1.3 Carotenoid precursor biosynthesis

Plants have evolved two distinct pathways for the synthesis of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), the universal precursor compounds of all isoprenoids. The well-characterized mevalonate (MVA) pathway, which is only found in fungi and animals, as well as in the cytoplasm of phototrophic organisms, generates cytosolic IPP, which is subsequently isomerized to DMAPP via IPP isomerase (IPPI). In contrast, plastidial IPP and DMAPP are produced via the more recently described 2-C-methyl-D-erythritol 4-phosphate (MEP) (non-mevalonate) pathway; this pathway is found in plant chloroplasts as well as in most bacteria (Schwender et al., 1996; Lichtenthaler et al., 1997; Orihara et al., 1997; Rodriguez-Concepcion and Boronat, 2002; Seemann et al., 2006).

The classical MVA pathway involves condensation of three units of acetyl-CoA to generate 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA). HMG-CoA is then reduced by 2 molecules of NADPH to yield mevalonic acid (MVA) (Lichtenthaler et al., 1997).

The MEP pathway-derived precursors include the main photosynthetic pigments (ie., chlorophylls and carotenoids) as well as other important compounds involved in photosynthesis such as the plastoquinones, phylloquinones, and

tocopherols. In addition, plant hormones such as gibberellins and abscisic acid (ABA) are synthesized from MEP-derived precursors (Rodriguez-Concepcion, 2010)

The first reaction of the MEP pathway involves condensation of the glycolytic products pyruvate and glyceraldehyde 3-phosphate (GAP) by the deoxyxylulose 5-phosphate synthase (DXS) enzyme to produce deoxyxylulose 5-phosphate (DXP). The next step of the pathway involves an intramolecular rearrangement and subsequent reduction of DXP via DXP reductoisomerase (DXR) to generate MEP, the first committed substrate of plastidial isoprenoid compounds. MEP is then converted to hydroxymethylbutenyl diphosphate (HMBPP) in four enzymatic steps. In the last step of the pathway, HMBPP reductase (HDR) then converts HMBPP into DMAPP and IPP (Carretero-Paulet et al., 2006).

1.4 Carotenoid biosynthesis

The biosynthetic pathway including enzymatic steps and structures are depicted in **Fig. 1-1**. GGPP, which is the immediate precursor for carotenoid biosynthesis, is produced via the sequential addition of three molecules of IPP to one molecule of the IPP isomer, DMAPP; these reactions are catalyzed by GGPP synthase (Cuttriss et al., 2011; Wurtzel et al., 2012).

Head-to-head condensation of two molecules of GGPP form the colorless compound phytoene, which is the first C_{40} carotenoid of the pathway. This reaction is catalyzed by the enzyme phytoene synthase (PSY). Phytoene subsequently undergoes a series of four desaturation reactions which extend the conjugated double

bond system that constitutes the chromophore in carotenoid pigments. These desaturations therefore lead to the transformation of the colorless phytoene to the pink/red pigmented lycopene. In plants, the four sequential desaturation reactions undergone by phytoene are catalyzed by two enzymes: phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS) (Cuttriss et al., 2011). In contrast, this entire process is performed by only one enzyme, CRTI, in bacteria. (Linden et al., 1991).

15-*cis* phytoene is converted to the pathway intermediates 9,15-*cis* phytofluene and 9,15,9′-tri-*cis*-ζ-carotene via two desaturation reactions performed by PDS. The *y9* allele from maize was identified as encoding a factor required for isomerase activity which functions in the dark or in dark-grown tissues in catalyzing the *cis*- to *trans*-conversion of the 15-*cis*-bond in 9,15,9′-tri-*cis*-ζ-carotene, the desaturation product of PDS, to generate 9,9′-di-*cis*-ζ-carotene (Li et al., 2007), the suitable geometrical isomer substrate for ζ-carotene desaturase (ZDS) (Beyer et al., 1989). This enzyme activity was termed Z-ISO (ζ-carotene isomerase) (Li et al., 2007). Mutational analysis subsequently led to the isolation of the *Z-ISO* gene in both maize and *Arabisopsis*. This identification was based on the phenotype of *Arabidopsis zic1* mutants which accumulated 9,15,9′-tri-*cis*-ζ-carotene in the dark. Functional complementation in *E. coli* also confirmed Z-ISO activity for both the maize and *Arabidopsis Z-ISO* gene copies (Chen et al., 2010).

The pathway continues with further desaturation of 9,9′-di-*cis*-ζ-carotene via ZDS to produce the intermediate 7,9,9′-tri-*cis*-neurosporene. Cyclization of the downstream ZDS product tetra-*cis* lycopene cannot occur until it is converted to all-

trans lycopene (Bartley et al., 1999). Characterization of the *tangerine* tomato mutant led to the cloning and identification of the carotenoid isomerase gene *CRTISO*, which was the first isomerase associated with these desaturation steps (Isaacson et al., 2002; Isaacson et al., 2004; Kato et al., 2004). It has been shown that CRTISO exclusively recognizes the products of ZDS and not those generated by PDS (Isaacson et al., 2004). These studies demonstrated that CRTISO functions in the conversion of 7,9,9′-tri-*cis* neurosporene to 9′-*cis* neurosporene and 7,9,7′,9′-tetra-*cis* lycopene to all-*trans* lycopene, however this enzyme is not involved in isomerization of the single 15 to 15′-*cis* double bond that is required for generating the acceptable ZDS substrate (Beyer et al., 1989; Bartley et al., 1999; Matthews et al., 2003). In summary, lycopene biosynthesis requires a two-step desaturation by PDS, with subsequent isomerization of the 15-*cis* bond via Z-ISO, which generates the acceptable di-*cis*-ζ-carotene substrate for ZDS. ZDS subsequently performs two desaturation steps in concert with CRTISO-mediated isomerization of each pair of double bonds generated by ZDS.

Photoisomerization of *cis*-carotenoids to *trans*-configured states may occur, particularly with respect to the central 15-15′-*cis*-double bond of phytoene, phytofluene, or zetacarotene (Bartley et al., 1999). In the dark, poly-*cis* lycopene accumulated in *Arabidopsis* CRTISO null mutants (Park et al., 2002), and maize y9 mutants conditioned accumulation of the 9,15,9′-tri-*cis*-ζ-carotene, the Z-ISO *cis* isomer substrate (Li et al., 2007). Nevertheless, light only partially compensates for lack of Z-ISO activity, as demonstrated by mutant phenotypes (Li et al., 2007).

The cyclization of the linear compound lycopene is a major branch point in the carotenoid biosynthetic pathway. The lycopene β-cyclase (LCYB) and lycopene ε-cyclase (LCYE) enzymes catalyze the formation of β- and ε-rings, respectively. In plant carotenoids, two types of cyclic end groups (and derivatives thereof) are commonly found: β-rings and ε-rings. Cyclization at both ends of lycopene by the action of LCYB produces β-carotene, while α-carotene synthesis requires the combined actions of both the LCYB and LCYE enzymes, which respectively add one β-ring and one ε-ring to the ends of the lycopene molecule. In terms of the number of specific ring-additions, an exception has been found with the LCYE enzyme from lettuce, which is uniquely capable of adding ε-rings to both ends of lycopene to form ε-ε-carotene (Cunningham Jr. et al., 1996; Cunningham Jr. and Gantt, 2001).

Hydroxylation of the β- and ε-rings of the provitamin A carotenes leads to biosynthesis of non-provitamin A xanthophylls (ie., lutein and zeaxanthin). Zeaxanthin and lutein are produced from their respective β- and α-carotene precursors by the action of carotenoid β- and ε-hydroxylases, which add hydroxyl groups to the third carbon of carotene β- and ε-rings, respectively. In plants, β-ring hydroxylations may involve either diiron-type (HYDB) or cytochrome P450-type (CYP97A) β-hydroxylases**.** The diiron-HYDs have been well-characterized from a variety of organisms, including bacterial, algal, and plants species (Sun et al., 1996; Bouvier et al., 1998; Tian et al., 2003; Tian and DellaPenna, 2004; Vallabhaneni et al., 2009). CYP97C is a cytochrome P450 ε-ring hydroxylase which was identified via genetics-based studies involving the *A. thaliana lut1* mutant (Tian et al., 2004). *Arabidopsis*

mutant studies subsequently indicated that CYP97A is a β-ring hydroxylase (Fiore et al., 2006; Kim and DellaPenna, 2006), and both CYP97A and CYP97C enzymes from rice were functionally demonstrated in *E. coli* **(Chapter 2)** (Quinlan et al., 2007).

Further oxygen additions via the enzyme zeaxanthin epoxidase (ZEP) convert zeaxanthin to epoxy-xanthophylls, such as antheraxanthin (an intermediate) and violaxanthin (**Fig. 1-1**) These conversions are reversible; under high-light stress violaxanthin de-epoxidase regenerates zeaxanthin in a process that is commonly referred to as the xanthophyll cycle (Cunningham and Gantt, 1998).

1.5 Carotenoid cleavage products: apocarotenoids

Apocarotenoids are important signaling and accessory molecules that function in an array of biological processes. Biologically important apocarotenoids include the plant hormone abscisic acid (ABA), and the visual molecule retinal (vitamin A). These compounds are formed via oxidative cleavage of carotenoid precursors such as β-carotene and zeaxanthin by the action of carotenoid cleavage dioxygenases (CCDs) (Auldridge et al., 2006; Schmidt et al., 2006).

The first CCD was identified via characterization of the ABA-deficient *Vp14* mutant from maize; the Vp14 gene product is a nine-*cis*-epoxy dioxygenase (NCED) involved in abscisic acid (ABA) biosynthesis in plants (Schwartz et al., 1997; Tan et al., 1997). *Vp14* is predicted to contain an N-terminal chloroplast targeting peptide sequence, and shares sequence homology with a *Pseudomonas paucimobilis*

lignostilbene dioxygenase that catalyzes an analogous oxidative cleavage (Cutler and Krochko, 1999).

CCD family members share several characteristics. They all require a Fe^{2+} for catalytic activity and harbor four conserved histidine residues thought to be involved in the coordination of iron-binding. In addition, they contain a conserved carboxyl terminus peptide sequence which has been identified as a signature sequence for the CCD family (Auldridge et al., 2006). Various different types of CCDs can be distinguished on the basis of their particular cleavage sites, since they exhibit a high degree of regio- and stereospecificity for specific double bond positions, which is in contrast to their frequently observed substrate promiscuity (Floss and Walter, 2009). For example, NCEDs can specifically cleave two 9-*cis*-epoxycarotenoids (ie., violaxanthin and neoxanthin) at their 11,12 double-bond positions. Other CCDs from plants that cleave all-*trans* carotenoid substrates cleave the 9,10 bond in an asymmetrical fashion, or symmetrically cleave 9,10 (9′,10′) bonds, or 7,8 (7′,8′) and 5,6 bonds in these substrates (Marasco et al., 2006). A total of nine CCD genes in *Arabidopsis* have been identified based on sequence homology to maize *VP14* (Tan et al., 2001). Of these genes, five *AtNCEDs* (2,3,5,6, and 9) are designated as nine-cis-epoxycarotenoid dioxygenases (NCEDs) since they are specifically active toward 9-*cis*-epoxycarotenoids and are involved in ABA biosynthesis, while the other four (AtCCD1, AtCCD4, AtCCD7, AtCCD8) cleave various *trans*-carotenoid substrates (Floss and Walter, 2009). A number of studies have indicated that the carotenoid cleavage reaction is a key regulatory step with respect to controlling stress-induced

ABA biosythesis (Qin and Zeevaart, 1999). The expression of *AtNCED3* in particular accounts for most of the stress-induced ABA biosynthesis in leaves; in contrast, the other four NCED genes are mainly regulated developmentally. AtNCED3, AtNCED5, AtNCED6, and AtNCED9, play roles in ABA synthesis in the developing seed. *AtNCED3* is predominantly expressed in seed maternal tissue, while *AtNCED5* and *AtNCED6* are expressed in male and female gametophytes just prior to fertilization, as well as in tissues of the endosperm and embryo (Auldridge et al., 2006).

While the majority of CCDs/NCEDs have been shown to be plastid-localized, AtCCD1 acts in the cytosol to produce C_{13} and C_{14} apocarotenoids. Of all the CCDs, CCD1 has been the best-studied due to its role in the generation of C13 apocarotenoid-based flower scent and fruit and wine aroma compounds. The cytosolic location of AtCCD1 is surprising given that the presumed carotenoid C40 substrate is plastid-localized. Gene-silencing studies involving mycorrhizal hairy roots of *Medicago truncatula* appeared to have solved this contradiction (Floss and Walter, 2009). This work showed that there was an accumulation of C27 apocarotenoids in mycorrhizal RNAi roots, which indicated that it is C27 rather than C40 derivatives which serve as the main substrates for CCD1 *in planta*. These data suggest that there is a consecutive two-step process, in which another plastid-localized CCD carries out the primary cleavage of C40 compounds to C27 compounds, followed by export of the C27 cleavage product and additional cleavage by CCD1 in the cytosol to generate the final C13/C14 apocarotenoid end-products

(Floss and Walter, 2009). In addition, *in vitro* analyses indicated that apocarotenoids were the main substrates of the rice OsCCD1 (Ilg et al., 2010).

However, studies in strawberry (*Fragaria ananassa*) showed that FaCCD1 expression levels correlated with a decrease in lutein content. These data suggest that lutein (a C40 carotenoid and not a C27 cleavage product) could be the principle natural substrate of FaCCD1 (Garcia-Limones et al., 2008). In addition, maize ZmCCD1 was demonstrated to efficiently cleave C40 carotenoids at the 9,10 position (Sun et al., 2008; Vogel et al., 2008). Furthermore, high CCD1 expression levels associated with maize alleles of the dominant *white cap1* (*wc1*) locus correlated with reduced endosperm carotenoids, and a maize inbred line containing high *ZmCCD1* copy number had concurrently low levels of carotenoids in endosperm; these data point to a possible gene dosage effect (Vallabhaneni et al., 2010). The significant variation in maize carotenoid content and composition makes this crop an important genetic resource for the identification of favorable *CCD1* alleles that promote retention of carotenoids in cereal endosperm (Wurtzel et al., 2012).

In addition to ABA, strigolactones are another group of carotenoid-derived phytohormones. Strigolactones were previously identified as germination stimulants for parasitic weeds and as signaling compounds that play a role in promoting hyphal branching of arbuscular mycorrhizal fungi (Akiyama et al., 2005; Matusova et al., 2005). More recent studies suggest there are additional roles for strigolactones in root tissue. A C18 B-apo-13-carotenone known as "D′orenone" prevents root hair growth

by interfering with PIN-2-mediated transport of auxin. The synthetic "D′orenone" is identical in structure to the predicted C18 apocarotenoid strigolactone precursor (Floss and Walter, 2009). Biosynthesis of strigolactone is thought to start with β-carotene and involve two consecutive cleavage steps in the plastid which are performed by CCD7 and CCD8 enzymes respectively. The CCD8 cleavage product, which is a C18 compound (or a derivative thereof) is exported to the cytosol where it is predicted to serve as a mobile strigolactone precursor. This precursor is further modified and subsequently transported to the shoot where it plays a role in the regulation of shoot branching (Floss and Walter, 2009). Recent *in vitro* studies by Alder et al., (2012) showed that the iron-binding protein, D27, which is a β-carotene isomerase, catalyzes the isomerization of the 9,10 double bond in all-*trans*-β-carotene to generate 9-*cis*-β-carotene, which is cleaved by CCD7 at the 9′,10′ position to produce a 9-*cis*-configured aldehyde. CCD8 introduces three oxygens into the aldehyde 9-*cis*-β-apo-10′-carotenal and carries out a molecular rearrangement, which links carotenoids with strigolactones to produce carlactone, a compound found to have strigolactone-like biological activities (Alder et al., 2012).

1.6 Carotenoid pathway localization

Nuclear-encoded carotenoid enzymes are targeted to various plastids including chloroplasts, chromoplasts, and amyloplasts, and carotenoid biosynthesis occurs on a variety of architecturally-distinct plastid membranes such as thylakoids

and envelopes (Cuttriss et al., 2011). The mechanisms controlling membrane targeting and metabolon assembly in different plastids are not well understood. Immunolocalization assays demonstrated that the pathway rate-limiting enzyme PSY1 localizes to amyloplast envelope membranes, which suggests that the envelope is the site of the carotenoid biosynthetic metabolon in these plastids (Li et al., 2008b). Proteomics studies also indicated that the *Arabidopsis* PDS, ZDS, CRTISO, lycopene β-cyclase (LCYB) and P450 β- and ε-ring hydroxylases (CYP97A and CYP97C), as well as zeaxanthin epoxidase (ZEP) localize to chloroplast envelope membranes, while PDS and ZEP were also found in thylakoid membranes, along with violaxanthin de-epoxidase (VDE) (Joyard et al., 2009). As expected, these studies revealed that both ZEP and VDE are located in thylakoids membranes since both epoxidase and de-epoxidase activities would be required in the reactions of the xanthophyll cycle, which is involved in the non-photochemcial quenching of singlet-excited chlorophyll in photosynthesis (Li et al., 2009).

In addition, biochemical evidence indicated that the chaperonins Hsp70 and Cpn60, whose expression is associated with carotenogenesis, may facilitate carotenoid enzyme localization in chromoplasts of daffodil flowers (Wurtzel, 2004). Moreover, carotenoids are known to be sequestered in lipid bodies known as plastoglobules during the transition of a chloroplast to a chromoplast (Tevini and Steinmuller, 1985). Plastoglobuli may provide these pigments with some protection from light damage as carotenoids localized in these lipid structures appear to have higher light stability than carotenoids resident in chloroplast membranes (Merzlyak

and Solovchenko, 2002). Carotenoids have also been shown to be sequestered in plastids as large sheets or crystals as demonstrated by studies of the cauliflower *Or* mutant, which exhibits an orange curd due to the accumulation of β-carotene in the plastids of this normally unpigmented tissue (Li et al., 2001; Lu et al., 2006).

In the nutritionally-important endosperm tissues of cereal crops such as maize, provitamin A carotenoids such as α- and β-carotene do not typically accumulate in amyloplasts, the plastids found in these tissues. Rather, in diverse cultivars of maize, the dominant carotenoids in endosperm are either lutein or zeaxanthin, or a combination of the two (Wurtzel et al., 2012). Wheat seed amyloplasts were also found to accumulate mainly lutein (Hentschel et al., 2002; Howitt et al., 2009).

Other plastids which accumulate carotenoids include leucoplasts and etioplasts. Trace levels of neoxanthin and violaxanthin accumulate in leucoplasts in root cells, which amount to less than 1% of the levels of these xanthophylls found in light-grown leaf tissue (Parry and Horgan, 1992). Etioplasts (found in dark-grown seedlings), harbor rudimentary membrane structures known as prolamellar bodies, which accumulate carotenoids such as lutein and violaxanthin that play vital roles in LHC assembly and photoprotection in mature photosynthetic chloroplast membranes (Joyard et al., 2009).

1.7 Carotenoid pathway regulation

Carotenoid accumulation in plants is regulated by various environmental and developmental cues. While no regulatory genes controlling carotenogenesis in any grass species have been characterized, it has been established that transcriptional regulation of the rate-limiting PSY in a variety of plant species is a major force driving carotenoid production (Wurtzel et al., 2012).

PSY is a single copy gene in *Arabidopsis*, but this enzyme is encoded by 3 genes in maize and other cereals such as rice and wheat (Wurtzel et al., 2012). Although functionally redundant, the PSY paralogs exhibit unique tissue specificities and differentially respond to abiotic stimuli. For example, while maize PSY1 drives endosperm carotenogenesis, it also plays a role in photosynthetic tissue where it is needed for carotenogenesis in the dark as well as for increased thermotolerance (Li et al., 2008b). Maize PSY3 expression was shown to be strongly influenced by salt and drought stress in root tissue; elevated PSY3 transcript levels correlated with an increase in carotenoid flux and ABA in roots (Li et al., 2008a). Rice PSY homologs exhibited similar responses (Welsch et al., 2008), which suggests that the maize data can be used to inform biofortification projects involving other cereals (Wurtzel et al., 2012).

Changes in PSY transcript levels are most marked during plastid development and differentiation (ie., from etioplast to chloroplast). During photomorophogenesis, PSY transcripts are up-regulated via a phytochrome-mediated (red-light) pathway (Welsch et al., 2000). *Arabidopsis* mutant studies revealed that in dark-grown plants the transcription factor phytochrome interacting factor 1 (PIF1), down-regulates

carotenogenesis by repressing PSY gene expression. Upon illumination, photoactivated phytochromes degrade PIFs during de-etiolation and PSY gene expression is subsequently induced, which leads to a rapid production of carotenoids in coordination with chlorophyll accumulation and chloroplast development (Toledo-Ortiz et al., 2010).

While PSY has been demonstrated to be the major rate-limiting enzyme of the pathway, other biosynthetic steps were thought to likely influence the size of the carotenoid pool. A maize germplasm diversity collection was analyzed to unveil other potential regulatory points in the pathway. These studies aimed to determine whether pathway gene transcript levels correlated with carotenoid content. Transcipt levels were statistically analyzed to identify gene family members that affect carotenoid content and composition as well as the time during endosperm development when this effect was observed (Li et al., 2008b; Vallabhaneni et al., 2009; Vallabhaneni and Wurtzel, 2009). These studies led to the identification of new targets for endosperm carotenoid biofortification efforts. A number of pathway bottlenecks for both carotenoid biosynthesis and for MEP precursor isoprenoid synthesis were identified in specific temporal windows of endosperm development. Transcript levels of genes which control steps that supply precursor isoprenoids including DXS3, DXR, HDR, and GGPPS1 positively correlated with carotenoid content in endosperm, while for carotenoid pathway enzymes, the transcripts of CrtISO, ZEP1, and ZEP2 inversely correlated with seed carotenoid content

(Vallabhaneni and Wurtzel, 2009); this is in contrast to the positive correlation of PSY1 transcripts (Li et al., 2008b; Vallabhaneni and Wurtzel, 2009).

The timing of expression for pathway controlling genes was shown to be a major factor in influencing carotenoid content. In maize endosperm, carotenoids accumulate continuously from 10 days after pollination (DAP) on as maturation of the seed occurs (Li et al., 2008b). Expression of MEP pathway genes which control steps that determine the supply of carotenoid pathway precursors correlated with carotenoid content at 25 DAP, while the carotenoid biosynthesis genes PSY1, CrtISO, and ZEP showed correlation earlier in development, at 20 DAP. This temporal difference may be due to the fact that in later stages of development, several pathways, including the carotenoid biosynthetic pathway, compete for isoprenoid precursors (Vallabhaneni and Wurtzel, 2009).

Another important source of carotenoid pathway regulation is allelic variation. The relative activities of the LCYE and LCYB enzymes regulate flux at the branch point of the pathway by modulating the ratio of lutein (on the α-branch of the pathway) to that of the β-branch carotenoids. Association analysis, linkage mapping, and expression analysis showed that variation at the *LCYE* locus resulted in changes in flux partitioning. The identification of four polymorphisms that controlled 58% of the variation between α- and β- branch carotenoid accumulation facilitates the selection of alleles which lead to enhanced provitamin A content for improved varieties of maize (Harjes et al., 2008).

Breeding strategies using select LCYE alleles to regulate increased flux through the β-branch of the pathway is only effective for generating crops with enhanced provitamin A content if there is a block in hydroxylation activity which leads to downstream non-provitamin A xanthophylls (ie., zeaxanthin). A maize germplasm diversity collection generated by "metabolite sorting" was used to identify target β-carotene hydroxylase genes. Two structurally distinct groups of hydroxylases (P450 CYP97 and diiron HYD) were found in maize to be encoded by a total of 8 genes (Vallabhaneni et al., 2009). Transcript levels of the carotene hydroxylase encoded by the *Hydroxylase 3 (HYD3)* locus were found to negatively correlate with high β-carotene levels and positively correlate with zeaxanthin levels. *HYD3* mapped close to a carotene quantitative trait locus (QTL). A PCR assay was developed to rapidly genotype 51 maize lines in order to test whether *HYD3* allelic variation could account for β-carotene variation in the maize germplasm collection. Three *HYD3* alleles in the 51 lines explained 36% of the variation and 4-fold difference in absolute levels of β-carotene. The identification of the *HYD3* locus as being responsible for the carotene QTL associated with β-carotene accumulation was also confirmed by association and linkage population studies in maize (Yan et al., 2010). Temperate maize varieties were found to contain the most favorable alleles; these alleles will be bred into tropical maize germplasm with the goal of alleviating vitamin A deficiency in developing countries (Wurtzel et al., 2012).

A systems biology approach using global microarray expression correlation analysis aimed to examine the mechanisms that regulate carotenoid biosynthesis as

well as the mechanisms which coordinate this synthesis with that of other isoprenoid-derived compounds, including ABA. These studies revealed that during later stages of seed development the coupled induction of the *Arabidopsis* carotene β-hydroxylase genes *βCHY1* and *βCHY2*, as well as *ABA1* with ABA biosynthesis genes coincide with ABA accumulation which suggests that *βCHY1* and *βCHY2*, and *ABA1* may serve to drive pathway intermediates towards β-xanthophyll and ultimately ABA production during these developmental stages (Meier et al., 2011). These data concord with the reduced *LCYE* expression observed in dry seeds as well as with studies which show that *Bchy1 Bchy2* double mutants have a reduced ability to generate ABA during drought stress (Tian et al., 2004). The *in silico* analysis by Meier et al., (2011) further revealed that while there is reduced expression of *βCHY1* and *βCHY2* and *ABA1* during imbibition and in dark grown plants, there is rapid induction of their expression upon illumination with expression of *βCHY1* reaching levels greater than two fold higher than that of *βCHY2*. The light-induced coupling of expression of these genes with other carotenoid pathway genes may point to the vital role they play in generating photoprotective β-xanthophylls (Meier et al., 2011).

1.8 Two classes of carotene hydroxylases: CYP97 and diiron HYD

As this dissertation focuses on the enzymes involved in carotene hydroxylation, the β- and ε-ring hydroxylases will be described here in more detail.

1.8.1 Structural distinctions and mechanisms of action

As mentioned previously, there are two classes of structurally distinct carotene β-hydroxylases in plants. β-ring hydroxylations may involve either non-heme diiron or heme-containing cytochrome P450-type β-hydroxylases. Diiron β-ring hydroxylases in plants are predicted to contain four transmembrane helices as well as four conserved histidine clusters which bind iron. The histidine residues coordinate the orientation of two iron atoms to form oxo-bridged diferric centers (Ryle and Hausinger, 2002), which are necessary for the formation of a substrate-based alkyl radical. The next step involves the hydroxylation of the substrate via an oxygen-rebound mechanism (Bertrand et al., 2005).

In plants, cytochrome P450 monooxygenases are involved in the production of various compounds such as phenylpropanoids, lipids, and phytohormones. Heme-containing NADPH-dependent monooxygenases are typically involved in catalyzing region- and stereospecific oxygenations of a variety of structurally diverse substrates through a mechanism that involves the activation of molecular oxygen. Due to their characteristic CO difference spectra with 450 nm Soret peaks, these hemoproteins are called cytochrome P450s (Inoue, 2004). The P450 enzymes have a P450 domain, and conserved oxygen-binding and heme-thiolate-binding signatures. The heme-thiolate-binding domain binds a single heme group with one iron atom in the center (Chapple, 1998; Danielson, 2002; Tian and DellaPenna, 2004). P450 hydroxylations, like diiron-type hydroxylations, are redox-sensitive. These reactions produce a substrate based alkyl-radical which is trapped by the OH· from the iron atom (Woggon, 2005).

1.8.2 Xanthophyll biosynthesis

Xanthophylls or oxygenated carotenoids are enzymatically synthesized oxidation products of the carotene substrates α- and β-carotene. Hydroxylation of the C3 and/or C3′ on one or both β-rings of β-carotene generates the intermediate β-cryptoxanthin (β-β-carotene-3-ol) and zeaxanthin (β-β-carotene-3,3′-diol). In contrast, hydroxylation of the ε-ring, β-ring, or both rings of α-carotene produces α-cryptoxanthin (β- ε-carotene-3′-ol), zeinoxanthin (β- ε-carotene-3-ol) and lutein (β- ε-carotene-3,3′-diol), respectively. The carotenes with β-rings have provitamin A activity, whereas their hydroxylated β-ring products do not. Since β-ring hydroxylation reduces vitamin A potential, β-hydroxylation activity is an important target for provitamin A biofortification projects (Cuttriss et al., 2011).

Lutein biosynthesis may involve the two classes of hydroxylases. Modification of the β-ionone ring of α-carotene may be performed via either a P450 or diiron-type β-ring hydroxylase, however ε-ring hydroxylation is catalyzed specifically by a P450 ε-hydroxylase. The two β-ionone rings of β-carotene may also be hydroxylated by either a P450 or diiron-type β-hydroxylase to form zeaxanthin (Quinlan et al., 2007).

It was generally thought that different enzymes catalyze ε- and β-ring hydroxylations since the chirality of the hydroxyl group on the ε-ring of lutein is opposite that of the hydroxyl group on the β-ring (Pogson et al., 1996; Sun et al., 1996; Cunningham Jr. and Gantt, 1998). An *Arabidopsis* non-heme diiron β-ring hydroxylase was first characterized via *E. coli* functional complementation studies by

(Sun et al., 1996). Subsequent genetic studies indicated that in *Arabidopsis*, two diiron-type hydroxylases (CHY1 and CHY2) primarily catalyze *in vivo* β-ring hydroxylations of β-β-carotenes (Tian and DellaPenna, 2004), while two heme-containing cytochrome P450 hydroxylases (CYP97A3 and CYP97C1) were preferentially active toward the β- and ε-rings respectively of β- ε carotenoids (ie., α-carotene) (Kim and DellaPenna, 2006). CYP97C1 (*LUT1* locus) and CYP97A3 (*LUT5* locus) were identified from phenotypes of knock-out *lut1* and *lut5* mutants in *Arabidopsis* (Tian et al., 2004; Kim and DellaPenna, 2006). In addition, enzyme activities for both CYP97A and CYP97C enzymes from rice were directly demonstrated in *E. coli* **(Chapter 2)** (Quinlan et al., 2007).

Genetic studies involving Arabidopsis quadruple mutants which lack all CYP97/HYD carotene hydroxylases indicated that these enzymes possibly have some overlapping activities (Kim et al., 2009). For example, triple mutants with only CYP97C1 (P450 ε-hydroxylase) or CYP97A3 (P450 β-hydroxylase) activity generated 74% and 6% of the WT levels of lutein. This research group inferred from these data that CYP97C1 is significantly active toward both β- and ε-rings of the α-carotene substrate, and that CYP97A3 also exhibits significant activity toward both the ε- and β-rings of α-carotene. These observations are not in agreement with results of *E. coli* complementation studies described here which indicated that the CYP97A4, when expressed as an individual enzyme (not in combination with the CYP97C2) in α- and β-carotene accumulating cells was only moderately active toward the β-rings of α-carotene to generate the intermediate zeinoxanthin; no lutein accumulated in

these cells (**Fig. 3-4**) (**Chapter 3**). In addition, the CYP97A4, when expressed alone in this complementation system, was preferentially active toward the β-rings of β-carotene to produce the end-product zeaxanthin (**Fig. 3-4**) (**Chapter 3**). Furthermore, the suggestion that the P450 ε-hydroxylase (CYP97C1) is efficiently active toward both β- and ε-rings does not comport with the findings in *E. coli*, in which preliminary studies showed that the rice P450 ε-hydroxylase (CYP97C2) was exclusively active towards ε-rings of ε- ε-carotene (**Fig. 2-6**) (**Chapter 2**) (Quinlan et al., 2007); subsequent complementation studies indicated that when expressed as an individual enzyme the CYP97C2 was barely active toward both the ε- and β-rings of α- and β-carotene to generate barely detectable amounts of the intermediates α-cryptoxanthin and β-cryptoxanthin respectively (**Fig. 3-4**) (**Chapter 3**); no lutein accumulated in cells expressing the CYP97C2 alone. The apparent overlapping carotene hydroxylase activities observed in the Arabidopsis triple mutants may be due to pleiotropic effects with respect to the activities of other enzymes in the pathway. In addition, these results may point to the existence of an as yet unidentified additional carotene hydroxylase which could potentially compensate for the loss of CYP97A or CYP97C function in the triple mutants. Perhaps this unknown hydroxylase interacts to form a protein complex with either the CYP97A or CYP97C enzyme which would significantly enhance their respective activities. Furthermore, this unknown hydroxylase may require interaction with one or more of the other known missing carotene hydroxylases in the quadruple mutant for activity, which would explain the lack of hydroxylated products in these mutants. The activities and

specificities, as well as the respective roles of the CYP97/HYD enzymes in metabolon biogenesis and substrate channeling in the carotenoid pathway are described in detail in chapter 3.

Both maize and rice contain one gene each for CYP97A and CYP97C, respectively, and phylogenetic analysis suggests that both the maize and rice CYP97 enzymes behave similarly (Chapter 2) (Quinlan et al., 2007; Vallabhaneni et al., 2009). Studies by Vallabhaneni et al., (2009) identified six unlinked paralogs which code for the diiron β-hydroxylases (*HYD*) in maize, while only three *HYD* genes have been identified in rice. Of the six maize paralogs, *HYD1* and *HYD2* were found to be pseudogenes, while *HYD3*, *HYD4*, *HYD5*, and *HYD6* code for enzymes that contain conserved hydroxylase domains and plastid-targeting signal peptide motifs. Maize *HYD3* and *HYD4* were found to be syntenous with rice *HYD1*. In addition, functional testing in *E. coli*, which successfully demonstrated function of CYP97A and CYP97C from rice as described previously, also showed that both *HYD3* and *HYD4* encode functional carotene β-hydroxylases (Vallabhaneni et al., 2009).

Since hydroxylation the β-ionone rings of carotenes reduces vitamin A potential, β-hydroxylase activity is an important provitamin A biofortification target in cereal endosperm. As described previously (section 1.7), favorable alleles of the *HYD3* locus in maize were discovered in a large germplasm collection for predicting enhanced β-carotene composition (Vallabhaneni et al., 2009).

1.8.3 Evolution

From an evolutionary standpoint, it is essential to consider the driving forces responsible for the emergence of the two distinct classes of carotene hydroxylases (P450- and diiron-type) in plants.

β-xanthophylls (ie. zeaxanthin) are ubiqitous in nature, and are found in all photosynthetic eukaryotes as well as in many photosynthetic and non-photosynthetic prokaryotic organisms. The diiron-type β-hydroxylases are thought to be primarily responsible for hydroxylation of β-carotene in these different organism groups. In green plant lineage, two *Arabidopsis* diiron β-hydroxylase isozymes, as well as the *Arabidopsis* P450-type hydroxylases (CYP97A and CYP97C), were shown to have relatively strong expression correlation in photosynthetic tissues of non-stressed plants (Kim et al., 2009); this suggests that both the diiron- and P450-type hydroxylase most recent common ancestor shared a similar activity and was expressed in tissues that performed photosynthesis. Double mutants for the diiron-hydroxylase isozymes exhibited negatively impacted NPQ as well as Fv/Fm, which indicated that the diiron β-hydroxylase likely played a vital role in photoprotection for the diiron-hydroxylase containing ancestor and that an intense selection pressure was at work for retention of this function throughout evolution (Tian and DellaPenna, 2001; Tian et al., 2003; Kim et al., 2009).

α-xanthophylls (ie., lutein) are synthesized exclusively in green algae and plants, as well as in some red algae. Lutein is the most abundant carotenoid in photosynthetic tissues of higher plants (Pogson et al., 1996; Galpaz et al., 2006). The fact that lutein accumulates in higher plants and not in cyanobacteria suggests that

evolution of the P450 CYP97 ε- and β-ring hydroxylases post-date that of the diiron-type β-ring hydroxylase. In addition, the conservation of α-xanthophyll biosynthesis during green algae and plant evolution indicates that powerful selection pressures were at work in the production and maintenance of the enzyme activities of the ε-cyclase (LCYE) and P450-type β- and ε-ring hydroxylases (CYP97A and CYP97C) which are required for α-xanthophyll synthesis (Kim et al., 2009).

Lutein plays an essential role in the assembly of light harvesting complexes (LHC) in the photosynthetic apparatus (Pogson et al., 1998; Lokstein et al., 2002; Dall'Osto et al., 2006; Dall'Osto et al., 2007). It is likely that the requirement for efficient photosystem structure and function would serve as the driving force for selection of P450 ε- and β-ring hydroxylase (CYP97-like) activity; this is in contrast to the photooxidation driven selection of the diiron-type β-ring carotene hydroxylases. In *Arabidopsis*, double diiron hydroxylase mutants showed negatively impacted NPQ which suggests that these enzymes have a vital photoprotective function and that retention of this function in the diiron hydroxylase-containing ancestor was at work throughout evolution. (Tian and DellaPenna, 2001; Tian et al., 2003; Kim et al., 2009).

In higher plants, the CYP97 P450 hydroxylases are ubiquitously encoded by single copy genes whereas there are multiple genes encoding the diiron-type hydroxylases in many species, including representatives of both monocots and eudicots. For example, in rice, as well as in *Arabidopsis*, there appear to be at least two diiron β-hydroxylases (Quinlan et al., 2007), and in maize there are 6 paralogs (2

functional and 4 pseudogenes) that code for the diiron β-hydroxylase enzyme (Vallabhaneni et al., 2009). Phylogenetic analysis indicates that the diiron β-hydroxylase gene duplications occurred relatively recently, after the monocot/dicot split (Kim et al., 2009). The products of these multiple diiron hydroxylase paralogs may be expressed in different tissues or in response to various environmental cues (Wurtzel, 2004; Castillo et al., 2005; Vallabhaneni et al., 2009). Genetic studies by Galpaz et al., (2006) of the tomato white flower mutant (*wf*), revealed that the wf phenotype is caused by mutations in a chromoplast-specific diiron β-hydroxylase (CrtR-b2); this is the second of two diiron β-ring hydroxylases found in tomato. The other diiron β-hydroxylase, CrtR-b1, is constitutively expressed in leaves. The flower-specific diiron β-hydroxylase CrtR-b2 revealed by the wf mutation, as well as flower- and fruit-specific GGPPS, PSY, and LCYB enzymes together define a chromoplast-specific carotenoid biosynthetic pathway, which points to the important role of gene duplication in such specialized pathways. This research group concluded that the duplication and preservation CrtR-b2, the second β-hydroxylase in tomato, was due to strong selection pressure for the generation of yellow pigments in flowers (Galpaz et al., 2006).

1.9 The xanthophyll cycle

In order to respond effectively to ever-changing light conditions, plants have evolved a variety of regulatory mechanisms. The xanthophyll cycle is of central importance in the regulation of photosynthesis to avoid photo-oxidative damage of

the photosynthetic apparatus due to the generation of reactive oxygen species under conditions of excess light . When plants absorb an excess of light energy that exceeds photosynthetic capacity, the xanthophyll violaxanthin, which is present in the photosynthetic apparatus, is reversibly de-epoxidized into zeaxanthin in the xanthophyll cycle. The de-epoxidation of violaxanthin into zeaxanthin is catalyzed by the enzyme violaxanthin de-epoxidase (VDE) which is located in the thylakoid lumen. Under conditions of high light, when energy input exceeds photosynthetic capacity, acidification of the thylakoid lumen activates VDE. This results in a shift in the xanthophyll balance from violaxanthin, which functions in light-harvesting, toward zeaxanthin, which plays a central role in the dissipation of excess absorbed energy as heat. When conditions allow for an increase in lumen pH (ie., low-light), zeaxanthin is converted back to violaxanthin via the action of zeaxanthin epoxidase. In this way the xanthophyll cycle is regulated so as to prevent unnecessary quenching of excitation energy in conditions of low light (Jahns and Holzwarth, 2012)

1.10 Objectives

The endosperm tissues of staple crops such as maize and rice are low in provitamin A content. In the developing world, the consumption of carotenoid-poor cereals is associated with vitamin A deficiency which is a global health burden that places up to 140 million children at risk annually for visual impairment, blindness,

and other health problems (ie., depressed immunological responses) (Sommer and Davidson, 2002).

Efforts to improve the provitamin A content of cereal endosperm in staples such as maize will require elucidation of the mechanisms involved in controlling the conversion of provitamin A carotenes such as α- and β-carotene to non-provitamin A xanthophylls (ie., lutein and zeaxanthin). In order to gain a better understanding of this conversion, the work described herein examined the activities and localization/interaction of both maize and rice representatives from the two known structurally distinct classes of carotene hydroxylases: the nonheme diiron-type hydroxylases from maize (HYD3 and HYD4) and the cytochrome P450-type hydroxylases CYP97A and CYP97C from rice.

1.10.1 Hypotheses

The P450 CYP97A and CYP97C enzymes are preferentially active on the β-ε "arm" of the carotenoid pathway to convert α-carotene to lutein. The diiron HYD enzymes are preferentially active on the β-β "arm" of the pathway to convert β-carotene to zeaxanthin.

1.10.2 Specific aims

The goals of this research were to identify the optimal substrate(s) for the CYP97/HYD carotene hydroxylases, to determine whether any of the CYP97/HYD enzymes require an interacting partner enzyme for optimal activity, and to establish

their respective subcellular/suborganellar locations. In order to achieve these aims, the following studies were conducted:

(1) Assessment of CYP97 activities and specificities (Chapter 2)

Preliminary investigations of the activities and substate specificites of the CYP97A4 and CYP97C2 enzymes were conducted via *E. coli* functional complementation assays which included substrates with only either β-rings (ie., β-carotene) or ε-rings (ie., ε- ε-carotene).

(2) Assessment of both individual and combined CYP97/HYD enzyme activities and specificities (Chapter 3)

The CYP97A4, CYP97C2, and diiron HYD3 and HYD4 enzymes were assayed individually, as well as in various combinations for their respective activities and specificities toward both α- and β-carotene substrates via bacterial functional complementation.

(3) Investigation of CYP97/HYD subcellular/suborganellar locations (Chapter 3)

Both *in vitro* chloroplast import and *in vivo* GFP fusion assays were used to determine whether these enzymes are imported into chloroplasts. In addition, import studies were performed to establish whether they form peripheral or integral associations with chloroplast membranes.

(4) Investigation of CYP97/HYD protein-protein interactions (Chapter 3)

Bimolecular Fluorescence Complementation (BiFC) assays were used to assess CYP97/HYD protein interactions.

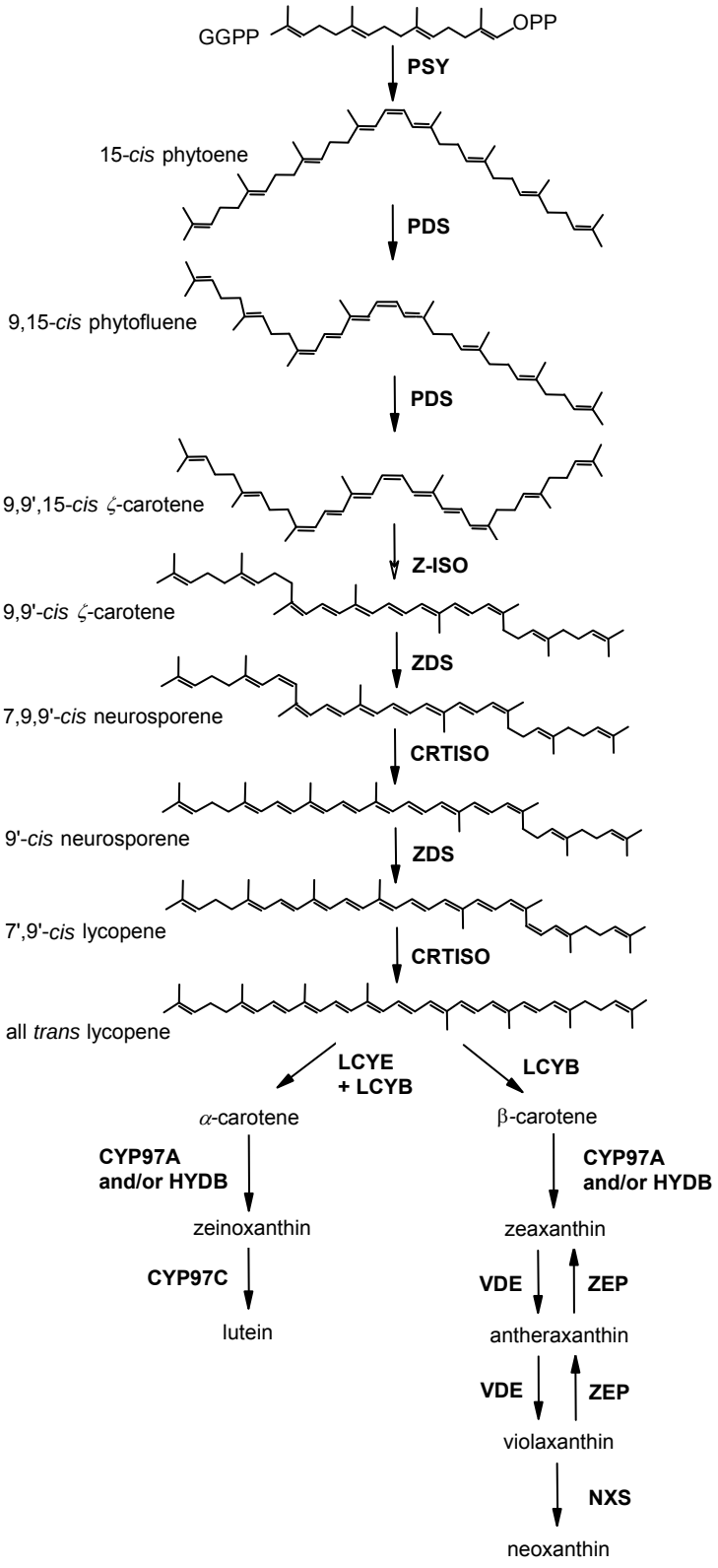


Figure 1-1. Carotenoid biosynthetic pathway in plants.
Figure prepared by Rena Quinlan. Cited in: (Matthews and Wurtzel, 2007)

CHAPTER 2

Escherichia coli as a platform for functional expression of plant P450 carotene hydroxylases

2.1 Abstract

Carotenoids and their derivatives are essential for growth, development, and signaling in plants and have an added benefit as nutraceuticals in food crops. Despite the importance of the biosynthetic pathway, there remain open questions regarding some of the later enzymes in the pathway. The CYP97 family of P450 enzymes was predicted to function in carotene ring hydroxylation, to convert provitamin A carotenes to nonprovitamin A xanthophylls. However, substrate specificity was difficult to investigate directly in plants, which mask enzyme activities by a complex and dynamic metabolic network. To characterize the enzymes more directly, cDNAs were amplified from a model crop, *Oryza sativa*, and functional complementation in *Escherichia coli* was performed to test activity and specificity of members of Clans A and C. This heterologous system will be valuable for further study of enzyme interactions and substrate utilization needed to understand better the role of CYP97 hydroxylases in plant carotenoid biosynthesis.

---

*This chapter was taken from: Quinlan R.F., Jaradat T.T., Wurtzel, E.T. (2007) *Escherichia coli* as a platform for functional expression of plant P450 carotene hydroxylases. Arch. Biochem. Biophys. **458:** 146-157 (www.sciencedirect.com) with permission from the publisher (Elsevier)

2.2 Introduction

Carotenoids are an abundant class of isoprenoid compounds found in all plants where they are synthesized in plastids by nuclear-encoded enzymes. Their numerous roles in plant growth and development include functions as accessory pigments in photosynthesis and as photoprotectors (Fraser and Bramley, 2004; Wurtzel, 2004; DellaPenna and Pogson, 2006). Carotenoids such as β-carotene and zeaxanthin, among others, serve as precursors to cleavage products including the hormone abscisic acid (ABA) (Milborrow, 2001) and to other apocarotenoids, some of which are gaining attention for their roles in internal and external signaling (Bouvier et al., 2003; Booker et al., 2004; Schwartz et al., 2004; Simkin et al., 2004; Castillo et al., 2005; Matusova et al., 2005; Moise et al., 2005). The importance of certain carotenoids in human health has led to efforts to breed or metabolically engineer carotenoid content and composition (Shewmaker et al., 1999; Mann et al., 2000; Rosati et al., 2000; Ye et al., 2000).

Plants are an important dietary source of provitamin A carotenoids; predictable efforts to use breeding of vitamin-rich crops to address human vitamin A deficiency will require elucidation of those mechanisms controlling conversion of provitamin A carotenes to nonprovitamin A xanthophylls. For example, in maize endosperm, there is wide variation in carotenoid content and composition (Kurilich and Juvik, 1999; Islam, 2004) making this tissue a target for improvement of carotenoid content and provitamin A levels (Wong et al., 2002; Wong et al., 2004;

Wurtzel, 2004). Like maize, wheat is another of many crops in the Poaceae that exhibit diversity in yellow seed color and are attractive breeding targets for manipulating endosperm carotenoids (Parker et al., 1998; Zhang et al., 2005). Endosperm carotenoids are synthesized and accumulate on amyloplast envelope membranes whereas in chloroplasts, carotenoids are found on envelope and thylakoid membranes. Targeting of carotenoid enzymes to specific plastid membranes and metabolon biogenesis/maintenance are poorly understood phenomena, especially given that many carotenoid enzymes are encoded by single copy genes and are destined for multiple suborganellar locations. Moreover, there is limited understanding of how pathway intermediates are channeled to downstream products including photosynthetic pigments or signaling molecule apocarotenoids.

In all plants, the biosynthesis of carotenoids begins with the formation of the 40-carbon phytoene, followed by desaturation steps leading to synthesis of lycopene, after which point the pathway diverges to form either β-carotene (having two β rings) or α-carotene (having one β-ring and one ε-ring) (**Fig. 2-1**). Further hydroxylation of the carotenes leads to biosynthesis of the xanthophylls. For example, hydroxylation of the C3 and/or C3′ on one or both β-rings of β-carotene leads to β-cryptoxanthin (β,β-carotene-3-ol) and zeaxanthin (β,β-carotene-3,3′-diol). In contrast, hydroxylation of the ε-ring, β-ring, or both rings of α-carotene generates α-cryptoxanthin (β,ε-carotene-3′-ol), zeinoxanthin (β,ε-carotene-3-ol), and lutein (β,ε-carotene-3,3′-diol), respectively (Cunningham Jr. and Gantt, 1998). The carotenes with β-rings have

provitamin A activity, whereas their hydroxylated β-ring products do not (von Lintig and Vogt, 2004).

In plants, the hydroxylation of provitamin A carotenes to form non-provitamin A xanthophylls is thought to be mediated by two structurally distinct classes of mixed function oxygenases. The first class of hydroxylases is comprised of the β-ring non-heme diiron monooxygenases. Members of this class were identified from a wide range of bacterial, algal, and plant species (Sun et al., 1996; Bouvier et al., 1998; Tian and DellaPenna, 2004). Plant diiron β-ring hydroxylases contain four transmembrane helices and four conserved iron-binding histidine clusters. Histidine residues in these clusters coordinate the orientation of two iron atoms to form oxo-bridged diferric centers (Ryle and Hausinger, 2002), which are involved in the formation of a substrate based alkyl-radical and the subsequent hydroxylation of the substrate via an oxygen-rebound mechanism (Bertrand et al., 2005). Enzyme activity of the diiron β-ring hydroxylases has been demonstrated in an *E. coli* functional complementation system (Sun et al., 1996). The second class consists of P450 enzymes from the CYP97 clan, which on the basis of genetic evidence, are hypothesized to hydroxylate carotene ε- and β- rings, respectively (Tian et al., 2004; Kim and DellaPenna, 2006). These predicted P450 heme-thiolate hydroxylases contain a single transmembrane anchoring sequence, typical of other eukaryotic P450 enzymes (Chapple, 1998; Danielson, 2002). In addition, each enzyme has a P450 domain, a conserved oxygen binding signature, and a conserved heme-thiolate binding signature that binds a single heme group with one iron atom in the center (Chapple, 1998; Danielson, 2002; Tian

and DellaPenna, 2004). Similar to diiron hydroxylation reactions, P450 hydroxylations are also redox-sensitive and involve the formation of a substrate based alkyl-radical that is immediately trapped by the HO• from the iron atom (Woggon, 2005). A major limitation for the genetic evidence used to support substrate specificity of the CYP97 carotene hydroxylases is that biochemical profiling of mutants is complicated by poorly understood compensatory mechanisms, including changes in transcriptional activity of genes encoding other carotenoid biosynthetic enzymes such as the nonheme diiron carotene hydroxylases, which may ultimately impact pathway flux (Tian et al., 2004; Kim and DellaPenna, 2006).

Progress in characterizing the full complement of enzymes and identifying the corresponding genes has been significantly advanced with use of *E. coli* to functionally test activity of the membrane bound carotenoid pathway enzymes, an approach adopted by the carotenoid community [e.g. (Misawa et al., 1990; Sun et al., 1996; Sandmann et al., 1999; Matthews and Wurtzel, 2000; Isaacson et al., 2002; Gallagher et al., 2004)]. The bacterial system could potentially provide a simple platform to manipulate and study the CYP97 carotene hydroxylases. Therefore, cDNAs were isolated from *Oryza sativa,* a representative member of the Poaceae, and *E. coli* complementation was used to demonstrate activity and investigate carotene ring specificity of the putative P450 β- and ε- ring hydroxylases. Since only a subset of P450 enzymes function in *E. coli*, the widely used heterologous bacterial system was not necessarily a feasible approach. (Bak et al., 1998; Naur et al., 2003; Schuler and Werck-Reichhart, 2003). This work demonstrates that these members of the

CYP97 clan function in *E. coli* and indicates that this system will be valuable in further dissecting the structural basis for ring specificity and other molecular applications. Given that β-ring hydroxylation is alternatively possible with either CYP97 or diiron enzymes raises the question of which are the appropriate genes/enzymes to choose as breeding targets or in metabolic engineering of this pathway; whether it is to achieve enhanced levels of endosperm provitamin A carotenoids in food crops, or increased abiotic and/or biotic resistance in plants in general.

2.3 Results

2.3.1 Phylogenetic analysis

Among all known CYP97 P450 enzyme family members, three distinctive, yet closely related, clans (*i.e.*, A, B, and C) can be discerned (Nelson et al., 2004). The amino acid alignment and phylogenetic tree construction for a selection of thirteen protein sequences from seven different plant species (monocots and dicots) is in agreement with the three clans designation (**Fig. 2-2A**). Clan A and C enzyme sequences are more closely related with three blocks of amino acid sequence insertions that differentiate members of Clan B from those in Clans A and C (**Fig. 2-2B**). Block 1 is about 100 residues downstream of the P450 domain; blocks 2 and 3 are 37 aa apart and located between the oxygen- and heme thiolate-binding signatures. Members of the CYP97 clan, Clan B, do not have any functionally

demonstrated representatives and as shown in the sequence comparisons, are more distant from Clans A and C. Clan B sequences do possess a potential chloroplast-targeting sequence (Emanuelsson et al., 1999) and transcripts are present in shoots as reflected by ESTs, suggesting localization in chloroplasts as predicted for Clan A and C enzymes. Clan B homologs are found in angiosperms (**Fig. 2-2**), in *Ginkgo biloba* (AAT28222) and the diatom, *Skeletonema costatum* (AAL73435) (Yang et al., 2003). This taxonomically broad distribution suggests that Clan B enzymes may function in all plants rather than catalyzing synthesis of taxon-specific metabolites.

To test carotene hydroxylase activity of Clan A and C enzymes, sequences that were most closely related to Arabidopsis LUT1 (CYP97C1) were identified using the genome of *Oryza sativa*, a representative of the Poaceae, and amplified cDNAs from *O. sativa* L. cv Nipponbare leaves. The *O. sativa* AK065689 and AK068163 cDNAs, correspond to CYP97C2 (P450 HydE, putative ε-ring carotene hydroxylase) and CYP97A4 (P450 HydB, putative β-ring carotene hydroxylase). They encode conceptually translated proteins sharing 84% and 76% similarity (excluding the non-conserved transit peptide sequences) with the *A. thaliana* CYP97C1 (LUT1) and CYP97A3 (LUT5) (Kim and DellaPenna, 2006) candidates for ε-ring and β-ring hydroxylases. Rice CYP97A4 and CYP97C2 map to chromosomes 2 and 10, respectively. CYP assignment of rice clones is based on BLAST searches using the Rice P450 BLAST Server, which implements the standardized system of cytochrome P450 nomenclature (Nelson et al., 2004).

2.3.2 P450 DNA sequence analysis

The *O. sativa* CYP97A4 (*P450 HydB)* is comprised of 15 exons and 14 introns, while the CYP97C2 (P450 *HydE*) gene consists of 10 exons and 9 introns (**Fig. 2-3A, B**). All intron boundaries are consistent with the conserved 5′-AG and 3′-GT flanking element rule (Hanley and Schuler, 1988). The CYP97C2 cDNA is described in the NCBI GenBank # AK065689 as a 1876 bp long cDNA with a 1683 bp open reading frame (ORF), a 104 bp 5′- untranslated region (5′-UTR), and an 89 bp 3′-UTR (Kikuchi et al., 2003). The CYP97A4 cDNA (GenBank #AK068163) is indicated as a 4217 bp long cDNA with a 1929 bp ORF, a 53 bp 5′-UTR, and a 2235 bp 3′-UTR (Kikuchi et al., 2003). However, analysis of AK068163 cDNA and corresponding genomic sequence (GenBank # NT_107182) support a 2414 bp full-length cDNA with a 1929 bp ORF, an 88 bp 5′-UTR, a 397 bp 3′-UTR with conserved near-upstream element (NUE; poly-A addition signal), far upstream element (FUE), and terminal "Pyr(T)A" poly-A site (**Fig. 2-3C**) (Kozak, 1987; Rothnie, 1996; Zhao et al., 1999). The modified full-length cDNA structure (**Fig. 2-3C**) is based on the finding that several rice 3′-end ESTs map around the proposed poly-A addition site, whereas there are no ESTs that map to the 3′-end of the model presented by Kikuchi and coworkers (Kikuchi et al., 2003).

2.3.3 Conserved domain characteristics of the CYP97C2 and CYP97A4 amino acid sequences

Significant *E* values from PSI- and PHI-BLAST searches against the conserved domain database (CDD) at NCBI confirm that CYP97C2 and CYP97A4 are P450 proteins (Marchler-Bauer et al., 2005). The conserved heme-thiolate-binding signature (FXXGXXXCXG) in plants corresponds to Phe_{488}-Ser-Gly-Gly-Pro-Arg-Lys-Cys-Val-Gly_{497} in CYP97C2 and to Phe_{534}-Gly-Gly-Gly-Pro-Arg-Lys-Cys-Val-Gly_{543} in CYP97A4 and as compared with those of other related plant enzymes (**Fig. 2-3D**). The underlined conserved cysteine in each of these motifs is an essential residue since it contributes the thiol group, which is the fifth ligand that binds the iron atom in these hemoproteins. Similarly, near the middle of the P450 conserved domain, a conserved oxygen binding signature (A/G)GX(D/E)T(T/S) is detected with the corresponding amino acid sequence of Ala_{361}-Gly-His-Glu-Thr-Thr_{366} for the CYP97C2 and Ala_{402}-Gly-His-Glu-Thr-Ser_{407} for the CYP97A4 (Fig. 3D) and as compared with those of other related plant enzymes. The underlined threonine in these respective sequences is involved in the binding of an oxygen molecule, which is essential for catalysis.

The overwhelming majority of eukaryotic P450 heme-thiolate enzymes are membrane bound through an anchoring trans-membrane helix. Conversely, all P450s in prokaryotes are cytosolic. CYP97C2 is predicted to be a 62.1 kDa (561 residue) preprotein with a 40 residue N-terminal transit peptide that is cleaved upon import to yield a chloroplast-localized 57.8 kDa, 521 residue, mature protein (Emanuelsson et al., 1999). CYP97A4 is predicted to be a 69.9 kDa (643 residue) preprotein containing a 41 residue N-terminal transit sequence that is removed upon chloroplast

import to yield a chloroplast-localized 65.9 kDa, 602 residue, mature protein. In addition, both enzymes harbor a trans-membrane helix sequence at the N-terminus of the P450 domain, as predicted by the HMMTOP server (Tusnady and Simon, 1998) to membrane anchor the mature P450 protein.

2.3.4 Heterologous expression of CYP97 genes in *E. coli* and substrate specificity

CYP97A4. For functional testing of β-ring carotene hydroxylase activity, *E. coli* cells were engineered as described in the methods to produce the β-carotene (β,β-carotene) substrate, which contains two β-rings for potential hydroxylation. For testing of a Clan A enzyme (a putative β-ring carotene hydroxylase), the full-length rice CYP97A4 cDNA (including its putative transit sequence) was amplified and cloned into pCOLADuet^{TM}-1. The resulting construct was introduced into the β-carotene-accumulating cells. Accumulated pigments from transformed cells with or without the cDNA (**Fig. 2-4A-e and 2-4A-d**) along with standards for β-carotene (**Fig. 2-4A-c, peak 3**) and hydroxylated products, β-cryptoxanthin (one hydroxylated ring, **Fig. 2-4A-b, peak 2**) and zeaxanthin (two hydroxylated rings, Fig. **2-4A-a, peak 1**) were analyzed by HPLC. Extracted pigments were identified based on HPLC retention times (**Fig. 2-4A**) and spectra (**Fig. 2-4B**) matching those of authentic standards. As seen in **Figure 2-4A-e**, only the doubly transformed cells accumulated β-cryptoxanthin (peak 2) and zeaxanthin (peak 1), in addition to the substrate β-carotene (peak 3). Cells that lacked the CYP97A4 cDNA did not have any hydroxylated β-carotene derivatives (**Fig. 2-4A-d**).

The effect of culturing temperature and IPTG concentration was then tested on CYP97A4 hydroxylation activity in *E. coli*. At 21°C and 10 mM IPTG, ~62% hydroxylated product accumulated, of which 52% was β-cryptoxanthin (monohydroxy intermediate) and 10% was zeaxanthin (di-hydroxy product) (**Fig. 2-5A**), a combined equivalent of 36% total hydroxylated β-rings (**Fig. 2-5B**). Increasing the culturing temperature from 21°C to 37°C increased production of β-carotene substrate but reduced the amount of hydroxylated products by 20% at 5 mM IPTG (data not shown) and by 40% at 10 mM IPTG (**Fig. 2-5**). Initial attempts using overnight cultures grown at 37 °C and lower concentration of IPTG (i.e., 0.4 mM and 1 mM) did not result in accumulation of detectable xanthophylls. Interestingly, the ratio of β-cryptoxanthin (one hydroxylated ring) to zeaxanthin (two hydroxylated rings) decreased from 5:1 for growth at 21°C to almost equimolar amounts for growth at 37°C. In summary, CYP97A4 is more effective at lower temperature (21°C) when expressed in *E. coli*, an organism having optimal growth at 37°C.

CYP97C2. Subsequent assays tested whether a Clan C enzyme, a putative ε-ring carotene hydroxylase, could utilize the β-ring substrate. The rice CYP97C2 cDNA was similarly amplified and cloned, missing only 7 codons of the predicted 40 residue transit sequence, into pCOLADuet^{TM}-1. This construct was introduced into β-carotene accumulating cells; the pigment profile was compared with that obtained from CYP97A4 double transformed cells as above. The HPLC chromatogram (**Fig. 2-6A**, left panel, bottom) shows that when CYP97C2 was introduced into β-carotene accumulating cells, β-carotene substrate (peak 3) was the sole accumulating pigment

as seen in the control (**Fig. 2-6A**, left panel, top, peak 3) and no hydroxylated products were detected. In comparison, introduction of CYP97A4 cDNA resulted in the mono- and di-hydroxylated products (left panel, middle, peaks 1 and 2).

In contrast, if the CYP97C2 cDNA was introduced into cells engineered with ε-ring substrates δ-carotene and ε,ε-carotene (**Fig. 2-6A** right panel, top, peaks 6 and 7), it was observed that CYP97C2 did confer hydroxylation of ε-rings. This was seen by accumulation of the more polar peaks including lactucaxanthin, a di-hydroxylated product (**Fig. 2-6A**, right panel, bottom, peak 4) with retention time and corresponding spectrum matching the authentic lactucaxanthin standard chromatographed in the same HPLC system (**Fig. 2-6B**). Hydroxylated products accumulated only when cells were grown at 37°C and not at lower temperatures (data not shown). When the β-ring hydroxylase CYP97A4 was introduced into cells with the ε-ring substrates, there was a barely detectable amount of accumulated lactucaxanthin (**Fig. 2-6** right panel, middle, peak 4). Spectra corresponding to the small bumps seen in the control (empty vector) do not match those for peak 4 nor for any carotenoid. Therefore, lactucaxanthin is only detected when the plant enzymes are present. These results demonstrate that the Clan A enzyme hydroxylates carotene β- rings and has weak activity towards carotene ε-rings while the Clan C enzyme only hydroxylates carotene ε-rings.

2.4 Discussion

2.4.1 Phylogenetic analysis of CYP97 genes

Predictive metabolic engineering or marker-based breeding of enhanced provitamin A levels depends on controlling conversion of provitamin A carotenes to non-provitamin A xanthophylls in endosperm and other tissues. To assess this conversion requires that there is a means to investigate activities and specificities of the putative hydroxylase enzymes. This work focused on the recently discovered CYP97 family encoding P450 enzymes for which enzyme activities have been implied but never demonstrated. An enzyme tree of CYP97 P450 enzymes was first constructed to identify candidate sequences for ε- and β-ring hydroxylases in rice, a member of the Poaceae and representative of many agronomically important crops. Enzymes clustered into three clans, A, B, and C, and appeared to be encoded by single copy genes based on the finding that in each case, only one orthologous sequence was found in each of four different plant species, two being among the most exhaustively sequenced genomes (i.e., *A. thaliana* and *O. sativa*).

2.4.2 Functional complementation results integrated with earlier genetic studies

These studies involved the isolation of cDNAs encoding both β- and ε-ring hydroxylases in the *Oryza sativa* CYP97 family and demonstrated that *E. coli* cells were suitable for expression of functional enzymes to allow testing of substrate specificity. A preliminary examination of substrate specificity was conducted with regard to β- and ε-ring-containing substrates. Using *E. coli* functional complementation, it was demonstrated that a Clan A enzyme, CYP97A4, is a β-ring

carotene hydroxylase with some minor activity towards ε-rings. This minor ε-ring activity of the Clan A enzyme likely contributes to the residual lutein that accumulated in Arabidopsis plants carrying a *lut1* null allele that blocks expression of a Clan C ε-ring hydroxylase (Kim and DellaPenna, 2006). The phenotype of the null *lut1* allele also indicates that the minor ε-ring activity observed in the *E. coli* system is consistent with minor activity in the plant; almost no lutein accumulates in the null mutant and therefore the minor activity of the Clan A enzyme is indeed minor and cannot compensate for the Clan C deficiency. Studies on the nonheme diiron β-ring hydroxylases also showed the enzymes to display some activity towards ε- rings (Sun et al., 1996). Given this observation in two structurally distinct enzymes (the P450 and diiron β-ring hydroxylases) suggests that this plasticity may be a function of the relationship between the extracted hydrogen and the configuration of the double bond in the β-ring as compared to the ε-ring (Tian et al., 2004).

Previous knockout mutagenesis in *A. thaliana* of the two genes encoding non-heme diiron β-carotene hydroxylases and one encoding the ε-ring hydroxylase (*LUT1*, a Clan C enzyme) (Tian et al., 2003) led the authors to suggest that the LUT1 ε-ring hydroxylase may be active towards β-rings. Using *E. coli,* it was shown that a Clan C enzyme, CYP97C2, lacked detectable activity towards β-rings. Therefore, these results do not support the hypothesis that further reduction in β-ring hydroxylation caused by a lesion in *Lut1* triple mutants compared to defects in the diiron enzymes alone, is attributed to a defect in the Clan C enzyme (which did not exhibit such activity under the conditions tested). Therefore, the observation in the

triple mutant must have some other explanation. One possibility is that the Clan C and A enzymes form a complex, and interference in the expression of one affects expression/stability of the other, and hence the additional impact on β-ring hydroxylation activity in the triple mutants.

When comparing the activity of the rice enzymes in *E. coli*, it was noted that the carotene ε-ring hydroxylation mediated by the Clan C enzyme was weak. One possibility is that the ε-ring substrate (one or two ε-rings) is not the best substrate as suggested by biochemical profiles of plant CYP97 mutants. Arabidopsis mutants defective in a Clan C ε-ring hydroxylase (*lut1*) conditioned accumulation of zeinoxanthin (hydroxylated β-ring of α-carotene) while plants defective in a Clan A β-ring hydroxylase (*lut5*) accumulated α-carotene and not α-cryptoxanthin (hydroxylated ε-ring of α-carotene) which would be expected if the available Clan C ε-ring hydroxylase could accept α-carotene directly. These combined observations suggest that the preferred *in planta* substrate of the Clan C ε-ring hydroxylases is zeinoxanthin (the mono-hydroxylated β-ring) and that the enzymes function in an ordered manner with the Clan A enzyme hydroxylating the carotene β-ring to produce zeinoxanthin followed by the Clan C enzyme hydroxylating the ε-ring (Fiore et al., 2006; Kim and DellaPenna, 2006). Therefore, Clan C enzyme activity in *E. coli* might be improved with another substrate, the monohydroxylated β-ring of α-carotene (zeinoxanthin) and/or association with the Clan A enzyme. Enhancement of Clan C activity, by co-expression of the Clan A enzyme needed to produce

zeinoxanthin, is now testable given demonstrated utility of the *E. coli* platform to express each individual enzyme.

It was also observed that expression of the Clan A enzyme in *E. coli* led to significant accumulation of β-cryptoxanthin, even when different conditions were tested; this compound is not usually found in Arabidopsis, although it is found in other tissues of other species such as maize (Islam, 2004). While this work demonstrated hydroxylation of β-rings in *E. coli* by Clan A enzymes using β-carotene (two β-rings), the Clan A enzyme might prefer α-carotene (mixed β-ring and ε-ring) as the substrate. Therefore, the *E. coli* platform will be valuable in testing different types of substrates and combinations of enzymes.

2.4.3 What other factors are required for successful function in E. coli?

P450 enzyme activity is generally predicated on availability of an electron donor partner (Bernhardt, 2006). Such partners include NADPH flavodoxin reductases of various forms that are found in plant plastids, among other locations, and which function as electron carries for photosynthetic and nonphotosynthetic processes such as sterol and fatty acid biosynthesis (Carrillo and Ceccarelli, 2003; Ceccarelli et al., 2004). In *E. coli*, a bacterium which lacks P450 enzymes, flavodoxin: NADPH flavodoxin reductase can serve as a replacement for the natural redox partner to facilitate function of some plant P450 enzymes but not others (Jenkins and Waterman, 1994; Carrillo and Ceccarelli, 2003). For example, sorghum CYP71E1, required for cyanogenic glycoside biosynthesis, could both function *in*

vivo in *E. coli* using the *E. coli* flavodoxin: NADPH flavodoxin reductase or reconstituted *in vitro* with a plant NADPH cytochrome P450 reductase (Bak et al., 1998). In comparison, CYP79B2, an enzyme required for indole glucosinolate biosynthesis in *Arabidopsis*, functioned in *E. coli* only with support of the plant reductase (Mikkelsen et al., 2000) as found in other cases (Schuler and Werck-Reichhart, 2003). The CYP97 Clan A and Clan C enzymes were functional in *E. coli* because they could accept electrons from the endogenous *E. coli* flavodoxin reductase (Jenkins and Waterman, 1994; Carrillo and Ceccarelli, 2003).

2.4.4 Why is CYP97 carotene hydroxylase activity suboptimal in E. coli?

Hydroxylase activity of the CYP97 enzymes in *E. coli* could be improved over the modest conversion of monohydroxylated intermediate to di-hydroxylated product. In addition to substrate choice as mentioned above, another factor that might limit enzyme activity might be the basal level expression of the endogenous reductase. Bacterial flavodoxin reductase, encoded by *E. coli fpr*, is induced about twenty-fold from basal levels as part of a global response to oxidative stress that is mediated by the *soxRS* regulon (Liochev et al., 1994; Krapp et al., 2002). In the absence of such induction, reductase expression may limit activity of the plant CYP. Other possible enhancements might include optimization of codon usage of the plant cDNA and/or removal of upstream plastid targeting sequences. However, posttranslational modification (e.g. glycosylation) is an unlikely prerequisite since the CYP97 enzymes did function in *E. coli*.

2.4.5 What is the structural basis for enzyme specificity?

Understanding the structural basis for β- and ε-ring specificity as demonstrated here will require modeling in combination with directed enzyme evolution. Modeling alone is insufficient to predict structural determinants of activity as seen in the characterization of other P450 enzymes. For example, directed evolution of *Bacillus subtilis* CYP102A2 enhanced its activity towards fatty acids and other aromatic substrates; affected amino acids were remote from the active site and unpredicted from modeling on related crystal structures to affect substrate specificity (Axarli et al., 2005). The *E. coli* platform used here will be valuable in addressing the issue of substrate specificity as well as to produce the protein crystals needed to elucidate a three dimensional structure of these crucial enzymes.

There are some recently described bacterial and fungal genes that encode P450 carotene ring hydroxylases. CrtS, is a 62.63 kDa (5.75 pI) fungal P450 that hydroxylates β-carotene to form astaxanthin. However, it is unrelated to the plant enzyme (Alvarez et al., 2006). A P450 β-ring hydroxylase gene was also found in the carotenoid-containing thermophilic bacterium, *Thermus thermophilus* and the crystal structure has been solved (Yano et al., 2003). The location of this single P450 gene on the thermophile's megaplasmid suggests that the gene may have been acquired through horizontal transfer and perhaps conferred some advantage to heat stress. The bacterial gene encodes a 44 kDa (9.72 pI) enzyme in comparison to the predicted mature 66 kDa (5.44 pI) plant Clan A enzyme (Blasco et al., 2004; Henne et al., 2004), though they share 21% similar residues and cluster in a neighbor-joining tree

(Inoue, 2004). While the plant and bacterial enzymes are sufficiently different, and perhaps cases of convergent evolution, they both functioned in *E. coli*.

2.4.6 Metabolic engineering in plants- challenges

From the results shown here and from prior studies, it is evident that higher plants contain one P450 ε-carotene hydroxylase that mediates carotene ε-ring hydroxylation and two structurally distinct (P450 and diiron) β-ring hydroxylases, both of which catalyze hydroxylation of carotene β-rings leading to xanthophyll formation. A fundamental issue to address is how do the P450 and non-P450 enzymes contribute to metabolon structure(s) whose biogenesis may be differentially controlled whether the pathway is located on the envelope or thylakoid membranes; each location leads to intermediates and products with multiple roles in plant development and physiology. In addition, how do these enzymes contribute to biogenesis and control of a pathway that also diverges to form alternate end-products, one via β-carotene (the carotenoid with highest provitamin A value) and the other via α-carotene? Indeed, to develop rational strategies for metabolic engineering of specific carotenoids such as the provitamin A β-carotene or any other of the numerous carotenoids found in nature, will require a deeper understanding of how these hydroxylases participate in metabolon biogenesis and substrate channeling.

In rice, as in *Arabidopsis*, in addition to CYP97C2, there appear to be at least two nonheme diiron β-carotene hydroxylases (GenBank #XM_473611, AK060559). Further investigation of their suborganellar membrane localization will help to

elucidate the specific roles of these β-ring hydroxylase classes on envelope and thylakoid membrane carotenoid biosynthetic pathways. Structurally they are unique: the diiron β-ring hydroxylases contain four transmembrane helices causing the enzymes to embed in the membrane while the P450 hydroxylases are peripheral membrane enzymes due to their single transmembrane anchoring sequence. Other structural distinctions seen in the enzymes are their isoelectric points (pI); the pI of the mature rice Clan A enzyme is 5.44, while the rice diiron enzymes are 6.98 and 9.04. Given the differences in enzyme structure and membrane association, the enzymes may respond differentially to pH gradients that form across the thylakoid lumen at high light irradiance (lumenal pH 5/stromal pH 8). If, as is the case of the xanthophyll cycle de-epoxidase enzyme, activities of the carotene hydroxylases are pH-dependent, localization/orientation on the thylakoid membranes may be critical. At high light irradiance, acidification of the thylakoid lumen activates membrane association of the violaxanthin de-epoxidase, an enzyme having a pI of 4.57 (Bugos and Yamamoto, 1996), to convert violaxanthin to zeaxanthin, a product which dissipates the high light energy; in the dark violaxanthin accumulates. Perhaps diiron carotene β-ring hydroxylases function in the dark or on different plastid membranes (including envelope) for apocarotenoid biosynthesis, while P450 β-ring hydroxylases are light activated or function on the thylakoid membranes in association with the photosynthetic apparatus.

If on the other hand, the two structurally distinct carotene hydroxylases exist on the same membrane system, it is compelling to consider how they contribute to

metabolon biogenesis and substrate channeling. Early modeling of Cunningham (Cunningham Jr. and Gantt, 1998) portrayed separate complexes for production of α-carotene and β-carotene, precursors of lutein and zeaxanthin, respectively. However, some observations agree and others conflict with this model and suggest that the enzymes might play some compensatory role. For example, in studies of potato tuber endosperm, overexpression of a bacterial phytoene synthase gene led to increases in lutein and β-carotene but not in zeaxanthin (Ducreux et al., 2005). The result suggests that there were adequate levels of the P450 CYP97 ε-ring and β-ring activities to accommodate the increased pathway flux to lutein. However, the Clan A P450 β-ring hydroxylase was not accessible for hydroxylation of the β-carotene needed to produce zeaxanthin which suggests that cells were limited for specific β-carotene hydroxylase activity required for zeaxanthin biosynthesis, presumably that of the nonheme diiron enzyme. Therefore, despite the fact that these studies demonstrate that the P450 Clan A enzyme could utilize β-carotene as a substrate in *E. coli*, it may be that the enzyme is not in a biochemical context to utilize this substrate in potato tuber. Our interpretation of the potato experiments is that they support Cunningham's view of separate submetabolons for the α-carotene and β-carotene sides of the pathway. However, if the Clan A enzyme cannot compensate for the diiron enzyme, then one would predict that a Clan A knockout would affect lutein accumulation only and not that of the β-carotene derived xanthophylls. This was not the case for the Arabidopsis *lut5* (Clan A) mutation which caused both reduced lutein and β-carotene derived xanthophylls (Kim and DellaPenna, 2006), thus suggesting that the Clan A enzyme is

not limited to α-carotene derived hydroxylation activity but may also play a role in hydroxylation of β-carotene. In another example, a double T-DNA knockout of the two genes encoding Arabidopsis diiron hydroxylases conferred no increase in β-carotene, an increase in lutein, and reduced β-carotene derived xanthophylls (Tian et al., 2003). Given that β-carotene was not increased and β-carotene derived xanthophylls were not completely eliminated, suggests that the P450 Clan A enzyme could partially compensate for the missing diiron enzymes to catalyze synthesis of zeaxanthin from β-carotene. Interestingly, the double knockout in the diiron enzymes did affect lutein levels in the seed, suggesting that differences in plastid membranes in the seed and in the leaf may be associated with different mechanisms with respect to compensation between the P450 CYP97 and diiron β-ring hydroxylases. In summary, the literature holds conflicting examples regarding interchangeability between the P450 Clan A and diiron carotene β-ring hydroxylases. Plant mutations may have a pleiotropic effect on expression of other genes in the pathway, thereby confusing the interpretation of the resulting biochemical profile. Moreover, the observations made in Arabidopsis leaves may not translate to similar effects in cereal endosperm where plastid structure and gene family structure is different (Gallagher et al., 2004; Wurtzel, 2004). However, despite individual limitations, with the combined use of the bacterial system reported here, together with the plant genetic studies, it is possible to derive a deeper understanding of plant carotenogenesis.

2.5 Conclusion

In conclusion, this work demonstrated feasibility of the *E. coli* system as a platform to assess substrate specificity for representative members of the CYP97 A and C clans. This heterologous system will be valuable for further study and manipulation of these enzymes. The roles and topologies of the carotenoid metabolons are still an open question and important topics of research in identifying components that may be key targets for metabolic engineering of either "arm" of the carotenoid biosynthetic pathway. For example, to confer accumulation of the pathway intermediate, β-carotene, would necessitate a block in β-carotene hydroxylase expression/ activity. At this point, it is not possible to predict the appropriate class of enzyme target, the P450 or the diiron. Future elucidation of components of such complexes will be especially significant for metabolic engineering of enhanced levels of carotenes, given that β-carotene and α-carotene have provitamin A activity, whereas their hydroxylated xanthophyll products zeaxanthin and lutein do not.

2.6 Materials and Methods

2.6.1 Phylogenetic and sequence analyses

Nucleotide and corresponding protein sequences, highly similar to the putative CYP97A4 coding mRNA from *Oryza sativa* L. (AK068163) (Kikuchi et al., 2003), were obtained using BLAST analyses from all available public databases in NCBI GenBank and Institute of Genomic Research (TIGR) gene indices: CYP97A3

(#AAL25587), CYP97B3 (#AAL32753), CYP97C1 (#AAR83120) (Tian et al., 2004), and CYP86A1 (#NM_116260) from *Arabidopsis thaliana* (L.) Heynh.; CYP97B4 (#XP_464306) and CYP97C2 (#AK065689) from *Oryza sativa* L.; #BJ234910 and #CA501638 from *Triticum aestivum*; # TC69886, #TC76166 and #BM816653 from *Hordeum vulgare* L.; CYP97B1; synonym CYP97A2 (#Z49263) from *Pisum sativum* L. (Baltrusch et al., 1997); CYP97B2 (#AAB94586) and #TC228439 from *Glycine max* (L.) Merr. (Siminszky et al., 1999); #BQ971938 from *Helianthus annuus* L.; #BE552887 from *Zea mays* L.; #TC101515 and #TC109838 from *Medicago truncatula* Gaertn; and #BT012891 from *Lycopersicon esculentum* Mill.

Protein sequences were screened for chloroplast targeting signal peptides using two prediction algorithms: GENOPLANTE™ PREDOTAR (Predotar for Prediction of organelle targeting sequences) at http://genoplante-info.infobiogen.fr/predotar/ and ChloroP 1.1 Server at http://www.cbs.dtu.dk/services/ChloroP/ (Emanuelsson et al., 1999). Rice protein sequences were compared against the P450 database at http://132.192.64.52/blast/P450.html using the P450 blast server. Prior to alignment, protein sequences were truncated to include the P450 conserved domain (CD) and exclude the chloroplast targeting sequence. Boundaries of conserved cytochrome P450 domains were identified using NCBI Conserved Domain Search tool (RPS-BLAST). Amino acid sequences were aligned by ClustalW, and a neighbor-joining

tree was constructed with a 500 bootstrap replication support using MEGA3 software (Kumar et al., 2001).

2.6.2 Total RNA extraction and Isolation of cDNAs

Total RNA was extracted from *Oryza sativa* L. cv Nipponbare (*japonica*) leaves (approx. 6 weeks old) using the RNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA). Approximately 2 μg of total RNA, in a 20 μl reaction, was used for cDNA synthesis using the SuperScript™ First-Strand Synthesis System for RT-PCR kit (Invitrogen Corporation, Carlsbad, CA). GenBank #AK068163 has a 1929 bp ORF between nucleotide number 54 and 1982. Genbank #AK065689 has a 1683 bp ORF between nucleotide number 105 and 1787. Gene specific primers, with terminal *Eco*RI and *Xho*I restriction sites, (forward, 5′-CCG GAA TTC A_{54}TG AGC TCA GCG ACG TCA GTG AGT_{77}-3′ and reverse, 5′-ACC GCT CGA GT_{1985}C AGA TTC GAG TTG CTG AGA CTT G_{1962}-3′, MWG-Biotech Inc**.,** Oaks Parkway, NC) were used to amplify the full-length coding sequence of CYP97A4 (#AK068163) (Kikuchi et al., 2003). The coding sequence (excluding the first 21 bp) of CYP97C2 (#AK065689) was amplified using gene specific primers with terminal *EcoRI* and *XhoI* restriction sites, (forward, 5′-CCG GAA TTC C_{105}CG TCC CGT GCG TAC CAT TC_{124} -3′ and reverse, 5′-ACC GCT CGA G_{1793}TC ATC TGG ACC CAC TGA GTG CA_{1774}-3′, MWG-Biotech Inc**.,** Oaks Parkway, NC) (50). 50 μl reactions each contained 4 μl cDNA, 1 μl of each primer (20 μM), 19 μl H_2O, and 25 μl of a master mix (50mM $MgCl_2$, 10mM dNTP mix, 50 units/ml *Taq* polymerase, 2X PCR buffer)

(Promega Corp., WI). Amplification conditions were: 1 cycle, 95°C, 3 min; 40 cycles, 95°C, 30 sec, 58.1°C, 40 sec, 72°C 2 min; 1 cycle, 72°C, 8 min.

2.6.3 Construction of expression vectors and functional analysis

The CYP97A4 (#AK068163) and CYP97C2 (#AK065689) sequences amplified as above were subcloned in-frame into *Eco*RI and *Xho*I sites of the pCOLADuet™-1 vector (Novagen, WI), renamed pRT-A4 and pTR-C2, respectively. For testing of β-ring substrate specificity, constructs were transformed into *E. coli* BL21 (DE3) cells (Novagen, WI) harboring pAC B-O4 which confers β-carotene accumulation (Sun et al., 1996). pRT-A4 and pTR-C2 ORFs were also introduced into *E. coli* BL21 (DE3) cells carrying the pACCRT-EIB (Cunningham Jr. et al., 1993) and plasmid y2 (Cunningham Jr. et al., 1996) which carries the Arabidopsis lycopene ε cyclase, which together confer accumulation of lycopene and δ-carotene (ε,ψ-carotene), along with some ε-carotene (ε,ε –carotene) (minor product)(Cunningham Jr. et al., 1996). For carotenoid analyses, saturated cultures in LB medium were diluted 100-fold into 25 ml fresh medium in 250 ml flasks, then grown in the dark at 300 r.p.m. at 21°C, 28°C, or 37°C (cultures containing pTR-C2 were grown at 37°C) until O.D. 0.6 at which point they were induced with 10 mM IPTG (unless otherwise noted) and further cultured for a total of three days.

2.6.4 Extraction of carotenoids from *E. coli* cells, HPLC separation and identification

Carotenoids were extracted in ethanol:ether (1:2 v/v), essentially, as described (Davison, 2002) Bacterial cell pellets were extracted in 10 volumes of the ethanol:ether mix using a Dounce homogenizer, incubated at room temperature for 5-10 minutes, and centrifuged at 10,000×*g*. Supernatants were frozen at -80°C for 30 minutes, centrifuged at 10,000×*g* at 4°C, supernatants dried under nitrogen gas, dissolved in HPLC mobile phase, and stored at -20°C.

Separation was carried out using a Waters HPLC system equipped with a 2695 Alliance separation module, a 996 photodiode array detector, a column heater, a fraction collector II, _Empower software (Millipore, Franklin, MA), and a Nucleosil 5 C_{18} (5 μm, 250 x 4.6mm) column with a Nucleosil C_{18} (5 μm, 4 x 3.0 mm) guard column (Phenomenex, Torrance, CA), mobile phase of 100 parts acetone: 4 parts water, column flow rate of 0.6 ml/min, and sample and column temperature of 30 ± 5°C. Peaks were identified on the basis of retention times/spectra matching those of authentic standards (INDOFINE Chemical Company, Inc., Sommerville, NJ), and standards purified from bacteria expressing genes encoding carotenoid biosynthetic enzymes (Cunningham Jr. et al., 1996). Authentic lactucaxanthin, purified from *Lactuca sativa* leaves (Siefermann-Harms et al., 1981) and distinguished as a novel HPLC peak in comparison to a maize leaf profile, exhibited a spectrum matching reported values (Britton et al., 2004). Identification of ε,ε-carotene and lactucaxanthin were confirmed by comparison with reported λ_{max} (wavelength of maximum absorption) spectral values (Britton et al., 2004) and comparison to standards chromatographed in the identical system. For quantification of extracted metabolites,

data were collected at λ_{max} for individual metabolites and integrated peak areas calculated.

2.7 Acknowledgements

We thank Dr. Francis Cunningham Jr. for *Erwinia* expression plasmids, Christina Murillo for technical support, and Faqiang Li and Ratnakar Vallabhaneni for technical advice. This research was supported by NIH (#S06-GM08225), PSC-CUNY, and New York State.

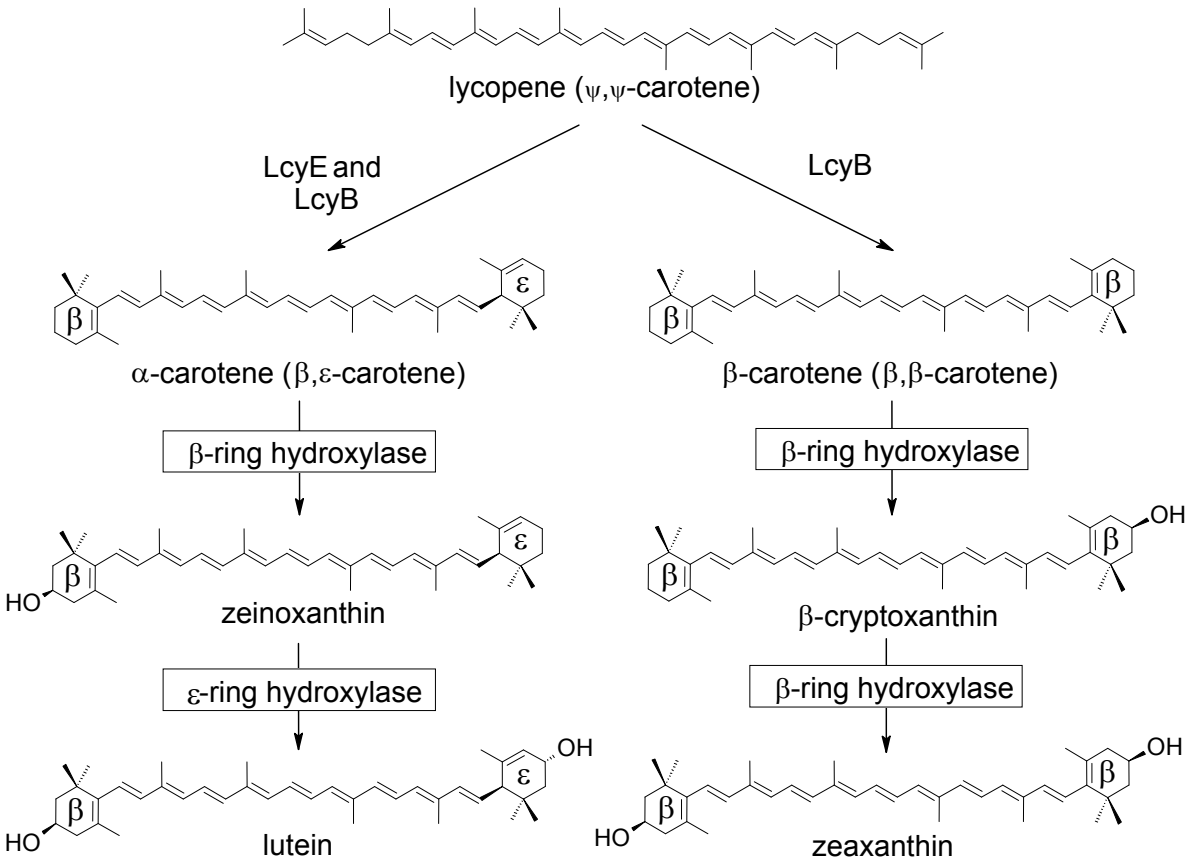
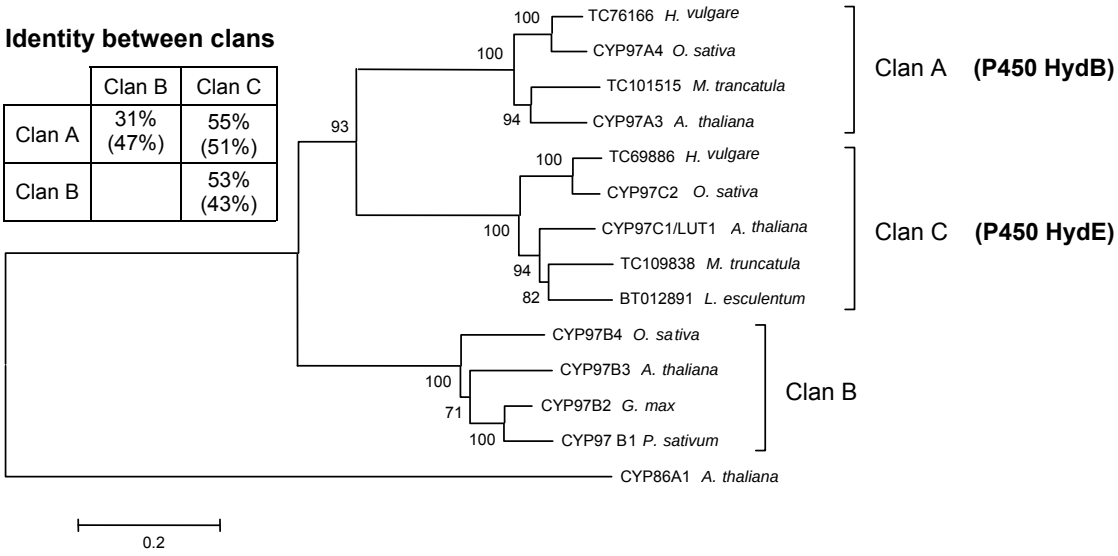


Figure 2-1. Schematic illustration of the cyclohexene ring hydroxylation in β- and α-carotene. Conversion of lycopene to β-carotene requires lycopene β-cyclase (LcyB) alone, while conversion of lycopene to α-carotene requires lycopene β- and ε-cyclases (LcyE). Formation of the hydroxylated xanthophylls zeaxanthin and lutein is mediated by two separate stereospecific β- and ε-ring hydroxylases. In hydroxylation of α-carotene, the two hydroxylases may act in the reverse order shown, to yield the mono-hydroxy intermediate α-cryptoxanthin, in which the ε-ring is hydroxylated.

A

Identity between clans

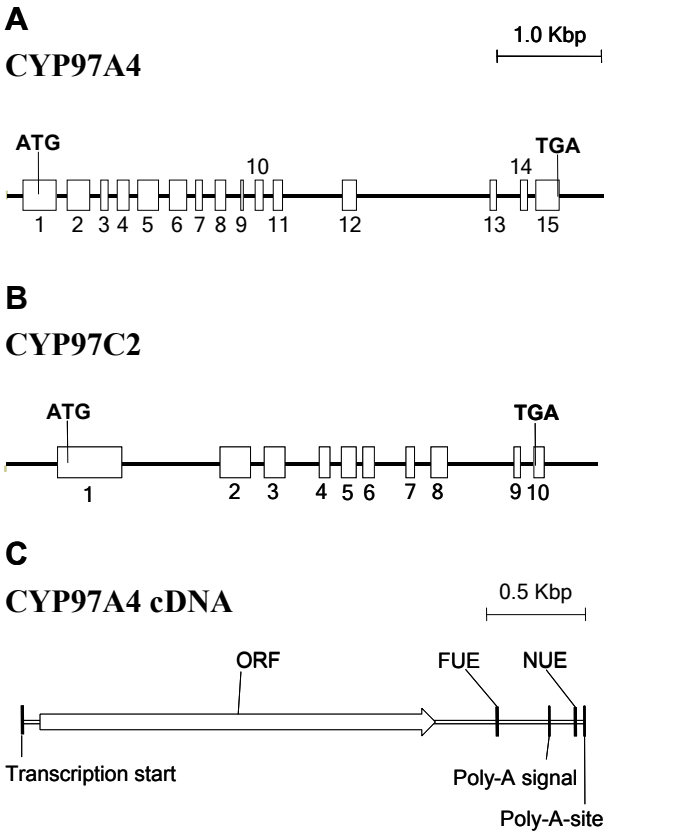
	Clan B	Clan C
Clan A	31% (47%)	55% (51%)
Clan B		53% (43%)



B

	Block 1	Block 2	Block 3	
CYP97B1 *P. sativum*	LLEG-EGRDGQKSVE	DVLPGGHKGDKDGYTI	LVQNNNEEVEGWAGFDPSRSPGALK	Clan B
CYP97B2 *G. max*	LLEG-EGYDGPDSIE	DVLPGGHKGEKDGYTI	LVQNKNEEIEGWAGLDPSRSPGALK	
CYP97B3 *A. thaliana*	LIREKETSSGEDTIE	ETLPGGHKGEKEGHKV	LRTKESNGIEGWAGFDPSRSPGALK	
CYP97B4 *O. sativa*	LIER--GEHGEKYTI	DKLPGGHNGAKEGYEI	SVPKKDESIEGWAGFDPDRSPGAMK	
BT012891 *L. esculentum*	PDAI-------SGSA	DVLPG-------NYKV	DLE-------------------GPV	Clan C
TC109838 *M. truncatula*	ADAV-------NGTA	DELPG-------DYKI	DLD-------------------GPV	
CYP97C1 *A. thaliana*	PYAE-------DGSA	DILPG-------NYKV	DID-------------------GAV	
CYP97C2 *O. sativa*	TSAL-------SGKP	DVLPG-------NYKI	DLE-------------------GPV	
TC69886 *H. vulgare*	TYAL-------SGEP	DVLPG-------NYKV	DLE-------------------GPI	
CYP97A3 *A. thaliana*	AAAL-------KGEE	DILG--------EYPI	PLD-------------------GPN	Clan A
TC101515 *M. trancatula*	TAAS-------DGED	DVLG--------EYPI	PLD-------------------GPK	
CYP97A4 O. sativa	KAAT-------DGED	DMLG--------GYPI	PLD-------------------GPK	
TC76166 *H. vulgare*	KAAS-------DGED	DMLG--------EYPI	PLD-------------------GPK	

Figure 2-2. Comparison between the conserved P450 domain of the three clans in the CYP97 family. A, A rooted neighbor-joining tree was constructed using the fatty acid ω-hydroxylase (CYP86A1) from *A. thaliana* as an outgroup. Numbers adjacent to branches are bootstrap values supporting the presented final tree. Average identities among clans are shown in table in comparison to pair wise comparisons between *A. thaliana* and rice sequences, which are enclosed in parentheses. B, Three sequence blocks distinguish members of Clan B from those in Clans A and C are shown in excerpts from the alignment of the conserved P450 domains



D

Conserved Domains

	Oxygen-binding	Heme-binding
Clan C		
CYP97C2 (*O. sativa*)	AGHETT	FSGGPRKCVG
CYP97C1 (*A. thaliana*)	AGHETT	FSGGPRKCVG
TC69886 (*H. vulgare*)	AGHETT	FSGGPRKCVG
Clan A		
CYP97A4 (*O. sativa*)	AGHETS	FGGGPRKCVG
CYP97A3 (*A. thaliana*)	AGHETS	FGGGPRKCIG
TC76166 (*H. vulgare*)	AGHETS	FGGGPRKCVG

Figure 2-3. Gene structure and P450 domains of CYP97 Clan A and C enzymes. A and B, Gene structures for rice CYP97A4 and CYP97C2, respectively. Open boxes indicate exons; C, Revised cDNA structure of CYP97A4 cDNA. The transcribed cDNA is a predicted 2414 bp fragment with an 88 bp 5′-UTR, a 1929 bp ORF, and a 397 bp 3′-UTR with conserved NUE, FUE, and poly-A site; D, Comparison of conserved oxygen and heme-binding domains for CYP97 Clan A and C enzymes of rice compared with those in other plant species.

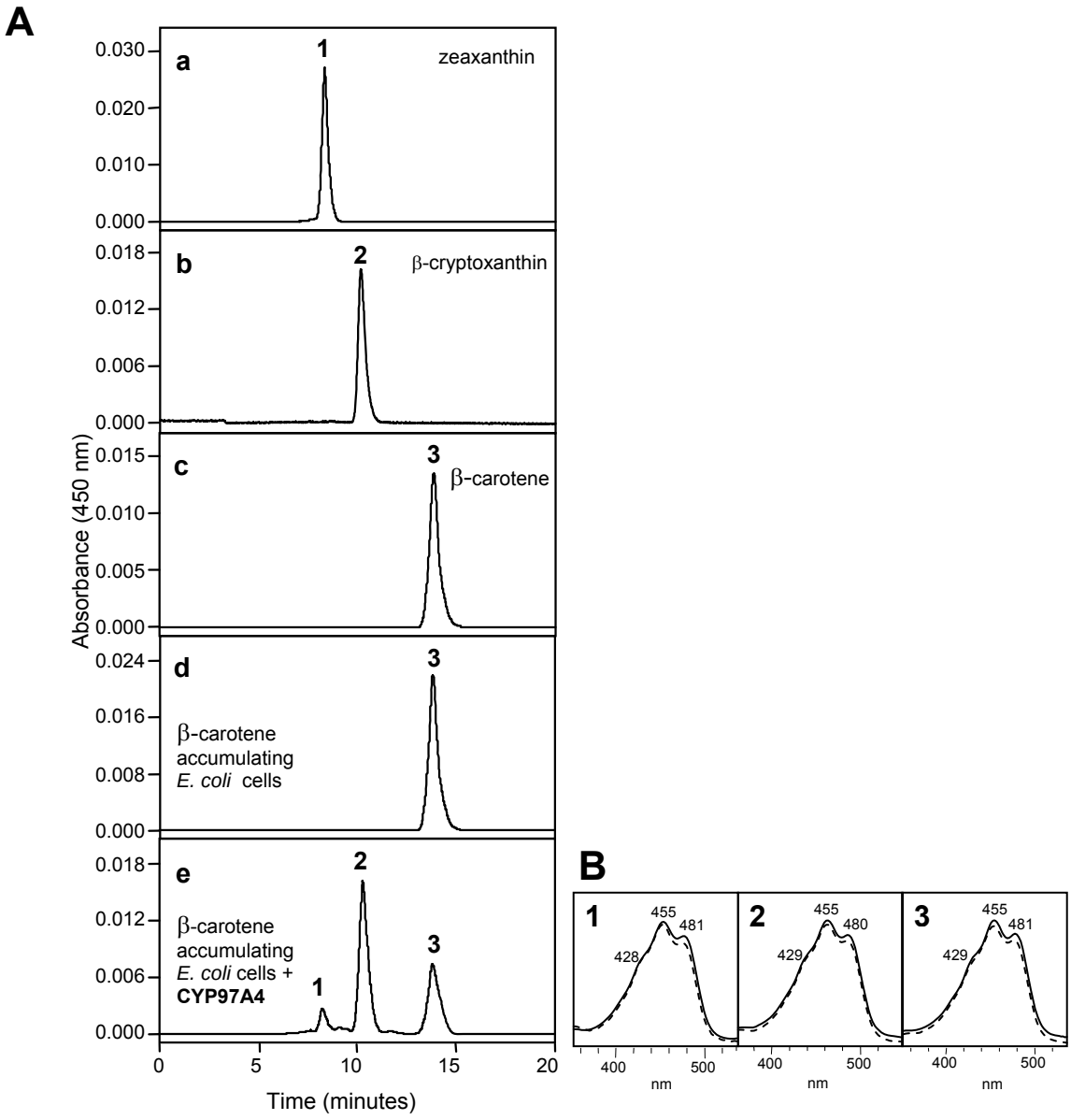


Figure 2-4. **Functional complementation of putative β-ring hydroxylase CYP97A4 cDNA in cells accumulating β-carotene.** Bacterial cells were grown at 21°C and extracted pigments were separated by reversed phase HPLC. A, Chromatograms showing elution profiles: (a-c) authentic standards: **1,** zeaxanthin; **2,** β-cryptoxanthin; **3,** β,β-carotene; (d) extracts from *E. coli* BL21 (DE3) cells accumulating β,β-carotene; (e) extracts from cells in (d) that were additionally transformed with the CYP97A4 cDNA. Peak numbers and identities in (d) and (e) correspond to those in (a-c). B, Spectral profiles of standards (solid line) shown in (a-c, peaks 1-3) are superimposed on corresponding spectra (dashed line) obtained from peaks 1-3 shown in (e).

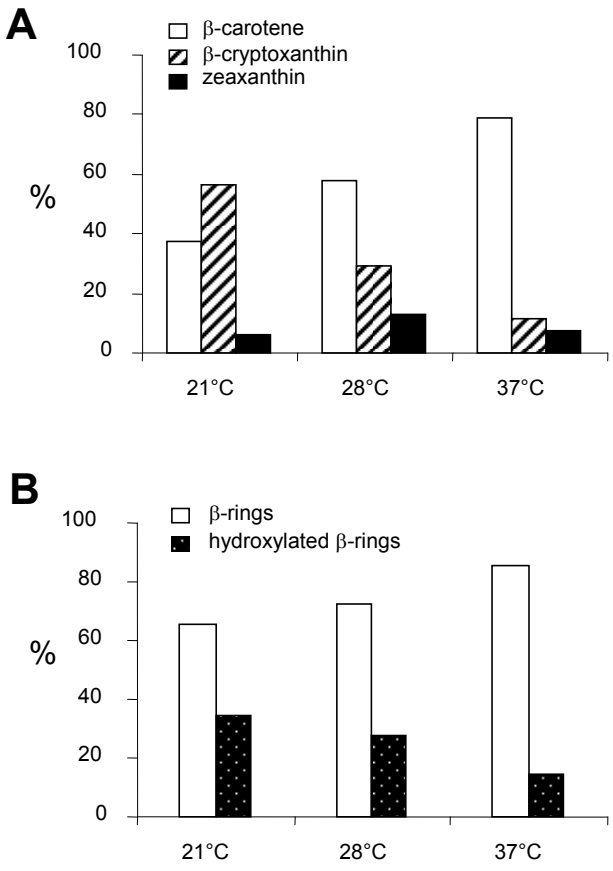


Figure 2-5. **Effect of temperature on product distribution and CYP97A4 enzyme activity *in vivo*.** A, Increasing the culturing temperature from 21°C to 37°C shifts the ratio of β-cryptoxanthin (one hydroxylated β-ring) to zeaxanthin (two hydroxylated β-rings) from 10:1 to ~1:1. B, Overall % hydroxylation activity of CYP97A4 decreases from 35% at 21°C to 15% at 37°C.

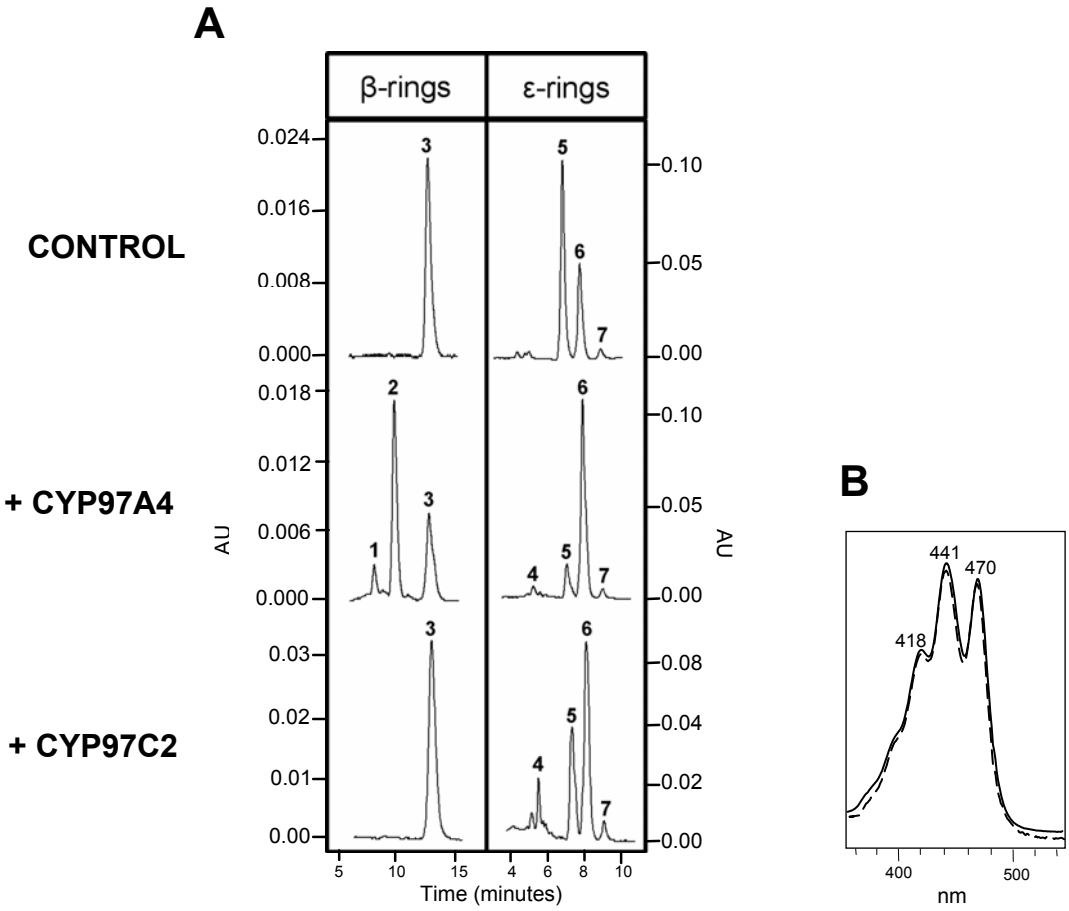


Figure 2-6. Substrate specificity of P450 carotene hydroxylases. A, *E. coli* cells engineered to accumulate carotenes with β rings (left panel) or ε-rings (right panel) were further transformed with test plasmids encoding CYP97A4 (putative β ring hydroxylase) or CYP97C2 (putative ε-ring hydroxylase). HPLC chromatograms (left panel, 450 nm; right panel, 470 nm) show composition of accumulating carotenoid substrates and products. Left panel, *E. coli* cells transformed with pAC B-O4 (control) plus indicated test gene; Right panel, *E. coli* cells transformed with pACCRT-EIB+y2 (control) plus indicated test gene. Peaks: **1**=zeaxanthin, **2**=β-cryptoxanthin, **3**=β,β-carotene, **4**=lactucaxanthin, **5**=lycopene, **6**=δ-carotene, **7**=ε,ε-carotene. B, Spectrum of peak 4 product of CYP97C2 (dashed line) overlayed with spectrum of authentic lactucaxanthin (solid line) chromatographed under identical conditions.

CHAPTER 3

Synergistic interactions between carotene ring hydroxylases drive lutein formation in plant carotenoid biosynthesis

3.1 Abstract

Lutein and zeaxanthin are non-provitamin A dihydroxy xanthophylls that are produced from their respective provitamin A carotene precursors by the action of cytochrome P450-type (CYP97) and diiron-type (HYD) carotenoid β- and ε-hydroxylases. In order to gain a better understanding of the respective roles of the CYP97 and HYD enzymes in this conversion, *E. coli* functional complementation assays were used to examine individual as well as combined enzyme activities and specificities, as well as *in vitro* chloroplast import and *in vivo* GFP fusion assays to assess subcellular/suborganellar locations for the monocot CYP97A4 (β-carotene hydroxylase) and CYP97C2 (ε-carotene hydroxylase), and for the HYD3 and HYD4 β-carotene hydroxylase enzymes. In addition, *in vivo* Bimolecular Fluorescence (BiFC) assays were performed to test for protein interaction. This work demonstrated that the CYP97A4 and CYP97C2 interact and function optimally when expressed together in an *E. coli* functional complementation system to efficiently convert α-carotene to lutein. These studies also demonstrated interaction between HYD3 and HYD4 to form a heterodimer, and the HYD4 was shown to form a homodimer complex.

---

*This chapter is being prepared for journal submission:
Rena F. Quinlan[1], Maria Shumskaya[1], Louis M.T. Bradbury, Jesús Beltrán, Edward J. Kennelly, Chunhui Ma, and Eleanore T. Wurtzel (2012) Synergistic interactions between carotene ring hydroxylases drive lutein formation in plant carotenoid biosynthesis. [1]co-first author

Furthermore, HYD4 functioned optimally when expressed individually to efficiently convert β-carotene to zeaxanthin in a bacterial complementation system. *In vitro* import studies also indicated that the CYP97 and HYD enzymes respectively form peripheral and integral associations with the chloroplast membrane. Taken together, these data offer new insights into the activities and specificities of the CYP97 and HYD enzymes as well as their respective roles in metabolon biogenesis and substrate channeling on the separate α- and β-carotene "arms" of the carotenoid pathway for the conversion of provitamin A carotenes to non-provitamin A xanthophylls. These insights will prove valuable in advancing the development of rational strategies for metabolic engineering and/or breeding efforts aimed at enhancing provitamin A levels in cereal crops.

3.2 Introduction

Carotenoids are a large class of isoprenoid pigments which are synthesized by all photosynthetic organisms, as well as some bacteria and fungi (Britton et al., 2004). In plants, carotenoids have essential roles in growth and development which include functions in light-harvesting in photosynthesis and protection against photooxidative damage (Jahns and Holzwarth, 2012). In addition, carotenoids such as β-carotene and zeaxanthin serve as precursors to various apocarotenoids including the phytohormone abscisic acid (ABA) and strigolactones which respectively function in regulating seed dormancy and in mediating symbiotic root/mycorrhizal fungi associations (Cuttriss,

2011). Furthermore, these apocarotenoids play roles in biotic and abiotic stress signaling pathways (Walter et al., 2010; Cao et al., 2011) .

Among the roughly 600 carotenoids found in nature, about 50 are obtained in the human diet. Some important physiological functions associated with these compounds include their activities as antioxidants and anti-cancer agents (van den Berg et al., 2000; von Lintig, 2010). In addition, carotenoids (ie., β-carotene) serve as precursors to nutritionally-essential compounds such as vitamin A, while the nonprovitamin A xanthophylls lutein and zeaxanthin are thought to aid in the prevention of light-induced retinal damage (von Lintig, 2010).

Efforts to improve the provitamin A content of cereal endosperm in staple crops such as maize will require elucidation of the mechanisms involved in controlling the conversion of provitamin A carotenes such as α- and β-carotene to non-provitamin A xanthophylls (ie., lutein and zeaxanthin). Xanthophylls are a class of oxygenated carotenoids that function in light-harvesting, photoprotection, and assembly of the light harvesting antenna complexes of photosynthetic membranes (Pogson and Rissler, 2000). The xanthophylls zeaxanthin and lutein are derived from their respective carotene precursors by the action of β- and ε-ring hydroxylase enzymes which add hydroxyl groups to the third carbon of carotene β- and ε-rings, respectively (**Fig. 3-1**) (Tian et al., 2003). Synthesis of lutein requires the hydroxylation of both the β- and ε-rings of α-carotene, while zeaxanthin biosynthesis involves hydroxylation of both β-rings of β-carotene (**Fig. 3-1**).

The cytochrome P450 CYP97 and diiron HYD enzymes, on the basis of both genetic and biochemical evidence, are hypothesized to function in carotene ring hydroxylation, to convert provitamin A carotenes (ie., α- and β-carotene) to nonprovitamin A xanthophylls (ie., lutein and zeaxanthin) (Sun et al., 1996; Fiore et al., 2006; Kim and DellaPenna, 2006; Quinlan et al., 2007). *Arabidopsis* genetic studies suggested that carotene β-ring hydroxylations were carried out by both a P450-type CYP97A enzyme as well as by diiron-type β-hydroxylases (HYD), while hydroxylations of ε-rings were performed by a P450-type CYP97C enzyme (Tian et al., 2004; Fiore et al., 2006; Kim and DellaPenna, 2006). Enzyme activities were also directly demonstrated for the rice CYP97A4 and CYP97C2 via *E. coli* complementation assays which showed that the CYP97C2 is specific for ε-rings while the CYP97A4 is primarily a β-ring hydroxylase, with minor activity towards ε-rings (Quinlan et al., 2007), as had previously been shown for the diiron-type β-hydroxylases(Sun et al., 1996).

Recent *Arabidopsis* genetic studies involving quadruple mutants which lack all CYP97/HYD enzymes indicated that in triple mutants containing only the CYP97C enzyme, 74% of lutein was generated. These genetic studies point to the CYP97C enzyme as being strongly active toward both the β- and ε-rings of α-carotene to efficiently convert this substrate to lutein (Kim et al., 2009). These results do not agree with earlier *E. coli* complementation studies which indicated that the CYP97C enzyme is exclusively an epsilon-ring hydroxylase; no activity toward β-rings was detected for this enzyme (**Chapter 2**) (Quinlan et al., 2007). Subsequent

complementation assays indicated that the CYP97C enzyme was only marginally active toward β-rings (**Chapter 3**). While genetic studies offer some insights into enzyme activities, there are limitations with this approach. It is difficult to investigate substrate specificities and enzyme activities directly using genetics-based assays, since mutations may have a pleiotropic effect on expression of other pathway genes, rendering an interpretation of the resulting biochemical profile problematic.

There remain open questions regarding CYP97 and diiron HYD enzyme activities and specificities, as well as with respect to the roles of these enzymes in the regulation of metabolite flow in the carotenoid pathway. It has been suggested that "metabolic channeling" of carotenoid intermediates in plastids may be facilitated by the simultaneous binding of membrane-anchored multi-enzyme protein complexes which would allow for more efficient processing of intermediates (Inoue, 2006). The fact that wild type *Arabidopsis* leaf tissue accumulates lutein (Pogson et al., 1996), and not the mono-hydroxylated intermediates α-cryptoxanthin (one hydroxylated ε-ring) or zeinoxanthin (one hydroxylated β-ring) points to a channeled flow of substrates in the pathway which is likely mediated by efficient multienzyme complexes in plastid membranes. Early studies by Lutzow and Beyer (1988) support the idea of the existence of multi-enzyme complexes in the carotenoid pathway. This research group described the so-called phytoene synthase (PSY) complex, which is a peripherally-membrane associated system, as consisting of at least two enzymes: geranylgeranyl diphosphate synthase (GGPPS) and PSY. In addition, it has been suggested that the membrane-associated phytoene desaturation process is mediated by the enzymes PDS and zeta-carotene desaturase (ZDS), which are thought to

closely interact, since there is no accumulation of the zeta-carotene intermediate *in vivo* (Lutzow and Beyer, 1988; Bonk et al., 1997).

Cunningham and Gantt proposed a hypothetical model which described a multienzyme carotenogenic complex in the thylakoid membranes and stroma of plant plastids (Cunningham Jr. and Gantt, 1998) This model was primarily based on studies which indicated that the active forms of PSY and PDS are either peripherally or tightly associated with internal membranes of plastids (Al-Babili et al., 1996; Schledz et al., 1996). In addition, further support for this model was obtained via bioinformatics predictions based on amino acids sequences for the lycopene β- and ε-cyclases (LCYB and LCYE, respectively) and carotene β- and ε-hydroxylases which suggested that these enzymes are associated with plastid membranes (Cunningham Jr. et al., 1996; Sun et al., 1996; Tian et al., 2004). Furthermore, *in vitro* chloroplast import studies have investigated the suborganellar localization of PSY, PDS, and LCYB (Bartley et al., 1991; Bonk et al., 1997; Mann et al., 2000). Recent proteomics studies however challenge the prevailing paradigm with respect to membrane localization of many of the carotenoid enzymes. It had previously been established that carotenoid pigments are part of the photosynthetic apparatus in the thylakoid membrane where they play roles in light-harvesting and photoprotection (Hirschberg, 2001). However, proteomics analyses indicated that many of the enzymes involved in carotenoid biosynthesis that were thought to be located on thylakoid membranes were actually found to be associated with envelope membranes (Joyard et al., 2009).

Though the individual pathway enzymes are known, there remains a major gap in understanding with respect to the nature of protein interactions and complexes involved in mediating carotenogenesis. The presence of enzyme complexes combined with the general absence of intermediates in the carotenoid pathway (Maudinas et al., 1977; Camara et al., 1982; Kreuz et al., 1982; Al-Babili et al., 1996; Bonk et al., 1997; Lopez et al., 2008) suggests that multi-enzyme complexes (metabolons) in plastid membranes may mediate the conversion of the provitamin A carotenes α- and β-carotene to the nonprovitamin A xanthophylls lutein and zeaxanthin. The biogenesis and maintenance of these putative metabolons in different plastids during various developmental processes (ie., photomorphogenesis) are poorly understood phenomena. In order to gain some insights into the mechanisms governing these phenomena the interactions and synergies of the CYP97A/CYP97C and HYD enzymes were examined. Both *in vitro* and *in vivo* approaches were used to examine the activities and specificities of the rice CYP97A4 and CYP97C2 and maize HYD3 and HYD4 enzymes, as well as their respective subcellular/suborganellar locations. In addition, these studies investigated whether any of these enzymes interact to form either heterodimer or homodimer complexes *in vivo*.

3.3 Results

3.3.1 Testing of CYP97 and HYD enzyme activities and substrate specificities via functional complementation

Based on earlier modeling by (Cunningham and Gantt, 1998) it was predicted that the CYP97 and diiron HYD enzymes respectively exist as separate complexes on the α-carotene (β-ε) and β-carotene (β-β) "arms" of the carotenoid pathway for the production of lutein and zeaxanthin. Preliminary investigations into the activities and substrate specificities of the CYP97A4 and CYP97C2 enzymes via *E. coli* complementation were previously conducted using cells accumulating substrates with either only β-rings (ie., β-carotene) or ε-rings (ie., ε-ε-carotene) (**Chapter 2**) (Quinlan et al., 2007), however these studies lacked cells accumulating the putatively preferred CYP97 substrate α-carotene. These studies demonstrated that the CYP97A4 enzyme was active toward the β-rings of β-carotene, converting this substrate to a modest amount of the end-product zeaxanthin. It was evident that CYP97A4 activity was suboptimal however as the mono-hydroxylated intermediate β-cryptoxanthin accumulated in these cells. CYP97A4 also showed minor activity toward the ε-rings of ε-ε-carotene, to produce a low amount of lactucaxanthin. The CYP97C2 showed modest activity toward the ε-rings of ε-ε-carotene, converting this substrate to a low amount of lactucaxanthin. No β-ring hydroxylation activity toward the β-carotene substrate was detected in cells expressing the CYP97C2 (**Chapter 2**) (Quinlan et al., 2007).

It was unknown whether the suboptimal activity observed for the CYP97 enzymes in the earlier complementation system was due to the lack of a partner enzyme which may be required for optimal activity, or whether the low enzyme activity was due to the unavailability of their putatively preferred substrate α-

carotene, as these assays included only either β-carotene or ε-ε-carotene accumulating *E. coli*. In order to address these issues, the studies described here investigated the activities of the CYP97 and diiron HYD enzymes when expressed individually as well as in various combinations in *E. coli* accumulating both α- and β-carotene. Accumulated carotenoids and standards for lutein, zeaxanthin, β-cryptoxanthin, and α- and β-carotene were analyzed via HPLC; α-cryptoxanthin and zeinoxanthin were analyzed by LC-MS.

When CYP97A4 and CYP97C2 were co-expressed in α- and β carotene accumulating *E. coli*, there was significant conversion of the α-carotene substrate to generate approximately 29% (% of total carotenoids) of the dihydroxy end-product lutein (one hydroxylated β-ring and one hydroxylated ε-ring of α-carotene) (**Fig. 3-1, 3-5, table 3-4**), while only approximately 3% (% of total carotenoids) of the end-product zeaxanthin (two hydroxylated β-rings of β-carotene) accumulated (**Fig. 3-1, 3-5, table 3-4**). α- and β-carotene accumulating cells transformed with an empty vector (control) did not generate any hydroxylated products (**Fig. 3-5, table 3-4**).

In contrast, when expressed as an individual enzyme, the CYP97C2 was only barely active toward both the α- and β-carotene substrates, respectively producing only roughly 0.7% and 1% (% of total carotenoids) of the pathway intermediates α-cryptoxanthin (one hydroxylated ε-ring of α-carotene) and β-cryptoxanthin (one hydroxylated ring of β-carotene) (**Fig. 3-1, 3-4, table 3-3**). While previous complementation assays did not detect any β-ring hydroxylation activity for the

CYP97C2 (**Chapter 2**) (Quinlan et al., 2007), it should be noted that double the amount of culture volume was used in these current assays (refer to Materials and Methods).

For identification of α-cryptoxanthin and zeinoxanthin LC-MS analysis was performed. The cryptoxanthin isomers all have the same mass (m/z 552.4); when subjected to positive ionization, the mass is 552.4 + 1 including the proton added; thus the m/z when positively ionized is 553.4 ($M+H^+$). It is not possible to distinguish between α-cryptoxanthin and zeinoxanthin on the basis of HPLC retention time or absorption spectra. However, compounds with an allylic hydroxyl group, such as α-cryptoxanthin, readily lose water when positively ionized, whereas compounds with non-allylic hydroxyl groups (ie., zeinoxanthin) do not (Kim and DellaPenna, 2006). LC-MS analysis confirmed a major ion fragment with a mass of 535.4 ($M+H^+-H_2O$) for α-cryptoxanthin. Based on this loss of water, it is apparent that its single hydroxyl group is bound to the ε-ring of α-carotene; therefore this compound is α-cryptoxanthin (**Fig. 3-6**).

When expressed individually, the CYP97A4 was moderately active toward the β-rings of α-carotene to generate approximately 14% (% of total carotenoids) of the intermediate zeinoxanthin (one hydroxylated β-ring of α-carotene); no activity towards the ε-ring of α-carotene was detected and no lutein accumulated (**Fig. 3-1, 3-4 , table 3-3**). LC-MS analysis confirmed an m/z value of 553.4 ($M+H^+$) for zeinoxanthin in agreement with its formula (**Fig. 3-6**). In addition, when expressed alone, the CYP97A4 was more active towards the β-carotene substrate than α-

carotene, as shown by the accumulation of the end product zeaxanthin (roughly 3% of total carotenoids) (**Fig. 3-4, table 3-3**), however enzyme activity was suboptimal as significantly more of the β-hydroxylated intermediate β-cryptoxanthin was generated (approximately 17% of total carotenoids). Taken together, these data suggest that the CYP97A4 and CYP97C2 enzymes function optimally when expressed together on the α-carotene (β-ε) "arm" of the carotenoid pathway to efficiently convert the α-carotene substrate to lutein.

Combinations of HYD3 + CYP97C2, and HYD4 + CYP97C2 were similarly assayed for enzyme activity in α- and β-carotene accumulating *E. coli*. Cells harboring the HYD3 + CYP97C2 combination accumulated 5% lutein and approx. 2% zeaxanthin (% of total carotenoids) (**Fig. 3-5, table 3-4**), while cells containing HYD4 + CYP97C2 produced only roughly 1.6% lutein and 3% zeaxanthin (**Fig. 3-5, table 3-4**). When compared with the efficient conversion of the α-carotene substrate to lutein observed in cells expressing CYP97A4 + CYP97C2, neither HYD + CYP97C2 combination showed strong activity toward the α-carotene substrate; in both cases (HYD3 + CYP97C2 and HYD4 + CYP97C2), only a relatively low amount of lutein accumulated. There was however, a roughly two–fold higher amount of lutein relative to zeaxanthin produced in cells co-expressing the CYP97C2 + HYD3 combination indicating that CYP97C2 activity is moderately enhanced when this enzyme is co-expressed with the HYD3. When expressed individually in cells accumulating either β-carotene only or both α- and β- carotene, the HYD4 efficiently converted the β-carotene substrate to approximately 29% and 31% zeaxanthin (% of

total carotenoids) (**Fig. 3-2, 3-4, tables 3-1, 3-3**). However, when co-expressed with the CYP97C2 in α- and β- carotene accumulating *E. coli*, the HYD4 exhibited markedly reduced activity toward the β-carotene substrate, generating only roughly 3% zeaxanthin (% of total carotenoids), although this was nearly two-fold higher than the amount of lutein generated (**Fig. 3-5, table 3-4**). By contrast, the HYD3 enzyme was not very active when expressed either individually or in combination with the CYP97C2 in cells accumulating a β-carotene substrate. The HYD3 conferred hydroxylation of the β-rings of β-carotene to generate roughly 1% of the end-product zeaxanthin when expressed individually (**Fig. 3-2, 3-4, tables 3-1, 3-3**) and 2% zeaxanthin when expressed in combination with the CYP97C2 in these cells (**Fig. 3-5, table 3-4**). In addition, the HYD3 exhibited moderate activity toward the β-rings of the α-carotene substrate in these cells, generating approximately 3% of the intermediate zeinoxanthin (when expressed alone) (**Fig. 3-4, table 3-3**) versus roughly 8% zeinoxanthin (when co-expressed with the CYP97C2) (**Fig. 3-5, table 3-4**). When expressed as individual enzymes, neither the HYD3 nor the HYD4 were able to convert the α-carotene substrate to the dihydroxy end product lutein; only moderate amounts of the β-hydroxylated intermediate zeinoxanthin accumulated (**Fig. 3-4, table 3-3**). Together, these data indicate that the HYD3 and HYD4 enzymes are preferentially active on the β-β (β-carotene) "arm" of the carotenoid pathway to convert β-carotene to the end-product zeaxanthin. HYD4 functions most efficiently as an individual enzyme, while there was no marked difference observed for HYD3 activity whether this enzyme was expressed alone or in combination with the

CYP97C2. In addition, CYP97C2 activity was moderately enhanced when this enzyme was co-expressed with the HYD3.

3.3.2 *In vitro* protein localization studies

3.3.2.1 *Chloroplast import assays*

Functional complementation in *E. coli* showed that the CYP97A4 and CYP97C2 enzymes function optimally when expressed together on the α-carotene side of the carotenoid pathway to convert the α-carotene substrate to lutein, while the HYD3 and HYD4 enzymes show preferential activity on the β-carotene side of the pathway to convert β-carotene to zeaxanthin. In light of these data, it would be expected that the CYP97A4 and CYP97C2 enzymes localize to the same subcellular/suborganellar location; as would be predicted for the HYD3 and HYD4.

The CYP97 and diiron HYD enzymes are nuclear-encoded and are predicted to be imported into chloroplasts, where they function in carotenoid biosynthesis (Hirschberg, 2001). In addition, the bioinformatics server Chlorop predicted these enzymes to each contain a chloroplast targeting signal peptide. In order to experimentally determine the subcellular and suborganellar locations of these enzymes chloroplast import assays were performed. Isolated pea chloroplasts were competent for the *in vitro* import of recombinant CYP97 and HYD 35S-methionine radio-labelled preproteins which were produced using a coupled reticulocyte in vitro transcription/translation system. Chloroplasts harboring imported proteins were then

re-isolated and subjected to thermolysin treatment in order to distinguish between proteins that were peripherally bound to the outer chloroplast envelope and those that had been imported and thus crossed the envelope membrane. As shown in **Figure 3-7**, the CYP97 and HYD proteins were protease-resistent, confirming import of these proteins into chloroplasts. SDS-PAGE analysis indicated that the CYP97A4 and CYP97C2 proteins were synthesized as precursors of about 69 kDa and 62 kDa, and then processed to approximately 64 and 59 kDa (**Fig. 3-7, table 3-5**) indicating transit peptide lengths of 5 and 4 kDa, respectively, which are roughly consistent with predicted values (**Table 3-5**). As shown in **Figure 3-7, table 3-5**, the HYD3 and HYD4 were synthesized as preproteins of roughly 35 and 34 kDa, and the processed proteins migrated to approximately 31 and 27 kDa respectively, indicating transit peptides of approximately 4 and 6 kDa. The observed transit peptide length of approximately 6 kDa for the HYD4 is roughly consistent with the predicted value (**Table 3-5**), however, the 4 kDa transit length observed for the HYD3 conflicts with the predicted value for this enzyme (**Table 3-5**); a prediction inaccuracy could explain this discrepancy.

In order to determine whether the imported CYP97 or HYD proteins are peripherally or integrally associated with the membrane, chloroplasts containing imported proteins were hyptonically lysed and fractionated into supernatant and pellet (membrane) fractions. The pellet fractions were then treated with alkaline to test whether any of these proteins were still associated with the membrane. As shown in **Figure 3-7**, most of the CYP97A4 and CYP97C2 proteins were dissociated from the

membrane upon alkaline treatment, indicating that these proteins are peripherally associated with the chloroplast membrane. In addition, a majority of the CYP97A4 protein was found in the soluble fraction, which also suggests that the protein is peripherally membrane-bound (**Fig. 3-7**). As a soluble control, the thylakoid lumen tpsOE16::GFP (Marques et al., 2003) was used; most of this protein as expected was contained in the soluble fraction (**Fig. 3-7**). These results corroborated with HMMTOP server predictions which indicated that the CYP97A4 and CYP97C2 enzymes each contain only a single transmembrane anchoring domain, which would only allow for a peripheral association with the membrane. These studies also indicate that the HYD3 and HYD4 proteins show a stronger association with the membrane than the CYP97 enzymes; HYD membrane fractions were still intact after treatment with alkaline (**Fig. 3-7**) suggesting an integral membrane association for these enzymes. The light-harvesting complex protein (LHCP), which is known to bind integrally to thylakoid membranes (Tan et al., 2001), was used as an intergral membrane control (**Fig. 3-7**). These data obtained from assays involving the HYD enzymes are concordant with earlier chloroplast import studies which showed that the diiron carotene β-ring hydroxylase from citrus (CitChyb) is integrally-bound to the chloroplast membrane (Inoue et al., 2006). In addition, HMMTOP predictions also indicated that the HYD3 and HYD4 enzymes each harbor four transmembrane helices which would render them significantly lipophilic, thereby enabling these enzymes to embed integrally in the chloroplast membrane.

Taken together, these studies indicate that, as expected, the CYP97A4 and CYP97C2, as well as the HYD3 and HYD4 enzymes which respectively have similar activities and specificities, localize to the same suborganellar/subcellular locations.

3.3.3 *In vivo* protein localization and protein interaction studies

3.3.3.1 *GFP fusion assays*

As described previously, the CYP97 and HYD enzymes were predicted by ChloroP to contain chloroplast targeting sequences that would enable these proteins to localize to chloroplasts. Chloroplast import assays confirmed import of these proteins *in vitro*. In order to confirm import *in vivo*, GFP fusion assays were performed.

The CYP97A4, CYP97C2, and HYD3 and HYD4 proteins were respectively fused to the GFP protein. Protoplasts were isolated from either etiolated or green leaf maize tissue and transformed with these GFP protein fusion constructs via Polyethylene Glycol (PEG) mediated transformation. Transformed protoplasts were examined by confocal microscopy. The GFP fluorescence observed in **Figures 3-8 and 3-10** co-localized with chlorophyll autofluorescence to confirm that the CYP97 and diiron HYD enzymes are localized to chloroplasts (**Fig. 3-8 and 3-10**).

GFP fluorescence in protoplasts isolated from both etiolated and green tissues showed distinct patterns for the CYP97 and HYD protein fusions. In protoplasts transformed with the CYP97A4 and CYP97C2 GFP fusions GFP fluorescence

appeared to be more diffuse than that observed with transformants harboring the HYD3 and HYD4 GFP fusions, which appeared to cluster in more discrete patches within the chloroplast (**Fig. 3-8 and 3-10**). The diffuse pattern of GFP fluorescence observed with the CYP97 enzymes is similar to the pattern observed with the GFP soluble control (**Fig. 3-8 and 3-10**), which would be expected due to the more soluble nature of the CYP97 enzymes. These observations are consistent with results of our import experiments as well as with HMMTOP predictions as described previously. The more discrete clusters of GFP fluorescence observed with the diiron HYD enzymes point to these proteins as being less soluble, which is concordant with the import data as well as with bioinformatics (HMMTOP) server predictions with respect to these enzymes. In protoplasts isolated from etiolated tissues in particular, both HYD3 and HYD4 GFP fusion fluorescence also accumulated in a ring at the chloroplast boundary suggesting envelope localization (**Fig. 3-8B**). The Toc34 protein was used as a chloroplast envelope control (Chen and Schnell, 1997); as expected, Toc34 GFP fusion fluorescence accumulated at the boundary of the chloroplast.

3.3.3.2 *Bimolecular Fluorescence Complementation (BiFC) assays*

As stated previously, one of the main goals of these studies was to determine whether any of the CYP97 and diiron HYD enzymes require an interacting partner enzyme for optimal activity. Bacterial complementation studies indicated that the CYP97A4 and CYP97C2 enzymes function optimally when expressed together to

efficiently convert α-carotene to lutein, however, there remained the open question as to whether these enzymes physically interact to form a heterodimer *in planta*.

In order to address this question, Bimolecular Fluorescence Complementation (BiFC) assays were performed, in which various combinations of CYP97/HYD proteins were tested for protein-protein interaction. The BiFC assay involves fusions of putative interacting proteins to non-fluorescent N-terminal (nYFP) and C-terminal (cYFP) fragments of the yellow fluorescent protein (YFP); if there is interaction between the two proteins, they bring together the non-fluorescent fragments thereby allowing for a restoration of fluorescence. These nYFP and cYFP sequences were introduced into sets of pSATN vectors (Citovsky et al., 2006).

Various combinations of the CYP97A4, CYP97C2, HYD3, and HYD4 enzymes were C-terminally fused to the N- and C-terminal halves of the YFP protein; the resulting constructs were transiently co-expressed in maize protoplasts from both etiolated and green leaf tissue and examined using confocal microscopy.

These studies demonstrated interaction for the CYP97A4 and CYP97C2 enzymes in protoplasts isolated from etiolated maize leaf tissue (**Fig. 3-9**). It was expected that the CYP97 enzymes would interact to form a heterodimer complex given that they functioned optimally when co-expressed in the conversion of α-carotene to lutein in the *E. coli* system; when expressed as individual enzymes, the CYP97A4 and CYP97C2 exhibited suboptimal activity. In addition, the HYD4 formed a homodimer complex in protoplasts isolated from both etiolated and green leaf tissues (**Fig. 3-9 and 3-11**). As seen with the complementation studies, the

HYD4 functioned optimally when expressed individually in both β-carotene only and α- and β-carotene accumulating *E. coli* to convert β-carotene to the end-product zeaxanthin (**Fig. 3-2 and 3-4**). Thus, it is reasonable to postulate that the robust enzyme activity observed with the HYD4 is due to this enzyme's ability to form a homodimer complex for the efficient conversion of the β-carotene substrate to zeaxanthin. The homodimer formation observed with the HYD4 was the only interaction demonstrated in protoplasts isolated from both etiolated and green leaf tissues.

Interestingly, the HYD3 and HYD4 enzymes also interacted to form a heterodimer in protoplasts isolated from etiolated leaf tissue (**Fig. 3-9**). It was not feasible to test combined HYD3 and HYD4 enzyme activities in the *E. coli* system as it would be impossible to discern individual enzyme activities in this case.

These studies did not indicate any homodimer formations for CYP97A4, CYP97C2, or HYD3, and no heterodimer complexes were observed for the CYP97A4 + HYD3, CYP97A4 + HYD4, CYP97C2 + HYD3, or CYP97C2 + HYD4 combinations tested in these assays **(Fig. 3-12, Tables 3-6 and 3-7)**.

In all cases where protein interaction was observed, YFP fluorescence co-localized with chlorophyll autofluorescence to confirm that these interacting enzymes are localized to chloroplasts. The ChrD protein from cucumber was also used as a control for both subcellular localization and protein interaction; ChrD is known to form homodimer complexes in plastids (Libal-Weksler et al., 1997) (**Fig. 3-9 and 3-11**). As seen with the GFP fusion assays, YFP fluorescence exhibited a diffuse

pattern throughout the chloroplast in assays demonstrating CYP97A4 and CYP97C2 interaction, in keeping with the soluble nature of these enzymes, while YFP fluorescence associated with the diiron HYD interactions appeared to be less diffuse, due to their less soluble nature. In protoplasts isolated from both etiolated and green leaf tissues, the HYD3 + HYD4 heterodimer and HYD4 homodimer complexes appeared to localize to the chloroplast boundary (**Fig. 3-9B and 3-11B**), indicating that these complexes are bound to the envelope membrane. However, as seen with the HYD GFP fusions expressed in protoplasts isolated from light grown tissue, envelope localization for the HYD4 homodimer complex was not as obvious; fluorescence was mainly observed in discrete patches within the chloroplast (**Fig. 3-11**).

3.4 Discussion

The development of rational strategies for metabolic engineering and/or marker-based breeding efforts aimed at enhancing provitamin A levels in cereal crops will require a better understanding of the mechanisms involved in controlling the conversion of provitamin A carotenes to non-provitamin A xanthophylls. In order to gain a better understanding of this conversion, these studies investigated the activities and localization/interaction of both maize and rice representatives from the two known structurally distinct classes of carotene hydroxylases: the nonheme diiron-type

hydroxylases from maize (HYD3 and HYD4) and the cytochrome P450-type hydroxylases CYP97A and CYP97C from rice.

3.4.1 *E. coli* functional complementation data integrated with earlier genetic studies

Functional complementation assays involving α- and β-carotene accumulating *E. coli* demonstrated that the CYP97A4 and CYP97C2 enzymes function optimally when expressed together and are preferentially active toward the α-carotene substrate for the production of lutein, while the diiron HYD enzymes are preferentially active toward the β-rings of β-carotene to convert this substrate to zeaxanthin.

The *E. coli* complementation studies described here, together with previous complementation assays which involved cells that accumulated substrates with either β-rings only (ie., β-carotene), or ε-rings rings only (ie., ε- ε-carotene), indicated that when expressed as individual enzymes, CYP97A4 and CYP97C2 activities were suboptimal, regardless of substrate choice (**Chapter 2**) (Quinlan et al., 2007). As shown in this current work, when expressed individually in α- and β-carotene accumulating *E. coli*, the CYP97A4 was unable to convert the α-carotene substrate to the pathway end-product lutein; only the monohydroxylated intermediate zeinoxanthin accumulated. These results do not comport with previous genetic data involving *Arabidopsis* quadruple mutants which lacked all *CYP97/HYD* genes; enzyme activities were inferred based on the phenotype of double and triple mutants. Triple mutants containing only the CYP97A enzyme accumulated 6% of WT lutein

levels (Kim et al., 2009). It's possible that another as yet unknown carotene hydroxylase exists which could be compensating for the loss of CYP97C activity in the triple mutants. In addition, it's also possible that this unknown carotene hydroxylase requires interaction with one of the other known CYP97 and/or HYD enzymes for activity, which would explain the lack of hydroxylated products in quadruple mutants lacking the other four known carotene hydroxylases.

It was also observed that, when expressed individually, the CYP97C2 was barely active toward the α-carotene substrate in *E. coli*. These results conflict with the aforementioned *Arabidopsis* genetics studies, which indicated that triple mutants containing only the CYP97C enzyme generated 74% of WT lutein (Kim et al., 2009). The genetic data point to the CYP97C enzyme as being efficiently active toward both the β- and ε-rings of α-carotene, which was not observed in the complementation systems described here. These *E. coli* studies indicated that the CYP97C2, when expressed as an individual enzyme in α- and β-carotene accumulating *E. coli,* was only marginally active toward both the ε-rings of α-carotene and the β-rings of β-carotene, to respectively produce very low amounts of the intermediates α-cryptoxanthin and β-cryptoxanthin. The triple mutant phenotype may be better explained by activity of an additional endogenous P450 carotene hydroxylase which may require interaction with the CYP97C for activity. Moreover, the mutant studies indicated that the diiron HYD enzyme cannot adequately compensate for lack of CYP97C activity, as double mutants containing only the two *Arabidopsis* HYD

enzymes (BCH1 and BCH2) generated only very low levels of lutein (approx. 1% of WT lutein levels) (Kim et al., 2009).

It can be inferred from these complementation studies that the enhanced CYP97C2 activity observed when this enzyme is co-expressed with the CYP97A4 in α- and β-carotene accumulating *E. coli* may in part be due to the ability of the CYP97A4 to accept the α-carotene substrate directly to produce zeinoxanthin, the putatively preferred CYP97C2 substrate. Similarly, the moderately enhanced CYP97C2 activity observed when this enzyme was co-expressed with HYD3 is likely also due to the ability of HYD3 to hydroxylate the β-ring of α-carotene first to generate the intermediate zeinoxanthin. These results are in agreement with earlier genetic studies involving *Arabidopsis* CYP97 mutants. These studies showed that mutants lacking a functional CYP97C enzyme *(lut1)* conditioned accumulation of the monohydroxylated intermediate zeinoxanthin, while plants defective in CYP97A *(lut5)* activity generated accumulation of α-carotene rather than the intermediate α-cryptoxanthin which would be expected if the CYP97C enzyme was active toward the α-carotene substrate (Kim and DellaPenna, 2006). These genetic studies, together with our *E. coli* complementation data, suggest that zeinoxanthin is the preferred CYP97C substrate, and that hydroxylation of α-carotene occurs in an ordered manner with the CYP97A hydroxylating the β-ring of α-carotene first to produce zeinoxanthin, which is then subsequently converted to the end-product lutein by the action of the CYP97C enzyme.

3.4.2 Protein localization and interaction studies integrated with functional complementation data

Earlier modeling described separate multi-enzyme complexes for the respective conversion of α- and β-carotene to lutein and zeaxanthin (Cunningham and Gantt, 1998). The functional complementation assays performed in this current work showed that in α- and β-carotene accumulating *E. coli* the CYP97A4 and CYP97C2 enzymes function optimally when expressed together to efficiently convert their preferred substrate, α-carotene to lutein. The diiron HYD3 and HYD4 enzymes were respectively preferentially active toward β-carotene to convert this substrate to the end-product zeaxanthin. In addition, the HYD4 functioned optimally when expressed as an individual enzyme for the efficient conversion of β-carotene to zeaxanthin. From these data, as well as the earlier hypothetical model for the existence of separate carotenogenic complexes described by Cunningham and Gantt (1998), it was reasonable to postulate that the CYP97A4 and CYP97C2 enzymes function in a membrane-bound topologically coordinated heterodimer complex on the α-carotene (ε-β) "arm" of the carotenoid pathway to efficiently channel the α-carotene substrate towards production of the end-product lutein. By contrast, the preferential activity towards the β-carotene substrate observed with the diiron HYD enzymes suggested that these enzymes function optimally on the β-carotene (β-β) "arm" of the pathway. In addition, when expressed as an individual enzyme in *E. coli* accumulating a β-carotene substrate, the HYD4 in particular appeared to be robustly active, which suggested that this enzyme may function optimally in a homodimer complex in the

chloroplast membrane to efficiently convert the β-carotene substrate to the end-product zeaxanthin. Furthermore, given their respective activities and specificities, as well as the previously described bioinformatics predictions by Chorop and HMMTOP, it was also expected that these two separate classes of carotene hydroxylases (CYP97 vs. HYD) would localize to chloroplasts, but form different associations with the chloroplast membrane.

To test these scenarios, both *in vitro* and *in vivo* approaches were used to examine CYP97 and diiron HYD protein localization and protein interaction. Bimolecular Fluorescence (BiFC) assays demonstrated that, as predicted, the CYP97A4 and CYP97C2 enzymes localize to chloroplasts and interact to form a heterodimer complex *in planta*. *In vitro* import studies also revealed that these soluble proteins are both peripherally associated with the chloroplast membrane; since these enzymes interact and together operate mainly on the α-carotene "arm" of the carotenoid pathway, a similar membrane association for these enzymes was expected.

This interaction between CYP97A4 and CYP97C2 appears to be required for optimal enzyme activity as the *E. coli* studies imply. Indeed, combined observations based on data from both *E. coli* complementation and protein interaction studies with respect to these two enzymes parallel those found *in planta* as it has been established that lutein accumulates in *Arabidopsis* leaf tissue, whereas the intermediates α-cryptoxanthin and zeinoxanthin do not (Pogson et al., 1996). These data point to a channeled flow of substrates mediated by an efficient CYP97 heterodimer complex

on the α-carotene side of the carotenoid pathway. Interaction between the CYP97A4 and CYP97C2 would facilitate channeling of the α-carotene substrate by allowing the enzymes to simultaneously be brought into close proximity to the β- and ε-rings of the substrate. The CYP97A4 is hypothesized to hydroxylate the β-ring of α-carotene first to generate zeinoxanthin, the preferred CYP97C2 substrate. The CYP97C2, by virtue of being constrained in a heterodimer complex with the CYP97A4, would then be in close proximity to the other end of the zeinoxanthin substrate which has an ε-ring available for hydroxylation; the CYP97C2 would then be in the proper orientation to efficiently confer hydroxylation of this ε-ring and convert this intermediate to the end-product lutein.

Similarly, the HYD4 enzyme, which was demonstrated via BiFC to form a homodimer complex, would be capable of simulataneouly hydroxylating the two β-rings at either end of β-carotene. Thus, this enzyme, which showed such robust activity when expressed as an individual enzyme toward the β-carotene substrate in our *E. coli* system, would, in the form of an integrally membrane-bound homodimer complex function to efficiently hydroxylate both β-rings of β-carotene thereby effectively channeling this substrate toward the production of the end-product zeaxanthin.

Interestingly, the HYD3 and HYD4 enzymes were also shown to interact to form a heterodimer complex. Transcript profiling indicated that the HYD3 and HYD4 are expressed at similar levels in maize leaf and root tissue (Vallabhaneni et al., 2009). Similar transcript levels would lead to eqi-molar amounts of HYD3 and

HYD4 protein which would facilitate efficient interaction between these two proteins. In addition, both enzymes were demonstrated via *in vitro* import to be integrally associated with the chloroplast membrane; this result is in agreement with previous chloroplast import studies which showed that a citrus diiron carotene β-ring hydroxylase (CitChyb) was integrally-bound to the chloroplast membrane (Inoue et al., 2006). It was expected that the HYD3 and HYD4 would share the same subcellular/suborganellar location given their similar activities and specificities as demonstrated by the *E. coli* complementation system, as well as by previous genetic and biochemical analyses (Sun et al., 1996; Tian et al., 2003; Tian et al., 2004).

Together, these studies indicate that the HYD3 and HYD4 enzymes form an integral association with the chloroplast membrane where they function in a heterodimer complex on the β-carotene side of the carotenoid pathway for the efficient channeling of the β-carotene substrate toward production of the pathway end-product zeaxanthin.

Complementation studies indicated that CYP97C2 activity was moderately enhanced when this enzyme was co-expressed with the HYD3 in α- and β-carotene accumulating *E. coli* (lutein accumulation was greater than 2-fold higher than zeaxanthin). This increase is relatively modest in comparison to the nearly 10-fold higher amount of lutein relative to zeaxanthin observed with cells co-expressing the CYP97A4 and CYP97C2. No interaction was observed between CYP97C2 and HYD3 via BiFC analysis. In terms of pathway metabolon assembly, the HYD3 and CYP97C2 enzymes would not be expected to be topologically coordinated partners,

given their predicted and experimentally determined differences with respect to membrane association (ie., integrally vs. peripherally-bound), as well as the fact that complementation studies indicated that these two structurally distinct classes of enzymes function optimally on the separate α-carotene and β-carotene sides of the carotenoid pathway for the production of lutein and zeaxanthin respectively. The modestly higher levels of lutein produced relative to zeaxanthin in the case of the CYP97C2 + HYD3 combination was likely due to the availability of the preferred CYP97C2 substrate, zeinoxanthin, which was generated by HYD3 activity towards the β-ring of α-carotene, rather than any protein interaction. HYD4 was also somewhat active toward the β-ring of α-carotene, which would make zeinoxanthin available for CYP97C2 ε-ring hydroxylation, however, only a low amount of lutein (roughly 1.6%) accumulated in cells co-expressing both HYD4 and CYP97C2. In addition, when co-expressed with the CYP97C2, HYD4 activity was severely inhibited; only a low amount of zeaxanthin was generated (approximately 3%). It's possible that the inhibited HYD4 activity observed may be due to substrate competition. The CYP97C2 was shown to have minor activity towards the β-rings of β-carotene, as shown **(Fig. 3-2 and 3-4)**. In the case of cells co-expressing CYP97C2 and HYD4, perhaps the CYP97C2 was binding to one of the β-rings of β-carotene but was not functioning efficiently in the hydroxylation of the ring, thus, the normal "bind and release" mechanism was impaired due to lack of proper β-ring hydroxylation function of the CYP97C2. The CYP97C2 enzyme may subsequently not have been released from the β-ring thereby occupying the site and preventing the

HYD4 from binding; this would necessarily inhibit HYD4 activity in the conversion of β-carotene to zeaxanthin. As stated previously, the HYD4 functioned optimally when expressed as an individual enzyme in the *E. coli* system; its ability to form a homodimer appears to enhance its activity toward β-carotene. By contrast, BiFC studies indicated that the HYD3 does not form a homodimer; its inability to form such a complex is likely responsible for the suboptimal activity observed toward the β-carotene substrate in our complementation assays.

3.4.3 BiFC studies involving protoplasts isolated from either etiolated or green leaf tissue

BiFC studies involving the use of protoplasts that had been isolated from etiolated maize leaf tissue demonstrated interactions between CYP97A4 and CYP97C2, as well as HYD3 and HYD4 (to form heterodimers). Transcript assays indicated that both the maize *CYP97A* and *CYP97C*, as well as the maize *HYD3* and *HYD4* gene pairs are similarly expressed in maize leaf tissue (although there was somewhat less similarity observed between *CYP97A* and *CYP97C* transcript levels) (Vallabhaneni et al., 2009). Similar transcript profiles in leaf tissue for both sets of enzymes would be expected given their respective heterodimer formations in protoplasts isolated from maize leaf tissue.

Although the HYD3 interacted with the HYD4 in protoplasts isolated from etiolated leaf tissue, no HYD3 homodimer formation was observed in protoplasts isolated from either of the two types of leaf tissue used. Transcript assays by

Vallabhaneni et al., (2009) showed that *HYD3* is more highly expressed in maize endosperm (which contains amyloplasts, not chloroplasts) than in any other tissue, including leaf tissue. In addition, these studies indicated a positive correlation between *HYD3* transcript levels and zeaxanthin accumulation in maize endosperm; that is, the higher the transcripts, the greater the level of zeaxanthin production, as would be expected from an increase in HYD3 hydroxylation activity. While no homodimer formation was observed for the HYD3 in protoplasts isolated from either etiolated or green leaf tissue, perhaps in endosperm, the HYD3 forms a homodimer complex on the amyloplast membrane for the efficient conversion of β-carotene to zeaxanthin. In addition, transcript profiling also indicated that both *HYD3* and *HYD4* are expressed in endosperm, although respective transcripts were at markedly dissimilar levels (*HYD3* transcripts were much higher than *HYD4* transcripts in this tissue) (Vallabhaneni et al., 2009). While interaction was demonstrated between HYD3 and HYD4 in protoplasts isolated from etiolated tissue, it's possible that these enzymes also form a heterodimer complex in endosperm amyloplasts; such an interaction may enhance HYD enzyme activities thereby compensating for the relative lack of HYD4 expression in this tissue; this heterodimer complex on the amyloplast membrane would facilitate an efficient conversion of β-carotene to zeaxanthin in maize endosperm.

Interestingly, only the HYD4 was able to form homodimer complexes in protoplasts isolated from both etiolated as well as green leaf tissues; these results point to the possibility that the HYD4 is able to promiscuously associate with a wider

variety of plastid membranes. Unlike green leaf tissue protoplasts, which would contain thylakoid membranes, protoplasts isolated from etiolated leaf tissue would have a membrane architecture that is somewhat similar to that of amyloplasts which lack internal membrane structures such as thylakoids. The ability of the HYD4 to form associations with various types of plastid membranes would facilitate an interaction between the HYD4 and the HYD3 in endosperm amyloplast membranes. Indeed, these studies showed that the HYD3 was able to interact with the HYD4 in protoplasts isolated from etiolated maize leaf tissue; such protoplasts would have a similar membrane structure to that of amyloplasts. In addition, while the proteomics analyses of Joyard et al., (2009), as well as the localization studies described in this work indicated that the CYP97 and HYD proteins repectively localize to the chloroplast envelope, it's possible that under certain stress conditions (ie., high-light), the HYD4 also localizes to thylakoid membranes where it forms a homodimer complex to function in the efficient production of zeaxanthin which has been shown to play a vital photoprotective role in the photosynthetic apparatus. Thylakoid membrane association of the downstream enzyme violaxanthin de-epoxidase (VDE), which converts violaxanthin to the photoprotective xanthophyll zeaxanthin, is activated by high-light induced acidification of the thylakoid lumen (lumenal pH 5/ stromal pH 8). While the isoelectric point (pI) of the mature HYD4 (pI 9.51) is markedly different to that of the VDE enzyme (pI 4.57), perhaps, given its integral association with the chloroplast membrane, the HYD4 is less sensitive to either lumenal or stromal pH than the peripherally membrane-bound CYP97A enzyme (pI

5.44). Thus, in such high-light conditions, the HYD4 would be capable of forming an association with the thylakoid membrane for the production of zeaxanthin, a carotenoid required for the dissipation of high-light energy.

Given that both classes of cartotene hydroxylase enzymes appear to localize to envelope membranes yet their respective hydroxylated end-products lutein and zeaxanthin play vital roles in LHC assembly and photoprotection respectively in thylakoid membranes (Pogson et al., 1998; Lokstein et al., 2002; Dall'Osto et al., 2006; Dall'Osto et al., 2007) raises the question as to what trafficking mechanisms are involved in transporting these enzymes and/or metabolites to the thylakoid membrane. These poorly understood phenomena are further complicated by the fact that the CYP97A and CYP97C enzymes are coded for by single-copy genes which may apparently be targeted for multiple suborganellar locations.

Previous electron microscopical data has indicated that plastid-generated lipid bodies referred to as plastoglobules, which typically contain galactolipids, plastoquinones, and carotenoids (Greenwood et al., 1963), are associated with the stromal surfaces of both thylakoid and inner chloroplast membranes (Kessler et al., 1999), and more recent studies suggested that plastoglobules may be involved in the trafficking of inner chloroplast envelope enzymes and/or metabolites to the thylakoid membrane (Vidi et al., 2006). Specifically, these latter studies by Vidi et al., (2006) point to the plastoglobule as being the site of the tocopherol cyclase (VTE1) reaction of vitamin E synthesis; the VTE1 substrate 2,3-dimethyl-5 phytyl-1, 4-hydroquinol, an inner envelope protein is transferred from the inner envelope to the plastoglobule

where the cyclase reaction takes place. VTE4 (γ-tocopherol methyl transferase) completes α-tocopherol biosynthesis, however this enzyme was not found in plastoglobules; thus the final methylation step does not occur in these lipid bodies. Therefore plastoglobules may not only be the site of tocopherol cyclase activity, but may serve as trafficking vehicles for tocopherol biosynthetic intermediates. Taken together, these data indicate that in addition to providing a transport pathway, plastoglobules are dynamic, metabolically active structures.

Similarly, plastoglobules may serve to transport the CYP97/HYD enzymes and/or their hydroxylated products lutein and zeaxanthin from the chloroplast envelope to the thylakoids. In addition, given that these lipid bodies have been demonstrated to have metabolic functions, it is possible that at least a portion of the carotenoid metabolites targeted for the thylakoid membrane, such as lutein and zeaxanthin, are synthesized in plastoglobules rather than on the chloroplast envelope.

Another possible mode of enzyme/metabolite transport may involve envelope-derived vesicles. It has been established that biogenesis and maintenance of thylakoids is largely dependent on metabolites such as galactolipids and carotenoids, which are mainly synthesized on the inner chloroplast envelope (Douce, 1974; Joyard et al., 1980). Early electron microscopic studies described vesicular structures in chloroplasts (von Wettstein, 1958; Muhlethaler and Frey-Wyssling, 1959), and more recent data has indicated that vesiculation of the envelope inner membrane may provide a transport pathway for inner envelope derived lipids to the thylakoids (Carde et al., 1982; Hoober et al., 1991; Morre et al., 1991). Inhibitor studies

presented definitive evidence for the existence of a protein (phosphatase-calmodulin)-mediated vesicular transport system that connects the inner-envelope and thylakoid membranes in chloroplasts (Westphal et al., 2001). In addition, mutational analysis in *Arabidopsis* involving the deletion of a single gene *VIPP1* (vesicle-inducing protein in plastids 1) which codes for the VIPP1 protein involved in the budding of vesicles from the inner envelope, showed that the lack of vesicle generation in these mutants is paralleled by the inhibition of thylakoid development. These data indicated that VIPP1 may be involved in a vesicle trafficking system that links the chloroplast inner envelope and thylakoid membranes (Kroll et al., 2001).

It's also possible that these envelope-derived vesicles may be involved in transporting the CYP97/HYD enzymes and/or their hydroxylated products lutein and zeaxanthin from the envelope to the thylakoids. Further study will be required in order to fully elucidate the specific mechanisms involved in the trafficking of carotenoid enzymes and/or metabolites between chloroplast membranes.

3.5 Conclusion

Taken together, these data represent a coherent set of studies that offers new insights into the roles of the two separate classes of carotene hydroxylase enzymes involved in the conversion of provitamin A carotenes to non-provitamin A xanthophylls. This seminal body of work lays the foundation for further investigations into the roles and topologies of the putative carotenoid metabolons, and should provide a solid knowledge base for the development of rational strategies for

metabolic engineering and/or breeding to improve provitamin A carotenoids in cereal endosperm.

3.6 Materials and Methods

3.6.1 Cloning of CYP97A4, CYP97C2, HYD3, and HYD4 into pCOLADuet and pCDFDuet vectors

The CYP97A4 (#AK068163) and CYP97C2 (#AK065689) ORFs were amplified from rice cDNA and cloned into the pCOLADuet^{TM}-1 vector (Novagen), and resulting constructs were respectively named pRT-A4 and pTR-C2 as described previously (**Chapter 2**) (Quinlan et al., 2007). The CYP97C2 ORF was amplified from pTR-C2 using the following primers with terminal NdeI and Acc65I restriction sites: (forward): 5′- ACCG CAT ATG GCC GTC CCG TGC GTA C – 3′, (reverse): 5′-GAGA GGT ACC TCA TCT GGA CCC ACT GAG- 3′. The HYD3 (#BM382572/AY844958) ORF was amplified from pTHYD3 (Vallabhaneni et al., 2009) using the following primers with terminal EcoRI and HindIII restriction sites: (forward): 5′-GAGA G AAT TCC ATG GCC GCC GCG ATG A-3′, (reverse): 5′-ACCG AAG CTT CTA GAA CTC ATT TGG CAC A-3′, and gene specific primers with terminal EcoRI and HindIII restriction sites (forward 5′- GAGA G AAT TCA ATG GCC GCC GGT CTG T-3′ and reverse 5′- ACCG AAG CTT_TCA GAT GGT

CCG GCC G-3′) were used to amplify the HYD4 (#BG320875/AY844956) ORF from pTHYD4 (Vallabhaneni et al., 2009)

PCR conditions to amplify the CYP97C2 ORF, HYD3 and HYD4 ORFs were as follows: 20 µl reactions contained 1µl plasmid DNA, 1 µl of each primer (20 µM), and 17 µl of a Platinum PCR Supermix High Fidelity master mix, 1.1X (2.4 mM $MgSO_4$, 220 mM dNTPs, 22 U/ml complexed recombinant Taq DNA polymerase, Pyrococcus species GB-D thermostable polymerase, and Platinum Taq Antibody; 66mM Trid- SO_4 (pH 8.9) and stabilizers) (Invitrogen Corp.). Amplifications conditions for CYP97C2: 1 cycle, 95° C, 3 min; 40 cycles, 95° C, 45s, 58° C, 45s, 72° C, 2:00 min; 1 cycle, 72° C, 10 min.

3.6.2 Construction of expression vectors and functional analysis

CYP97C2 was cloned in frame into the NdeI and Acc65I sites (MSC2) of the pCDFDuet-1 vector (Novagen) and renamed pRQ-C2. HYD3 ws cloned in frame into the EcoRI and HindIII sites of the pCOLADuet™-1 vector and renamed pRQ-H3. HYD4 was cloned in frame into the EcoRI and HindIII sites of the pCOLADuet ™-1 vector and renamed pRQ-H4. For testing of substrate specificity for individual enzymes, pRT-A4 (carries the CYP97A4 ORF), pRQ-C2 (carries the CYP97C2 ORF), pRQ-H3 (carries the HYD3 ORF) or pRQ-H4 (carries the HYD4 ORF) were respectively transformed into *E. coli* BL21 (DE3) cells (Novagen) harboring either pACBETA-At only, which confers β-carotene accumulation (Cunningham et al., 2007), pACCRT-EIB (which confers accumulation of lycopene) (Cunningham et al.,

1993) + plasmid y2 (Cunningham Jr. et al., 1996) (which carries the *Arabidopsis* lycopene ε-cyclase); together these constructs confer accumulation of δ- and ε-ε-carotene, or pACBETA-At + plasmid y2, which together confer accumulation of α- and β- carotene (Cunningham et al., 2007). For testing of substrate specificity for enzyme combinations the pRT-A4 + pRQ-C2, pRQ-C2 + pRQ-H3, and pRQ-C2 + pRQ-H4 constructs were co-transformed into *E. coli* BL21 (DE3) cells harboring both pACBETA-At + plasmid y2 which together confer accumulation of α- and β-carotene. For carotenoid analyses, saturated cultures in LB medium were diluted 50-fold into 50 ml fresh medium in 500 ml flasks, then grown in the dark at 250 rpm at 37°C until OD 0.6 at which point they were induced with 10 mM IPTG and further cultured for a total of three days.

3.6.3 Extraction of carotenoids from *E. coli* cells, HPLC and LC-MS analysis

50 ml cultures were centrifuged at 3,500 rpm for 10 min. Bacterial cell pellets were extracted in 5 ml of methanol using a Sonicator (Vibra Cell), and centrifuged at 3,500 rpm for 10 min. to pellet disrupted cells. Supernatants were transferred to 100 ml Pyrex flasks and evaporated under nitrogen gas**.** Once dried, 300 µl of Methanol was added to dissolve samples. Samples were then frozen at -80°C for 30 minutes, centrifuged at 14,000 rpm at 4°C, and the supernatants transferred to HPLC vials (Waters).

Separation was carried out using a Waters HPLC system equipped with a 2695 Alliance separation module, a 996 photodiode array detector, a column heater, a

fraction collector II, Empower software (Millipore, Franklin, MA), and a Develosil C30 RP-Aqueous (5 μm, 250 x 4.6mm) column (Phenomenex, Torrance, CA), with a Nucleosil C_{18} (5 μm, 4 x 3.0 mm) guard column (Phenomenex, Torrance, CA), with mobile phase consisting of mixtures of acetonitrile:methanol:water (84:2:14 v/v/v (A) and methanol:ethyl acetate (68:32 v/v (B), with a gradient to obtain 100% B at 60 min (flow rate 0.6 ml/min), 100% B at 71 min with flow rate changing to 1.2 ml/min, followed by 100% A (flow rate 1.2 ml/min) at 110 min. Peaks were identified on the basis of retention times/spectra matching those of authentic standards (INDOFINE Chemical Company, Inc., Sommerville, NJ). All data were collected at lambda max of 450nm.

LC-MS was performed on a Waters 2695 HPLC equipped with a 2998 PDA detector coupled to a Waters LCT Premiere XE Time of Flight (TOF) Mass Spectrometer system using electrospray ionization in positive ion mode. Separation was performed using a Develosil C30 RP-Aqueous (5 μm, 250 x 4.6mm) column (Phenomenex, Torrance, CA), with mobile phase consisting of mixtures of acetonitrile:methanol:water (84:2:14 v/v/v (A) and methanol:ethyl acetate (68:32 v/v (B), with a gradient to obtain 100% B at 60 min (flow rate 0.6 ml/min), 100% B at 71 min with flow rate changing to 1.2 ml/min, followed by 100% A (flow rate 1.2 ml/min) at 110 min.

3.6.4 Cloning of CYP97A4, CYP97C2, HYD3, and HYD4 into the pTnT vector

The CYP97A4 ORF from rice was amplified from pRT-A4 via PCR using the following primers with terminal XhoI and XbaI restriction sites: primer 2175 (forward): 5′-ACCG CTC GAG GCC ACC ATG AGC TCA GCG ACG TCA GTG AGT G -3′, primer 2176 (reverse): 5′-GAGA TCT AGA TCA GAT TCG AGT TGC TGA GAC TTG C-3.′ The CYP97C2 ORF was amplified from rice cDNA using the following primers with terminal XhoI and XbaI restriction sites: primer 2140 (forward): 5′- GAGA CTC GAG AAT CCA TCT CGA ATC CCT AGC– 3′, primer 2168 (reverse): 5′-ACCG TCT AGA TCA TCT GGA CCC ACT GAG TG- 3.′ Primers with terminal XhoI/XbaI restriction sites used to amplify the maize diiron HYD3 and HYD4 ORFs from pTHYD3 and pTHYD4 respectively were: primer 2163 (forward): 5′- ACCG CTC GAG GCC ACC ATG GCC GCC GCG ATG ACC-3′, primer 2164 (reverse): 5′-GAGA TCT AGA CTA GAA CTC ATT TGG CAC ACT -3,′ and primer 2165 (forward): 5′-ACCG CTC GAG GCC ACC ATG GCC GCC GGT CTG TCC-3,′ primer 2166 (reverse): 5′-GAGA TCT AGA TCA GAT GGT CCG GCC GAT T-3.′

PCR conditions to amplify the CYP97A4, CYP97C2, maize HYD3 and maize HYD4 ORFs from were as follows: 20 µl reactions contained 1µl plasmid DNA, 1 µl of each primer (20 µM), and 17 µl of a Platinum PCR Supermix High Fidelity master mix, 1.1X (2.4 mM $MgSO_4$, 220 mM dNTPs, 22 U/ml complexed recombinant Taq DNA polymerase, Pyrococcus species GB-D thermostable polymerase, and Platinum Taq Antibody; 66mM Trid- SO_4 (pH 8.9) and stabilizers) (Invitrogen Corp.). Amplifications conditions for CYP97A4: 1 cycle, 95° C, 3 min; 40 cycles, 95° C, 45s, 58° C, 45s, 72° C, 2:30 min; 1 cycle, 72° C, 10 min.

CYP97A4, CYP97C2, maize HYD3 and maize HYD4 were cloned in frame into the XhoI and XbaI sites of the pTnT vector (Promega) and respectively named pTnT-A4, pTnT-C2, pTnT-H3, pTnT-H4.

3.6.5 Chloroplast Isolation

Chloroplasts used in import assays were isolated from 10-14 day old pea plants as described by (Bruce et al., 1994). Approx. 25g of leaves were homogenized at 4°C with a blender in 75ml of cold grinding buffer (50mM HEPES pH 8, 0.33M sorbitol, 1mM $MgCl_2$, 1mM $MnCl_2$, 2mM Na_2EDTA pH 8, 0.1% BSA, 0.1% Na-ascorbate) by 3-5 bursts of 1s each. All further operations were performed on ice using cold buffers. The homogenate was filtered through 2 layers of cheesecloth and 1 layer of Nylon mesh (60 μm) and the filtrate was centrifuged at 2000 g for 2 min. Pellets were carefully resuspended in 1ml of grinding buffer and overlaid on top of two 36 ml Percoll gradients (prepared by centrifugation of 50% Percoll (Sigma) in grinding buffer, 40000 g, 30 min, 4°C), and centrifuged at 12000 g, 11 min, 4°C. Intact chloroplasts in the lower band were gently collected with a pipette, washed with 3 volumes of import buffer (50mM HEPES pH 8, 0.33 M sorbitol) and then pelleted at 2000 g, 2 min, 4°C. Washed intact chloroplasts were resuspended in import buffer to yield chloroplast concentration of 0.5 mg/ml and kept on ice until use.

3.6.6 *In vitro* chloroplast import

The plasmid constructs previously described (pTnT-A4, pTnT-C2, pTnT-H3, pTnT-H4) were used as templates for in vitro transcription/translation performed with the TnT Coupled Reticulocyte Lysate System (Promega) in the presence of [^{35}S]-methionine according to the manufacturer's instructions.

Reaction mixtures were prepared containing purified chloroplasts (0.5 mg/ml), 1X import buffer, 4mM methionine, 4 mM ATP, 4 mM $MgCl_2$, 10 mM KAc, 10 mM $NaHCO_3$ and 10 µl of reticulocyte lysate translation product in a total volume of 150 µl. Reactions mixtures were incubated for 25 min at 25°C in light. Import reactions were stopped by adding 500 µl of 1X import buffer and samples were centrifuged at 800 g for 2 min at 4°C to obtain pellet of intact chloroplasts. Pellets were resuspended in 200 µl import buffer, supplemented by 1mM $CaCl_2$ and each reaction mixture was divided into two equal aliquots. Thermolysin was added to one of the two aliquots to a concentration of 125 ng/ul and incubated for 30 min at 4°C. The reaction was terminated by addition of EDTA to a concentration of 10 mM. For fractionation experiments after import reaction intact chloroplasts were washed twice with import buffer, then diluted with HL buffer (10 mM HEPES-KOH, 10 mM $MgCl_2$, pH=8); the total mixture was frozen in liquid nitrogen/thawed 3 times, and then centrifuged (16 000 g, 20 min). Alkaline treatment of membrane fractions were performed with 200 mM Na_2CO_3, pH>10, 10 min on ice, and pellets containing treated membranes were separated from the supernatant by centrifugation (16 000 g, 20 min). All fractions including soluble, membrane, and purified membrane pellets were analyzed by SDS-PAGE. Radiolabelled protein bands were visualized using a Storm Phosphoimager (Amersham Biosciences). A prestained molecular weight marker (Benchmark Pre-stained

protein ladder, Invitrogen) was included on each gel and used to determine the sizes of the radiolabeled bands.

3.6.7 Cloning of CYP97/HYDs into pUC35S-sGFP-Nos vectors

The CYP97A4 ORF was amplified from pRT-A4 via PCR using the following primers with terminal XbaI and BamHI restriction sites: (forward): 5′- ACCG TCT AGA ATG AGC TCA GCG ACG TCA GTG AG-3′ and (reverse): 5′- GAGA GGA TCC GAT TCG AGT TGC TGA GAC TTG CC-3′, the CYP97C2 ORF was amplified from pGEMT-C2 (pGEM®T-Easy vector (Promega) harboring full-length cDNA of CYP97C2) via PCR using the following primers with terminal XbaI and BclI restriction sites (forward): 5′- ACCG TCT AGA ATG GCC GCC GCC GCC GCC GCC GCC-3′ and (reverse): 5′-GAGATGATCATCTGGACCCACTGAGTGCAAAATCAG-3.′ Gene-specific primers with terminal XbaI and BamHI restriction sites (forward): 5′- ACCG_TCT AGA ATG GCC GCC GCG ATG ACC-3′ and reverse 5′- GAGA GGA TCC GAA CTC ATT TGG CAC ACT CT-3′ were used to amplify the full-length coding sequence of HYD3 via PCR from the pRQ-H3 vector. The HYD4 ORF was amplified from the pRQ-H4 vector via PCR using the following primers with XbaI and BamHI restriction sites (forward): 5′-ACCG_TCT AGA ATG GCC GCC GGT CTG TCC-3′ and (reverse): 5′-GAGA GGA TCC GAT GGT CCG GCC GAT TCG-3.′

PCR conditions to amplify the CYP97A4, CYP97C2, maize HYD3 and maize HYD4 ORFs from were as follows: 20 µl reactions contained 1µl plasmid DNA, 1 µl of each primer (20 µM), and 17 µl of a Platinum PCR Supermix High Fidelity master mix, 1.1X (2.4 mM $MgSO_4$, 220 mM dNTPs, 22 U/ml complexed recombinant Taq DNA polymerase, Pyrococcus species GB-D thermostable polymerase, and Platinum Taq Antibody; 66mM Trid- SO_4 (pH 8.9) and stabilizers) (Invitrogen Corp.). Amplifications conditions for CYP97A4 and CYP97C2: 1 cycle, 95° C, 3 min; 35 cycles, 95° C, 30s, 58° C, 45s, 72° C, 2:30 min; 1 cycle, 72° C, 10 min.

CYP97A4, CYP97C2, HYD3, and HYD4 were cloned in frame into the XbaI and BamHI (or BclI) sites of the pUC35S-sGFP-Nos vector (based on pUC35S-GUS-Nos and pBIG121 vectors) (Okada et al., 2000) and respectively named A4-GFP, C2-GFP, H3-GFP, H4-GFP.

3.6.8 Cloning of CYP97/HYDs into BiFC vectors

For cloning into the pSAT 2236 vectors (Citovsky et al., 2006):

The CYP97A4 ORF was amplified from the pRT-A4 vector via PCR using the following primers with terminal XhoI and EcoRI restriction sites: (forward): 5′-ACCGCTCGAGGCAACAATGAGCTCAGCGACGTCAGTGAG-3′, (reverse): 5′-GAGAGAATTCGATTCGAGTTGCTGAGACTTGCC-3′. The CYP97C2 ORF was amplified from pGEMT-C2 (pGEM®T-Easy vector (Promega) harboring full-length cDNA of CYP97C2) via PCR using the following primers with terminal XhoI and

EcoRI restriction sites: (forward) 5′-ACCG CTCGAGATGGCCGCCGCCGCC GCCGCCGCC-3′ and (reverse) GAGAGAATTC TCTGGACCCACTGAGTGCAAAATCAG-3.′ The HYD3 ORF was amplified from the pRQ-H3 vector using the following primers with terminal XhoI and EcoRI restriction sites: (forward): 5′- ACCGCTCGAGATGGCCGCCGCGATGACC -3′ and (reverse) 5′-GAGAGAATTCGAACTCATTTGGCACACT CT- 3′. The HYD4 ORF was amplified from the pRQ-H4 vector using the following primers with terminal XhoI and EcoRI restriction sites: (forward): 5′-ACCGCTCGAGATGGCC GCCGGTCTGTCC-3′ and (reverse): 5′-GAGAGAATTCGATGGTCCGGCCGAT TCG-3′.

For cloning into the pSAT 1476 vectors (Citovsky et al., 2006):

The CYP97A4 ORF was amplified from the pRT-A4 vector via PCR using the following primers with terminal XhoI and EcoRI restriction sites: (forward): 5′-ACCG CTCGAG ATG AGC TCA GCG ACG TCA GTG AG-3′ and (reverse): 5′-GAGA GAA TTC GAT TCG AGT TGC TGA GAC TTG CC-3′. The CYP97C2 ORF was amplified from pGEMT-C2 (pGEM®T-Easy vector (Promega) harboring full-length cDNA of CYP97C2) via PCR using the following primers with terminal NcoI and EcoRI restriction sites: (forward): 5′-ACCGCCATGGCCGCCGCCGCC-3′ and (reverse): 5′-GAGAGAATTCTCTGGACCCACTGAGTGC-3′. The HYD3 ORF was amplified from the pRQ-H3 vector via PCR using the following primers with terminal BspHI and EcoRI restriction sites: (forward): 5′- ACCGTCATGA TGGCCGCCGCGATGACCAG and (reverse): 5′- GAGA GAA TTC

GAACTCATTTGGCACACTCTGGC-3′. The The HYD4 ORF was amplified from the pRQ-H4 vector via PCR using the following primers with terminal BspHI and EcoRI restriction sites: (forward): 5′- ACCG TCA TG A TGGCCGCCGGTCTGTCCGG -3′ and (reverse): 5′- GAGA GAA TTC GATGGTCCGGCCGATTCGCG-3′.

PCR conditions to amplify the CYP97A4, CYP97C2, maize HYD3 and maize HYD4 ORFs from were as follows: 20 µl reactions contained 1µl plasmid DNA, 1 µl of each primer (20 µM), and 17 µl of a Platinum PCR Supermix High Fidelity master mix, 1.1X (2.4 mM $MgSO_4$, 220 mM dNTPs, 22 U/ml complexed recombinant Taq DNA polymerase, Pyrococcus species GB-D thermostable polymerase, and Platinum Taq Antibody; 66mM Trid- SO_4 (pH 8.9) and stabilizers) (Invitrogen Corp.). Amplifications conditions for CYP97A4 and CYP97C2: 1 cycle, 95° C, 3 min; 35 cycles, 95° C, 30s, 58° C, 45s, 72° C, 2:30 min; 1 cycle, 72° C, 10 min.

CYP97A4, CYP97C2, HYD3, and HYD4 were cloned into the XhoI and EcoRI sites of pSAT vector 2236 (pSat4(A)-nEYFP-N1) (Citovsky et al., 2006), and respectively named A4_2236, C2_ 2236, H3_2236, H4_2236.

CYP97A4 was cloned into the XhoI and EcoRI sites of pSAT vector 1476 (pSAT6-cEYFP-N1) (Citovsky et al., 2006) and named A4_1476. CYP97C2 was cloned into the NcoI and EcoRI sites of the pSAT vector 1476 (pSAT6-cEYFP-N1) and named C2_1476, and HYD3 and HYD4 were respectively cloned into the BspHI and EcoRI sites of the pSAT vector 1476 (pSAT6-cEYFP-N1) and named H3_1476 and H4_1476.

3.6.9 Isolation and transformation of maize protoplasts

Isolation and transformation of maize protoplasts were performed according to (Sheen, 1991; van Bokhoven et al., 1993) with modifications. Maize var. B73 plants were grown in the dark at 26°C for 12 days (12 h day, 12 h night in an Avantis growth chamber (Conviron)). Middle parts of 2nd leaves of 20 plants were cut into razor thin sections and transferred to 500 ml Erlenmeyer flask containing 50 ml of Ca/Mannitol solution (10mM $CaCl_2$, 0.6M Mannitol, 20mM MES pH 5.7) to which was added 1% cellulase (*Trichoderma viride*), 0.3% pectinase (*Rhizopus sp.)* (Sigma), 5mM β-mercaptoethanol (Sigma), and 0.1% BSA (Sigma). Vacuum was applied for 5 min followed by shaking at 60 rpm at RT in the dark for 3 hours. The supernatant was filtered by 60 µm nylon mesh, and collected in a 50 ml Falcon centrifuge tube. Protoplasts were pelleted at 60g for 5 min at RT and then washed with 25 ml Ca/Mannitol solution (repeated 3 times). Protoplasts were aliquoted into portions of 10^6 in 150 µl. To each reaction 10 µg of ice-cold plasmid DNA was added. Protoplasts were then mixed with 500 µl of polyethylene glycol solution (40% PEG 6000, 0.5M Mannitol, 0.1 M $Ca(NO_3)_2$) for 10 seconds followed by addition of 4.5 ml of Mannitol/MES solution (15 mM $MgCl_2$, 0.1% MES, pH 5.5, 0.5M Mannitol) and incubated at RT for 25 min. The suspension was then centrifuged at 60 g for 5 min at RT and the supernatant was discarded. The sediment was washed with Ca/Mannitol solution and pelleted at 60 g for 5 min at RT. The supernatant was discarded and protoplasts were re-suspended in 1 ml Ca/Mannitol solution.

Protoplasts were transferred to a 24 well plate and incubated overnight at 25°C under dim light. Transformational efficiency for protoplasts was 80-90%.

.

3.6.10 Confocal microscopy

Transient expression of GFP, YFP and RFP fusion proteins was visualized using a DM16000B inverted confocal microscope with TCS SP5 system (Leica Microsystems CMS, Germany). Oil or water immersion objective (63X) was used in all cases. A 488 nm argon laser was used to excite the fluorescence of GFP and chlorophyll and a 543 laser was used to excite RFP. Chlorophyll autofluorescence was detected between 664 and 696 nm, and GFP fluorescence was detected between 500 and 539 nm and always confirmed by recording the emission spectrum by wavelength scanning (lambda scan) between 500 and 600 nm with a 3-nm detection window. The RFP fluorescence was detected between 600 and 650 nm, and emission spectrum confirmed by scanning between 570 and 700 nm. A 514 nm line of argon laser was used to excite the fluorescence of YFP, and the emission spectrum was detected and confirmed by lambda scan between 524 and 575 nm.

LAS AF software (Leica Microsystems CMS, Germany) was used for image acquisition. Images were obtained by combining several confocal Z-planes.

3.7 Acknowledgements

We thank Dr. Francis Cunningham for the *Erwinia* expression plasmids, Dr. Vitaly Citovsky for the pSATN vectors, and Drs. Edward Kennelly, Akira

Kawamura, Chunhui Ma, Gema Flores, and Keyvan Dastmalchi for technical advice in LC-MS analysis.

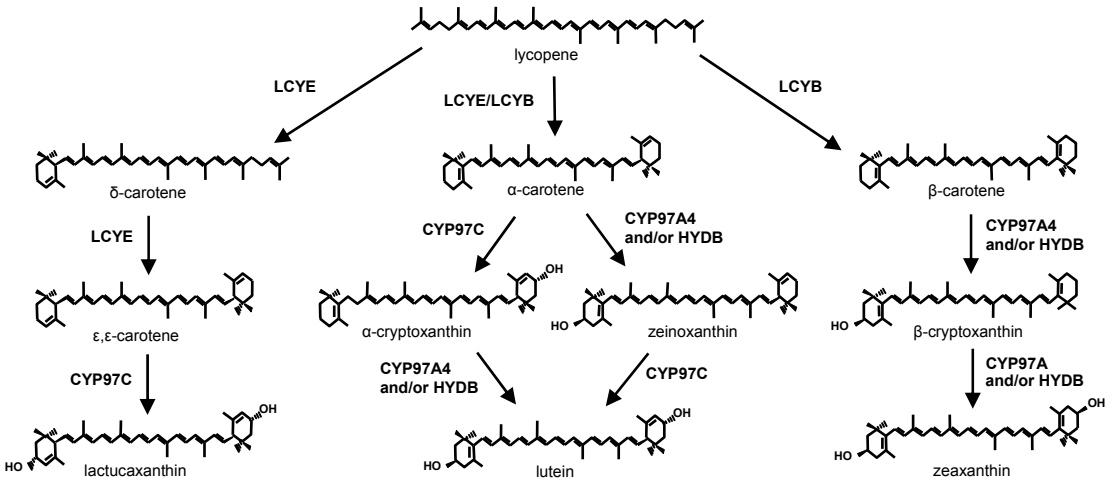


Figure. 3-1. Carotene hydroxylases convert provitamin A carotenes (ie., β- and α-carotene) to non-provitamin A xanthophylls (ie., lutein and zeaxanthin). Enzymes: CYP97A, P450 β-hydroxylase; CYP97C, P450 ε-hydroxylase; HYDB, diiron β-hydroxylase. Conversion of lycopene to β-carotene requires lycopene β-cyclase (LCYB) alone, while conversion of lycopene to α-carotene requires lycopene β- and ε-cyclases (LCYE). Conversion of lycopene to δ-carotene (intermediate) and ε,ε-carotene requires LCYE. Formation of the hydroxylated xanthophylls zeaxanthin, lutein, and lactucaxanthin is mediated by two separate stereospecific β- and ε-ring hydroxylases (CYP97/diiron HYD).

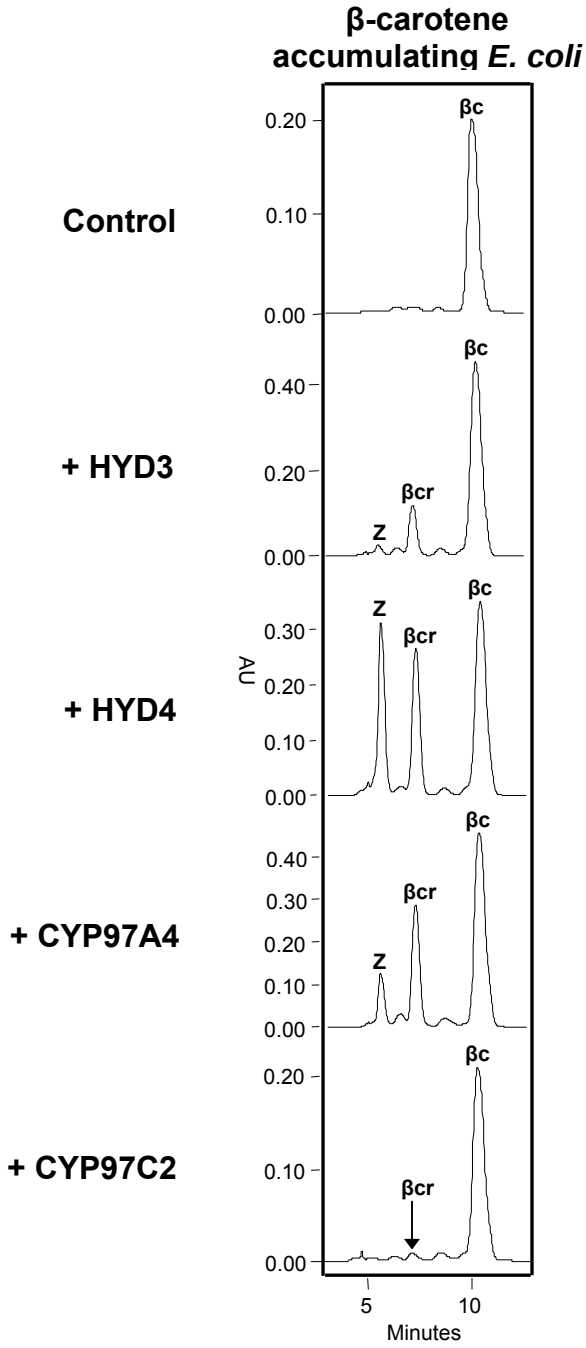


Figure 3-2. Functional complementation of CYP97/HYD enzymes in cells accumulating β-carotene. Extracted pigments were separated by reversed phase HPLC. *E. coli* cells engineered to accumulate a carotene with β-rings (β-carotene) were further transformed with test plasmids encoding HYD3, HYD4, CYP97A4, or CYP97C2. HPLC chromatograms (450 nm) show composition of accumulated substrates and products. **Control**, empty pColaDuet; **z,** zeaxanthin; **βcr**, β-cryptoxanthin; **βc**, β,β-carotene

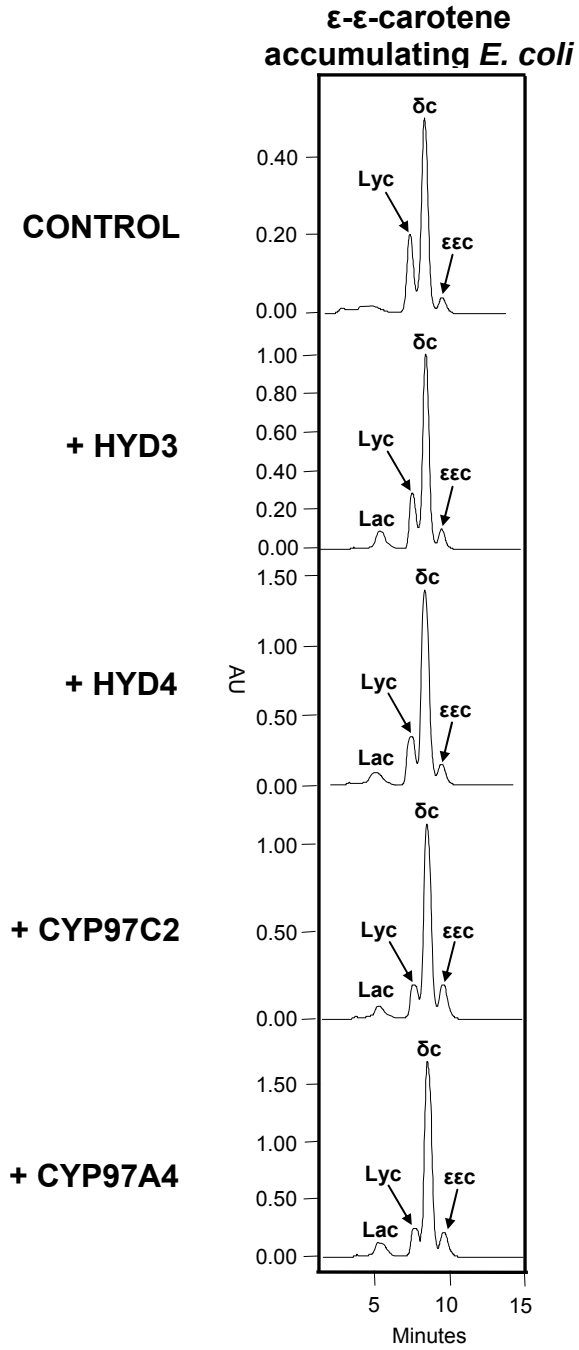


Figure 3-3. Functional complementation of CYP97/HYD enzymes in cells accumulating ε-ε-carotene. Extracted pigments were separated by reversed phase HPLC. *E. coli* cells engineered to accumulate a carotene with ε-rings (ε-ε-carotene) were further transformed with test plasmids encoding HYD3, HYD4, CYP97A4, or CYP97C2. HPLC chromatograms (450 nm) show composition of accumulated substrates and products. **Control**, empty pColaDuet; **Lac,** lactucaxanthin; **Lyc**, lycopene; **δc,** δ-carotene, **εεc,** ε-ε-carotene

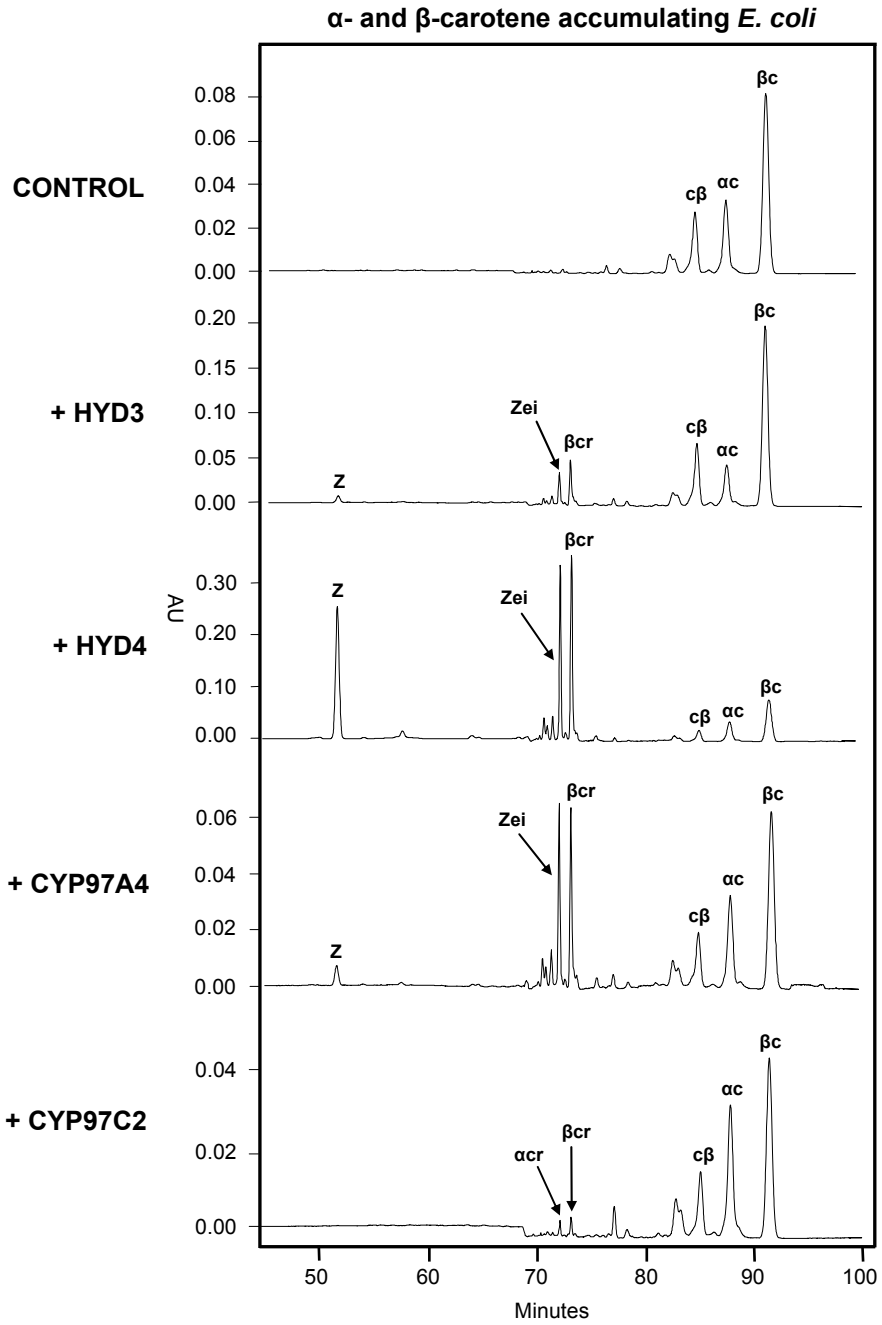


Figure 3-4. Functional complementation of CYP97/HYD enzymes in cells accumulating α- and β-carotene. Extracted pigments were separated by reversed phase HPLC. *E. coli* cells engineered to accumulate carotenes with β-rings (β-carotene) only as well as both α- and β-rings (α-carotene) were further transformed with test plasmids encoding HYD3, HYD4, CYP97A4, or CYP97C2. HPLC chromatograms (450 nm) show composition of accumulated substrates and products. **Control**, empty pColaDuet; **Z,** zeaxanthin; **Zei,** zeinoxanthin; **αcr**, α-cryptoxanthin; **βcr**, β-cryptoxanthin; **cβ,** 13-cis β-carotene; **αc;** α-carotene; **βc**, β-carotene

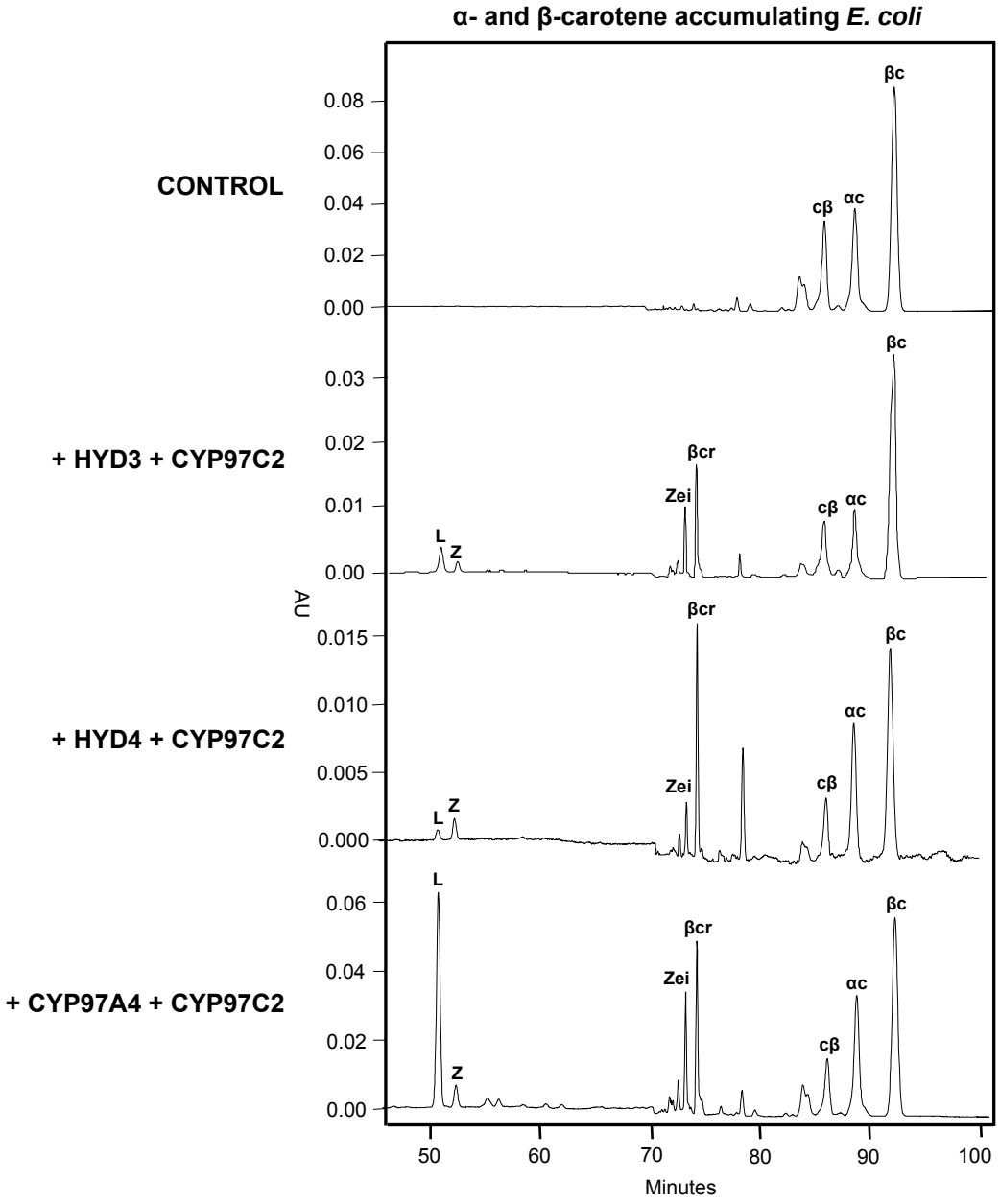


Figure 3-5. Functional complementation of CYP97/HYD enzyme combinations in cells accumulating α- and β-carotene. Extracted pigments were separated by reversed phase HPLC. *E. coli* cells engineered to accumulate carotenes with β-rings (β-carotene) only as well as both α- and β-rings (α-carotene) were further transformed with the following combinations of test plasmids: HYD3 + CYP97C2, HYD4 + CYP97C2, and CYP97A4 + CYP97C2. HPLC chromatograms (450 nm) show composition of accumulated substrates and products. **Control**, empty pColaDuet; **L,** lutein; **Z,** zeaxanthin; **Zei,** zeinoxanthin; **βcr**, β-cryptoxanthin; **cβ,** 13-cis β-carotene; **αc;** α-carotene; **βc**, β-carotene

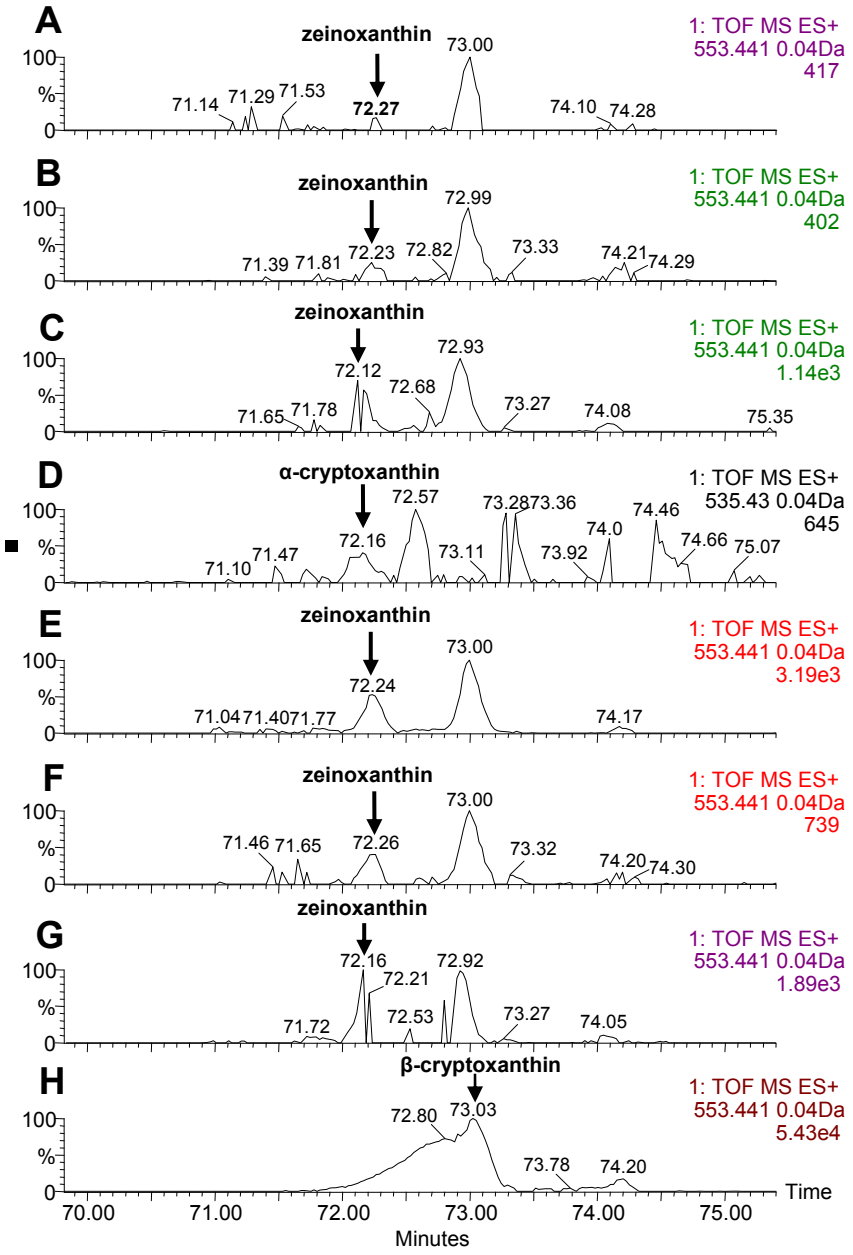


Figure 3-6. **Mass traces corresponding to the major quasimolecular ions (or fragmented ion) for carotenoid extracts from *E. coli* accumulating α- and β-carotene. A-H were separated via HPLC and peaks identified as in (Kim and DellaPenna, 2006). A,** CYP97C2 + HYD4, **B,** CYP97C2 + HYD3, **C,** CYP97C2 + CYP97A4 (peaks labeled **zei**, **FIG. 3**), **D;** CYP97C2 (peak labeled **αcr**, **FIG. 2**) **E,** HYD4; **F**, HYD3; **G,** CYP97A4 (peaks labeled **zei**, **FIG. 2**). **H**; mass trace of β-cryptoxanthin standard. The masses of the major quasimolecular ions (or fragmented ion) for the indicated carotenoids are: zeinoxanthin ($[MH^+]$ = 553.4), β-cryptoxanthin ($[MH^+]$ = 553.4), α-cryptoxanthin ($[MH^+-H_2O]$ = 535.4).

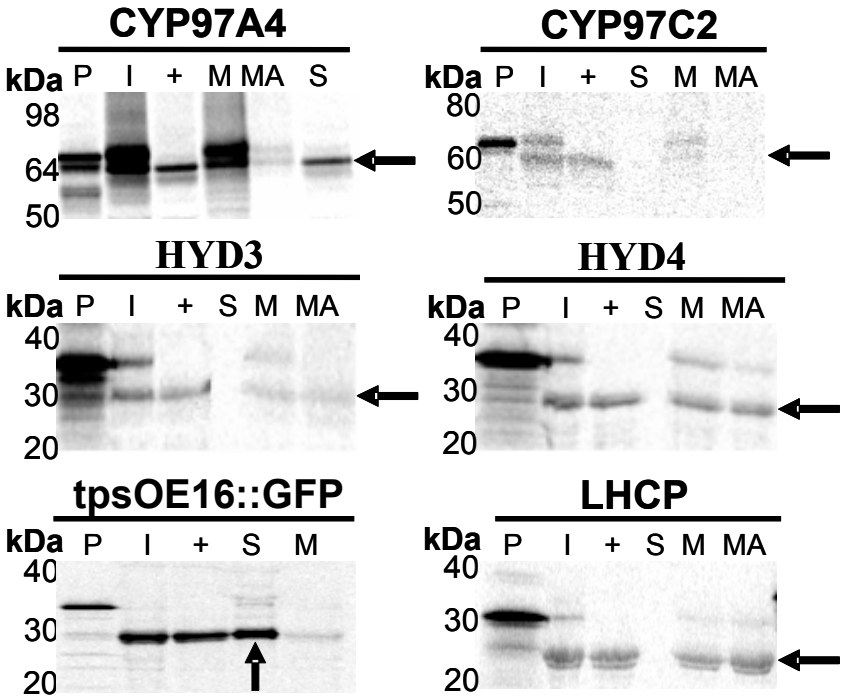


Figure 3-7. Chloroplast import assays of CYP97 and diiron HYD proteins. Isolated pea chloroplasts were used for *in vitro* import of ^{35}S-methionine radio-labeled protein precursors. Chloroplasts harboring imported proteins were then re-isolated and subjected to thermolysin treatment to distinguish between proteins that were peripherally bound to the outer chloroplast envelope, and those that had been imported and thus processed to remove the transit peptide. The mature proteins appear as protease-resistant forms (arrow), confirming import of these proteins into chloroplasts. Chloroplasts containing imported proteins were hyptonically lysed and fractionated into soluble and membrane fractions. The pellet fractions were then treated with an alkaline buffer to remove peripherally-associated membrane proteins. Purity of fractions was determined by import and fractionation analysis of a chloroplast lumen protein, tpsOE16::GFP (Marques et al., 2003) and integral membrane-bound protein, LHCP (Tan et al., 2001). SDS-PAGE analysis indicated that the CYP97A4 and CYP97C2 were synthesized as precursors of about 69 kDa and 62 kDa, and then processed to 64 and 59 kDa respectively. HYD3 and HYD4 were synthesized as precursors of roughly 35 and 34 kDa, and processed to 31 and 27 kDa respectively. P, translation products; I, imported protein; (+), thermolysin treatment; S, soluble proteins; M, membrane proteins; MA, alkaline-treated membrane fraction.

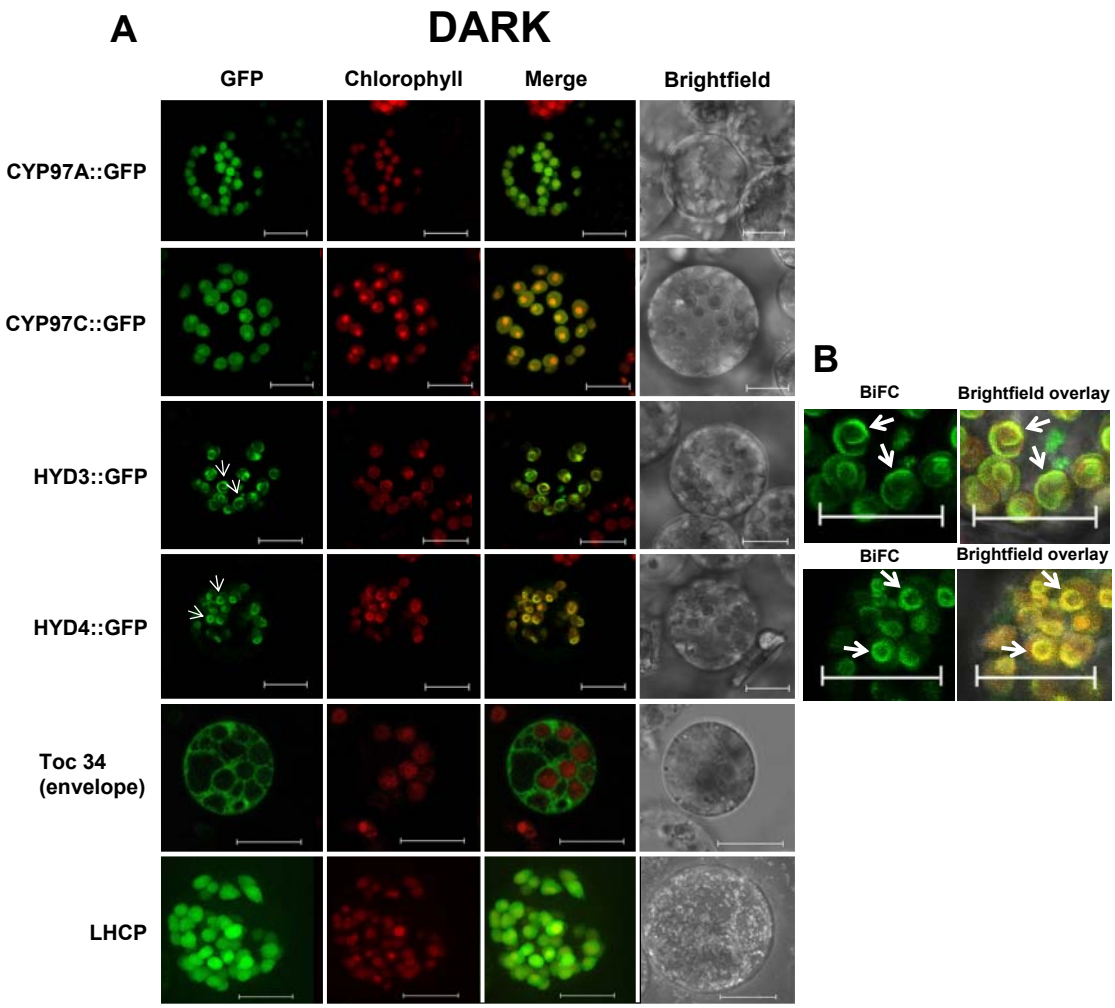


Figure 3-8. A, GFP protein fusion studies involving protoplasts isolated from etiolated (dark-grown) maize leaf tissue. Scale bar = 10µm. Envelope localization indicated by arrows. Controls: Toc34 (34 kDa outer membrane transport complex protein) (Chen and Schnell, 1997), LHCP (integral thylakoid membrane protein) (Tan et al., 2001); *GFP, Chlorophyll,* and *Merge* indicate GFP fluorescence, chlorophyll autofluorescence, and the superposition of both fluorescent signals, respectively. **B,** magnified images to show envelope localization (indicated by arrows) for HYD3::GFP and HYD4::GFP protein fusions.

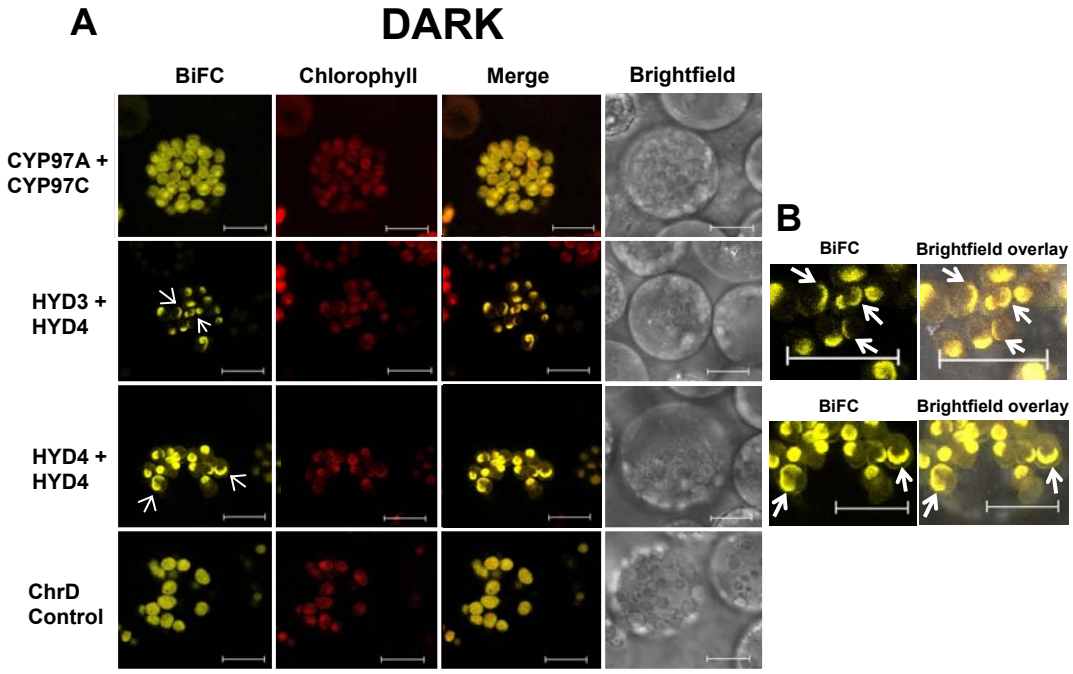


Figure 3-9 . A, BiFC detection of protein-protein interaction in protoplasts isolated from etiolated (dark-grown) maize leaf tissue. Scale bar = 10µm. Envelope localization indicated by arrows. Control: ChrD protein from cucumber is known to form homodimers in plastids (Libal-Weksler et al., 1997); *BiFC, Chlorophyll,* and *Merge* indicate YFP fluorescence, chlorophyll autofluorescence, and the superposition of both fluorescent signals, respectively. **B,** magnified images to show envelope localization (indicated by arrows) for HYD3 + HYD4 and HYD4 + HYD4 interacting proteins.

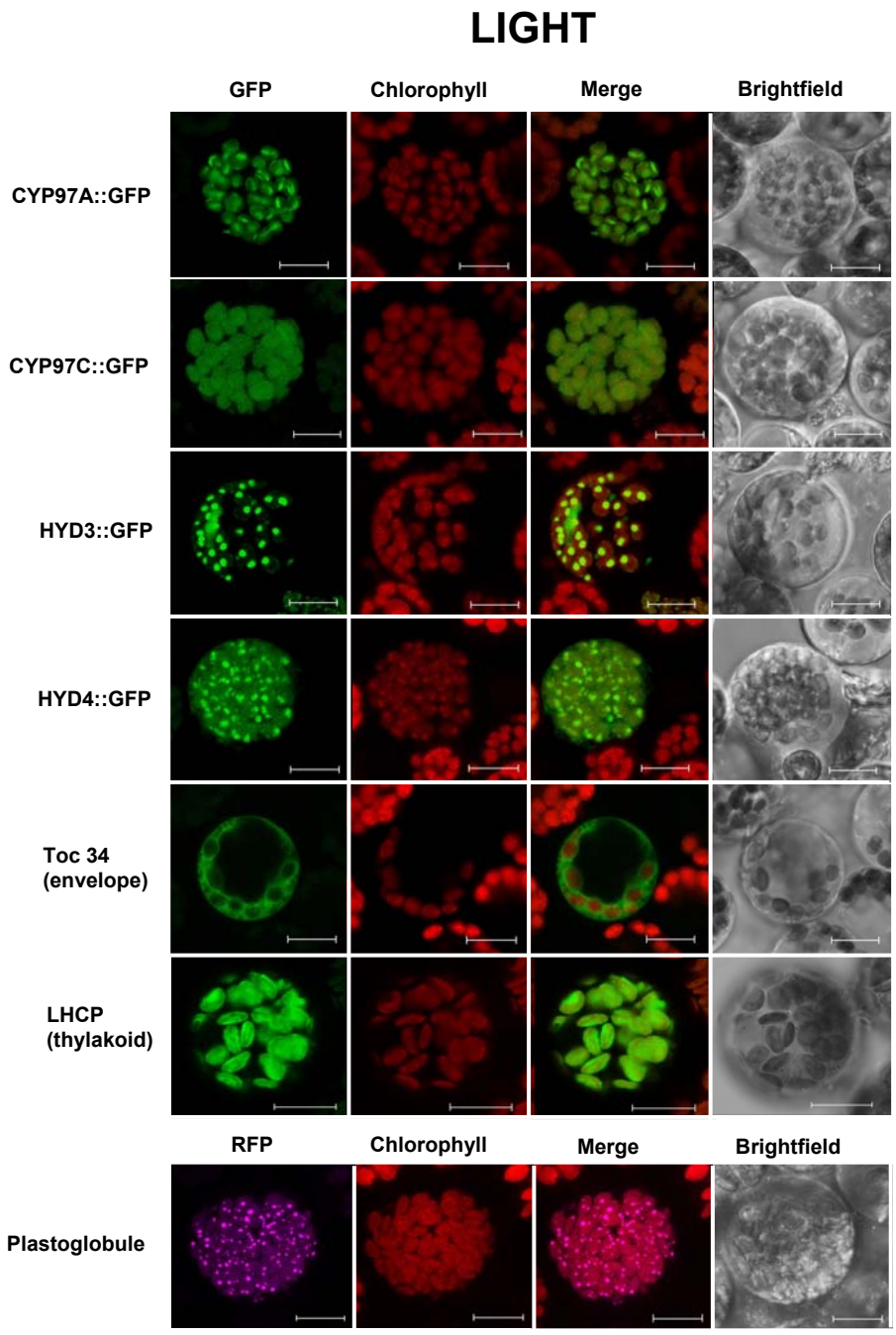


Figure 3-10. A, GFP protein fusion studies involving protoplasts isolated from green (light-grown) maize leaf tissue. Scale bar = 10µm. Controls: Toc34 (34 kDa outer membrane transport complex protein) (Chen and Schnell, 1997), LHCP (integral thylakoid membrane protein) (Tan et al., 2001), Plastoglobule (Plastoglobulin PGL-2 protein); *GFP (or RFP), Chlorophyll,* and *Merge* indicate GFP or RFP fluorescence, chlorophyll autofluorescence, and the superposition of both fluorescent signals, respectively.

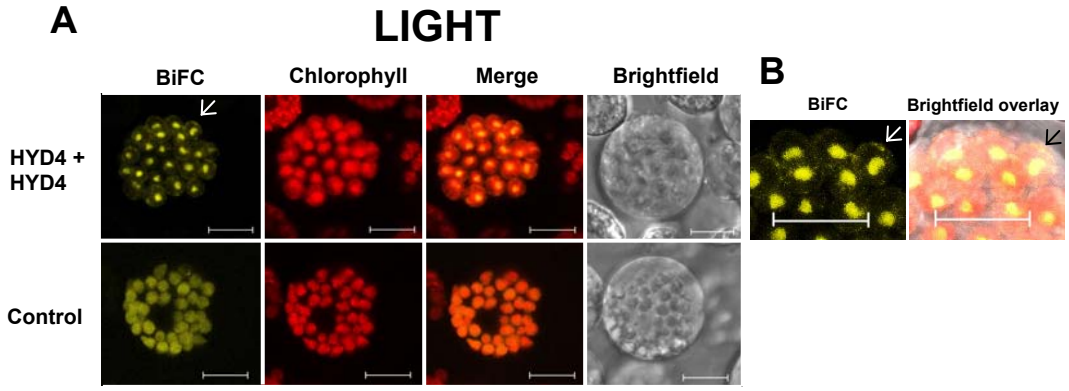


Figure 3-11. A, BiFC detection of protein-protein interaction in protoplasts isolated from green (light-grown) maize leaf tissue. Scale bar = 10µm. Envelope localization indicated by arrow. Control: ChrD protein from cucumber is known to form homodimers in plastids (Libal-Weksler et al., 1997); *BiFC, Chlorophyll,* and *Merge* indicate YFP fluorescence, chlorophyll autofluorescence, and the superposition of both fluorescent signals, respectively. **B,** magnified images to show envelope localization (indicated by arrows) for HYD4 + HYD4 interacting proteins.

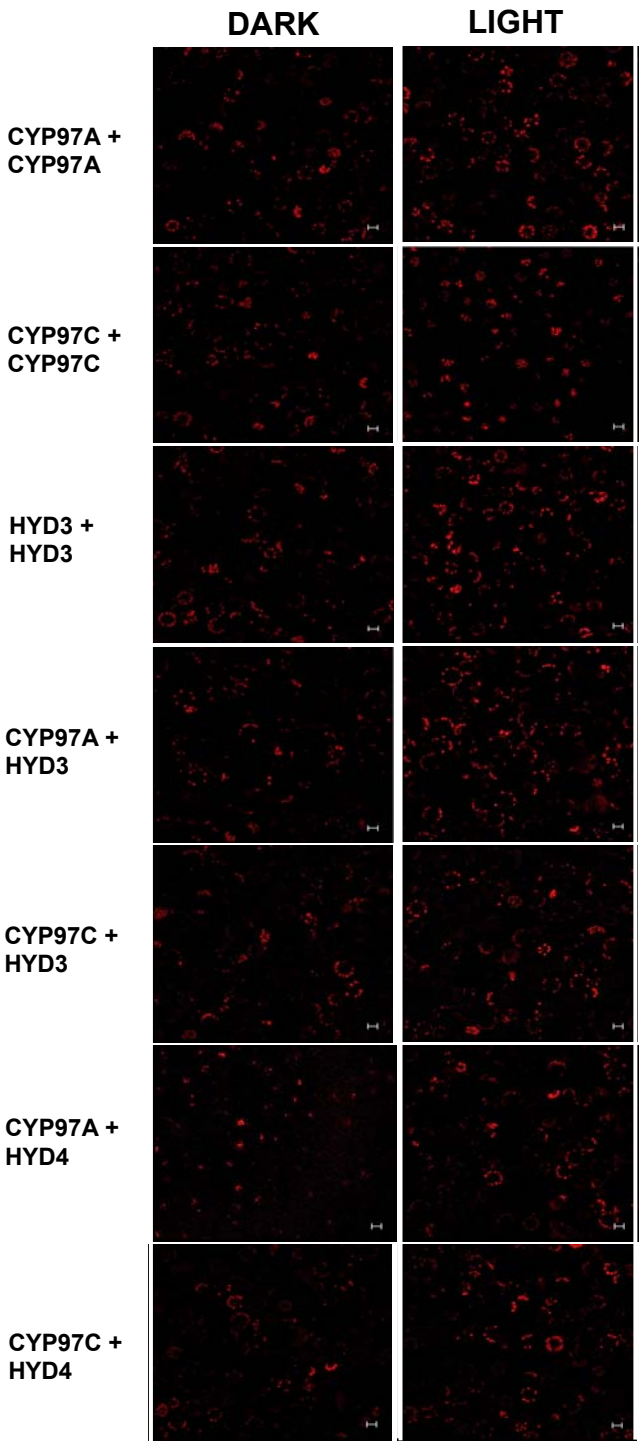


Figure 3-12. BiFC studies demonstrating no protein interaction for the indicated CYP97/HYD combinations tested using protoplasts isolated from both etiolated (dark-grown) and green (light-grown) maize leaf tissue. Scale bar = 10μm. Negative results show only chlorophyll autofluorescence.

Table 3-1. % Major products in β-carotene accumulating *E. coli*

	zeaxanthin	β-cryptoxanthin
CYP97A4	11.08 ± 1.21	26.19 ± 0.53
CYP97C2	N.D	0.78 ± 0.12
HYD3	1.54 ± 0.35	11.30 ± 2.11
HYD4	29.34 ± 3.86	24.14 ± 1.92
Empty vector control	ND	ND

Carotenoids are expressed as a percentage of total carotenoids. Each value is the mean result of 3 replicates ± SD. ND, not detectable.

Table 3-2. % Major products in ε-ε-carotene accumulating *E. coli*

	lactucaxanthin
CYP97A4	6.30 ± 1.19
CYP97C2	4.61 ± 0.29
HYD3	6.26 ± 1.29
HYD4	6.21 ± 0.94
Empty vector control	ND

Carotenoids are expressed as a percentage of total carotenoids. Each value is the mean result of 3 replicates ± SD. ND, not detectable.

Table 3-3. % Major products in α- and β-carotene accumulating *E. coli*

	zeaxanthin	α-cryptoxanthin	zeinoxanthin	β-cryptoxanthin
CYP97A4	3.38 ± 0.27	ND	13.63 ± 2.97	16.76 ± 2.14
CYP97C2	ND	0.71 ± 0.21	ND	1.14 ± 0.30
HYD3	1.07 ± 0.36	ND	2.84 ± 0.92	5.07 ± 1.73
HYD4	30.74 ± 1.85	ND	23.03 ± 2.72	24.03 ± 0.36
Empty vector control	ND	ND	ND	ND

Carotenoids are expressed as a percentage of total carotenoids. Each value is the mean result of 3 replicates ± SD. ND, not detectable.

Table 3-4. % Major products in α- and β-carotene accumulating *E. coli*

	lutein	zeaxanthin	α-cryptoxanthin	zeinoxanthin	β-cryptoxanthin
CYP97A4 + CYP97C2	28.99 ± 2.90	2.98 ± 0.44	ND	7.86 ± 1.28	13.32 ± 1.90
HYD3 + CYP97C2	5.00 ± 0.70	2.09 ± 0.10	ND	7.59 ± 2.42	11.90 ± 0.70
HYD4 + CYP97C2	1.58 ± 0.14	3.16 ± 0.13	ND	3.49 ± 0.47	17.93 ± 1.57
Empty vector control	ND	ND	ND	ND	ND

Carotenoids are expressed as a percentage of total carotenoids. Each value is the mean result of 3 replicates ± SD. ND, not detectable.

Table 3-5. Predicted vs. calculated chloroplast target peptide (CTP) values

	Precursor/mature kDa	Transit kDa	Observed residues	Predicted residues	Predicted kDa	Predicted unprocessed kDa	Predicted mature kDa	Predicted Transit kDa
HYD3	34.9/30.96	3.9	36	69	7.6	34.6	27.5	7.1
HYD4	33.91/27.58	6.3	58	68	7.5	33.5	26.5	7
CYP97C2	62.9/59.14	3.8	34	40	4.3	62.1	57.8	4.3
CYP97A4	68.93/63.6	5.3	48	41	4.0	69.9	65.9	4

Table 3-6. CYP97/HYD combinations tested for protein interaction via BiFC (protoplasts isolated from etiolated tissue)

	CYP97A4	CYP97C2	HYD3	HYD4
CYP97A4	–	+	–	–
CYP97C2	+	–	–	–
HYD3	–	–	–	+
HYD4	–	–	+	+

+ indicates protein interaction between tested combination
– indicates no protein interaction between tested combination

Table 3-7. CYP97/HYD combinations tested for protein interaction via BiFC (protoplasts isolated from light-grown tissue)

	CYP97A4	CYP97C2	HYD3	HYD4
CYP97A4	–	–	–	–
CYP97C2	–	–	–	–
HYD3	–	–	–	–
HYD4	–	–	–	+

+ indicates protein interaction between tested combination
– indicates no protein interaction between tested combination

CHAPTER 4

Summary and future perspectives

4.1 Summary

In these studies, the activities and localization of the CYP97 and HYD carotene hydroxylases were investigated using both bacterial complementation as well as *in vitro*/*in vivo* localization/interaction assays. The well-studied diiron HYD β-carotene hydroxylases had previously been characterized from a wide variety of organisms (Sun et al., 1996; Bouvier et al., 1998; Tian and DellaPenna, 2004). More recently, the CYP97A (β-hydroxylase) and CYP97C (ε-hydroxylase) enzyme activities and substrate specificities had been inferred based on genetic data (Tian et al., 2004; Kim and DellaPenna, 2006) and their respective activities were directly demonstrated in *E. coli*, as described in this work **(Chapter 2)**, (Quinlan et al., 2007). The *E. coli* functional complementation assays demonstrated that the CYP97A4 is primarily a β-ring hydroxlase, with minor activity toward ε-rings, while the CYP97C2 appeared to be exclusively active toward ε-rings **(Chapter 2)**, (Quinlan et al., 2007). These complementation studies were preliminary however, in that they included cells accumulating only either β-carotene or ε- ε-carotene substrates; these assays lacked cells accumulating the putatively preferred CYP97 substrate α-carotene. In addition, it was unknown whether any of the CYP97/HYD enzymes required an interacting partner enzyme for optimal acitivity or whether they localized to the same subcellular/suborganellar locations. The presence of enzyme complexes

as well as the general absence of intermediates in the carotenoid pathway (Maudinas et al., 1977; Camara et al., 1982; Kreuz et al., 1982; Al-Babili et al., 1996; Bonk et al., 1997; Lopez et al., 2008) led to the prediction that the CYP97 and HYD enzymes localize to chloroplast membranes where they function in separate multienzyme complexes to respectively convert α- and β-carotene to lutein and zeaxanthin.

Here, via combined *E. coli* functional complementation and *in vitro/in vivo* import and Bimolecular Fluorescence Complementation (BiFC) studies, the CYP97A4 and CYP97C2 enzymes were shown to function optimally when expressed together and that they interact to form a heterodimer complex in a peripheral association with the chloroplast membrane to efficiently convert their preferred substrate α-carotene to lutein **(Chapter 3).** When expressed as individual enzymes in these complementation systems CYP97A4 and CYP97C2 activities were suboptimal regardless of substrate choice (ie., ε-ε-carotene, α-carotene, or β-carotene substrates). Given their similar activities and specificities it was expected that the CYP97A4 and CYP97C2 enzymes would interact and localize to the same subcellular/suborganellar locations for the efficient channeling of the α-carotene substrate toward production of lutein.

These complementation assays also showed that the diiron HYD3 and HYD4 enzymes are preferentially active toward β-carotene **(Chapter 3),** as had been previously demonstrated via both biochemical and genetic studies (Sun et al., 1996; Tian and DellaPenna, 2004). When expressed as an individual enzyme in α- and β-

carotene *E. coli*, the HYD4 functioned optimally to efficiently convert β-carotene to zeaxanthin. However, co-expression with the CYP97C2 appeared to significantly lower HYD4 activity. HYD3 activity was suboptimal whether this enzyme was expressed alone or in combination with the CYP97C2. *In vitro* import assays showed that the HYD3 and HYD4 enzymes localize to chloroplasts and are integrally-bound to the membrane **(Chapter 3).** In addition, in vivo BiFC studies demonstrated interaction for the HYD4 enzyme which forms a homodimer, while the HYD3 and HYD4 enzymes were shown to form a heterodimer complex **(Chapter 3).** It is reasonable to postulate that the robust enzyme activity observed with the HYD4 toward β-carotene in the complementation assays is due to the ability of this enzyme to form a homodimer complex for the efficient channeling of the β-carotene substrate toward production of zeaxanthin. By contrast, the low enzyme activity observed with the HYD3 may be due to the fact that this enzyme does not form a homodimer complex. Interestingly, the HYD3 interacted with the HYD4 in protoplasts isolated from etiolated maize leaf tissue to form a heterodimer; transcript assays showed that these enzymes are expressed at similar levels in both leaf and root tissue. Together, these data imply that the HYD3 and HYD4 interact in these tissues for the channeling of β-carotene toward generation of zeaxanthin.

In addition, transcript assays indicated that HYD3 is expressed in maize endosperm tissue, and a direct correlation was shown between an increase in HYD3 transcripts and an increase in zeaxanthin accumulation in endosperm. HYD4 is also expressed in maize endosperm, albeit at much lower levels than HYD3 (Vallabhaneni

et al., 2009). Given that the HYD3 was able to form an association with the HYD4 in membranes of protoplasts isolated from etiolated tissue, which in terms of membrane architecture would be similar to amyloplasts in that they both lack thylakoids, as well as the fact that the HYD3 showed elevated transcripts in endosperm, which contain amyloplasts, it is possible that the HYD3 forms a homodimer complex in endosperm amyloplasts for the efficient conversion of β-carotene to zeaxanthin.

Recent proteomics studies (Joyard et al., 2009), as well as the localization studies described in this work (**Chapter 3**) respectively indicated that the CYP97 and HYD enzymes are localized to chloroplast envelope membranes. Interestingly, while protein interaction was demonstrated for the CYP97A4 + CYP97C2, HYD3 + HYD4, and HYD4 + HYD4 combinations in maize protoplasts isolated from etiolated leaf tissue, the only interaction observed in protoplasts isolated from green tissue was for the HYD4 which formed a homodimer. The ability of the HYD4 to form such a complex in protoplasts isolated from both types of tissues may be due to the capacity of this enzyme to associate with different types of plastid membranes since unlike protoplasts isolated from etiolated tissue, green leaf tissue protoplasts would also contain thylakoid membranes. Although the GFP fusion assays described here indicated that the HYD enzymes are localized to the envelope (**Chapter 3**), it is possible that under certain conditions, such as high-light stress, that the HYD4 can also form a homodimer in association with thylakoid membranes for the effective

channeling of β-carotene to zeaxanthin, a xanthophyll known to play a major role in photoprotection of the photosynthetic apparatus in these membranes.

4.2 Future Perspectives

While this research offered some insights into the activities and specificities of the CYP97 and diiron HYD enzymes with respect to their respective roles in the separate metabolons on the α- and β-carotene "arms" of the carotenoid pathway, the roles and topologies of the carotenoid metabolons have yet to be fully elucidated. Predictive metabolic engineering of the carotenoid pathway for enhanced provitamin A content will require a better understanding of the mechanisms controlling the localization of multi-enzyme complexes in plastids with different membrane structures. Given that plastids have different membrane architectures, it is expected that there are both tissue- and plastid-specific differences with respect to targeting of pathway metabolons to plastid membranes. Metabolons targeted to chloroplasts would have a choice in forming membrane associations with either envelope or thylakoid membrane structures; this would not be the case with amyloplasts, which only contain envelope membranes. The carotenoid pathway is localized on both envelope and thylakoid membranes, (Vishnevetsky et al., 1999; Joyard et al., 2009) which suggests that the pathway may have membrane-specific functions. These phenomena are not well understood and are further complicated by the fact that many carotenoid pathway enzymes are coded for by single-copy genes that may potentially be targeted for multiple suborganellar locations. If the single-copy encoded CYP97A

and CYP97C enzymes only localize to chloroplast envelope membranes, as indicated by the protemics studies of Joyard et al., (1999) to convert α-carotene to lutein, the question arises as to what mechanisms are involved in the trafficking of enzymes and/or metabolites such as lutein to the thylakoid membrane since this xanthophyll is essential for LHC assembly in these membranes (Pogson et al., 1998; Lokstein et al., 2002; Dall'Osto et al., 2006; Dall'Osto et al., 2007). Similarly, if the HYD enzymes localize to chloroplast envelope membranes, as suggested by the localization studies described here **(Chapter 3),** yet the hydroxylated product zeaxathin is required for photoprotection of the photosynthetic apparatus in thylakoid membranes (Goss et al., 2008; Arnoux et al., 2009; Jahns and Holzwarth, 2012), the same question with respect to transporting of either the HYD enzymes or the essential metabolite zeaxanthin to thylakoid membranes applies. Further investigations into these trafficking mechanisms will be required for a more complete understanding of carotenoid metabolon biogenesis/maintenance in various plastid membrane structures.

Finally, it should be noted that while provitamin A carotenoids, such as β-carotene, play a vital role in maintaining human health, other non-provitamin A carotenoids have been shown to have important health benefits. The non-provitamin A xanthophyll lutein, for example, has been shown to aid in the prevention of retinal damage due to light exposure (von Lintig, 2010). While the maize *Hydroxylase3* locus has been identified as an important provitamin A biofortification target in cereal endosperm (Vallabhaneni et al., 2009), CYP97A and CYP97C activities have been shown to be of central importance for the efficient conversion of α-carotene to the

nutritionally-valuable non-provitamin A carotenoid lutein in *E.coli* **(Chapter 3)**. The successful expression of these enzymes in bacteria is potentially promising for the economically feasible production of lutein on an industrial scale.

Appendix 1: A first-authored publication related to this dissertation[1]

ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect
Archives of Biochemistry and Biophysics 458 (2007) 146–157

ABB
www.elsevier.com/locate/yabbi

Escherichia coli as a platform for functional expression of plant P450 carotene hydroxylases

Rena F. Quinlan [a,b], Tahhan T. Jaradat [a], Eleanore T. Wurtzel [a,b,*]

[a] *Department of Biological Sciences, Lehman College, The City University of New York, 250 Bedford Park Boulevard West, Bronx, NY 10468, USA*
[b] *The Graduate School and University Center-CUNY, 365 Fifth Ave., New York, NY 10016-4309, USA*

Received 16 August 2006, and in revised form 2 November 2006
Available online 3 December 2006

Abstract

Carotenoids and their derivatives are essential for growth, development, and signaling in plants and have an added benefit as nutraceuticals in food crops. Despite the importance of the biosynthetic pathway, there remain open questions regarding some of the later enzymes in the pathway. The CYP97 family of P450 enzymes was predicted to function in carotene ring hydroxylation, to convert provitamin A carotenes to non-provitamin A xanthophylls. However, substrate specificity was difficult to investigate directly in plants, which mask enzyme activities by a complex and dynamic metabolic network. To characterize the enzymes more directly, we amplified cDNAs from a model crop, *Oryza sativa*, and used functional complementation in *Escherichia coli* to test activity and specificity of members of Clans A and C. This heterologous system will be valuable for further study of enzyme interactions and substrate utilization needed to understand better the role of CYP97 hydroxylases in plant carotenoid biosynthesis.

Keywords: Carotene; P450; CYP97; Hydroxylases; Vitamin A; *Oryza sativa*; Metabolic engineering; Xanthophylls; Plants; *Escherichia coli* functional complementation

Carotenoids are an abundant class of isoprenoid compounds found in all plants where they are synthesized in plastids by nuclear-encoded enzymes. Their numerous roles in plant growth and development include functions as accessory pigments in photosynthesis and as photoprotectors [1–3]. Carotenoids such as β-carotene and zeaxanthin, among others, serve as precursors to cleavage products including the hormone abscisic acid (ABA) [4] and to other apocarotenoids, some of which are gaining attention for their roles in internal and external signaling [5–11]. The importance of certain carotenoids in human health has led to efforts to breed or metabolically engineer carotenoid content and composition [12–15].

Plants are an important dietary source of provitamin A carotenoids; predictable efforts to use breeding of vitamin-rich crops to address human vitamin A deficiency will require elucidation of those mechanisms controlling conversion of provitamin A carotenes to non-provitamin A xanthophylls. For example, in maize endosperm, there is wide variation in carotenoid content and composition [16,17] making this tissue a target for improvement of carotenoid content and provitamin A levels [1,18,19]. Like maize, wheat is another of many crops in the Poaceae that exhibit diversity in yellow seed color and are attractive breeding targets for manipulating endosperm carotenoids [20,21]. Endosperm carotenoids are synthesized and accumulate on amyloplast envelope membranes whereas in chloroplasts, carotenoids are found on envelope and thylakoid membranes. Targeting of carotenoid enzymes to specific plastid membranes and metabolon biogenesis/maintenance are poorly understood phenomena, especially given that many carotenoid enzymes are encoded by single copy genes and are destined for multiple suborganellar locations. Moreover, there is limited understanding of

* Corresponding author. Fax: +1 718 960 8236.
E-mail address: wurtzel@lehman.cuny.edu (E.T. Wurtzel).

0003-9861/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2006.11.019

[1] This appendix is a publication of Quinlan et al., (2007) Arch Biochem Biophys. 458: 146-157, (www.sciencedirect.com) with permission from the publisher (Elsevier)

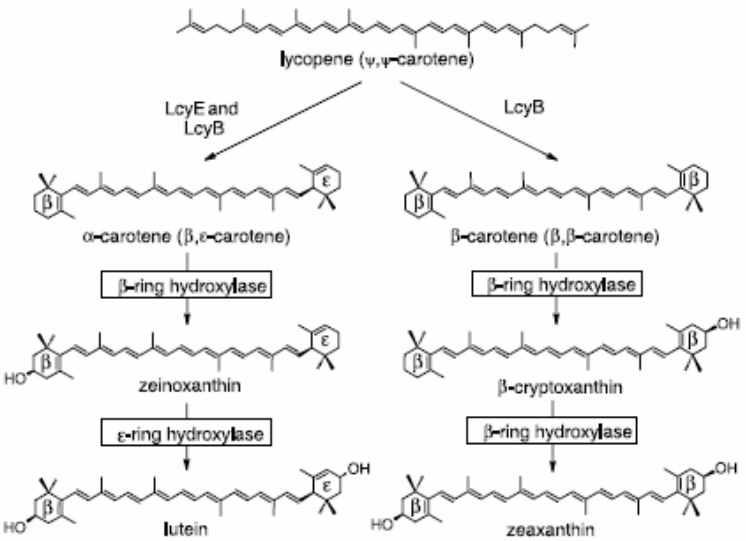


Fig. 1. Schematic illustration of the cyclohexene ring hydroxylation in β- and α-carotene. Conversion of lycopene to β-carotene requires lycopene β-cyclase (LcyB) alone, while conversion of lycopene to α-carotene requires lycopene β- and ε-cyclases (LcyE). Formation of the hydroxylated xanthophylls zeaxanthin and lutein is mediated by two separate stereospecific β- and ε-ring hydroxylases. In hydroxylation of α-carotene, the two hydroxylases may act in the reverse order shown, to yield the mono-hydroxy intermediate α-cryptoxanthin, in which the ε-ring is hydroxylated.

how pathway intermediates are channeled to downstream products including photosynthetic pigments or signaling molecule apocarotenoids.

In all plants, the biosynthesis of carotenoids begins with the formation of the 40-carbon phytoene, followed by desaturation steps leading to synthesis of lycopene, after which point the pathway diverges to form either β-carotene (having two β rings) or α-carotene (having one β-ring and one ε-ring) (Fig. 1). Further hydroxylation of the carotenes leads to biosynthesis of the xanthophylls. For example, hydroxylation of the C3 and/or C3′ on one or both β-rings of β-carotene leads to β-cryptoxanthin (β,β-carotene-3-ol) and zeaxanthin (β,β-carotene-3,3′-diol). In contrast, hydroxylation of the ε-ring, β-ring, or both rings of α-carotene generates α-cryptoxanthin (β,ε-carotene-3′-ol), zeinoxanthin (β,ε-carotene-3-ol), and lutein (β,ε-carotene-3,3′-diol), respectively [22]. The carotenes with β-rings have provitamin A activity, whereas their hydroxylated β-ring products do not [23].

In plants, the hydroxylation of provitamin A carotenes to form non-provitamin A xanthophylls is thought to be mediated by two structurally distinct classes of mixed function oxygenases. The first class of hydroxylases is comprised of the β-ring non-heme diiron monooxygenases. Members of this class were identified from a wide range of bacterial, algal, and plant species [24–26]. Plant diiron β-ring hydroxylases contain four transmembrane helices and four conserved iron-binding histidine clusters. Histidine residues in these clusters coordinate the orientation of two iron atoms to form oxo-bridged diferric centers [27], which are involved in the formation of a substrate based alkyl-radical and the subsequent hydroxylation of the substrate via an oxygen-rebound mechanism [28]. Enzyme activity of the diiron β-ring hydroxylases has been demonstrated in an *Escherichia coli*[1] functional complementation system [24]. The second class consists of P450 enzymes from the CYP97 clan, which on the basis of genetic evidence, are hypothesized to hydroxylate carotene ε- and β-rings, respectively [29,30]. These predicted P450 heme-thiolate hydroxylases contain a single transmembrane anchoring sequence, typical of other eukaryotic P450 enzymes [31,32]. In addition, each enzyme has a P450 domain, a conserved oxygen binding signature, and a conserved heme-thiolate-binding signature that binds a single heme group with one iron atom in the center [25,31,32]. Similar to diiron hydroxylation reactions, P450 hydroxylations are also redox-sensitive and involve the formation of a substrate based alkyl-radical that is immediately trapped by the HO˙ from the iron atom [33]. A major limitation for the genetic evidence used to support substrate specificity of the CYP97 carotene hydroxylases is that biochemical profiling of mutants is complicated by poorly understood compensatory mechanisms, including changes in transcriptional activity of genes encoding other carotenoid biosynthetic enzymes such as the non-heme

[1] *Abbreviations used*: *E. coli*, *Escherichia coli*; CD, conserved domain; ORF, open reading frame; NUE, near-upstream element; FUE, far upstream element; CDD, conserved domain database; UTR, untranslated region.

diiron carotene hydroxylases, which may ultimately impact pathway flux [29,30].

Progress in characterizing the full complement of enzymes and identifying the corresponding genes has been significantly advanced with use of *E. coli* to functionally test activity of the membrane bound carotenoid pathway enzymes, an approach adopted by the carotenoid community [e.g. [24,34–38]]. The bacterial system could potentially provide a simple platform to manipulate and study the CYP97 carotene hydroxylases. Therefore, we isolated cDNAs from *Oryza sativa*, a representative member of the Poaceae, and used *E. coli* complementation to demonstrate activity and investigate carotene ring specificity of the putative P450 β- and ε- ring hydroxylases. Since only a subset of P450 enzymes function in *E. coli*, the widely used heterologous bacterial system was not necessarily a feasible approach. [39–41]. Our demonstration that these members of the CYP97 clan function in *E. coli* indicates that this system will be valuable in further dissecting the structural basis for ring specificity and other molecular applications. Given that β-ring hydroxylation is alternatively possible with either CYP97 or diiron enzymes raises the question of which are the appropriate genes/enzymes to choose as breeding targets or in metabolic engineering of this pathway; whether it is to achieve enhanced levels of endosperm provitamin A carotenoids in food crops, or increased abiotic and/or biotic resistance in plants in general.

Materials and methods

Phylogenetic and sequence analyses

Nucleotide and corresponding protein sequences, highly similar to the putative CYP97A4 coding mRNA from *O. sativa* L. (AK068163) [42], were obtained using BLAST analyses from all available public databases in NCBI GenBank and Institute of Genomic Research (TIGR) gene indices: CYP97A3 (#AAL25587), CYP97B3 (#AAL32753), CYP97C1 (#AAR83120) [29], and CYP86A1 (#NM_116260) from *Arabidopsis thaliana* (L.) Heynh.; CYP97B4 (#XP_464306) and CYP97C2 (#AK065689) from *Oryza sativa* L.; #BJ234910 and #CA501638 from *Triticum aestivum*; # TC69886, #TC76166 and #BM816653 from *Hordeum vulgare* L.; CYP97B1; synonym CYP97A2 (#Z49263) from *Pisum sativum* L. [43]; CYP97B2 (#AAB94586) and #TC228439 from *Glycine max* (L.) Merr. [44]; #BQ971938 from *Helianthus annuus* L.; #BE552887 from *Zea mays* L.; #TC101515 and #TC109838 from *Medicago truncatula* Gaertn; and #BT012891 from *Lycopersicon esculentum* Mill.

Protein sequences were screened for chloroplast targeting signal peptides using two prediction algorithms: GENOPLANTE™ PREDOTAR (Predotar for Prediction of organelle targeting sequences) at http://genoplante-info.infobiogen.fr/predotar/ and ChloroP 1.1 Server at http://www.cbs.dtu.dk/services/ChloroP/ [45]. Rice protein sequences were compared against the P450 database at http://132.192.64.52/blast/P450.html using the P450 blast server. Prior to alignment, protein sequences were truncated to include the P450 conserved domain (CD) and exclude the chloroplast targeting sequence. Boundaries of conserved cytochrome P450 domains were identified using NCBI Conserved Domain Search tool (RPS-BLAST). Amino acid sequences were aligned by ClustalW, and a neighbor-joining tree was constructed with a 500 bootstrap replication support using MEGA3 software [46].

Total RNA extraction and isolation of cDNAs

Total RNA was extracted from *Oryza sativa* L. cv Nipponbare (*japonica*) leaves (approx. 6 weeks old) using the RNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA). Approximately 2 μg of total RNA, in a 20 μl reaction, was used for cDNA synthesis using the SuperScript™ First-Strand Synthesis System for RT-PCR kit (Invitrogen Corporation, Carlsbad, CA). GenBank #AK068163 has a 1929 bp ORF between nucleotide number 54 and 1982. GenBank #AK065689 has a 1683 bp ORF between nucleotide number 105 and 1787. Gene specific primers, with terminal *Eco*RI and *Xho*I restriction sites, (forward, 5′-CCG GAA TTC A_{54}TG AGC TCA GCG ACG TCA GTG AGT_{77}-3′ and reverse, 5′-ACC GCT CGA GT_{1985}C AGA TTC GAG TTG CTG AGA CTT G_{1962}-3′, MWG-Biotech Inc., Oaks Parkway, NC) were used to amplify the full-length coding sequence of CYP97A4 (#AK068163) [42]. The coding sequence (excluding the first 21 bp) of CYP97C2 (#AK065689) was amplified using gene specific primers with terminal *Eco*RI and *Xho*I restriction sites, (forward, 5′-CCG GAA TTC C_{105}CG TCC CGT GCG TAC CAT TC_{124}-3′ and reverse, 5′-ACC GCT CGA G_{1793}TC ATC TGG ACC CAC TGA GTG CA_{1734}-3′, MWG-Biotech Inc., Oaks Parkway, NC) (50). Fifty microlitre reactions each contained 4 μl cDNA, 1 μl of each primer (20 μM), 19 μl H_2O, and 25 μl of a master mix (50 mM $MgCl_2$, 10 mM dNTP mix, 50 U/ml *Taq* polymerase, and 2× PCR buffer) (Promega Corp., WI). Amplification conditions were: 1 cycle, 95 °C, 3 min; 40 cycles, 95 °C, 30 s, 58.1 °C, 40 s, 72 °C 2 min; 1 cycle, 72 °C, 8 min.

Construction of expression vectors and functional analysis

The CYP97A4 (#AK068163) and CYP97C2 (#AK065689) sequences amplified as above were subcloned in-frame into *Eco*RI and *Xho*I sites of the pCOLADuet™-1 vector (Novagen, WI), renamed pRT-A4 and pTR-C2, respectively. For testing of β-ring substrate specificity, constructs were transformed into *E. coli* BL21 (DE3) cells (Novagen, WI) harboring pAC BETA-O4 which confers β-carotene accumulation [24]. pRT-A4 and pTR-C2 ORFs were also introduced into *E. coli* BL21 (DE3) cells carrying the pACCRT-EIB [47] and plasmid y2 [48] which carries the *Arabidopsis* lycopene ε cyclase, which together confer accumulation of lycopene and δ-carotene (ε,ψ-carotene), along with some ε-carotene (ε,ε-carotene) (minor product) [48]. For carotenoid analyses, saturated cultures in LB medium were diluted 100-fold into 25 ml fresh medium in 250 ml flasks, then grown in the dark at 300 rpm at 21 °C, 28 °C, or 37 °C (cultures containing pTR-C2 were grown at 37 °C) until OD 0.6 at which point they were induced with 10 mM IPTG (unless otherwise noted) and further cultured for a total of three days.

Extraction of carotenoids from E. coli cells, HPLC separation and identification

Carotenoids were extracted in ethanol:ether (1:2 v/v), essentially, as described [49]. Bacterial cell pellets were extracted in 10 volumes of the ethanol:ether mix using a Dounce homogenizer, incubated at room temperature for 5–10 min, and centrifuged at 10,000*g*. Supernatants were frozen at −80 °C for 30 min, centrifuged at 10,000*g* at 4 °C, supernatants dried under nitrogen gas, dissolved in HPLC mobile phase, and stored at −20 °C.

Separation was carried out using a Waters HPLC system equipped with a 2695 Alliance separation module, a 996 photodiode array detector, a column heater, a fraction collector II, Empower software (Millipore, Franklin, MA), and a Nucleosil 5 C_{18} (5 μm, 250 × 4.6 mm) column with a Nucleosil C_{18} (5 μm, 4 × 3.0 mm) guard column (Phenomenex, Torrance, CA), mobile phase of 100 parts acetone: 4 parts water, column flow rate of 0.6 ml/min, and sample and column temperature of 30 ± 5 °C. Peaks were identified on the basis of retention times/spectra matching those of authentic standards (INDOFINE Chemical Company Inc., Sommerville, NJ), and standards purified from bacteria expressing genes encoding carotenoid biosynthetic enzymes [48]. Authentic lactucaxanthin, purified

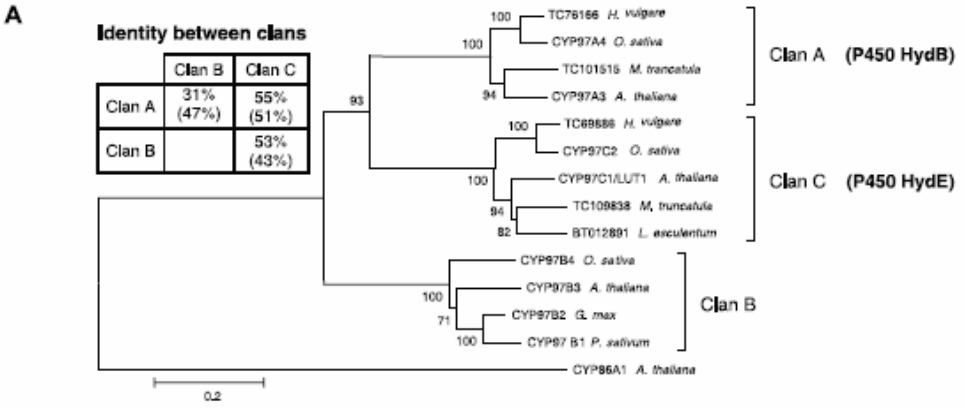
from *Lactuca sativa* leaves [50] and distinguished as a novel HPLC peak in comparison to a maize leaf profile, exhibited a spectrum matching reported values [51]. Identification of ε,ε-carotene and lactucaxanthin were confirmed by comparison with reported λ_{max} (wavelength of maximum absorption) spectral values [51] and comparison to standards chromatographed in the identical system. For quantification of extracted metabolites, data were collected at λ_{max} for individual metabolites and integrated peak areas calculated.

Results

Phylogenetic analysis

Among all known CYP97 P450 enzyme family members, three distinctive, yet closely related, clans (i.e., A, B, and C) can be discerned [52]. The amino acid alignment and phylogenetic tree construction for a selection of thirteen protein sequences from seven different plant species (monocots and dicots) is in agreement with the three clans designation (Fig. 2A). Clan A and C enzyme sequences are more closely related with three blocks of amino acid sequence insertions that differentiate members of Clan B from those in Clans A and C (Fig. 2B). Block 1 is about 100 residues downstream of the P450 domain; blocks 2 and 3 are 37 aa apart and located between the oxygen- and heme thiolate-binding signatures. Members of the CYP97 clan, Clan B, do not have any functionally demonstrated representatives and as shown in the sequence comparisons, are more distant from Clans A and C. Clan B sequences do possess a potential chloroplast targeting sequence [45] and transcripts are present in shoots as reflected by ESTs, suggesting localization in chloroplasts as predicted for Clan A and C enzymes. Clan B homologs are found in angiosperms (Fig. 2), in *Ginkgo biloba* (AAT28222) and the diatom, *Skeletonema costatum* (AAL73435) [53]. This taxonomically broad distribution suggests that Clan B enzymes may function in all plants rather than catalyze synthesis of taxon-specific metabolites.

To test carotene hydroxylase activity of Clan A and C enzymes, we identified sequences that were most closely related to *Arabidopsis* LUT1 (CYP97C1) using the genome of *Oryza sativa*, a representative of the Poaceae, and amplified cDNAs from *O. sativa* L. cv Nipponbare leaves. The



B

	Block 1	Block 2	Block 3	
CYP97B1 *P. sativum*	LLEG-EGRDGQKSVE	DVLPGGHKGDKDGYTI	LVQNNNEEVEGWAGFDPSRSPGALY	Clan B
CYP97B2 *G. max*	LLEG-EGYDGPDSIE	DVLPGGHKGEKDGYTI	LVQNKNEEIEGWAGLDPSRSPGALY	
CYP97B3 *A. thaliana*	LIREKETSSGEDTIE	ETLPGGHKGEKEGHKV	LRTKESNGIEGWAGFDPSRSPGALY	
CYP97B4 *O. sativa*	LIER—GEHGEKYTI	DKLPGGHNGAKEGYEI	SVPKKDESIEGWAGFDPDRSPGAMY	
BT012891 *L. esculentum*	PDAI------SGSA	DVLPG------NYKV	DLE----------GPV	Clan C
TC109838 *M. truncatula*	ADAV------NGTA	DELPG------DYKI	DLD----------GPV	
CYP97C1 *A. thaliana*	PYAE------DGSA	DILPG------NYKV	DID----------GAI	
CYP97C2 *O. sativa*	TSAL------SGKP	DVLPG------NYKI	DLE----------GPV	
TC69886 *H. vulgare*	TYAL------SGEP	DVLPG------NYKV	DLE----------GPI	
CYP97A3 *A. thaliana*	AAAL------KGEE	DILG-------EYPI	PLD----------GPN	Clan A
TC101515 *M. trancatula*	TAAS------DGED	DVLG-------EYPI	PLD----------GPN	
CYP97A4 *O. sativa*	KAAT------DGED	DMLG-------GYPI	PLD----------GPN	
TC76166 *H. vulgare*	KAAS------DGED	DMLG-------EYPI	PLD----------GPN	

Fig. 2. Comparison between the conserved P450 domain of the three clans in the CYP97 family. (A) A rooted neighbor-joining tree was constructed using the fatty acid ω-hydroxylase (CYP86A1) from *A. thaliana* as an outgroup. Numbers adjacent to branches are bootstrap values supporting the presented final tree. Average identities among clans are shown in the table in comparison to pair wise comparisons between *A. thaliana* and rice sequences, which are enclosed in parentheses. (B) Three sequence blocks distinguish members of Clan B from those in Clans A and C and are shown in excerpts from the alignment of the conserved P450 domains.

O. sativa AK065689 and AK068163 cDNAs, correspond to CYP97C2 (P450 HydE, putative ε-ring carotene hydroxylase) and CYP97A4 (P450 HydB, putative β-ring carotene hydroxylase). They encode conceptually translated proteins sharing 84% and 76% similarity (excluding the non-conserved transit peptide sequences) with the *A. thaliana* CYP97C1 (LUT1) and CYP97A3 (LUT5) [30] candidates for ε- and β-ring hydroxylases. Rice CYP97A4 and CYP97C2 map to chromosomes 2 and 10, respectively. CYP assignment of rice clones is based on BLAST searches using the Rice P450 BLAST Server, which implements the standardized system of cytochrome P450 nomenclature [52].

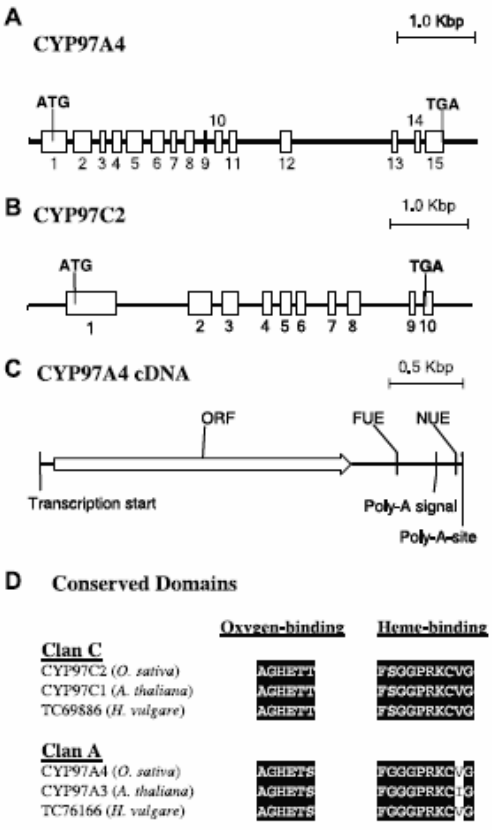


Fig. 3. Gene structure and P450 domains of CYP97 Clan A and C enzymes. A and B, gene structures for rice CYP97A4 and CYP97C2, respectively. Open boxes indicate exons; C, revised cDNA structure of CYP97A4 cDNA. The transcribed cDNA is a predicted 2414 bp fragment with an 88 bp 5′-UTR, a 1929 bp ORF, and a 397 bp 3′-UTR with conserved NUE, FUE, and poly-A site; D, comparison of conserved oxygen and heme-binding domains for CYP97 Clan A and C enzymes of rice compared with those in other plant species.

P450 DNA sequence analysis

The *O. sativa* CYP97A4 (P450 *HydB*) is comprised of 15 exons and 14 introns, while the CYP97C2 (P450 *HydE*) gene consists of 10 exons and 9 introns (Figs. 3A and B). All intron boundaries are consistent with the conserved 5′-AG and 3′-GT flanking element rule [54]. The CYP97C2 cDNA is described in the NCBI GenBank # AK065689 as a 1876 bp long cDNA with a 1683 bp open reading frame (ORF), a 104 bp 5′-untranslated region (5′-UTR), and an 89 bp 3′-UTR [42]. The CYP97A4 cDNA (GenBank #AK068163) is indicated as a 4217 bp long cDNA with a 1929 bp ORF, a 53 bp 5′-UTR, and a 2235 bp 3′-UTR [42]. However, analysis of AK068163 cDNA and corresponding genomic sequence (GenBank # NT_107182) support a 2414 bp full-length cDNA with a 1929 bp ORF, an 88 bp 5′-UTR, a 397 bp 3′-UTR with conserved near-upstream element (NUE; poly-A addition signal), far upstream element (FUE), and terminal "Pyr(T)A" poly-A site (Fig. 3C) [55–57]. The modified full-length cDNA structure (Fig. 3C) is based on the finding that several rice 3′-end ESTs map around the proposed poly-A addition site, whereas there are no ESTs that map to the 3′-end of the model presented by Kikuchi and coworkers [42].

Conserved domain characteristics of the CYP97C2 and CYP97A4 amino acid sequences

Significant *E* values from PSI- and PHI-BLAST searches against the conserved domain database (CDD) at NCBI confirm that CYP97C2 and CYP97A4 are P450 proteins [58]. The conserved heme-thiolate-binding signature (FXXGXXXCXG) in plants corresponds to Phe_{488}-Ser-Gly-Gly-Pro-Arg-Lys-Cys-Val-Gly_{497} in CYP97C2 and to Phe_{534}-Gly-Gly-Gly-Pro-Arg-Lys-Cys-Val-Gly_{543} in CYP97A4 and as compared with those of other related plant enzymes (Fig. 3D). The underlined conserved cysteine in each of these motifs is an essential residue since it contributes the thiol group, which is the fifth ligand that binds the iron atom in these hemoproteins. Similarly, near the middle of the P450 conserved domain, a conserved oxygen binding signature (A/G)GX(D/E)T(T/S) is detected with the corresponding amino acid sequence of Ala_{361}-Gly-His-Glu-Thr-Thr_{366} for the CYP97C2 and Ala_{402}-Gly-His-Glu-Thr-Ser_{407} for the CYP97A4 (Fig. 3D) and as compared with those of other related plant enzymes. The underlined threonine in these respective sequences is involved in the binding of an oxygen molecule, which is essential for catalysis.

The overwhelming majority of eukaryotic P450 heme-thiolate enzymes are membrane bound through an anchoring transmembrane helix. Conversely, all P450s in prokaryotes are cytosolic. CYP97C2 is predicted to be a 62.1 kDa (561 residue) preprotein with a 40 residue N-terminal transit peptide that is cleaved upon import to yield a chloroplast-localized 57.8 kDa, 521 residue, mature protein [45]. CYP97A4 is predicted to be a

69.9 kDa (643 residue) preprotein containing a 41 residue N-terminal transit sequence that is removed upon chloroplast import to yield a chloroplast-localized 65.9 kDa, 602 residue, mature protein. In addition, both enzymes harbor a transmembrane helix sequence at the N-terminus of the P450 domain, as predicted by the HMMTOP server [59] to membrane anchor the mature P450 protein.

Heterologous expression of CYP97 genes in E. coli and substrate specificity

CYP97A4. For functional testing of β-ring carotene hydroxylase activity, we engineered *E. coli* cells as described in the methods to produce the β-carotene (β,β-carotene) substrate, which contains two β-rings for potential hydroxylation. For testing of a Clan A enzyme (a putative β-ring carotene hydroxylase), we amplified and cloned the full-length rice CYP97A4 cDNA (including its putative transit sequence) into pCOLADuet™-1, and introduced the resulting construct into the β-carotene accumulating cells. Accumulated pigments from transformed cells with or without the cDNA (Figs. 4A(e) and (d)) along with standards for β-carotene (Fig. 4A(c), peak 3) and hydroxylated products, β-cryptoxanthin (one hydroxylated ring, Fig. 4A(b), peak 2) and zeaxanthin (two hydroxylated rings, Fig. 4A(a), peak 1) were analyzed by HPLC. Extracted pigments were identified based on HPLC retention times (Fig. 4A) and spectra (Fig. 4B) matching those of authentic standards. As seen in Fig. 4A(e), only the doubly transformed cells accumulated β-cryptoxanthin (peak 2) and zeaxanthin (peak 1), in addition to the substrate β-carotene (peak 3). Cells that lacked the CYP97A4 cDNA did not have any hydroxylated β-carotene derivatives (Fig. 4A(d)).

We next tested the effect of culturing temperature and IPTG concentration on CYP97A4 hydroxylation activity in *E. coli*. At 21 °C and 10 mM IPTG, ~62% hydroxylated product accumulated, of which 52% was β-cryptoxanthin (mono-hydroxy intermediate) and 10% was zeaxanthin (di-hydroxy product) (Fig. 5A), a combined equivalent of 36% total hydroxylated β-rings (Fig. 5B). Increasing the culturing temperature from 21 °C to 37 °C increased production of β-carotene substrate but reduced the amount of hydroxylated products by 20% at 5 mM IPTG (data not shown) and by 40% at 10 mM IPTG (Fig. 5). Our initial attempts using overnight cultures grown at 37 °C and lower concentration of IPTG (i.e., 0.4 and 1 mM) did not result in accumulation of detectable xanthophylls. Interestingly, the ratio of β-cryptoxanthin (one hydroxylated ring) to zeaxanthin (two hydroxylated rings) decreased from 5:1 for growth at 21 °C to almost equimolar amounts for growth at 37 °C. In summary, CYP97A4 is more effective at lower temperature (21 °C) when expressed in *E. coli*, an organism having optimal growth at 37 °C.

CYP97C2. We next tested whether a Clan C enzyme, a putative ε-ring carotene hydroxylase, could utilize the β-ring substrate. We similarly amplified and cloned the rice

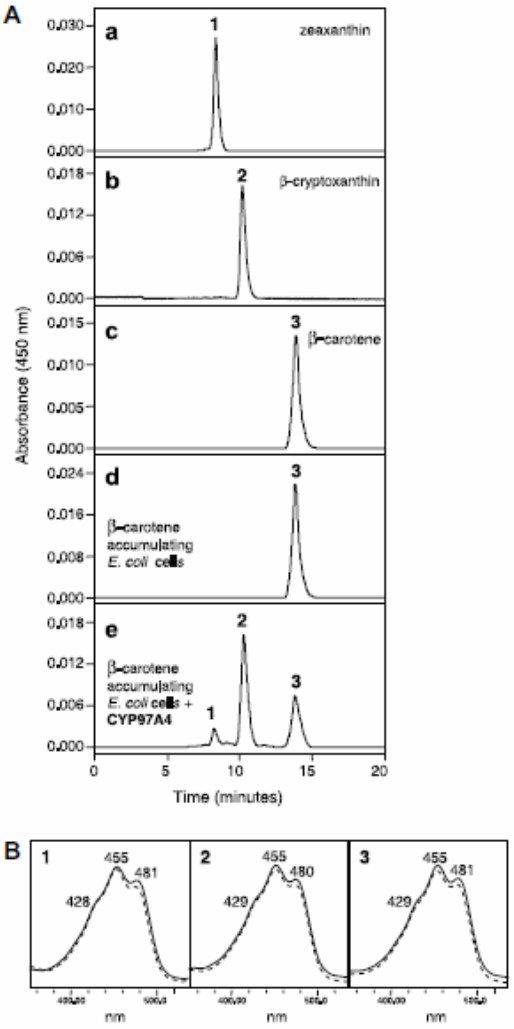


Fig. 4. Functional complementation of putative β-ring hydroxylase CYP97A4 cDNA in cells accumulating β-carotene. Bacterial cells were grown at 21 °C and extracted pigments were separated by reversed phase HPLC. (A) Chromatograms showing elution profiles: (a–c) authentic standards: 1, zeaxanthin; 2, β-cryptoxanthin; 3, β,β-carotene; (d) extracts from *E. coli* BL21 (DE3) cells accumulating β,β-carotene; (e) extracts from cells in (d) that were additionally transformed with the CYP97A4 cDNA. Peak numbers and identities in (d) and (e) correspond to those in (a–c). (B) Spectral profiles of standards (solid line) shown in (a–c, peaks 1–3) are superimposed on corresponding spectra (dashed line) obtained from peaks 1–3 shown in (e).

CYP97C2 cDNA, missing only 7 codons of the predicted 40 residue transit sequence, into pCOLADuet™-1 and introduced this construct into β-carotene accumulating cells; the pigment profile was compared with that obtained

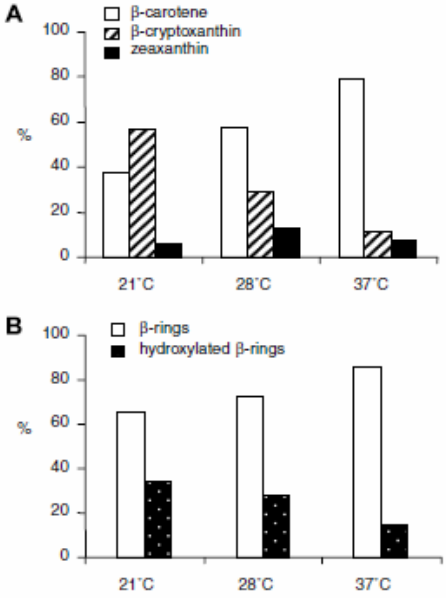


Fig. 5. Effect of temperature on product distribution and CYP97A4 enzyme activity *in vivo*. (A) Increasing the culturing temperature from 21 °C to 37 °C shifts the ratio of β-cryptoxanthin (one hydroxylated β-ring) to zeaxanthin (two hydroxylated β-rings) from 10:1 to ~1:1. (B) Overall % hydroxylation activity of CYP97A4 decreases from 35% at 21 °C to 15% at 37 °C.

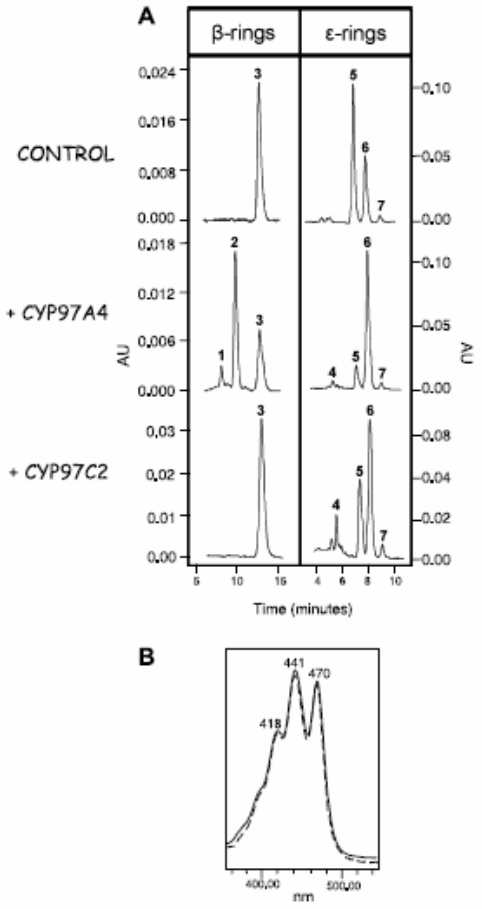


Fig. 6. Substrate specificity of P450 carotene hydroxylases. (A) *E. coli* cells engineered to accumulate carotenes with β rings (left panel) or ε-rings (right panel) were further transformed with test plasmids encoding CYP97A4 (putative β-ring hydroxylase) or CYP97C2 (putative ε-ring hydroxylase). HPLC chromatograms (left panel, 450 nm; right panel, 470 nm) show composition of accumulating carotenoid substrates and products. Left panel, *E. coli* cells transformed with pAC Beta-O4 (control) plus indicated test gene; Right panel, *E. coli* cells transformed with pACCRT-EIB+y2 (control) plus indicated test gene. Peaks: 1 = zeaxanthin, 2 = β-cryptoxanthin, 3 = β,β-carotene, 4 = lactucaxanthin, 5 = lycopene, 6 = δ-carotene, 7 = ε,ε-carotene. (B) Spectrum of peak 4 product of CYP97C2 (dashed line) overlayed with spectrum of authentic lactucaxanthin (solid line) chromatographed under identical conditions.

from CYP97A4 double transformed cells as above. The HPLC chromatogram (Fig. 6A, left panel, bottom) shows that when CYP97C2 was introduced into β-carotene accumulating cells, β-carotene substrate (peak 3) was the sole accumulating pigment as seen in the control (Fig. 6A, left panel, top, peak 3) and no hydroxylated products were detected. In comparison, introduction of CYP97A4 cDNA resulted in the mono- and di-hydroxylated products (left panel, middle, peaks 1 and 2).

In contrast, if the CYP97C2 cDNA was introduced into cells engineered with ε-ring substrates δ-carotene and ε,ε-carotene (Fig. 6A right panel, top, peaks 6 and 7), we found that CYP97C2 did confer hydroxylation of ε-rings. This was seen by accumulation of the more polar peaks including lactucaxanthin, a di-hydroxylated product (Fig. 6A, right panel, bottom, peak 4) with retention time and corresponding spectrum matching the authentic lactucaxanthin standard chromatographed in the same HPLC system (Fig. 6B). Hydroxylated products accumulated only when cells were grown at 37 °C and not at lower temperatures (data not shown). When the β-ring hydroxylase CYP97A4 was introduced into cells with the ε-ring substrates, there was a barely detectable amount of accumulated lactucaxanthin (Fig. 6 right panel, middle, peak 4). Spectra corresponding to the small bumps seen in the control (empty vector) do not match those for peak 4 nor for any carotenoid. Therefore, lactucaxanthin is only detected when the plant enzymes are present. These results demonstrate that the Clan A enzyme hydroxylates carotene β-rings and has weak activity towards carotene ε-rings while the Clan C enzyme only hydroxylates carotene ε-rings.

Discussion

Phylogenetic analysis of CYP97 genes

Predictive metabolic engineering or marker-based breeding of enhanced provitamin A levels depends on controlling conversion of provitamin A carotenes to non-provitamin A xanthophylls in endosperm and other tissues. To assess this conversion requires that we have a means to investigate activities and specificities of the putative hydroxylase enzymes. We focused on the recently discovered CYP97 family encoding P450 enzymes for which enzyme activities have been implied but never demonstrated. We first constructed an enzyme tree of CYP97 P450 enzymes to identify candidate sequences for ε- and β-ring hydroxylases in rice, a member of the Poaceae and representative of many agronomically important crops. Enzymes clustered into three clans, A, B, and C, and appeared to be encoded by single copy genes based on the finding that in each case, only one orthologous sequence was found in each of four different plant species, two being among the most exhaustively sequenced genomes (i.e., *A. thaliana* and *O. sativa*).

Functional complementation results integrated with earlier genetic studies

We isolated cDNAs encoding both β- and ε-ring hydroxylases in the *Oryza sativa* CYP97 family and demonstrated that *E. coli* cells were suitable for expression of functional enzymes to allow testing of substrate specificity. We conducted a preliminary examination of substrate specificity with regard to β- and ε-ring-containing substrates. Using *E. coli* functional complementation, we demonstrated that a Clan A enzyme, CYP97A4, is a β-ring carotene hydroxylase with some minor activity towards ε-rings. This minor ε-ring activity of the Clan A enzyme likely contributes to the residual lutein that accumulated in *Arabidopsis* plants carrying a *lut1* null allele that blocks expression of a Clan C ε-ring hydroxylase [30]. The phenotype of the null *lut1* allele also indicates that the minor ε-ring activity observed in the *E. coli* system is consistent with minor activity in the plant; almost no lutein accumulates in the null mutant and therefore the minor activity of the Clan A enzyme is indeed minor and cannot compensate for the Clan C deficiency. Studies on the non-heme diiron β-ring hydroxylases also showed the enzymes to display some activity towards ε- rings [24]. Given this observation in two structurally distinct enzymes (the P450 and diiron β-ring hydroxylases) suggests that this plasticity may be a function of the relationship between the extracted hydrogen and the configuration of the double bond in the β-ring as compared to the ε-ring [29].

Previous knockout mutagenesis in *A. thaliana* of the two genes encoding non-heme diiron β-carotene hydroxylases and one encoding the ε-ring hydroxylase (*LUT1*, a Clan C enzyme) [60] led the authors to suggest that the LUT1 ε-ring hydroxylase may be active towards β-rings. Using *E. coli* we showed that a Clan C enzyme, CYP97C2, lacked detectable activity towards β-rings. Therefore, our results do not support the hypothesis that further reduction in β-ring hydroxylation caused by a lesion in *Lut1* triple mutants compared to defects in the diiron enzymes alone, is attributed to a defect in the Clan C enzyme (which did not exhibit such activity under the conditions tested). Therefore, the observation in the triple mutant must have some other explanation. One possibility is that the Clan C and A enzymes form a complex, and interference in the expression of one affects expression/stability of the other, and hence the additional impact on β-ring hydroxylation activity in the triple mutants.

When comparing the activity of the rice enzymes in *E. coli*, we noted that the carotene ε-ring hydroxylation mediated by the Clan C enzyme was weak. One possibility is that the ε-ring substrate (one or two ε-rings) is not the best substrate as suggested by biochemical profiles of plant CYP97 mutants. *Arabidopsis* mutants defective in a Clan C ε-ring hydroxylase (*lut1*) conditioned accumulation of zeinoxanthin (hydroxylated β-ring of α-carotene) while plants defective in a Clan A β-ring hydroxylase (*lut5*) accumulated α-carotene and not α-cryptoxanthin (hydroxylated ε-ring of α-carotene) which would be expected if the available Clan C ε-ring hydroxylase could accept α-carotene directly. These combined observations suggest that the preferred *in planta* substrate of the Clan C ε-ring hydroxylases is zeinoxanthin (the mono-hydroxylated β-ring) and that the enzymes function in an ordered manner with the Clan A enzyme hydroxylating the carotene β-ring to produce zeinoxanthin followed by the Clan C enzyme hydroxylating the ε-ring [30,61]. Therefore, Clan C enzyme activity in *E. coli* might be improved with another substrate, the mono-hydroxylated β-ring of α-carotene (zeinoxanthin) and/or association with the Clan A enzyme. Enhancement of Clan C activity, by co-expression of the Clan A enzyme needed to produce zeinoxanthin, is now testable given demonstrated utility of the *E. coli* platform to express each individual enzyme.

We also observed that expression of the Clan A enzyme in *E. coli* led to significant accumulation of β-cryptoxanthin, even when different conditions were tested; this compound is not usually found in *Arabidopsis*, although it is found in other tissues of other species such as maize [16]. While we demonstrated hydroxylation of β-rings in *E. coli* by Clan A enzymes using β-carotene (two β-rings), the Clan A enzyme might prefer α-carotene (mixed β-ring and ε-ring) as the substrate. Therefore, the *E. coli* platform will be valuable in testing different types of substrates and combinations of enzymes.

What other factors are required for successful function in E. coli?

P450 enzyme activity is generally predicated on availability of an electron donor partner [62]. Such partners include NADPH flavodoxin reductases of various forms that are found in plant plastids, among other locations, and which function as electron carries for photosynthetic and non-photosynthetic processes such as sterol and fatty acid biosynthesis [63,64]. In *E. coli*, a bacterium which lacks P450 enzymes, flavodoxin: NADPH flavodoxin reductase can serve as a replacement for the natural redox partner to facilitate function of some plant P450 enzymes but not others [63,65]. For example, sorghum CYP71E1, required for cyanogenic glycoside biosynthesis, could both function *in vivo* in *E. coli* using the *E. coli* flavodoxin: NADPH flavodoxin reductase or reconstituted *in vitro* with a plant NADPH cytochrome P450 reductase [40]. In comparison, CYP79B2, an enzyme required for indole glucosinolate biosynthesis in *Arabidopsis*, functioned in *E. coli* only with support of the plant reductase [66] as found in other cases [39]. The CYP97 Clan A and Clan C enzymes were functional in *E. coli* because they could accept electrons from the endogenous *E. coli* flavodoxin reductase [63,65].

Why is CYP97 carotene hydroxylase activity suboptimal in E. coli?

Hydroxylase activity of the CYP97 enzymes in *E. coli* could be improved over the modest conversion of monohydroxylated intermediate to di-hydroxylated product. In addition to substrate choice as mentioned above, another factor that might limit enzyme activity might be the basal level expression of the endogenous reductase. Bacterial flavodoxin reductase, encoded by *E. coli fpr*, is induced about 20-fold from basal levels as part of a global response to oxidative stress that is mediated by the *soxRS* regulon [67,68]. In the absence of such induction, reductase expression may limit activity of the plant CYP. Other possible enhancements might include optimization of codon usage of the plant cDNA and/or removal of upstream plastid targeting sequences. However, posttranslational modification (e.g., glycosylation) is an unlikely prerequisite since the CYP97 enzymes did function in *E. coli*.

What is the structural basis for enzyme specificity?

Understanding the structural basis for β- and ε-ring specificity as demonstrated here will require modeling in combination with directed enzyme evolution. Modeling alone is insufficient to predict structural determinants of activity as seen in the characterization of other P450 enzymes. For example, directed evolution of *Bacillus subtilis* CYP102A2 enhanced its activity towards fatty acids and other aromatic substrates; affected amino acids were remote from the active site and unpredicted from modeling on related crystal structures to affect substrate specificity [69]. The *E. coli* platform used here will be valuable in addressing the issue of substrate specificity as well as to produce the protein crystals needed to elucidate a three dimensional structure of these crucial enzymes.

There are some recently described bacterial and fungal genes that encode P450 carotene ring hydroxylases. CrtS, is a 62.63 kDa (5.75 p*I*) fungal P450 that hydroxylates β-carotene to form astaxanthin. However, it is unrelated to the plant enzyme [70]. A P450 β-ring hydroxylase gene was also found in the carotenoid-containing thermophilic bacterium, *Thermus thermophilus* and the crystal structure has been solved [71]. The location of this single P450 gene on the thermophile's megaplasmid suggests that the gene may have been acquired through horizontal transfer and perhaps conferred some advantage to heat stress. The bacterial gene encodes a 44 kDa (9.72 p*I*) enzyme in comparison to the predicted mature 66 kDa (5.44 p*I*) plant Clan A enzyme [72,73], though they share 21% similar residues and cluster in a neighbor-joining tree [74]. While the plant and bacterial enzymes are sufficiently different, and perhaps cases of convergent evolution, they both functioned in *E. coli*.

Metabolic engineering in plants-challenges

From the results shown here and from prior studies, it is evident that higher plants contain one P450 ε carotene hydroxylase that mediates carotene ε-ring hydroxylation and two structurally distinct (P450 and diiron) β-ring hydroxylases, both of which catalyze hydroxylation of carotene β-rings leading to xanthophyll formation. A fundamental issue to address is how do the P450 and non-P450 enzymes contribute to metabolon structure(s) whose biogenesis may be differentially controlled whether the pathway is located on the envelope or thylakoid membranes; each location leads to intermediates and products with multiple roles in plant development and physiology. In addition, how do these enzymes contribute to biogenesis and control of a pathway that also diverges to form alternate end-products, one via β-carotene (the carotenoid with highest provitamin A value) and the other via α-carotene? Indeed, to develop rational strategies for metabolic engineering of specific carotenoids such as the provitamin A β-carotene or any other of the numerous carotenoids found in nature, will require a deeper understanding of how these hydroxylases participate in metabolon biogenesis and substrate channeling.

In rice, as in *Arabidopsis*, in addition to CYP97C2, there appear to be at least two non-heme diiron β-carotene hydroxylases (GenBank #XM_473611, AK060559). Further investigation of their suborganellar membrane localization will help to elucidate the specific roles of these β-ring hydroxylase classes on envelope and thylakoid membrane carotenoid biosynthetic pathways. Structurally they are unique: the diiron β-ring hydroxylases contain four transmembrane helices causing the enzymes to embed in

the membrane while the P450 hydroxylases are peripheral membrane enzymes due to their single transmembrane anchoring sequence. Other structural distinctions seen in the enzymes are their isoelectric points (p*I*); the p*I* of the mature rice Clan A enzyme is 5.44, while the rice diiron enzymes are 6.98 and 9.04. Given the differences in enzyme structure and membrane association, the enzymes may respond differentially to pH gradients that form across the thylakoid lumen at high light irradiance (lumenal pH 5/stromal pH 8). If, as is the case of the xanthophyll cycle de-epoxidase enzyme, activities of the carotene hydroxylases are pH-dependent, localization/orientation on the thylakoid membranes may be critical. At high light irradiance, acidification of the thylakoid lumen activates membrane association of the violaxanthin de-epoxidase, an enzyme having a p*I* of 4.57 [75], to convert violaxanthin to zeaxanthin, a product which dissipates the high light energy; in the dark violaxanthin accumulates. Perhaps diiron carotene β-ring hydroxylases function in the dark or on different plastid membranes (including envelope) for apocarotenoid biosynthesis, while P450 β-ring hydroxylases are light activated or function on the thylakoid membranes in association with the photosynthetic apparatus.

If on the other hand, the two structurally distinct carotene hydroxylases exist on the same membrane system, it is compelling to consider how they contribute to metabolon biogenesis and substrate channeling. Early modeling of Cunningham [22] portrayed separate complexes for production of α-carotene and β-carotene, precursors of lutein and zeaxanthin, respectively. However, some observations agree and others conflict with this model and suggest that the enzymes might play some compensatory role. For example, in studies of potato tuber endosperm, overexpression of a bacterial phytoene synthase gene led to increases in lutein and β-carotene but not in zeaxanthin [76]. The result suggests that there were adequate levels of the P450 CYP97 ε-ring and β-ring activities to accommodate the increased pathway flux to lutein. However, the Clan A P450 β-ring hydroxylase was not accessible for hydroxylation of the β-carotene needed to produce zeaxanthin which suggests that cells were limited for specific β-carotene hydroxylase activity required for zeaxanthin biosynthesis, presumably that of the non-heme diiron enzyme. Therefore, despite the fact that we demonstrated that the P450 Clan A enzyme could utilize β-carotene as a substrate in *E. coli*, it may be that the enzyme is not in a biochemical context to utilize this substrate in potato tuber. Our interpretation of the potato experiments is that they support Cunningham's view of separate submetabolons for the α-carotene and β-carotene sides of the pathway. However, if the Clan A enzyme cannot compensate for the diiron enzyme, then one would predict that a Clan A knockout would affect lutein accumulation only and not that of the β-carotene derived xanthophylls. This was not the case for the *Arabidopsis lut5* (Clan A) mutation which caused both reduced lutein and β-carotene derived xanthophylls [30], thus suggesting that the Clan A enzyme is not limited to α-carotene derived hydroxylation activity but may also play a role in hydroxylation of β-carotene. In another example, a double T-DNA knockout of the two genes encoding *Arabidopsis* diiron hydroxylases conferred no increase in β-carotene, an increase in lutein, and reduced β-carotene derived xanthophylls [60]. Given that β-carotene was not increased and β-carotene derived xanthophylls were not completely eliminated, suggests that the P450 Clan A enzyme could partially compensate for the missing diiron enzymes to catalyze synthesis of zeaxanthin from β-carotene. Interestingly, the double knockout in the diiron enzymes did affect lutein levels in the seed, suggesting that differences in plastid membranes in the seed and in the leaf may be associated with different mechanisms with respect to compensation between the P450 CYP97 and diiron β-ring hydroxylases. In summary, the literature holds conflicting examples regarding interchangeability between the P450 Clan A and diiron carotene β-ring hydroxylases. Plant mutations may have a pleiotropic effect on expression of other genes in the pathway, thereby confusing the interpretation of the resulting biochemical profile. Moreover, the observations made in *Arabidopsis* leaves may not translate to similar effects in cereal endosperm where plastid structure and gene family structure are different [1,37]. However, despite individual limitations, with the combined use of the bacterial system reported here, together with the plant genetic studies, we can derive a deeper understanding of plant carotenogenesis.

Conclusion

In conclusion, we demonstrated feasibility of the *E. coli* system as a platform to assess substrate specificity for representative members of the CYP97 A and C clans. This heterologous system will be valuable for further study and manipulation of these enzymes. The roles and topologies of the carotenoid metabolons are still an open question and important topics of research in identifying components that may be key targets for metabolic engineering of either "arm" of the carotenoid biosynthetic pathway. For example, to confer accumulation of the pathway intermediate, β-carotene, would necessitate a block in β-carotene hydroxylase expression/activity. At this point, we cannot predict the appropriate class of enzyme target, the P450 or the diiron. Future elucidation of components of such complexes will be especially significant for metabolic engineering of enhanced levels of carotenes, given that β-carotene and α-carotene have provitamin A activity, whereas their hydroxylated xanthophyll products zeaxanthin and lutein do not.

Acknowledgments

We thank Dr. Francis Cunningham Jr. for *Erwinia* expression plasmids, Christina Murillo for technical support, and Faqiang Li and Ratnakar Vallabhaneni for

technical advice. This research was supported by NIH (#S06-GM08225), PSC-CUNY, and New York State.

References

[1] E.T. Wurtzel, in: J. Romeo (Ed.), Recent Advances in Phytochemistry, Elsevier Ltd., 2004, pp. 85–110.
[2] P.D. Fraser, P.M. Bramley, Prog. Lipid Res. 43 (2004) 228–265.
[3] D. DellaPenna, B.J. Pogson, Annu. Rev. Plant Biol. 57 (2006) 711–738.
[4] B.V. Milborrow, J. Exp. Bot. 52 (2001) 1145–1164.
[5] A.J. Simkin, S.H. Schwartz, M. Auldridge, M.G. Taylor, H.J. Klee, Plant J. 40 (2004) 882–892.
[6] R. Castillo, J.-A. Fernandez, L. Gomez-Gomez, Plant Physiol. 139 (2005) 674–689.
[7] A.R. Moise, J. von Lintig, K. Palczewski, Trends Plant Sci. 10 (2005) 178–186.
[8] S.H. Schwartz, X. Qin, M.C. Loewen, J. Biol. Chem. 279 (2004) 46940–46945.
[9] R. Matusova, K. Rani, F.W.A. Verstappen, M.C.R. Franssen, M.H. Beale, H.J. Bouwmeester, Plant Physiol. 139 (2005) 920–934.
[10] J. Booker, M. Auldridge, S. Wills, D. McCarty, H. Klee, O. Leyser, Curr. Biol. 14 (2004) 1232–1238.
[11] F. Bouvier, C. Suire, J. Mutterer, B. Camara, Plant Cell 15 (2003) 47–62.
[12] X. Ye, S. Al-Babili, A. Kloti, J. Zhang, P. Lucca, P. Beyer, I. Potrykus, Science 287 (2000) 303–305.
[13] C.K. Shewmaker, J.A. Sheehy, M. Daley, S. Colburn, D.Y. Ke, Plant J. 20 (1999) 401–412.
[14] V. Mann, M. Harker, I. Pecker, J. Hirschberg, Nat. Biotechnol. 18 (2000) 888–892.
[15] C. Rosati, R. Aquilani, S. Dharmapuri, P. Pallara, C. Marusic, R. Tavazza, F. Bouvier, B. Camara, G. Giuliano, Plant J. 24 (2000) 413–419.
[16] S.N. Islam, Masters Thesis in Crop Sciences, University of Illinois, Urbana-Champaign, 2004, 93.
[17] A. Kurilich, J. Juvik, J. Agric. Food Chem. 47 (1999) 1948–1955.
[18] J.C. Wong, R.J. Lambert, T.R. Rocheford, Proc. 38th Ann. Illinois Corn Breeders School, 2002, pp. 145–170.
[19] J.C. Wong, R.J. Lambert, E.T. Wurtzel, T.R. Rocheford, Theor. Appl. Genet. 108 (2004) 349–359.
[20] W. Zhang, A.J. Lukaszewski, J. Kolmer, M.A. Soria, S. Goyal, J. Dubcovsky, TAG. Theor. Appl. Genet. 111 (2005) 573–582.
[21] G.D. Parker, K.J. Chalmers, A.J. Rathjen, P. Langridge, Theor. Appl. Genet. 97 (1998) 238–245.
[22] F.X. Cunningham Jr., E. Gantt, Annu. Rev. Plant Physiol. Plant Mol. Biol. (1998) 557–583.
[23] J. von Lintig, K. Vogt, J. Nutr. 134 (2004) 251S–256S.
[24] Z. Sun, E. Gantt, F.X. Cunningham Jr., J. Biol. Chem. 271 (1996) 24349–24352.
[25] L. Tian, D. DellaPenna, Arch. Biochem. Biophys. 430 (2004) 22–29.
[26] F. Bouvier, R.A. Backhaus, B. Camara, J. Biol. Chem. 273 (1998) 30651–30659.
[27] M.J. Ryle, R.P. Hausinger, Curr. Opin. Chem. Biol. 6 (2002) 193–201.
[28] E. Bertrand, R. Sakai, E. Rozhkova-Novosad, L. Moe, B.G. Fox, J.T. Groves, R.N. Austin, J. Inorg. Biochem. 99 (2005) 1998–2006.
[29] L. Tian, V. Musetti, J. Kim, M. Magallanes-Lundback, D. DellaPenna, Proc. Natl. Acad. Sci. USA 101 (2004) 402–407.
[30] J. Kim, D. DellaPenna, Proc. Natl. Acad. Sci. USA 103 (2006) 3474–3479.
[31] P.B. Danielson, Curr. Drug Metab. 3 (2002) 561–597.
[32] C. Chapple, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49 (1998) 311–343.
[33] W.D. Woggon, Acc. Chem Res. 38 (2005) 127–136.
[34] G. Sandmann, M. Albrecht, G. Schnurr, O. Knorzer, P. Boger, Trends Biotechnol. 17 (1999) 233–237.
[35] P.D. Matthews, E.T. Wurtzel, Appl. Microbiol. Biotechnol. 53 (2000) 396–400.
[36] N. Misawa, M. Nakagawa, K. Kobayashi, S. Yamano, Y. Izawa, K. Nakamura, K. Harashima, J. Bacteriol. 172 (1990) 6704–6712.
[37] C.E. Gallagher, P.D. Matthews, F. Li, E.T. Wurtzel, Plant Physiol. 135 (2004) 1776–1783.
[38] T. Isaacson, G. Ronen, D. Zamir, J. Hirschberg, Plant Cell 14 (2002) 333–342.
[39] M.A. Schuler, D. Werck-Reichhart, Annu. Rev. Plant Biol. 54 (2003) 629–667.
[40] S. Bak, R.A. Kahn, H.L. Nielsen, B.L. Møller, B.A. Halkier, Plant Mol. Biol. 36 (1998) 393–405.
[41] P. Naur, C.H. Hansen, S. Bak, B.G. Hansen, N.B. Jensen, H.L. Nielsen, B.A. Halkier, Arch. Biochem. Biophys. 409 (2003) 235–241.
[42] S. Kikuchi, K. Satoh, T. Nagata, N. Kawagashira, K. Doi, N. Kishimoto, J. Yazaki, M. Ishikawa, H. Yamada, H. Ooka, I. Hotta, K. Kojima, T. Namiki, E. Ohneda, W. Yahagi, K. Suzuki, C.J. Li, K. Ohtsuki, T. Shishiki, Y. Otomo, K. Murakami, Y. Iida, S. Sugano, T. Fujimura, Y. Suzuki, Y. Tsunoda, T. Kurosaki, T. Kodama, H. Masuda, M. Kobayashi, Q. Xie, M. Lu, R. Narikawa, A. Sugiyama, K. Mizuno, S. Yokomizo, J. Niikura, R. Ikeda, J. Ishibiki, M. Kawamata, A. Yoshimura, J. Miura, T. Kusumegi, M. Oka, R. Ryu, M. Ueda, K. Matsubara, J. Kawai, P. Carninci, J. Adachi, K. Aizawa, T. Arakawa, S. Fukuda, A. Hara, W. Hashizume, N. Hayatsu, K. Imotani, Y. Ishii, M. Itoh, I. Kagawa, S. Kondo, H. Konno, A. Miyazaki, N. Osato, Y. Ota, R. Saito, D. Sasaki, K. Sato, K. Shibata, A. Shinagawa, T. Shiraki, M. Yoshino, Y. Hayashizaki, A. Yasunishi, Science 301 (2003) 376–379.
[43] M. Baltrusch, M. Fukla, F.-P. Wolter, E. Heinz, Plant Physiol. 114 (1997) 1568.
[44] B. Siminszky, F.T. Corbin, E.R. Ward, T.J. Fleischmann, R.E. Dewey, Proc. Natl. Acad. Sci. USA 96 (1999) 1750–1755.
[45] O. Emanuelsson, H. Nielsen, G. von Heijne, Protein Sci. 8 (1999) 978–984.
[46] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, Bioinformatics 17 (2001) 1244–1245.
[47] F.X. Cunningham Jr., D. Chamovitz, N. Misawa, E. Gantt, J. Hirschberg, FEBS Lett. 328 (1993) 130–138.
[48] F.X. Cunningham Jr., B. Pogson, Z. Sun, K.A. McDonald, D. DellaPenna, E. Gantt, Plant Cell 8 (1996) 1613–1626.
[49] P.A. Davison, C.N. Hunter, P. Horton, Nature 418 (2002) 203–206.
[50] D. Siefermann-Harms, S. Hertzberg, G. Borch, S. Liaaen-Jensen, Phytochemistry 20 (1981) 85–88.
[51] G. Britton, S. Liaaen-Jensen, H. Pfander (Eds.), Carotenoids Handbook, Birkhäuser Verlag, Basel, 2004.
[52] D.R. Nelson, M.A. Schuler, S.M. Paquette, D. Werck-Reichhart, S. Bak, Plant Physiol. 135 (2004) 756–772.
[53] S. Yang, R.S.S. Wu, H.O.L. Mok, Z.P. Zhang, R.Y.C. Kong, J. Phycol. 39 (2003) 555–560.
[54] B.A. Hanley, M.A. Schuler, Nucleic Acids Res. 16 (1988) 7159–7176.
[55] M. Kozak, Nucleic Acids Res. 15 (1987) 8125–8148.
[56] H.M. Rothnie, Plant Mol. Biol. 32 (1996) 43–61.
[57] J. Zhao, L. Hyman, C. Moore, Microbiol. Mol. Biol. Rev. 63 (1999) 405–445.
[58] A. Marchler-Bauer, J.B. Anderson, P.F. Cherukuri, C. DeWeese-Scott, L.Y. Geer, M. Gwadz, S. He, D.I. Hurwitz, J.D. Jackson, Z. Ke, C.J. Lanczycki, C.A. Liebert, C. Liu, F. Lu, G.H. Marchler, M. Mullokandov, B.A. Shoemaker, V. Simonyan, J.S. Song, P.A. Thiessen, R.A. Yamashita, J.J. Yin, D. Zhang, S.H. Bryant, Nucleic Acids Res. 33 (2005) D192–D196.
[59] G.E. Tusnady, I. Simon, J. Mol. Biol. 283 (1998) 489–506.
[60] L. Tian, M. Magallanes-Lundback, V. Musetti, D. DellaPenna, Plant Cell 15 (2003) 1320–1332.
[61] A. Fiore, L. Dall'Osto, P.D. Fraser, R. Bassi, G. Giuliano, FEBS Lett. 580 (2006) 4718–4722.
[62] R. Bernhardt, J. Biotechnol. 124 (2006) 128–145.
[63] N. Carrillo, E.A. Ceccarelli, Eur. J. Biochem. 270 (2003) 1900–1915.

[64] E.A. Ceccarelli, A.K. Arakaki, N. Cortez, N. Carrillo, Biochim. Biophys. Acta (BBA). Proteins & Proteomics 1698 (2004) 155–165.
[65] C. Jenkins, M. Waterman, J. Biol. Chem. 269 (1994) 27401–27408.
[66] M.D. Mikkelsen, C.H. Hansen, U. Wittstock, B.A. Halkier, J. Biol. Chem. 275 (2000) 33712–33717.
[67] S. Liochev, A. Hausladen, W. Beyer Jr., I. Fridovich, Proc. Natl. Acad. Sci. USA 91 (1994) 1328–1331.
[68] A.R. Krapp, R.E. Rodriguez, H.O. Poli, D.H. Paladini, J.F. Palatnik, N. Carrillo, J. Bacteriol. 184 (2002) 1474–1480.
[69] I. Axarli, A. Prigipaki, N.E. Labrou, Biomol. Eng. 22 (2005) 81–88.
[70] V. Alvarez, M. Rodriguez-Saiz, J.L. de la Fuente, E.J. Gudina, R.P. Godio, J.F. Martin, J.L. Barredo, Fungal Genet. Biol. 43 (2006) 261–272.
[71] J.K. Yano, F. Blasco, H. Li, R.D. Schmid, A. Henne, T.L. Poulos, J. Biol. Chem. 278 (2003) 608–616.
[72] F. Blasco, I. Kauffmann, R.D. Schmid, Appl. Microbiol. Biotechnol. 64 (2004) 671–674.
[73] A. Henne, H. Bruggemann, C. Raasch, A. Wiezer, T. Hartsch, H. Liesegang, A. Johann, T. Lienard, O. Gohl, R. Martinez-Arias, C. Jacobi, V. Starkuviene, S. Schlenczeck, S. Dencker, R. Huber, H.-P. Klenk, W. Kramer, R. Merkl, G. Gottschalk, H.-J. Fritz, Nat. Biotech. 22 (2004) 547–553.
[74] K. Inoue, Trends Plant Sci. 9 (2004) 515–517.
[75] R.C. Bugos, H.Y. Yamamoto, Proc. Natl. Acad. Sci. USA 93 (1996) 6320–6325.
[76] L.J. Ducreux, W.L. Morris, P.E. Hedley, T. Shepherd, H.V. Davies, S. Millam, M.A. Taylor, J. Exp. Bot. 56 (2005) 81–89.

Appendix 2: A co-authored publication related to this dissertation[2]

Metabolite Sorting of a Germplasm Collection Reveals the *Hydroxylase3* Locus as a New Target for Maize Provitamin A Biofortification[1][C][W][OA]

Ratnakar Vallabhaneni, Cynthia E. Gallagher, Nicholas Licciardello, Abby J. Cuttriss, Rena F. Quinlan, and Eleanore T. Wurtzel*

Department of Biological Sciences, Lehman College, City University of New York, Bronx, New York 10468 (R.V., C.E.G., N.L., A.J.C., R.F.Q., E.T.W.); and Graduate School and University Center, City University of New York, New York, New York 10016–4309 (R.V., R.F.Q., E.T.W.)

Vitamin A deficiency, a global health burden, can be alleviated through provitamin A carotenoid biofortification of major crop staples such as maize (*Zea mays*) and other grasses in the Poaceae. If regulation of carotenoid biosynthesis was better understood, enhancement could be controlled by limiting β-carotene hydroxylation to compounds with lower or no nonprovitamin A activity. Natural maize genetic diversity enabled identification of hydroxylation genes associated with reduced endosperm provitamin A content. A novel approach was used to capture the genetic and biochemical diversity of a large germplasm collection, representing 80% of maize genetic diversity, without having to sample the entire collection. Metabolite data sorting was applied to select a 10-line genetically diverse subset representing biochemical extremes for maize kernel carotenoids. Transcript profiling led to discovery of the *Hydroxylase3* locus that coincidently mapped to a carotene quantitative trait locus, thereby prompting investigation of allelic variation in a broader collection. Three natural alleles in 51 maize lines explained 78% of variation and approximately 11-fold difference in β-carotene relative to β-cryptoxanthin and 36% of the variation and 4-fold difference in absolute levels of β-carotene. A simple PCR assay to track and identify *Hydroxylase3* alleles will be valuable for predicting nutritional content in genetically diverse cultivars found worldwide.

Vitamin A deficiency is a global health problem affecting 140 to 250 million children and accounts for increased childhood mortality and disease (World Health Organization, 1995; Underwood, 2004; Black et al., 2008). Humans and animals are unable to synthesize their own vitamin A and rely on dietary provitamin A carotenoid pigments; plant-derived carotenes are metabolized to produce vitamin A, two molecules from β-carotene and only one from α-carotene or β-cryptoxanthin. Seed endosperm tissue of maize (*Zea mays*) and other grasses (Poaceae) represents 70% of worldwide food production (Chandler and Brendel, 2002), but has limited provitamin A value. Improvement of provitamin A content in staple crops is therefore a critical step toward alleviating vitamin A deficiency worldwide. Maize, a major food staple in vitamin A-deficient Sub-Saharan Africa and Latin America, exhibits a wide range of carotenoid content and composition. The underlying basis for this variation has been the subject of quantitative trait locus (QTL) and association studies (Palaisa et al., 2003; Wong et al., 2004; Chander et al., 2007; Harjes et al., 2008). Given the genetic and biochemical diversity of maize, opportunity exists for developing breeding markers, based on candidate genes of the carotenoid biosynthetic pathway, to select for predicted nutritional content.

The carotenes are intermediates in the (recently revised) carotenoid biosynthetic pathway (Li et al., 2007; Matthews and Wurtzel, 2007), leading to the predominant nonprovitamin A xanthophylls in maize kernels (Fig. 1), which is why maize kernels generally have limited provitamin A content despite being rich in carotenoids (unlike rice [*Oryza sativa*], which does not naturally accumulate endosperm carotenoids; Islam, 2004). Lycopene, the immediate precursor of provitamin A carotenes, represents a branch point in the carotenoid biosynthetic pathway. The linear lycopene molecule is further modified by lycopene β-cyclase and lycopene ε-cyclase (LCYE) enzymes that catalyze formation of terminal β-ionone rings and ε-rings, respectively. Depending on relative activity levels of the cyclases, the products are β-carotene with two β-ionone rings, the most potent dietary source of vitamin A, and α-carotene, with only one β-ionone ring and therefore half the provitamin A potential compared to β-carotene; the hydroxylation intermedi-

[1] This work was supported by the National Institutes of Health (grant nos. GM08225 and GM081160 to E.T.W.).

* Corresponding author; e-mail wurtzel@lehman.cuny.edu.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Eleanore T. Wurtzel (wurtzel@lehman.cuny.edu).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.109.145177

[2]This appendix is a publication of Vallabhaneni et al., (2009) Plant Physiol. 151: 1635-1645, (www.plantphysiol.org) with permission from the publisher (American Society of Plant Biologists)

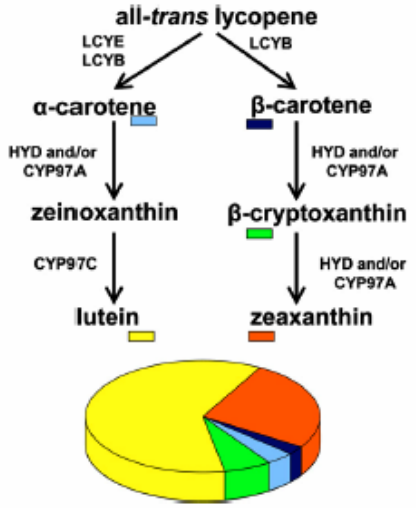


Figure 1. Conversion of provitamin A carotenes to nonprovitamin xanthophylls. Relative abundance of mature kernel carotenoids in maize inbred B73. LCYB, lycopene β-cyclase; HYD, β-carotene hydroxylase (diiron); CYP97A, β-carotene hydroxylase (P450); CYP97C, ε-carotene hydroxylase (P450). [See online article for color version of this figure.]

ate β-cryptoxanthin, with only one β-ionone ring, has half the provitamin A activity compared to β-carotene. The pathway continues with hydroxylation of the carotenes that depletes the provitamin A pool by converting provitamin A compounds to nonprovitamin A xanthophylls (Matthews and Wurtzel, 2007). Therefore, pathway branching and hydroxylation are key determinants in controlling provitamin A levels. Both of these aspects have been targets for metabolic engineering of endosperm carotenoids in other species (Diretto et al., 2006, 2007). However, transgenic solutions to manipulate the carotenoid biosynthetic pathway, as achieved in Golden Rice (Ye et al., 2000) and other plants (Giuliano et al., 2008), are not always acceptable as observed by the overwhelming public resistance to genetically modified organism food crops. Moreover, transgene incorporation into one genetic variety used for laboratory transformation produces variable carotenoid phenotypes (Aluru et al., 2008), and the desired transgene-produced trait is not predictably transferred to other genetic backgrounds of cultivars worldwide. This lack of predictability could be overcome given a better understanding of the rate-controlling steps and tools to genotype and/or phenotype varieties for predicting the outcome of transgene introduction as we report here. Given the extensive natural diversity inherent in maize, it is conceivable to predictably breed high provitamin A maize in a wide range of genotypes given a thorough understanding of pathway bottlenecks and development of corresponding breeding alleles.

Our collaborative effort in the recent development of *LCYE*-based breeding markers for maize demonstrated feasibility of a nontransgenic, traditional breeding approach to control the pathway branching step and force pathway flux toward β-carotene and its nonprovitamin A derivatives (the β-branch; Harjes et al., 2008). However, unless hydroxylation is also controlled, nonprovitamin A xanthophyll compounds will predominate. Given that lycopene cyclization can be forced toward the β-branch, the next challenge in breeding high β-carotene in maize endosperm is to block β-carotene hydroxylation to increase levels of β-carotene relative to β-cryptoxanthin and downstream xanthophylls. Therefore, we embarked on a study capitalizing on maize germplasm resources to characterize the maize β-carotene hydroxylase genes and to discover breeding markers for enhancing the relative levels of seed β-carotene.

RESULTS AND DISCUSSION

Two structurally distinct classes of carotene hydroxylases are known: the P450 heme-thiolate CYP97A and CYP97C enzymes and the nonheme diiron monooxygenases (for review, see Matthews and Wurtzel, 2007; Giuliano et al., 2008). As found in rice (Quinlan et al., 2007), maize contained one gene each for CYP97A and CYP97C, respectively. A total of six unlinked maize paralogs encoding nonheme diiron β-carotene hydroxylases (*HYD*) were identified in contrast to three rice genes (Fig. 2; Supplemental Tables S1 and S2); phylogenetic analysis suggested that the gene duplications found in the grasses occurred after the monocot and dicot evolutionary split. Maize *HYD1* and *HYD2* are pseudogenes, while *HYD3*, *HYD4*, *HYD5*, and *HYD6* encode enzymes with characteristic hydroxylase domains and plastid-targeting signals (Sun et al., 1996; Supplemental Fig. S1). *HYD3* and *HYD4* are syntenous with rice *HYD1* (Fig. 3) and are predicted to encode proteins with markedly different pIs, but functional testing confirmed that both encode carotene β-ring hydroxylases (Supplemental Fig. S2). The *CYP97A* and *CYP97C* genes have been functionally tested in rice (Quinlan et al., 2007); phylogenetic analysis of this ancient evolutionary clade would suggest that the maize encoded enzymes behave similarly.

In the maize inbred B73 that has been used extensively for investigating regulation of endosperm carotenogenesis (Li et al., 1996, 2008a, 2008b, 2009; Matthews et al., 2003; Vallabhaneni and Wurtzel, 2009), mRNA abundance for the six functional carotene hydroxylases genes varied among tissues and during endosperm development, where comparison was also made with carotenoid accumulation (Fig. 4). To test which, if any, carotene hydroxylase gene

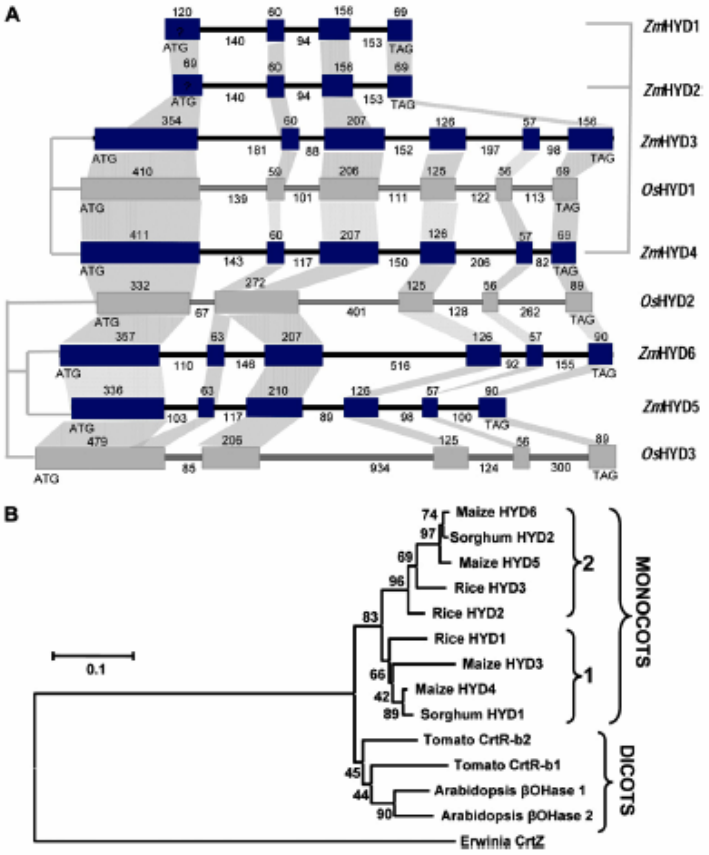


Figure 2. HYD gene family structure and phylogeny. A, Gene organizational structure showing paralogs and orthologs of the *HYD* gene family in maize (*Zm*) and rice (*Os*) species representing two subfamilies of the Poaceae (grasses). Boxes and lines indicate exons and introns, respectively, sizes for which are in bp. Vertical lines connecting paralogs represent a subfamily. B, Phenetic analysis showing the evolution and duplication of carotenoid hydroxylase amino acid sequences. Monocot sequences fall into two major groups labeled 1 and 2. GenBank accessions are in parentheses: *Erwinia* (*crtZ* AAA64983), Arabidopsis (βOHase1, AT4G25700; βOHase2, AT5G52570), tomato (*Solanum lycopersicum*; crtR-b1, Y14809; crtR-b2, Y14810); "Materials and Methods" and Supplemental Table S1 list monocot accession numbers. [See online article for color version of this figure.]

showed a statistical correlation between mRNA levels and endosperm provitamin A β-carotene content, transcript abundance was quantified in a genetically diverse maize germplasm collection (Liu et al., 2003; Islam, 2004; Harjes et al., 2008). However, it was impractical to screen endosperm developmental samples in staged, hand-pollinated plants for the entire collection, representing 80% of maize genetic diversity. Therefore, a uniquely selected subset of lines for testing was chosen from 148 lines with known carotenoid content and composition (Islam, 2004). From existing metabolite data, we sorted lines for highest total carotenoid content, highest ratio of β-carotene to β-cryptoxanthin, highest ratio of β-cryptoxanthin to zeaxanthin, highest ratio of β-carotene to zeaxanthin, highest ratio of α-carotene to lutein, and highest ratio of lutein to zeaxanthin. The resulting 10-line subset (Supplemental Tables S3 and S4) exhibited the most extreme carotenoid biochemical phenotypes and likely contained the most favorable and informative alleles for breeding carotenoid content and composition. For example, a block in carotene hydroxylase activity would be predicted to manifest as a high ratio of β-carotene to β-cryptoxanthin, while factors that controlled flux would be predicted to influence levels of total carotenoids. The lines are genetically diverse, spanning four major genetic diversity groups and eight subgroups (Supplemental Table S4) among the 260 Goodman diversity lines (Liu et al., 2003). The 10 lines were then used for quantitative transcript profiling of endosperm samples at different developmental stages (days after pollination [DAP]). To demonstrate adequacy of sample size and to validate use of this biochemically diverse subset to infer pathway regulation, we applied Pearson correlation analysis to show that *PSY1* but not *PSY2* or *PSY3* transcript levels showed statistically significant correlation with seed carotenoid content (Li et al., 2008b; Vallabhaneni and Wurtzel, 2009). The findings were consistent with molecular and genetic studies associating *PSY1* and control of pathway flux in endosperm (Randolph and Hand, 1940; Palaisa et al., 2003; Wong et al., 2004;

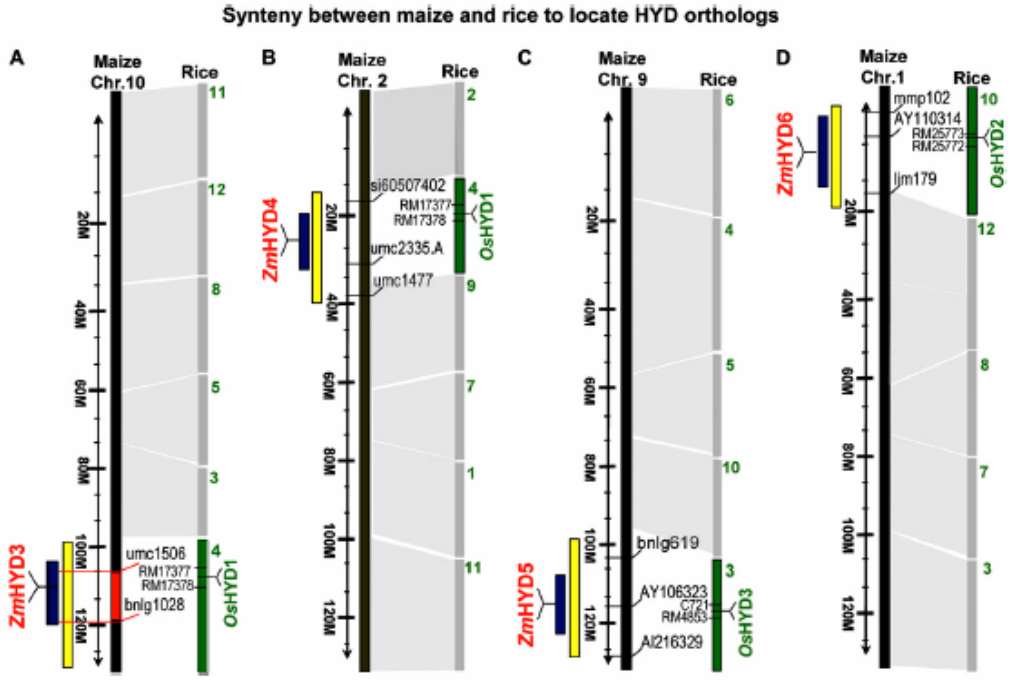


Figure 3. Synteny between maize and rice to map maize *HYD* orthologs. Orthologous genes encoding maize and rice nonheme diiron β-carotene hydroxylases were compared by synteny, where each chromosome of maize is syntenous to multiple rice chromosome regions numbered in green. Each maize chromosome is marked by bin (yellow box) and contig (dark-blue box), with its respective maize gene and linked genetic markers. Maize *HYD3* (*ZmHYD3*) maps to chromosome *10* between DNA markers *umc1506* and *bnlg1028* that flank a β-carotene QTL (Chander et al., 2007) colored in red. A, *ZmHYD3* (BAC clone no. c0156B23; contig no. 415; chromosome *10.05*). B, *ZmHYD4* (BAC clone no. c0112K16; contig no. 73; chromosome *2.03*). C, *ZmHYD5* (BAC clone no. b0589L12; contig no. 391; chromosome *9.07*). D, *ZmHYD6* (BAC clone no. b0072L05; contig no. 03; chromosome *1.01*). [See online article for color version of this figure.]

Pozniak et al., 2007; Li et al., 2008b). The subset has also been recently tested to examine additional genes that encode enzymes for steps that represent carotenoid biosynthetic pathway bottlenecks and those steps that do not (Vallabhaneni and Wurtzel, 2009). For example, the finding of carotenoid bottlenecks in the upstream isoprenoid biosynthetic pathway is consistent with other studies in the literature. The validated germplasm subset was then used to assess the possible role of carotene hydroxylase gene expression in controlling seed β-carotene composition. Transcript abundance was quantitatively measured, over a range of endosperm developmental stages, for the five carotene β-ring hydroxylase genes, *HYD3*, *HYD4*, *HYD5*, *HYD6*, and *CYP97A*, and the one ε-ring hydroxylase gene, *CYP97C*. Time points chosen were based on earlier studies that investigated carotenoid accumulation during endosperm development (Li et al., 2008b). *HYD3* was the only gene for which transcripts were found to correlate with carotenoids (Supplemental Table S3). Interestingly, some of the lines showed a steep reduction in *HYD3* transcripts between 15 and 25 DAP, ranging from 0.7- to >5,000-fold. A Pearson correlation analysis comparing β-carotene content and transcript levels for all carotene hydroxylase genes was conducted and *HYD3* was the only gene for which there was a statistically significant ($P \leq 0.05$) correlation between transcript levels and provitamin A β-carotene content (Fig. 5). In the correlation plots shown, where the colored lines denote a 95% confidence interval, the fold reduction of *HYD3* transcripts over endosperm development (15–25 DAP) positively correlated with seed β-carotene (μg/g) levels ($r = 0.79$, $P = 0.005$; Fig. 5A). As would be predicted for a critical endosperm hydroxylase gene, *HYD3* transcripts at 25 DAP also positively correlated with zeaxanthin (μg/g) levels ($r = 0.78$, $P = 0.007$; Fig. 5B). That is, the greater the reduction in transcripts over development, the higher the β-carotene content, as would be predicted from a correspondingly lower level of hydroxylase

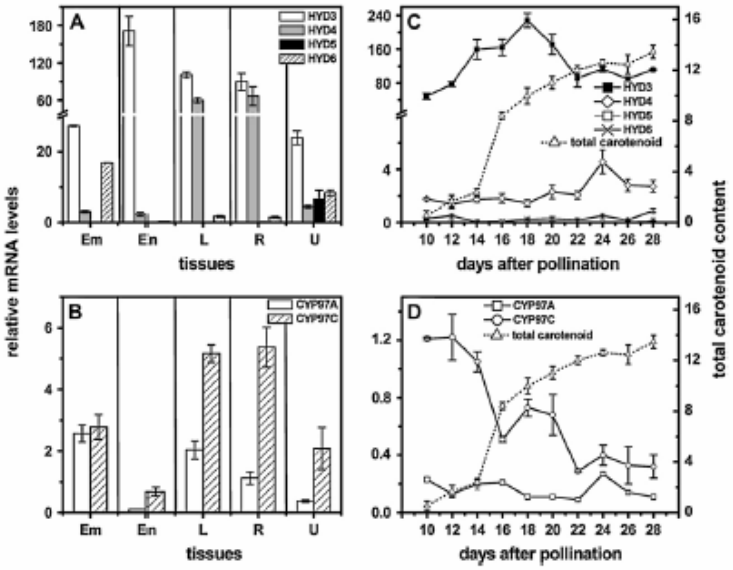


Figure 4. Transcript profiling of hydroxylase gene family members in maize standard, B73 inbred. Transcript levels of *HYD* (A) and P450-type (B) hydroxylase genes in different maize tissues: embryo (Em) and endosperm (En) collected at 20 DAP from field-grown plants; leaf (L) and root (R) samples collected from seedlings at the six-leaf stage; unfertilized ovule (U). Transcript levels were expressed relative to endosperm levels at 20 DAP. Transcript abundance in developing endosperm is compared to accumulation of total carotenoids (C and D). Mean of three replicates ± SD.

activity; conversely, the higher the transcripts at 25 DAP, the greater the level of zeaxanthin, as would be predicted from increased hydroxylase activity.

We noted the presence of a QTL for endosperm β-carotene content (Chander et al., 2007) that maps together with *HYD3*, between markers *umc1506* and *bnlg1028* on maize chromosome 10.05 (Fig. 3). Therefore, we sequenced all *HYD3* alleles in the 10-line subset. Variation consistently seen in the high β-carotene lines was found in an approximately 40-bp region adjacent to the transcript start site (Fig. 6, blue box). A conserved transcript start site and first ATG were mapped by aligning available paralog-specific ESTs and genomic DNA for maize B73 *HYD3* and sister paralog *HYD4*, and rice synteny partner *HYD1* (Fig. 6). B73 and the other six low β-carotene lines had identical sequence in this region (variant A); A619 had sequence variant B, while CI.7 and DE3 shared variation C. B and C in the high β-carotene lines appear to contain a duplicated sequence from a downstream region that replaced the progenitor sequence seen in A, although allele C shows a more complete match (14/17 nucleotides) as compared to allele B (10/17 nucleotides). Diversity analysis (Liu et al., 2003) of the 260 Goodman lines placed A619 and CI.7, which carried different *HYD3* variations, in two genetically distinct subgroups of the maize non-Stiff stalk lines; such diversity grouping is consistent with finding that they do not share the same polymorphism. The identical C polymorphism seen in DE3 and CI.7 may be a result of a possible overlapping pedigree since DE3 was placed in a mixed group as an indication of genetic structure shared with more than one maize diversity group. However, further sequencing of the entire *HYD3* gene of DE3 and CI.7 revealed some minor sequence differences between *HYD3* alleles in these two inbreds (Supplemental Fig. S3).

To test whether *HYD3* allelic variation could explain β-carotene variation, a PCR assay was developed to rapidly genotype a broader maize collection. The assay distinguished between A and C or between A and B alleles and generated a common *HYD3* paralog-specific product that could be sequenced to reveal new alleles; the assay correctly genotyped all 10 lines for their *HYD3* allele (Fig. 7). We then selected an additional 41 lines spanning maize genetic diversity and range of carotenoid content and composition (Supplemental Table S4). The combined group of 51 lines included nine new lines identified by metabolite sorting to exhibit high ratios of β-carotene to β-cryptoxanthin, as possible indicator of a block in β-carotene hydroxylation. The common *HYD3* paralog-specific product was amplified and sequenced to characterize *HYD3* alleles in this broader maize collection (Fig. 7). In the combined 51 lines, we identified 14 high-carotene alleles: seven C alleles, including one new C-like allele, and seven B alleles. The 14 B and C alleles were widely distributed among genetic diversity groups, as were the 37 A alleles (Supplemental Table S4). One-way ANOVA showed that the mean ratios of β-carotene to β-cryptoxanthin exhibited a statistically significant difference among the three *HYD3* alleles ($F_{2,48} = 89.3$, $P < 0.0001$) and that the *HYD3* allele explained 78% of the variation (R^2) in the β-carotene-to-β-cryptoxanthin ratio (Fig. 8). The absolute amount of β-carotene (μg/gm) also showed a statistically significant difference

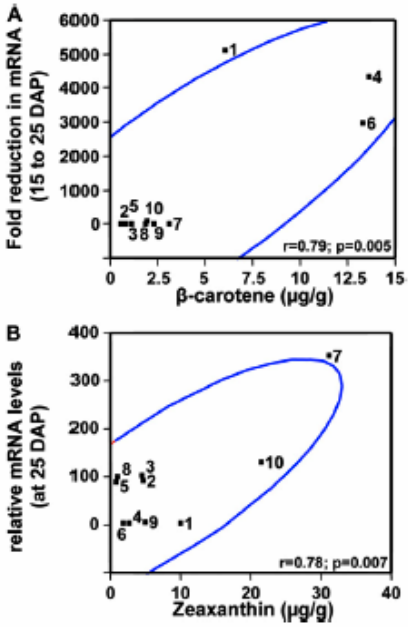


Figure 5. Correlation analysis showing *HYD3* mRNA levels correlate with endosperm provitamin A content. A, Provitamin A β-carotene levels correlate with the fold reduction in *HYD3* transcripts between 15 and 25 DAP. B, Nonprovitamin A zeaxanthin levels correlate with *HYD3* transcripts at 25 DAP. Pearson correlation (*r*) and statistical significance (*P*) for comparison of *HYD3* transcript levels and kernel carotenoid composition (taken from Supplemental Table S3) in maize inbred lines (Harjes et al., 2008) performed using JMP v. 5.1.2 (SAS Institute) to test the statistical significance ($P \leq 0.05$) of the relationship at a 95% confidence interval (blue line). Inbred lines are: 1, A619; 2, B73; 3, B37; 4, CI.7; 5, C131A; 6, DE3; 7, KUI2007; 8, NC300; 9, SD44; 10, TZI18. [See online article for color version of this figure.]

among the three alleles (A–C) in one-way ANOVA ($F_{2,48} = 13.7$, $P < 0.0001$) and *HYD3* allele explained 36% of total variation (R^2) in level of β-carotene (μg/gm). It is reasonable that *HYD3* allele will explain less of the variation in absolute levels of β-carotene compared to ratio of β-carotene to β-cryptoxanthin. Absolute leaves of β-carotene will be influenced not only by LCYE and HYD3 activities, but also by multiple factors controlling pathway flux (Vallabhaneni and Wurtzel, 2009). In comparison, β-carotene-to-β-cryptoxanthin ratio will be influenced most directly by the HYD3 enzyme activity.

We next compared the means of the three alleles using the Tukey-Kramer honestly significant difference test at a 0.05 level. We observed a significant difference ($P < 0.0001$) between alleles C and B, and C and A for the ratio of β-carotene and β-cryptoxanthin and β-carotene alone. No significant difference was observed between alleles B and A for both the ratio of β-carotene and β-cryptoxanthin and β-carotene alone.

The most favorable *HYD3* C allele was associated with approximately 11-fold higher levels of β-carotene relative to β-cryptoxanthin and 4-fold higher levels of β-carotene (μg/gm). The most likely explanation is that during endosperm development a reduction in *HYD3* transcripts leads to reduced conversion of β-carotene to downstream xanthophylls, causing β-carotene to accumulate. To test this possibility, we selected CI.7 and DE3, lines carrying the C allele, and measured β-carotene at each of three developmental stages and compared to *HYD3* transcript level for triplicate replicates of endosperms collected at 15, 20, and 25 DAP. As shown in Figure 9, levels of β-carotene increased during endosperm development (for equal amounts of extracted carotenoids), as *HYD3* mRNA levels decreased in the same endosperm developmental samples. Therefore, there is a direct link between *HYD3* transcript level and accumulated β-carotene.

Enhancement of provitamin A levels necessitates control of both pathway branching and carotene hydroxylation in addition to elevating pathway flux and minimizing product degradation. Previous development of *LCYE* markers (Harjes et al., 2008) provided the first step in tracking alleles to control pathway branching through traditional nontransgenic breeding selection. However, the *LCYE* markers were insufficient. Even if pathway branching could be controlled through selection of optimal *LCYE* alleles, there remained the problem that the provitamin A carotenes produced in the pathway would automatically be hydroxylated to nonprovitamin A compounds, as typically seen for maize endosperm composition of zeaxanthin or lutein (Kurilich and Juvik, 1999). Other upstream steps may influence carotenoid content (Vallabhaneni and Wurtzel, 2009), but enhancement of provitamin A composition will specifically require the combined use of tools that target both *LCYE* and *HYD3*.

The C allele can be monitored by the simple PCR assay and its presence in genetically diverse lines will facilitate use in different geographical regions. This PCR assay may now be used to test predictability of the *HYD3* allele in controlling β-carotene content in diverse cultivars. With 51 lines analyzed for *HYD3* allele, *LCYE* allele data was available for 48 of those lines (Harjes et al., 2008). However, none carried optimal alleles for both *HYD3* and *LCYE*. Therefore, the *HYD3* assay may be used in combination with the previously described *LCYE* assays (Harjes et al., 2008) to select parental lines containing optimal *HYD3* or *LCYE* alleles and to screen progeny at the seedling stage and identify those that are homozygous for optimal alleles of both genes. Thus it is predicted that the combination of *HYD3* and *LCYE* alleles will lead to higher β-carotene levels in maize endosperm than having optimal alleles of either gene alone.

Many ongoing studies have investigated pathway regulation through transgene expression in one culti-

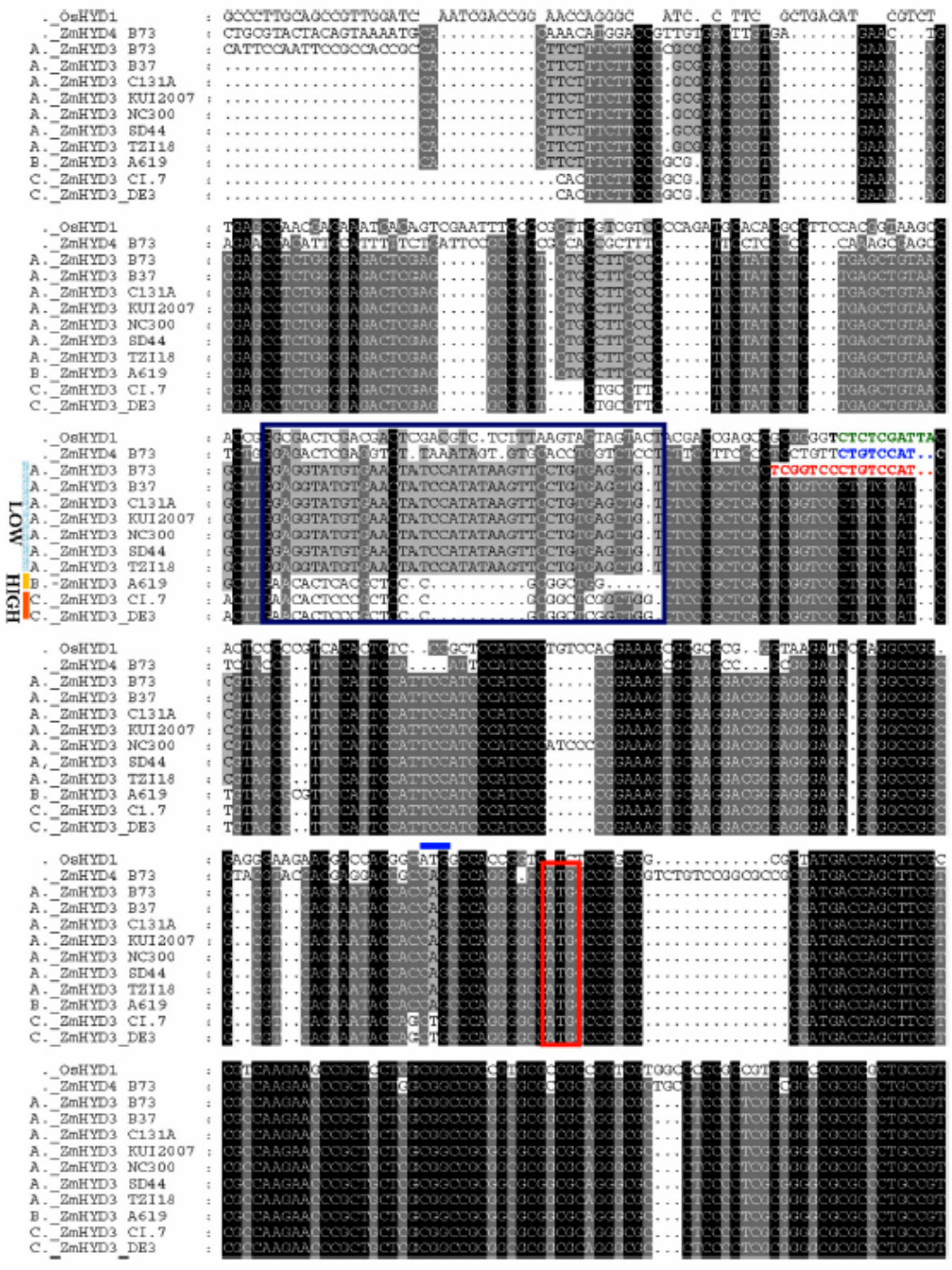


Figure 6. (*Legend appears on following page.*)

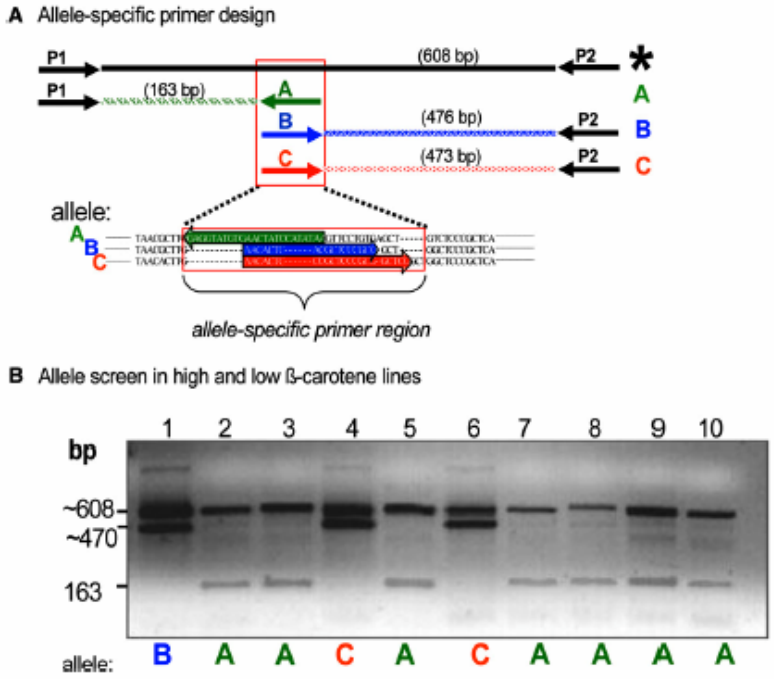


Figure 7. Multiplex PCR diagnosis of *HYD3* alleles in high and low β-carotene lines. The sequence variation in *HYD3* (Fig. 6, blue box) was used to develop a multiplex PCR assay to track *HYD3* alleles (B and C) in the high β-carotene lines and to distinguish these from the A allele in the low β-carotene line. With this simple PCR assay, natural alleles can be tracked using traditional, nontransgenic breeding. The assay contained four primers to (1) amplify a control, *HYD3*-specific product (*), which was conserved in all alleles and encompassed the allele-specific region (Figs. 7A and 6 [blue box]), and (2) to amplify and distinguish specific alleles A from C or A from B. The multiplex PCR assay was next used to genotype all of the 10 inbred lines used in the correlation study (Fig. 7); as predicted, all lines having B or C alleles also had a high β-carotene content (21%–55% of total carotenoids) as compared to maize B73 (A allele) where kernels have only 3% β-carotene. A, Allele-specific primer design. Top, *HYD3* primers (P1 and P2) amplify all alleles, as product (*), the size of which is shown based on the A allele (B and C alleles generate a * product that differs by a few bp); allele-specific primers (A, B, or C) amplify maize allele A from low β-carotene line B73, allele B from high β-carotene line A619, or allele C from high β-carotene lines CI.7 and DE3, respectively. Bottom, The boxed region used to design the allele-specific primers is shown in more detail. Sequences and amplification direction are shown in corresponding colors for the three alleles. PCR product sizes are based on the A allele; products from the other alleles differ slightly. B, Screening of alleles in the maize diversity lines. Primers used were P1/P2/A plus either B (lane 1) or C (lanes 2–10). Inbred lines are: 1, A619; 2, B73; 3, B37; 4, CI.7; 5, C131A; 6, DE3; 7, KUI2007; 8, NC300; 9, SD44; 10, TZI18. [See online article for color version of this figure.]

var (Giuliano et al., 2008); the resulting phenotype is dependent on the genotype used and the resulting data are not predictably transferred to other genotypes, perhaps because of limited understanding of pathway regulation. Here we took advantage of natural variation inherent in a large germplasm collection; by judicious sampling of the collection, we were able to pinpoint a specific gene family member for which favorable alleles were discovered for predicting enhanced β-carotene composition. Metabolite-based sorting of a germplasm collection provided accessibility to a valuable resource that yielded critical knowledge and tools that will be needed for breeding high provitamin A in maize, an important food staple in

Figure 6. Alignment of maize sister paralogs *HYD3* and *HYD4* plus their rice synteny partner *HYD1*. DNA sequences were aligned for *HYD3* and *HYD4*, their rice synteny partner *HYD1*, and the *HYD3* sequences of all of the maize inbreds used in this study. Transcription start sites (colored nucleotides) for B73 *HYD3*, B73 *HYD4*, and rice *HYD1* were estimated by aligning ESTs. The first ATG for *HYD1* is denoted by a blue overline and contained within a red box for the remaining *HYD3* and *HYD4* sequences. A blue box encompasses the variant region that distinguishes the high and low β-carotene lines. Black, dark-gray, and light-gray shading indicate degree of conserved nucleotides where black is the highest (100%) match. [See online article for color version of this figure.]

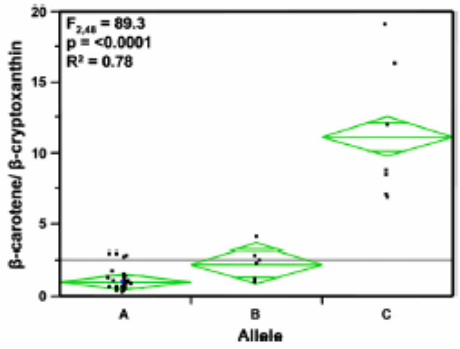


Figure 8. One-way ANOVA between *HYD3* alleles (A, B, C) and the ratio of β-carotene to β-cryptoxanthin in 51 genetically diverse maize lines. The mean ratios of β-carotene to β-cryptoxanthin showed a statistically significant difference among the three *HYD3* alleles in one-way ANOVA ($F_{2,48}$ = 89.3, P < 0.0001), and the *HYD3* allele explained 78% of the variation (R^2) in the β-carotene-to-β-cryptoxanthin ratio. Three different alleles denoted as A, B, and C are plotted on the *x* axis and ratios of β-carotene (μg) to β-cryptoxanthin (μg) are plotted on the *y* axis. Gray line indicates the grand mean of the entire sample. The green diamonds indicate 95% confidence interval of the mean, and the dots represent each inbred line in study. [See online article for color version of this figure.]

vitamin A-deficient Sub-Saharan Africa and Latin America, where vitamin A deficiency is a serious health problem. This approach can be easily adapted to other metabolite targets in other species to rapidly discover pathway regulation and to develop tools for predictive breeding.

MATERIALS AND METHODS

Plant Materials

Maize (*Zea mays*) inbred lines (A619, B73, B37, CI.7, CI31A, DE3, KUI2007, NC300, SD44, and TZI18; Harjes et al., 2008) were field grown in Bronx, NY, and sibling pollinated. Unfertilized ear, embryo (20 DAP), and endosperm (10–30 DAP) were collected from field-grown plants, whereas leaf and root samples were collected from six-leaf-stage seedlings grown in a greenhouse (16-h day at 25°C). Samples were stored at −80°C prior to use.

Sequence Analysis and Chromosome Mapping

Cloning and DNA Sequence Analysis

A maize B73 genomic bacterial artificial chromosome (BAC) library (Gallagher et al., 2004) was probed with the maize *HYD4* cDNA, GenBank AY844956. Eight BAC clones representing three groups were identified and a representative of each group was sequenced by primer walking (DNA Sequencing Facility, University of Chicago Research Center) and the data deposited in GenBank (*HYD1*, EU638325; *HYD2*, EU638326; *HYD3*, AY844957). The *HYD3* cDNA sequence (GenBank no. AY844958) was used as a query to identify additional maize paralogs and orthologs from rice (*Oryza sativa*; www.gramene.org) and *Sorghum bicolor* (www.phytozome.org). Maize methyl-filtered contig sequences and corresponding BAC clones harboring carotene hydroxylase genes (diiron type) were identified (Supplemental Table S1). Gene models were drawn using Genscan (www.genscan.com), Softberry (www.softberry.com), and Vector NTI Suite 9.0 (Invitrogen). Transcription start sites were estimated by EST and promoter analysis (www.softberry.com). Maize ESTs corresponding to CYP97A (DV169913) and CYP97C (BE552887) were found by searching plantGDB maize database using orthologous gene sequences of CYP97A (LOC_Os02g57290) and CYP97C (LOC_Os10g39930) from rice (Quinlan et al., 2007). As found for rice and sorghum, only one maize ortholog was found for each gene encoding these two CYP97 enzymes (Supplemental Table S1).

Mapping and Synteny

Chromosomal positions of maize genes in the maize B73 inbred were mapped either by utilizing tools available at Web AGCoL package (Pampanwar et al., 2005) or MaizeGDB. Orthologous genes from rice Japonica were identified through synteny comparisons with maize (www.tigr.org/tdb/synteny/).

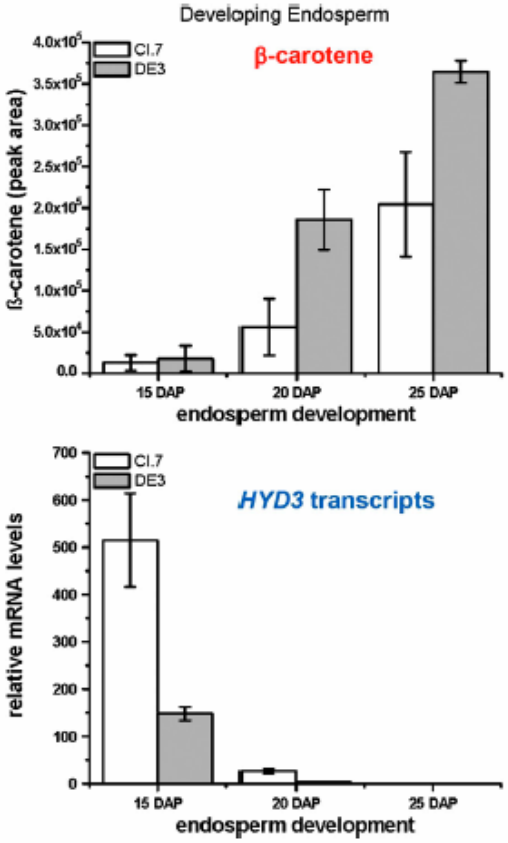


Figure 9. Relative amount of β-carotene increases and *HYD3* transcript levels decrease during development of maize endosperm in lines carrying a C allele of *HYD3*. Total carotenoids were extracted and equal amounts of total carotenoids subjected to UPLC separation and quantification of β-carotene in triplicate. Samples from the same endosperm samples were subjected to quantitative RT-PCR (see Supplemental Table S3). Means of three replicates are shown with SD. [See online article for color version of this figure.]

Phenetic and Protein Analysis

HYD amino acid sequences were aligned using ClustalW and a neighbor-joining tree was constructed with bacterial *crtZ* from *Erwinia herbicola* (GenBank no. AAA64983) as an out group with 500 bootstrap replication support using MEGA3 software (Kumar et al., 2001). Chloroplast transit peptide signal and transmembrane analysis were predicted using ChloroP 1.1 (Emanuelsson et al., 1999) and TMHMM 2.0 (Krogh et al., 2001), respectively.

Transcript and Total Carotenoid Analysis

RNA extraction and quantitative reverse transcription (RT)-PCR were performed using gene-specific primers (Supplemental Table S1) and normalized to actin, as described (Li et al., 2008a). Values are expressed as the mean of three RT-PCR replicates ±SD. Total carotenoids were extracted as previously described (Li et al., 2008b) and the concentration (shown in Fig. 4) was calculated using the Lambert-Beer equation (Schiedt and Liaaen-Jensen, 1995).

Plasmids and Functional Complementation

Escherichia coli BL21 (DE3) cells (Novagen) containing pAC-BETA-04 accumulate β-carotene and were used to test hydroxylase function (Sun et al., 1996). BM382572 (*HYD3*) and BG320875 (*HYD4*) cDNAs were subcloned for functional analysis into the *Eco*RI/*Not*I and *Eco*RI/*Xho*I sites of the pET23c expression vector (Novagen), and renamed as pTHYD3 and pTHYD4, respectively. Transformants carrying the test plasmids or empty vector were grown on selective medium and pigments extracted and analyzed by HPLC as described (Gallagher et al., 2004).

Total Carotenoid Content Measurement

The carotenoid extraction procedure was based on Kurilich and Juvik (1999). Five hundred milligrams of maize endosperm was ground in ethanol and incubated for 6 min (6 mL ethanol, 0.1% butylated hydroxytoluene) at 85°C, followed by 10 min saponification with 120 μL (1 g/mL) KOH. Samples were vortexed, placed on ice, and 4 mL cold distilled water added. Three milliliters of 2:1 petroleum ether:diethyl ether (PE:DE) (v/v) were added to each sample, vortexed, and centrifuged for 10 min at 3,500 rpm. The upper layer was retrieved and the separation was repeated twice with 3 mL of 2:1 PE:DE (v/v). The combined fractions were made up to 10 mL, and carotenoids were measured spectrophotometrically at OD 450 nm using a Lambda UV/VIS spectrophotometer (Perkin Elmer Life Sciences), and the concentration of carotenoids were calculated using the Lambert-Beer equation (Schiedt and Liaaen-Jensen, 1995).

Measurement of β-Carotene in Developing Endosperm of High β-Carotene Lines

The carotenoid extraction procedure was based on Kurilich and Juvik (1999). Five hundred micrograms of maize endosperm was ground in ethanol and incubated for 6 min (6 mL ethanol, 0.1% butylated hydroxytoluene) at 85°C, followed by 10 min saponification with 120 μL (1 g/mL) KOH. Samples were vortexed, placed on ice, and 4 mL cold distilled water added. Three milliliters of 2:1 PE:DE (v/v) were added to each sample, vortexed, and centrifuged for 10 min at 3,500 rpm. The upper layer was retrieved and transferred to a tube with a known weight. The separation was repeated twice with 3 mL of 2:1 PE:DE (v/v) and upper layers combined and dried under nitrogen gas. Total weight of the tube was measured to obtain the net weight of dried extract in each tube. All extractions and measurements were performed in triplicates, and finally dried extract was redissolved appropriately in HPLC grade isopropanol to maintain equal concentration (100 μg/μL) in all samples. Four microliters of each sample was injected onto an Acquity ultra performance LC (UPLC) BEH C18 1.7 μm 2.1 × 100 mm (part no. 186002352) column kept at 65°C and attached to a Waters Acquity UPLC with a photodiode array detector. Separation was carried out using a linear gradient from 100% acetonitrile to 80% acetonitrile/20% isopropanol over 6 min, followed by 100% acetonitrile at 6.5 min. The solvent flow rate was 0.3 mL/min. Data were collected for 7 min total and analyzed with Empower Pro software to determine β-carotene peak area for each sample. All samples were run in triplicates and SD calculated.

Statistical Analyses

Pearson correlation analysis of transcript and carotenoid composition from 10-line subset of maize inbreds was performed using JMP v. 5.1.2 (SAS Institute) to test the statistical significance ($P \leq 0.05$) of the relationship. One-way ANOVA and Tukey-Kramer honestly significant difference test for comparisons between the alleles (A–C) and ratio of β-carotene to β-cryptoxanthin or absolute levels of β-carotene in 51 lines was performed using JMP 5.1 (SAS institute Inc.).

Multiplex PCR Assay for Tracking *HYD3* Alleles

A multiplex PCR assay based on *HYD3* promoter variation (A, B, or C) was developed to distinguish between *HYD3* allele A (in the low-carotene B73 inbred) and the *HYD3* alleles B and C found in the high-carotene inbreds A619 (allele B), CI.7 (allele C), and DE3 (allele C). The assay involved amplification of one *HYD3* paralog-specific control PCR product produced by all alleles and a second product that distinguished allele A from C or allele A from B, in both homozygous and heterozygous samples. The PCR assay utilized four primers: (1) two external primers P1 (no. 1,595, GACTTGTGAGCAAGGGGAAG) and P2 (no. 1,592, GACGTGACTCCGAGGCTAGA) for amplification of a control product that was conserved in all alleles; and (2) two internal primers added for amplification of the *HYD3* specific alleles. To track alleles C and A, the specific allele forward primer was C (no. 2,111, AACACTCCCGCTCCCGCCGCTCG, allele C), and the reverse primer was A (no. 2,116, TTATATGGATAGTTCACATACCTC, allele A). Alternatively, to distinguish allele B from A, the forward primer used was primer B (no. 2,109, AACACTCACGCTCCCGCG, allele B) together with reverse primer A. The 50-μL PCR reaction contained: 1× PCRx amplification buffer (Invitrogen), 0.2 mM deoxyribonucleotide triphosphates (Invitrogen), 1.5 mM $MgSO_4$ (Invitrogen), 2.5 units iTaq DNA polymerase (BIO-RAD), 0.6 μm each of two external primers; 0.4 μm each of two internal primers, 0.5 μg genomic DNA template, 3× final concentration of PCRx enhancer solution (Invitrogen). Thermal cycling conditions were: one cycle (94°C for 3 min) followed by 40 cycles (94°C for 45 s, 54°C for 45 s, and 72°C for 1 min) and one cycle (72°C for 10 min). Amplified products were separated on 1% agarose gels and validated by sequencing. Maize inbred sequences used for developing the multiplex primers were: FJ228406, FJ228407, FJ228408, FJ228409, FJ228410, FJ228411, FJ228412, FJ228413, FJ228414, and FJ228415.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers AY844956, AY844957, AY844958, EU638325, EU638326, FJ228406, FJ228407, FJ228408, FJ228409, FJ228410, FJ228411, FJ228412, FJ228413, FJ228414, and FJ228415.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Conserved motifs in HYD proteins of maize and rice.

Supplemental Figure S2. Functional complementation of *HYD* genes.

Supplemental Figure S3. Alignment of maize *HYD3* in 51 lines.

Supplemental Table S1. Primer and gene accession numbers.

Supplemental Table S2. Predicted structural characteristics of HYD proteins.

Supplemental Table S3. Transcript profiling of *HYD* genes in maize diversity lines and Pearson correlation values.

Supplemental Table S4. Genotyping of *HYD3* alleles and comparison with carotenoid composition of diverse maize germplasm accessions.

ACKNOWLEDGMENTS

We thank Dr. Dwight Kincaid (Lehman College, City University of New York, New York) for advice on statistical analysis and Wurtzel lab members (Dr. Louis Bradbury, Dr. Faqiang Li, Dr. Maria Shumskaya, and Oren Tzfadia)

for helpful advice. Maize lines were obtained from the National Germplasm Collection, except for the initial 10-line subset (Li et al., 2008b).

Received July 22, 2009; accepted September 8, 2009; published September 18, 2009.

LITERATURE CITED

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

Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, Mathers C, Rivera J, Maternal Child Undernutrition Study Group (2008) Maternal and child undernutrition: global and regional exposures and health consequences. Lancet **371:** 243–260

Chander S, Guo YQ, Yang XH, Zhang J, Lu XQ, Yan JB, Song TM, Rocheford TR, Li JS (2007) Using molecular markers to identify two major loci controlling carotenoid contents in maize grain. Theor Appl Genet **116:** 223–233

Chandler VL, Brendel V (2002) The Maize Genome Sequencing Project. Plant Physiol **130:** 1594–1597

Diretto G, Tavazza R, Welsch R, Pizzichini D, Mourgues F, Papacchioli V, Beyer P, Giuliano G (2006) Metabolic engineering of potato tuber carotenoids through tuber-specific silencing of lycopene epsilon cyclase. BMC Plant Biol **6:** 13

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

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci **8:** 978–984

Gallagher CE, Matthews PD, Li F, Wurtzel ET (2004) Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses (Poaceae). Plant Physiol **135:** 1776–1783

Giuliano G, Tavazza R, Diretto G, Beyer P, Taylor MA (2008) Metabolic engineering of carotenoid biosynthesis in plants. Trends Biotechnol **26:** 139–145

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, et al (2008) Natural genetic variation in *lycopene epsilon cyclase* tapped for maize biofortification. Science **319:** 330–333

Islam SN (2004) Survey of carotenoid variation and quantitative trait loci mapping for carotenoid and tocopherol variation in maize. Master's thesis. University of Illinois at Urbana-Champaign, Urbana-Champaign

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol **305:** 567–580

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:** 1244–1245

Kurilich A, Juvik J (1999) Quantification of carotenoid and tocopherol antioxidants in *Zea mays*. J Agric Food Chem **47:** 1948–1955

Li F, Murillo C, Wurtzel ET (2007) Maize Y9 encodes a product essential for 15-cis zetacarotene isomerization. Plant Physiol **144:** 1181–1189

Li F, Tzfadia O, Wurtzel ET (2009) The *Phytoene Synthase* gene family in the grasses: subfunctionalization provides tissue-specific control of carotenogenesis. Plant Signal Behav **4:** 208–211

Li F, Vallabhaneni R, Wurtzel ET (2008a) *PSY3*, a new member of the phytoene synthase gene family conserved in the Poaceae and regulator of abiotic-stress-induced root carotenogenesis. Plant Physiol **146:** 1333–1345

Li F, Vallabhaneni R, Yu J, Rocheford T, Wurtzel ET (2008b) The maize phytoene synthase gene family: overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress-tolerance. Plant Physiol **147:** 1334–1346

Li ZH, Matthews PD, Burr B, Wurtzel ET (1996) Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. Plant Mol Biol **30:** 269–279

Liu K, Goodman M, Muse S, Smith JS, Buckler E, Doebley J (2003) Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. Genetics **165:** 2117–2128

Matthews PD, Luo R, Wurtzel ET (2003) Maize phytoene desaturase and zetacarotene desaturase catalyze a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. J Exp Bot **54:** 2215–2230

Matthews PD, Wurtzel ET (2007) Biotechnology of food colorant production. *In* C Socaciu, ed, Food Colorants: Chemical and Functional Properties. CRC Press, Boca Raton, FL, pp 347–398

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell **15:** 1795–1806

Pampanwar V, Engler F, Hatfield J, Blundy S, Gupta G, Soderlund C (2005) FPC web tools for rice, maize, and distribution. Plant Physiol **138:** 116–126

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor Appl Genet **114:** 525–537

Quinlan R, Jaradat T, Wurtzel ET (2007) *Escherichia coli* as a platform for functional expression of plant P450 carotene hydroxylases. Arch Biochem Biophys **458:** 146–157

Randolph LF, Hand DB (1940) Relation between carotenoid content and number of genes per cell in diploid and tetraploid corn. J Agric Res **60:** 51–64

Schiedt K, Liaaen-Jensen S (1995) Chapter 5: isolation and analysis. *In* G Britton, S Liaaen-Jensen, H Pfander, eds, Carotenoids—Isolation and Analysis, Vol 1A. Birkhäuser, Basel, pp 81–108

Sun Z, Gantt E, Cunningham JFX (1996) Cloning and functional analysis of the β-carotene hydroxylase of *Arabidopsis thaliana*. J Biol Chem **271:** 24349–24352

Underwood BA (2004) Vitamin A deficiency disorders: international efforts to control a preventable "pox". J Nutr **134:** 231S–236S

Vallabhaneni R, Wurtzel ET (2009) Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. Plant Physiol **150:** 562–572

Wong JC, Lambert RJ, Wurtzel ET, Rocheford TR (2004) QTL and candidate genes phytoene synthase and zetacarotene desaturase associated with the accumulation of carotenoids in maize. Theor Appl Genet **108:** 349–359

World Health Organization (1995) Global Prevalence of Vitamin A Deficiency. Micronutrient Deficiency Information System. Working Paper No. 2. WHO/NUT/95.3. World Health Organization, Geneva

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

Appendix 3: Plasmids used in this study and laboratory clone records

Clone Name	**Wurtzel Database #**	**Brief Description**	**Insert Size (bp)**	**Accession # / Reference**	**Chapter**
pACCRT-EIB	8	*Erwinia* crt cassette (crtE, crtI, crtB).	3.8Kb	(Cunningham et al., 1993)	2,3
y2	399	*Arabidopsis* LCYE	1575	U50738/ (Cunningham et al., 1996)	2,3
pACBETA-04	17	*E. herbicola crtEIB* + *Arabidopsis* LCYB + ipp isomerase	4.2Kb	(Sun et al., 1996)	2
pRT-A4	438	Rice P450 β-hydroxylase	1932	AK068163/ (Quinlan et al., 2007)	2,3
pTR-C2	392	Rice P450 epsilon-hydroxylase	1686	AK065689/ (Quinlan et al., 2007)	2,3
pGEMT-C2	664	Rice P450 epsilon-hydroxylase	1725	AK065689	3
pRQ-C2	432	Rice P450 epsilon-hydroxylase	1686	AK065689	3
pRQ-H3	644	Maize diiron β-hydroxylase 3	960	BM382572	3
pRQ_H4	645	Maize diiron β-hydroxylase 4	930	BG320875	3
pTHYD3	203	Maize diiron β-hydroxylase 3	1.2Kb	BM382575/ (Vallabhaneni et al., 2009)	3
pTHYD4	299	Maize diiron β-hydroxylase 4	1.38Kb	BG320875/ (Vallabhaneni	3

				et al., 2009)	
pACBETA-At	437	*E. herbicola crtEIB* + *Arabidopsis* LCYB	11.3Kb	(Cunningham et al., 2007)	3
pTnT-A4	433	Rice P450 β-hydroxylase	1932	AK068163	3
pTnT-C2	434	Rice P450 epsilon-hydroxylase	1725	AK065689	3
pTnT-H3	435	Maize diiron β-hydroxylase 3	960	BM382572	3
pTnT-H4	436	Maize diiron β-hydroxylase 4	930	BG320875	3
prLHCP	293	LHCP (integral thylakoid membrane protein)	1993	(Tan et al., 2001)	3
16/EGFP	295	tpsOE16 ::GFP (chloroplast lumen protein)	N/A	(Marques et al., 2003)	3
A4_GFP	640	Rice P450 β-hydroxylase	1932	AK068163	3
C2_GFP	641	Rice P450 epsilon-hydroxylase	1686	AK065689	3
H3_GFP	642	Maize diiron β-hydroxylase 3	960	BM382572	3
H4_GFP	643	Maize diiron β-hydroxylase 4	930	BG320875	3
pUC35S-Toc34-sGFP-Nos	569	34 kDa outer membrane transport complex protein	933	(Chen and Schnell, 1997)	3
pUC35S-LHCP-sGFP-Nos	567	Major chlorophyll a/b-binding polypeptide	800	(Tan et al., 2001)	3
p-SAT-M-PGL-RFP	564	Plastoglobulin PGL-2	957	BT039786_1	3
A4_2236	486	Rice P450 β-hydroxylase	1932	AK068163	3
C2_2236	574	Rice P450 epsilon-hydroxylase	1686	AK065689	3
H3_2236	489	Maize diiron β-hydroxylase 3	960	BM382572	3
H4_2236	488	Maize diiron β-hydroxylase 4	930	BG320875	3
A4_1476	573	Rice P450 β-hydroxylase	1932	AK068163	3
C2_1476	487	Rice P450 epsilon-hydroxylase	1686	AK065689	3
H3_1476	575	Maize diiron β-hydroxylase 3	960	BM382572	3
H4_1476	576	Maize diiron β-hydroxylase 4	930	BG320875	3

ChrD_1476	662	Cucumber ChrD protein	564	DQ117526/ (Libal-Weksler et al., 1997)	3
ChrD_2236	663	Cucumber ChrD protein	564	DQ117526/ (Libal-Weksler et al., 1997)	3
ChrD_2485	490	Cucumber ChrD protein	564	DQ117526/ (Libal-Weksler et al., 1997)	3
ChrD_3038	491	Cucumber ChrD protein	564	DQ117526/ (Libal-Weksler et al., 1997)	3
pCOLADuet-1	395	Expression vector (Novagen)	3719		2,3
pCDFDuet-1	431	Expression vector (Novagen)	3781		3
pTNT	340	Expression vector (for *in vitro* transcription/translation) (Promega)	2871	AF479322	3
pUC35S-sGFP-Nos	326	Expression vector	4500	(Okada et al., 2000)	3
pSAT-1476 (pSAT6-cEYFP –N1)	572	Expression vector	4378	(Citovsky et al., 2006)	3
pSAT-2236 (pSAT4(A)-nEYFP –N1)	423	Expression vector	4378	DQ169003/ (Citovsky et al., 2006)	3

WURTZEL LAB
CLONE INFORMATION

Date Today: 2 / 10 / 07 (MONTH) / (DAY)/ (YEAR) Entered into database yes: 8 / no

CLONE NAME: pACCRT-EIB, updated RQ

Lab Clone Number/Name: _pACCRT-EIB_____Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _____

Clone Description: Confers accumulation of lycopene. Contains *Erwinia* cassette genes crtE, crtI, crtB

Constructed by: _____Date rec'd_____
Purified by: _____
DNA Location (-20°C) Box Number: ___6__Position: ____E8__Conc. _0.5 µg/µl
Tube labeled as: _____
Strain Location (-80°C) Box Number: _____15_____Position___C7,C8___
Tube labeled as: _pACCRT-EIB_____ Strain: BL21(DE3)_____

Cited in journal: Cunningham Jr. FX et al., (1993) FEBS Lett 328:130-138

Lab Notebook to reference: ___Book 5_____Date: ___2-2-07_____

Original clone name (if different): _____

Organism source of gene: _Erwinia uredovora___Variety_____

Cloning vector used: __pAC184__Vector size: 4.26Kb___Insert size: 3.8Kb _____

Antibiotic markers: _____amp; _____tet; ___x___chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _____

Sequence verified: yes___x__no_____ Junction verified: yes__x____no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: __Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today: 08 /15 / 07 (MONTH) / (DAY)/ (YEAR) Entered into database yes: 399 / no _

CLONE NAME: y2 (Arabidopsis LCYe cDNA), updated RQ

Lab Clone Number/Name: __y2_____Alternate name(s): pAtLCYe__
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: U50738___

Clone Description: y2 carries the *Arabidopsis* LCYe gene (epsilon-cyclase) which confers the accumulation of delta-carotene.
Refer to Cunningham paper: Plant Cell 8(9): 1613-1626, 1996
Also refer to Kieber paper: Cell 72: 427-441. 1993. Kieber paper describes which enzymes release LCYe insert

Constructed by: Dr. Cunningham_____Date rec'd_2/1/05__
Purified by: _____
DNA Location (-20°C) Box Number: __6__Position: _E3_____Conc. _60 ng/µl
Tube labeled as: _y2_____
Strain Location (-80°C) Box Number: ___16_____Position___E6, E7_____
Tube labeled as: y2_____ Strain: _DH5α_____

Cited in journal: Plant Cell 8(9): 1613-1626, 1996 and Cell 72: 427-441

Lab Notebook to reference: ____Rena (Book 5)___ Date: ___6/21/07_____
Original clone name (if different): _____

Organism source of gene: _Arabidopsis thaliana_____Variety_____

Cloning vector used: pBluescript SK-__Vector size: _2960 bp Insert size: _1575 bp___

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other

Restriction enzyme(s) to release insert: _EcoRV/NotI_____

Sequence verified: yes_x____no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No___
Hybridization results attached: Yes____ No___
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: _Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today:_08_/_21_/97 _ Entered into database yes: 17_ / no __
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pAC-BETA-04, updated RQ

Lab Clone Number/Name: __pAC-BETA-04_____Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression___x_____ Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _____

Clone Description:
Accumulates β-carotene; has ipp isomerase to give deeper color.

Constructed by: __Dr. Cunningham_____Date rec'd_____
Purified by: _____
DNA Location (-20°C) Box Number: _6_____Position: A9__Conc. _32 ng/µl
Tube labeled as: _pAC-BETA-04_____
Strain Location (-80°C) Box Number: __15__Position__A4_____
Tube labeled as: __pAC-BETA-04_____ Strain: __BL21 (DE3)_____

Cited in journal: J. Biol. Chem 271: 24349-24352, 1996

Lab Notebook to reference: _____Date: _____

Original clone name (if different): _____

Organism source of gene: Erwinia uredovora and H. pluvialis (IPP) Variety__

Cloning vector used: _pACYC 184_____Vector size: __4.2 Kb__Insert size: _____

Antibiotic markers: _____amp; _____tet; __x ___chloramphenicol;____other (_____)

Restriction enzyme(s) to release insert: ___Hind III_____

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____

Other: A cDNA encoding an isopentenyl pyrophosphate (IPP) isomerase in the green alga *Haematococcus pluvialis* was excised from the library cloning vector (pBluescript SK +) with BamHI and KpnI sites of the pTrcHis A (invitrogen). Digestion of the resulting plasmid with EcoRV and KpnI produces a fragment of 2.0 Kb containing the *H. pluvialis* IPP isomerase downstream of the strong bacterial trc promoter. This fragment was cloned in the blunted HindIII site of pAC-BETA to produce pAC-BETA-04.

Depositor's Name: __Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today: 5/12/09 Entered into database yes_438_ / no ___
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pRT-A4 (rice P450 beta-hydroxylase (CYP97A4))

Lab Clone Number/Name: pRT-A4 Alternate name(s): _P450 HYDB___
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK068163

Clone Description: Two primers (forward1018 and reverse 1019) were used to amplify the CYP97A4 (AK068163) ORF. EcoRI and XhoI restriction sites were added to primers 1018 and 1019, respectively. The amplified CYP97A4 ORF and pColaDuet vector were cut with both restriction enzymes (EcoRI and XhoI). Resulting fragments were ligated and co-transformed with pAC BETA-04 into BL21 (DE3) cells for expression.
Primer 1018: 5′-CCG GAATTC ATGAGCTCAGCGACGTCAGTGAGT-3′
EcoRI site
Primer 1019: 5′-ACCG CTCGAG TCAGATTCGAGTTGCTGAGACTTG-3′
XhoI site

Constructed by: Tahhan Jaradat and Rena Quinlan_____Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: _6__Position: _H1____Conc. _0.2µg/µl
Tube labeled as: __pRT-A4_____
Strain Location (-80°C) Box Number: __18_ Position__E6_____
Tube labeled as: __pRT-A4_____ Strain: _BL21 (DE3)___

Cited in journal:

Lab Notebook to reference: _Rena-Book 3__ Date: _3/29/05_____

Organism source of gene: __Oryza sativa_____Variety:_Nipponbare_____

Cloning vector used: pCOLADuet-1_____Vector size: _3719bp Insert size: 1932bp

Antibiotic markers: _____amp; _____tet; ___X___chloramphenicol;___x__other (Kan)

Restriction enzyme(s) to release insert: _EcoRI/XhoI_____

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan____

WURTZEL LAB
CLONE INFORMATION

Date Today: 11/22/05 Entered into database yes: 392 / no___
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pTR-C2 (rice P450 epsilon-hydroxylase (CYP97C2))

Lab Clone Number/Name: pTR-C2_____Alternate name(s): ___P450 HYDE
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK065689

Clone Description:
Two primers (forward 1024 and reverse 1017) were used to amplify the CYP97C2 (AK065689) cds. EcoRI and XhoI restriction sites were added to primers 1024 and 1017, respectively. The amplified AK065689 cds and pColaDuet vector were cut with both restriction enzymes (EcoRI and XhoI). Resulting fragments were ligated and co-transformed with pACCRT-EIB (accumulates lycopene) and Y2 (accumulates delta-carotene) into BL21 (DE3) cells for expression.
Primer 1024: 5′- CCG GAA TTC CCG TCC CGT GCG TAC CAT TC -3′
EcoRI site
Primer 1017: 5′- ACCG CTCGAG TC ATC TGG ACC CAC TGA GTG CA -3′
XhoI site

Constructed by: Tahhan Jaradat and Rena Quinlan_____Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: _6__Position: _B2_____Conc. _0.2µg/µl
Tube labeled as: __pTR-C2_____
Strain Location (-80°C) Box Number: _15_ Position:____B6_____
Tube labeled as: __pTR-C2_____Strain: __BL21 (DE3)_

Lab Notebook to reference: _Rena-Book 3__ Date: _3/29/05_____

Original clone name (if different): _____

Organism source of gene: __Oryza sativa_____Variety:_Nipponbare_____

Cloning vector used: pCOLADuet-1_____Vector size: _3719bp Insert size: 1686bp

Antibiotic markers: _____amp; _____tet; _____chloramphenicol;_____other (Kan)

Restriction enzyme(s) to release insert: _EcoRI/XhoI_____

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan____

WURTZEL LAB
CLONE INFORMATION

Date Today: 08/22/08 Entered into database yes: 664 / no ____
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pGEMT-C2 (rice P450 epsilon-hydroxylase (CYP97C2))

Lab Clone Number/Name: pGEMT-C2 Alternate name(s): P450 HYDE
Clone type: Genomic: _____
cDNA: Expression_____Non-expression_____x_____

New GenBank Accession: _____Original GenBank Accession: AK065689

Clone Description: Two primers (forward 2138 and reverse 2139) were used to amplify the CYP97C2 (AK065689) entire ORF plus 39bp of 5' UTR from rice cDNA, and the amplified product was ligated into the pGEMT-Easy (Promega) vector. The resulting construct was then transformed into TOP10F' cells.

Primer 2138: AAT CCA TCT CGA ATC CCT AGC
Primer 2139: TCA TCT GGA CCC ACT GAG TG

Constructed by: Rena Quinlan Date rec'd_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: 7 Position: I6 Conc. 0.15µg/µl
Tube labeled as: pGEMT-C2
Strain Location (-80°C) Box Number: 17 Position: H1
Tube labeled as: pGEMT-C2 Strain: TOP10F'

Cited in journal:

Lab Notebook to reference: Rena-Book 7 Date: 8/03/08

Original clone name (if different): _____

Organism source of gene: Oryza sativa Variety: Nipponbare

Cloning vector used: pGEMT-Easy Vector size: 3015bp Insert size: 1725bp

Antibiotic markers: x amp; _____tet; _____chloramphenicol;_____other

Restriction enzyme(s) to release insert: _____

Sequence verified: yes x no_____ Junction verified: yes x no_____

GenBank sheet attached: Yes_____No x
Hybridization results attached: Yes____ No x
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: 5/12/09 Entered into database yes_432_ / no ____
(MONTH)/(DAY)/(YEAR)

CLONE NAME: pRQ_C2 (rice P450 epsilon hydroxylase (HYDE))

Lab Clone Number/Name: pRQ-C2 Alternate name(s): CYP97C2
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK065689

Clone Description:
Two primers (forward 2370 and reverse 2371) were used to amplify the CYP97C2 ORF (minus 21bp of CTP) from rice cDNA. NdeI and Acc65I restriction sites were added to primers 2370 and 2371, respectively. The amplified CYP97C2 ORF (minus 21bp of CTP) and pCDFDuet vector were cut with restriction enzymes (NdeI and Acc65I). Resulting fragments were ligated and co-transformed with pRT-A4 (contains rice P450 HYDB) pAC-BETA-At (accumulates β-carotene) and y2 (accumulates delta-carotene) into BL21 (DE3) cells for expression.
Primer 2370: 5′- ACCG CAT ATG GCC GTC CCG TGC GTA C – 3′
NdeI
Primer 2371: 5′- GAGA GGT ACC TCA TCT GGA CCC ACT GAG- 3′
Acc65I

Constructed by: Rena Quinlan Date rec′d_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: _6_ Position: _G6_____Conc. _0.37µg/µl
Tube labeled as: __pRQ-C2
Strain Location (-80°C) Box Number: _18_ Position:____E1_____
Tube labeled as: __pRQ-C2____Strain: __BL21 (DE3)_

Lab Notebook to reference: _Rena-Book 8__ Date: _4-8-09_____

Organism source of gene: __Oryza sativa_____Variety:_Nipponbare_____

Cloning vector used: pCDFDuet-1_____Vector size: _3781bp Insert size: 1665bp

Antibiotic markers: ____amp; _____tet; ____chloramphenicol;___X___other (Streptomycin)

Restriction enzyme(s) to release insert: _NdeI/Acc65I_____

Sequence verified: yes__x _no_____ Junction verified: yes__x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan____

WURTZEL LAB
CLONE INFORMATION

Date Today: 05/20/08 Entered into database yes: 644 / no ____
(MONTH)/(DAY)/(YEAR)

CLONE NAME: pRQ_H3 (maize diiron beta-hydroxylase 3)

Lab Clone Number/Name: pRQ_H3 Alternate name(s):_____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BM382572

Clone Description:
Two primers (forward 1940 and reverse 1941) were used to amplify the BM382572 cds. EcoRI and HindIII restriction sites were added to primers 1940 and 1941, respectively. The amplified BM382572 cds and pCOLADuet vector were cut with both restriction enzymes (EcoRI and HindIII). Resulting fragments were ligated and transformed into BL21 (DE3) cells.
(1940) Forward: GAGA G AAT TCC ATG GCC GCC GCG ATG A
EcoRI
(1941) Reverse: ACCG AAG CTT CTA GAA CTC ATT TGG CAC A
HindIII

Constructed by: Rena Quinlan Date rec'd_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: 9 Position: J5 Conc. 0.08µg/µl
Tube labeled as: pRQ_H3
Strain Location (-80°C) Box Number: 20 Position: E10
Tube labeled as: pRQ_H3 Strain: BL21 (DE3)

Lab Notebook to reference: Rena-Book 6 Date: 04/18/08

Original clone name (if different): _____

Organism source of gene: Zea Mays Variety: B73

Cloning vector used: pCOLADuet-1 Vector size: 3719bp Insert size: 960bp

Antibiotic markers: _____amp; _____tet; _____chloramphenicol;_____other (Kan)

Restriction enzyme(s) to release insert: EcoRI/HindIII

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: 05/20/08 Entered into database yes: 645 / no ____
(MONTH)/(DAY)/(YEAR)

CLONE NAME: pRQ_H4 (maize diiron beta-hydroxylase 4)

Lab Clone Number/Name: pRQ_H4 Alternate name(s):_____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BG320875

Clone Description:
Two primers (forward 1932 and reverse 1933) were used to amplify the BM382572 cds. EcoRI and HindIII restriction sites were added to primers 1932 and 1933, respectively. The amplified BG320875 cds and pCOLADuet vector were cut with both restriction enzymes (EcoRI and HindIII). Resulting fragments were ligated and transformed into BL21 (DE3) cells.
(1932) Forward: 5′- GAGA G AAT TCA ATG GCC GCC GGT CTG T-3′
EcoRI
(1933) Reverse: 5-′ ACCG AAG CTT TCA GAT GGT CCG GCC G -3′
HindIII

Constructed by: Rena Quinlan Date rec′d_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: 9 Position: J6 Conc. 0.05µg/µl
Tube labeled as: pRQ_H4
Strain Location (-80°C) Box Number: 20 Position: F1
Tube labeled as: pRQ_H4 Strain: BL21 (DE3)

Lab Notebook to reference: Rena-Book 6 Date: 04/18/08

Organism source of gene: Zea Mays Variety: B73

Cloning vector used: pCOLADuet-1 Vector size: 3719bp Insert size: 930bp

Antibiotic markers: ___amp; _____tet; _____chloramphenicol;_____other (Kan)

Restriction enzyme(s) to release insert: EcoRI/HindIII

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence

Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: 05/08/2008 Entered into database : 203

Clone Number/Name: **pTHYD3**

Lab Clone Number/Name: __pTHYD3_____Alternate name(s): Carotene Hydroxylase 3
Clone type: Genomic: _____
cDNA: Expression_____**x**_____Non-expression_____

New GenBank Accession: _____AY844958_____Original GenBank Accession: _BM382575___

Clone Description:
The maize HYD3 cDNA, dervived from the BM382575 plasmid, was excised by restriction digestion with EcoR*I* and Not*1* and cloned into the corresponding sites of the pET23c. This construct complements the pAC-Beta-04 successfully.

Constructed by: C.E.Gallagher
Purified by: C.E.Gallagher
Storage Location (-20°C) Box Number: 6 Position: A2 Conc. 1.0_ug/ul
Tube labeled as: pTHYD3
Storage Location (-80°C) Box Number: 14 Position: A2
Tube labeled as: pTHYD3 Strain: E.*coli* Nova Blue

Cited in journal:

Lab Notebook to reference: C.E.Gallagher, Note book 2 Page 140
Original clone name (if different): _____

Organism source of gene: Maize B73

Cloning vector used: pET23C Vector size: 3.6 kb Insert size: 1.2 kb

Antibiotic markers: amp

Restriction enzyme(s) to release insert: *EcoRI/Not1*

Sequence verified: yes Junction verified: yes

GenBank sheet attached: Yes_____No__x___
Hybridization results attached: Yes____ No__x__
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

WURTZEL LAB
CLONE INFORMATION

Date Today: 05/08/2008 Entered into database 299

Clone Number/Name: **pTHYD4**

Lab Clone Number/Name: __pTHYD4_____Alternate name(s): Carotene Hydroxylase 4
Clone type: Genomic: _____
cDNA: Expression_____**x**_____Non-expression_____

New GenBank Accession: _____AY844956_____Original GenBank Accession: _BG320875___

Clone Description:
The maize HYD4 cDNA, dervived from the BG320875 plasmid, was excised by restriction digestion with EcoRI and XhoI and cloned into the corresponding sites of the pET23c. This construct complements the pAC-Beta-04 successfully.

Constructed by: Nicholas Licciardello
Purified by: Nicholas Licciardello
Storage Location (-20°C) Box Number: 6 Position: A1 Conc. 1.0_ug/ul
Tube labeled as: pTHYD4
Storage Location (-80°C) Box Number: 14 Position: A1
Tube labeled as: pTHYD4 Strain: E.*coli* Nova Blue

Cited in journal:

Lab Notebook to reference: Nicholas Licciardello note book#3 June 4, 2004
Original clone name (if different): _____

Organism source of gene: Maize CO328

Cloning vector used: pET23C Vector size: 3.6 kb Insert size: 1.38 kb

Antibiotic markers: amp

Restriction enzyme(s) to release insert: *EcoRI/XhoI*

Sequence verified: yes Junction verified: yes

GenBank sheet attached: Yes_____No__x___
Hybridization results attached: Yes____ No__x__
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

WURTZEL LAB
CLONE INFORMATION

Date Today:_08_/_15_/07 _ Entered into database yes: 437 / no __
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pAC-BETA-At

Lab Clone Number/Name: __pAC-BETA-At_____Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression___x_____ Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _____

Clone Description:
Accumulates β-carotene using the Erwinia herbicola carotenoid pathway cassette genes (except the lycopene β-cyclase is from Arabidopsis). A 1.8 kb NotI-NotI fragment containing LCYb (lycopene β-cyclase) of A. thaliana was excised from pAtLCYbSK, blunted, and inserted in HindIII-digested, Klenow-blunted pAC-LYC (Refer to Cunningham paper: Eukaryot Cell **6:** 533-545, 2007. Also refer to Cunningham paper: The Plant Cell **8:** 1613-1626 for reference to crt cassette from *E. herbicola*

Constructed by: __Dr. Cunningham_____Date rec'd__6/21/07_____
Purified by: ___Rena Quinlan_____
DNA Location (-20°C) Box Number: ___6___Position: E2__Conc. _0.15 ng/µl
Tube labeled as: ___pAC-BETA-At_____
Strain Location (-80°C) Box Number: __16__Position__F5,F6_____
Tube labeled as: __pAC-BETA-At_____Strain: __BL21 (DE3)___

Cited in journal: Eukaryot Cell **6:** 533-545, 2007

Lab Notebook to reference: __Rena (Book 5)_____Date: ___6/21/07_____

Original clone name (if different): _____

Organism source of gene: *Erwinia herbicola* and *Arabidopsis thaliana* Variety__

Cloning vector used: _pACYC 184_____Vector size: _11.3 Kb__Insert size: _LCYB CDS=1506bp

Antibiotic markers: _____amp; _____tet; __x ___chloramphenicol;____other (_____)

Restriction enzyme(s) to release insert: _Cannot release insert (it was Klenow-blunted) – refer to Cunningham paper: Eukaryot Cell 6: 533-545, 2007_____

Sequence verified: yes__x___no_____ Junction verified: yes_____no_____

GenBank sheet attached: Yes_____No____
Hybridization results attached: Yes____ No___
Other:

Depositor's Name: _Rena Quinlan____

WURTZEL LAB
CLONE INFORMATION

Date Today: 5/11/09 Entered into database yes: 433 / no __
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pTnT-A4 (rice P450 beta-hydroxylase (CYP97A4))

Lab Clone Number/Name: pTnT-A4 Alternate name(s): P450 HYDB
Clone type: Genomic: _____
cDNA: Expression x Non-expression

New GenBank Accession: _____ Original GenBank Accession: AK068163

Clone Description: Two primers (forward 2175 and reverse 2176) were used to amplify the CYP97A4 ORF from pRT-A4. XhoI (+ Kozak sequence) and XbaI restriction sites were added to primers 2175 and 2176, respectively. The amplified CYP97A4 ORF and pTNT vector were cut with both restriction enzymes (XhoI and XbaI). Resulting fragments were ligated and contruct was used as a template for in vitro transcription/translation using either TNT Coupled Reticulocyte Lysate or TNT Coupled Wheat Germ Extract systems.

Primer 2175: 5′-ACCG CTC GAG GCC ACC ATG AGC TCA GCG ACG TCA GTG AGT G -3′
XhoI Kozak seq.
Primer 2176: 5′-GAGA TCT AGA TCA GAT TCG AGT TGC TGA GAC TTG C-3.′
XbaI

Constructed by: Rena Quinlan Date rec′d_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: 6 Position: G7 Conc. 0.8µg/µl
Tube labeled as: pTnT-A4
Strain Location (-80°C) Box Number: 18 Position E2
Tube labeled as: pTnT-A4 Strain: TOP10F′

Lab Notebook to reference: Rena-Book 7 Date: 9/14/08

Original clone name (if different): _____

Organism source of gene: Oryza sativa Variety: Nipponbare

Cloning vector used: pTNT Vector size: 2871bp Insert size: 1932bp

Antibiotic markers: X amp; _____tet; _____chloramphenicol;_____other

Restriction enzyme(s) to release insert: XhoI/XbaI

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: 5/11/09 Entered into database yes: 434 / no ____
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pTnT-C2 (rice P450 epsilon-hydroxylase (CYP97C2))

Lab Clone Number/Name: pTnT-C2 Alternate name(s): P450 HYDE
Clone type: Genomic: _____
cDNA: Expression x Non-expression

New GenBank Accession: _____ Original GenBank Accession: AK065689

Clone Description: Two primers (forward 2140 and reverse 2168) were used to amplify the CYP97C2 ORF (plus 39bp
of 5′ UTR) from rice cDNA. XhoI and XbaI restriction sites were added to primers 2140 and 2168, respectively. The amplified CYP97C2 ORF (plus 39bp of 5′ UTR) and pTNT vector were cut with restriction enzymes XhoI and XbaI. Resulting fragments were ligated and contruct was used as a template for in vitro transcription/translation using either TNT Coupled Reticulocyte Lysate or TNT Coupled Wheat Germ Extract systems.
Primer 2140: 5′- GAGA CTC GAG AAT CCA TCT CGA ATC CCT AGC– 3′
XhoI
Primer 2168: 5′- ACCG TCT AGA TCA TCT GGA CCC ACT GAG TG- 3′
XbaI

Constructed by: Rena Quinlan Date rec′d_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: 6 Position: G8 Conc. 0.38µg/µl
Tube labeled as: pTnT-C2
Strain Location (-80°C) Box Number: 18 Position E3
Tube labeled as: pTnT-C2 Strain: TOP10F′

Lab Notebook to reference: Rena-Book 7 Date: 9/5/08

Original clone name (if different): _____

Organism source of gene: Oryza sativa Variety: Nipponbare

Cloning vector used: pTNT Vector size: 2871bp Insert size: 1725 bp

Antibiotic markers: X amp; _____tet; _____chloramphenicol;_____other

Restriction enzyme(s) to release insert: XhoI/XbaI

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: 5/11/09 Entered into database yes: 435__ / no ____
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pTnT-H3 (maize diiron beta-hydroxylase 3)

Lab Clone Number/Name: pTnT-H3 Alternate name(s): maize HYD3
Clone type: Genomic: _____
cDNA: Expression x Non-expression

New GenBank Accession: _____ Original GenBank Accession: BM382572

Clone Description: Two primers (forward 2163 and reverse 2164) were used to amplify the maize HYD3 ORF from the pSLIP vector. XhoI (+ Kozak sequence) and XbaI restriction sites were added to primers 2163 and 2164, respectively. The amplified maize HYD3 ORF and pTNT vector were cut with restriction enzymes XhoI and XbaI. Resulting fragments were ligated and contruct was used as a template for in vitro transcription/translation using either TNT Coupled Reticulocyte Lysate or TNT Coupled Wheat Germ Extract systems.

Primer 2163: 5′- ACCG CTC GAG GCC ACC ATG GCC GCC GCG ATG ACC -3′
XhoI Kozak seq.
Primer 2164: 5′-GAGA TCT AGA CTA GAA CTC ATT TGG CAC ACT -3′
XbaI

Constructed by: Rena Quinlan Date rec′d_____
Purified by: Rena Quinlan
DNA Location (-20°C) Box Number: 6 Position: G9 Conc. 0.2µg/µl
Tube labeled as: pTnT-H3
Strain Location (-80°C) Box Number: 18 Position E4
Tube labeled as: pTnT-H3 Strain: TOP10F′

Lab Notebook to reference: Rena-Book 7 Date: 9/5/08

Original clone name (if different): _____

Organism source of gene: Zea mays Variety: B73

Cloning vector used: pTNT Vector size: 2871bp Insert size: 960 bp

Antibiotic markers: X amp; _____tet; _____chloramphenicol;_____other

Restriction enzyme(s) to release insert: XhoI/XbaI

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: 5/11/09 Entered into database yes: 436__ / no ___
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pTnT-H4 (maize diiron beta-hydroxylase 4)

Lab Clone Number/Name:__pTnT-H4_____Alternate name(s): maize HYD4
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BG320875

Clone Description: Two primers (forward 2165 and reverse 2166) were used to amplify the maize HYD4 ORF from the pBluescript SK (-) vector. XhoI (+ Kozak sequence) and XbaI restriction sites were added to primers 2165 and 2166, respectively. The amplified maize HYD4 ORF and pTNT vector were cut with restriction enzymes XhoI and XbaI. Resulting fragments were ligated and contruct was used as a template for in vitro transcription/translation using either TNT Coupled Reticulocyte Lysate or TNT Coupled Wheat Germ Extract systems.

Primer 2165: 5′-ACCG CTC GAG GCC ACC ATG GCC GCC GGT CTG TCC-3′
XhoI Kozak seq.
Primer 2166: 5′-GAGA TCT AGA TCA GAT GGT CCG GCC GAT T-3.′
XbaI

Constructed by: __Rena Quinlan__Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: _6__Position: _G10____Conc. _0.25µg/µl
Tube labeled as: __pTnT-H4_____
Strain Location (-80°C) Box Number: __18_ Position__E5_____
Tube labeled as: __pTnT-H4_____ Strain: _TOP10F′__

Lab Notebook to reference: _Rena-Book 7__ Date: _9/5/08_____

Original clone name (if different): _____

Organism source of gene: __ Zea mays_____Variety:_ __B73_____

Cloning vector used: pTNT____Vector size: _2871bp Insert size: 930 bp

Antibiotic markers: ___X__amp; _____tet; _____chloramphenicol;_____other

Restriction enzyme(s) to release insert: _XhoI/XbaI_____

Sequence verified: yes_x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_x____
Hybridization results attached: Yes____ No___x_
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan____

WURTZEL LAB
CLONE INFORMATION

Date Today: 07 16 03
(MONTH)/ (DAY)/ (YEAR)

Clone Number/Name: prLHCP

Constructed by: __Dr. Kenneth Cline_____
Purified by: __Dr. Kenneth Cline _____
Date constructed: _____/_____/_ ____
(MONTH)/ (DAY)/ (YEAR)

Storage Location-
Box Number: ____7_____ Position: _____A2_____

Clone description: Please refer to the attached papers.
Cited in journal: Plant J, 2001, 27(5):373-382
Cloning vector used:____unknown_____ Vector size: _____
Organism source of gene: ____Pea_____ Insert size: ___1993bp_____
Restriction enzyme(s) to release insert:_____unknown _____
Lab Notebook to
Concentration:____1μg/μl _____
E. coli Antibiotic markers: -X-amp; __tet; _ chloramphenicol; __ other (_____)
Strain Transformed into: Top10, Box 15, D6

Cartoon of construct:

WURTZEL LAB
CLONE INFORMATION

Date Today: 07 16 03
(MONTH)/ (DAY)/ (YEAR)

Clone Number/Name: 16/EGFP

Constructed by: Dr. JP Marques
Purified by: Dr. JP Marques
Date constructed: _____/_____/_____
(MONTH)/ (DAY)/ (YEAR)

Storage Location-
Box Number: 7 Position: A5

Clone description: Please refer to the attached papers.
Cited in journal: Mol Genet Genomics, 2003,269(3):381-387.
Cloning vector used: pBSCM13 _____ Vector size: _____
Organism source of gene: Spinach Insert size: _____
Restriction enzyme(s) to release insert: unknown _____
Lab Notebook to 8/17/3
Concentration:_____
E. coli Antibiotic markers: -X-amp; __tet; _ chloramphenicol; __ other (_____)
Strain Transformed into: Top10, Box 15, D9

Cartoon of construct:

'

WURTZEL LAB
CLONE INFORMATION

Date Today: 03/16/11 Entered into database yes: 640 / no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: CYP97A4::GFP (Green Fluorescent Protein vector with CYP97A4)

Lab Clone Number/Name: A4::GFP Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK068163 (CYP97A4)

Clone Description: Two primers (forward 2634 and reverse 2635) were used to amplify the CYP97A4 ORF. XbaI and BamHI restriction sites were added to primers 2634 and 2635, respectively. The amplified CYP97A4 ORF and the pUC35S-GFP-Nos vector were cut with XbaI and BamHI. Resulting fragments were ligated and transformed into TOP10F′ cells. pUC35S-GFP contains the N-terminus of the Green Fluorescent Protein (GFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of GFP.

2634 Forward: ACCG TCT AGA ATG AGC TCA GCG ACG TCA GTG AG
XbaI

2635 Reverse: GAGA GGA TCC GAT TCG AGT TGC TGA GAC TTG CC
BamHI

Refer to Okada et al., 2000, Plant Physiol, 122:1045-1056

Constructed by: ____Rena Quinlan_____ Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: __9_____Position: _J1___Conc. ___0.2__µg/µl
Tube labeled as: __A4_GFP_____
Strain Location (-80°C) Box Number: _____20____Position_____E6_____
Tube labeled as: __A4_GFP_____ Strain: TOP10F′___

Cited in journal: Okada et al., 2000, Plant Physiol, 122:1045-1056

Lab Notebook to reference: _Rena (book 9)_____Date: _____1/19/10_____

Original clone name (if different): _____

Organism source of gene: _Oryza sativa_____Variety__Nipponbare_____

Cloning vector used: pUC35S-GFP-Nos _Vector size: _3500bp_Insert size: 1932bp___

Antibiotic markers: __X__amp; _____tet; ____chloramphenicol;_____other (__)

Restriction enzyme(s) to release insert: ____XbaI/BamHI____

Sequence verified: yes__X__no_____ Junction verified: yes___X__no_____

GenBank sheet attached: Yes_____No__x__
Hybridization results attached: Yes____ No__x__
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan___

WURTZEL LAB
CLONE INFORMATION

Date Today: 04/30/11 Entered into database yes: 641 / no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: CYP97C2::GFP (Green Fluorescent Protein vector with CYP97C2)

Lab Clone Number/Name: C2::GFP Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK065689 (CYP97C2)

Clone Description: Two primers (forward 2879 and reverse 2880) were used to amplify the CYP97C2 ORF. XbaI and BclI restriction sites were added to primers 2879 and 2880, respectively. The amplified CYP97C2 ORF was cut with XbaI and BclI and the pUC35S-GFP-Nos vector was cut with restriction enzymes XbaI and BamHI. (CYP97C2 ORF contains a BamHI site - therefore BclI was used instead since both BamHI and BclI have the same overhang). Resulting fragments were ligated and transformed into TOP10F′ cells. pUC35S-GFP contains the N-terminus of the Green Fluorescent Protein (GFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of GFP.

2879 Forward: ACCG TCT AGA ATGGCCGCCGCCGCCGCCGCCGCC
XbaI

2880 Reverse: GAGA TGATCA TCTGGACCCACTGAGTGCAAAATCAG
BclI

Refer to Okada et al., 2000, Plant Physiol, 122:1045-1056

Constructed by: ____Rena Quinlan_____ Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: __9_____Position: _J2___Conc. ___0.5__µg/µl
Tube labeled as: __C2_GFP_____
Strain Location (-80°C) Box Number: _____20____Position_____E7_____
Tube labeled as: __C2_GFP_____ Strain: TOP10F′___

Cited in journal: Okada et al., 2000, Plant Physiol, 122:1045-1056

Lab Notebook to reference: _Rena (book 10)_____Date: _____4/14/11____

Organism source of gene: _Oryza sativa_____Variety__Nipponbare_____

Cloning vector used: pUC35S-GFP-Nos _Vector size: _3500bp_____Insert size: 1686bp___

Antibiotic markers: __X__amp; _____tet; ____chloramphenicol;_____other (__)

Restriction enzyme(s) to release insert: ____XbaI/BclI____

Sequence verified: yes__X__no_____ Junction verified: yes___X__no_____

GenBank sheet attached: Yes_____No__X__
Hybridization results attached: Yes____ No____
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan_

WURTZEL LAB
CLONE INFORMATION

Date Today: 03/16/11 Entered into database yes: 642 / no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: HYD3::GFP (Green fluorescent protein vector with HYD3)

Lab Clone Number/Name: _HYD3::GFP___Alternate name(s): H3_GFP_____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BM382572 (HYD3)

Clone Description: Two primers (forward 2638 and reverse 2639) were used to amplify the HYD3 ORF. XbaI and BamHI restriction sites were added to primers 2638 and 2639, respectively. The amplified HYD3 ORF and the pUC35S-GFP-Nos vector were cut with XbaI and BamHI. Resulting fragments were ligated and transformed into TOP10F′ cells. pUC35S-GFP contains the N-terminus of the Green Fluorescent Protein (GFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of GFP.

2638 Forward: ACCG TCT AGA ATG GCC GCC GCG ATG ACC
XbaI

2639 Reverse: GAGA GGA TCC GAA CTC ATT TGG CAC ACT CT
BamHI

Refer to Okada et al., 2000, Plant Physiol, 122:1045-1056

Constructed by: ____Rena Quinlan_____ Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: __9_____Position: _J3___Conc. ___0.2__µg/µl
Tube labeled as: __H3_GFP_____
Strain Location (-80°C) Box Number: _____20____Position_____E8_____
Tube labeled as: __H3_GFP_____ Strain: TOP10F′___

Cited in journal: Okada et al., 2000, Plant Physiol, 122:1045-1056

Lab Notebook to reference: _Rena (book 9)_____Date: _____2/4/10_____

Original clone name (if different): _____

Organism source of gene: _Zea mays_____Variety:__B73_____

Cloning vector used: pUC35S-GFP-Nos _Vector size: _3500bp____Insert size: 960bp___

Antibiotic markers: __X__amp; _____tet; ____chloramphenicol;_____other (__)

Restriction enzyme(s) to release insert: ____XbaI/BamHI____

Sequence verified: yes__X__no_____ Junction verified: yes___X__no_____

GenBank sheet attached: Yes_____No__X__
Hybridization results attached: Yes____ No____
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan____

WURTZEL LAB
CLONE INFORMATION

Date Today: 03/16/11 (MONTH)/(DAY)/(YEAR) Entered into database yes: 643 / no___

CLONE NAME: HYD4::GFP (Green fluorescent protein vector with HYD4)

Lab Clone Number/Name: _HYD4::GFP____Alternate name(s): H4_GFP_____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BG320875 (HYD4)

Clone Description: Two primers (forward 2640 and reverse 2641) were used to amplify the HYD4 ORF. XbaI and BamHI restriction sites were added to primers 2640 and 2641, respectively. The amplified HYD4 ORF and the pUC35S-GFP-Nos vector were cut with XbaI and BamHI. Resulting fragments were ligated and transformed into TOP10F′ cells. pUC35S-GFP contains the N-terminus of the Green Fluorescent Protein (GFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of GFP.

2640 Forward: ACCG TCT AGA ATG GCC GCC GGT CTG TCC
XbaI

2641 Reverse: GAGA GGA TCC GAT GGT CCG GCC GAT TCG
BamHI

Refer to Okada et al., 2000, Plant Physiol, 122:1045-1056

Constructed by: ____Rena Quinlan_____ Date rec′d_____
Purified by: _____Rena Quinlan_____
DNA Location (-20°C) Box Number: __9_____Position: _J4___Conc. ___0.2__µg/µl
Tube labeled as: __H4_GFP_____
Strain Location (-80°C) Box Number: _____20____Position_____E9_____
Tube labeled as: __H4_GFP_____ Strain: TOP10F′___

Cited in journal: Okada et al., 2000, Plant Physiol, 122:1045-1056

Lab Notebook to reference: _Rena (book 9)_____Date: _____2/4/10_____

Original clone name (if different): _____

Organism source of gene: _Zea mays_____Variety:__B73_____

Cloning vector used: pUC35S-GFP-Nos _Vector size: _3500bp____Insert size: 930bp___

Antibiotic markers: __X__amp; _____tet; ____chloramphenicol;_____other (__)

Restriction enzyme(s) to release insert: ____XbaI/BamHI____

Sequence verified: yes__X__no_____ Junction verified: yes___X__no_____

GenBank sheet attached: Yes_____No__X__
Hybridization results attached: Yes____ No____
Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor′s Name: __Rena Quinlan__

WURTZEL LAB
CLONE INFORMATION

Date Today: 03/03 / 10 **Entered into database #**569
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pUC35S-Toc34-sGFP-Nos

Lab Clone Number/Name: **Alternate name(s):**
Clone type: Genomic:
cDNA: Expression **Non-expression**

New GenBank Accession: **Original GenBank Accession**:

Clone Description: A full copy of Toc34 (34 kDa outer membrane transport complex protein from *P. sativum*) without a stop codon was amplified from pToc34 (Chen et al., 1997, JBC **272**(10): 6614-6620)) with a forward primer 2579, containing XbaI site, and a reverse primer 2580, containing BamHI site, and inserted into pUC35S-sGFP-Nos vector, digested with XbaI/BamHI, giving pUC35S-Toc34-sGFP-Nos plasmid.
2579: 5' ATC**TCTAGA**ATGGCTTCACAACAACAAACTGTTCG
2580: 5' ATC**GGATCC**CTTCCGATCACCTACATCGCGAGTC

Constructed by: M. Shumskaya **Date rec'd** 08/28/09
Purified by: M. Shumskaya
DNA Location (-20°C) Box Number: 50 **Position:** D7 **Conc**. 0.4 µg/µl
Tube labeled as: Toc34-GFP
Strain Location (-80°C) Box Number: 51 **Position** C5
Tube labeled as: Toc34 GFP **Strain:**

Cited in journal:

Lab Notebook to reference: Maria **Date**: 08/11/09

Original clone name (if different):

Organism source of gene: *P.sativum* **Variety**

Cloning vector used: **Vector size**: 4500bp **Insert size**: 933 bp

Antibiotic markers: amp

Restriction enzyme(s) to release insert: XbaI/BamHI

Sequence verified: yes **Junction verified**: yes

GenBank sheet attached: No
Hybridization results attached: Yes____ No____
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)
Toc34-GFP with 2436 primer
Ngccatccnctatcctcgcagacccttcctctatataaggaagttcatttcatttggagagaacacgggggactctagaatggcttcacaacaacaaactgttcgtgaatggtcaggaatcaatacattcgctcctgctacacagactaagt
tgcttgaactcttgggaaaccttaaacaagaggatgtgaactccttaaccatactttgtgatgggaaaaggtggcgttgggaagtcttcaactgtgaactccatcattggagaaagggtggtttcaattagtccctttcagtctgaagggccaag
acctgttatggtgtcacgatcaagggcaggttttacattgaacattatcgacactcctggtcttattgaaggggatacatcaatgatatggcgcttaatataataaaaagtttccttctggacaagaccatagatgttctgctttacgtggaccgc
ttagatgcgtatcgagtagacaacttggataagttagtcgccaaagctataactgatagttttggcaaaggaatatggaacaaggctatagtagcactcacacatgcccaattctctccaccagatgggttgccttatgatgaattcttctctaa
aagatccgaggctctcttgcaagttgtaagatcaggtgcctccttaaagaaggatgctcaggcttctgacattcctgttgttttgatcgagaatagtgggagatgcaacaaaaatgacagtgatgaaaaggttcttccaaatgggattgcatg
gattcctcatctagtccaaacaatcacagaagttgcattgaacaaaagtgaatctatttttgttgacaagaatttgattgacggaccgaatccaaatcaagagaggaaaattatggattcctctcatatttgctctacaatacttgttcctcncga
agccaatagaagcnctaatcaggagagactttgctactgnaacaaaccacctggggan
Toc34-GFP with 2437 primer
ctttngngataggaagagggtcnccaaccncntngtaaagaagcntnntngagtnctctncgccgggtccaaaganggncccccccccgngnnnttgggaaaaaaaggnttccaccnctntcaaagcaagggatgatgngnnttt
tccctgactaggggatgangcncaatcccatttcttngcaagaccctncnctatntaaggaagttcatttcatttggggagaacncggggnctcttgaatggcttccccaccaacaacctgttcgtgaatggtcaggaatccaaacattcgct
cntgctacacagactaagttgcttgaactcttggggaaccttaaacaagaggatgtgaactcctacccatacttgtgatgggaaaaggtggcgttgggaagtcttcaactgtgaactccatcattggagaaaggggtggtttcaattagtcc
ctttcagtctgaagggccaagacctgttatggtgtcacgatcaagggcaggttttacattgaacattatcgacactcctggtcttattgaaggggatacatcaatgatatggcgcttaatataataaaaagtttccttctggacaagaccatag
atgttctgctttacgtggaccgcttagatgcgtatcgagtagacaacttggataagttagtcgccaaagctataactgatagttttggcaaaggaatatggaacaaggctatagtagcactcacacatgcccaattctctccaccagatgggtt
gccttatgatgaattcttctctaaaagatccgaggctctcttgcaagttgtaagatcaggtgcctccttaaagaaggatgctcaggcttctgacattcctgttgttttgatcgagaatagtgggagatgcaacaaaaatgacagtgatgaaaag
gttcttccaaatgggattgcatggattcctcatctagtccaaacaatcacagaagttgcattgaacaaaagtgaatctatttttgttgacaagaatttgattgacggaccgaatccaaatcagagaggaaaattatggattcctctcatatttgctc
tacaatacttgttcctcgcgaagccaatagaagcactaatcaggagagacattgctactgaaacaaaaccagcatgggagactcgcgatgtaggtgatcggaagggatccatggtgagcaagggcgaggagcgtcaccggg

Depositor's Name: M. Shumskaya

WURTZEL LAB
CLONE INFORMATION

Date Today: 07/27 / 09 **Entered into database** #567
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pUC35S-LHCP-sGFP-Nos

Lab Clone Number/Name: **Alternate name(s):**
Clone type: Genomic:
cDNA: Expression **Non-expression**

New GenBank Accession: **Original GenBank Accession**:

Clone Description: A full copy of LHCP (major chlorophyll a/b-binding polypeptide from *P. sativum*) without a stop codon was amplified from prLHCP (Tan et. al, 2001, Plant J **27:** 373-382) with the forward primer 2419, containing XbaI site, and a reverse primer 2420, containing BamHI site, and inserted into pUC35S-sGFP-Nos vector, digested with XbaI/BamHI, giving pUC35S-LHCP-sGFP-Nos plasmid.
2419: 5' ATC**TCTAGA**ATGGCCGCTTCATCC
2420: 5' ATC**GGATCC**CTTTCCGGGAACAAAGTTGGTAGC

Constructed by: M. Shumskaya **Date rec'd** 08/28/09
Purified by: M. Shumskaya
DNA Location (-20°C) Box Number: plasmid box 7 **Position:** B7 **Conc**. 2 µg/µl
Tube labeled as: B7, pUC35S-LHCP-GFP-Nos
Strain Location (-80°C) Box Number: 51 **Position** D8
Tube labeled as: LHCP GFP **Strain:** Top10

Cited in journal:

Lab Notebook to reference: Maria **Date**: 08/28/09

Original clone name (if different):

Organism source of gene: P.sativum **Variety**

Cloning vector used: **Vector size**: 4500bp **Insert size**: 800 bp

Antibiotic markers: amp

Restriction enzyme(s) to release insert: XbaI/BamHI

Sequence verified: yes **Junction verified**: yes

GenBank sheet attached: No
Hybridization results attached: Yes____ No____
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)
LHCP-GFP with 2436 primer

gtnagcnatcccactatccttcgcagacccttcctctatataaggaagttcatttcatttggagagaacacgggggactctagaatggccgcttcatcatcatcatccatggctctctcttctccaaccttggctggcaagcaactcaagctgaa
cccatcaagccaagaattgggagctgcaaggttcaccatgaggaagtctgctaccaccaagaaagtagcttcctctggaagcccatggtacggaccagaccgtgttaagtacttaggcccattctccggtgagtctccatcctacttgact
ggagagttccccggtgactacggttgggacactgccggactctctgctgacccagagacattctccaagaaccgtgagcttgaagtcatccactccagatgggctatgttgggtgctttgggatgtgtcttcccagagcttttgtctcgcaacg
gtgttaaattcggcgaagctgtgtggttcaaggcaggatctcaaatctttagtgagggtggacttgattacttgggcaacccaagcttggtccatgctcaaagcatccttgccatatgggccactcaggttatcttgatgggagctgtcgaaggt
taccgtattgccggtgggcctctcggtgaggtggttgatccacctttacccaggtggaagctttgatccattgggcttagctgatgatccagaagcattcgcagaattgaaggtgaaggaactcaagaacggtagattagccatgttctcaatgt
ttggattcttcgttcaagctattgtaactggaaagggtcctttggagaaccttgctgatcatcttgcagacccaggtcaacaacaatgcatgggtcatatgctaccaactttgttcccggaaaggggattcatggntgaagcaanggcgaag
aactgtttcacccgggggtggtgnccattcctggtcgagctgggannggnnancgtaaaacggccncaanttcaaccgtgttccggcgagggggnagggcgatncccccctacgnnaagcttgaccctgaaattcatntggcncaac
gggaaancttgcccgggccctggnccaacccnngggacccctttacctangggggggangggtttagccgcttccccgaaaaantaaaaannccaatttttaaatccccccngcccnaaggtnntnccagaag

Depositor's Name: M. Shumskaya

WURTZEL LAB
CLONE INFORMATION

Date Today: 01/31 / 11 **Entered into database** yes__#564

CLONE NAME: p-SAT-M-PGL-RFP
CLONE TYPE: cDNA clone in vector for transient expression in plant protoplasts

New GenBank Accession: **Original GenBank Accession**: BT039786_1

Clone Description: A full copy of maize plastoglobulin PGL-2 gene without a stop codon was amplified from clone ZM_BFc0045D12 (ordered from maize stock collection in http://www.genome.arizona.edu/orders/) with a forward primer 3020, containing EcoRI site, and a reverse primer 3017, containing site BclI and extra G to clone the PGL protein in frame with following RFP, and inserted into vector pSAT6(A)-RFP-N1-pchrD-cuc (3038) (obtained from Rena Quinlan), which was digested with same restrictases to remove chrD sequence from the vector, giving plasmid pSAT-M-PGL-RFP.
3020: 5' atc**GAATTC**atggcgtcctccgcgttcctcaacg
3017: 5' atcTGATCAggtatagaagagtacttccc

Constructed by: M. Shumskaya **Date rec'd** 02/23/10
Purified by: M. Shumskaya
DNA Location (-20°C) Box Number: 50 **Position:** D1 **Conc**. 100 ng/µl
Tube labeled as: PG2-RFP
Strain Location (-80°C) Box Number: 51 **Position** D7
Tube labeled as: PG2-RFP **Strain:**

Cited in journal:

Lab Notebook to reference: Maria **Date**: 02/23/10

Original clone name (if different):

Organism source of gene: *Zea mays* **Variety** B73

Cloning vector used: pSAT6(A)-N1-RFP **Vector size**: 4534 bp **Insert size**: 957 bp

Antibiotic markers: amp

Restriction enzyme(s) to release insert: EcoRI/BclI

Sequence verified: yes **Junction verified**: yes

GenBank sheet attached: No
Hybridization results attached: Yes____ No____
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)
M-PGL-RFP with 2619 primer
nnnnnnnnnnnnttcnnttggagaggacgtcgagagttctcaacacaacatatacaaaacaaacgaatctcaagcaatcaagcattctacttctattgcagcaatttaaatcatttctttttaaagcaaaagcaattttctgaaaattttcacc
atttacgaacgatagagatctcgagctcaagcttcgaattcatggcgtcctccgcgttcctcaacgctctcgtcctccgatcgccatcaccgtcgctctctgtcctcgtctaaccgccgcagcaccggcgccttcgcgttcccgaacccaccgc
gcttcccatgcctgcgttctgcgcgtcgactcgtccttgcgagggcggcggcagcgccggcgccgggcgatggccccgaggacgagtgggggccggagcccgagggcggttccgctgtcacgggggccgcggtggcggaggcgc
cggaggcgagagaggtcgcggagctcaaggcgcagctaaaggacgcgctgtacggcacggagcgggggcctgcgggcgtccagcgagtcacgagccaaagtgctggagctcatcacgcagctcgagacgcgcaaccccacgc
cggctcctacggaggcgcttacgctccttaacgggaagtggatcctcgcgtacacatcattttcccaactattcccactattgggggtttggtaatctgcctcagctggtgaaggtggaggaaatatcacaaaccattgattctgagaacttcac
agtgcaaaactgtatcaagttttcaggacctttggctacaacctcagttgccactaatgcaaaatttgaaattagaagccccaagcgtgtgcagatcaaatttgatgaagggattgttggtacaccacagttgaccgactccattgtgctacca
gagaagtttgagttatttggacagaacatcgacctgagtccactgaaaggcatattttcttccatcgaaaatgcagcatcctcggttgcaaagaccatatctgagcagcctccgctcaaaatacccgatcaggannnannatgctgaatcnn
ngnngcttacnannntaccttgncgaananctcnnntnntccannngcgatggcagcannatctttnnnctnnnnnnnnggnantacnnnttcnnnnnnnnatcctgannggnntcnnncnangncntcatcnagnnnnnnan
gnnntncnnggnnnnnnnngnnnnnnnnnnnnnnnnnnnnnnnnnnnnaantnnnnnggggnnnnnggnnnannnnn
M-PGL-RFP with 3016 primer
nnnnnnnnnnnnnntcgcnatcaccgtcgctctcgtcctcgtctaaccgccgcagcaccggcgccttcgcgttcccgaacccaccgcgcttcccatgcctgcgttctgcgcgtcgactcgtccttgcgagggcggcggcagcgccggcg
ccgggcgatggccccgaggacgagtgggggccggagcccgagggcggttccgctgtcacgggggccgcggtggcggaggcgccggaggcgagagaggtcgcggagctcaaggcgcagctaaaggacgcgctgtacggcac
ggagcgggggcctgcgggcgtccagcgagtcacgagccaaagtgctggagctcatcacgcagctcgagacgcgcaaccccacgccggctcctacggaggcgcttacgctccttaacgggaagtggatcctcgcgtacacatcattttc
ccaactattcccactattgggggtttggtaatctgcctcagctggtgaaggtggaggaaatatcacaaaccattgattctgagaacttcagtgcaaaactgtatcaagttttcaggacctttggctacaacctcagttgccactaatgcaaaat
ttgaaattagaagccccaagcgtgtgcagatcaaatttgatgaagggattgttggtacaccacagttgaccgactccattgtgctaccagagaagtttgagttatttggacagaacatcgacctgagtccactgaaaggcatattttcttccat
cgaaaatgcagcatcctcggttgcaaagaccatatctgagcagcctccgctcaaaataccgatcaggacaaacaatgctgaatcgtggttgcttacgacatacttgacgaagagctcntctccagaggcgatggcagcagcatctttgt
gctgttcaaggagggaagtactcttctatacctgatcctgatggcctcctccgagnacgtcatcaaggagttcatgcgcttcaagnngcgcatgnagggctcgtgaacggnnacnagttnnagntcnagggccgaggncnaaggnnn
nncnnnnnngggcnnccnnnanannnnnnncnnaaggnnnncnnggnngccccnnnncntnnccnngnnntctgtnccnnnnnnncnagnnnngctccaggnnnnnnnnnancnn

Depositor's Name: M. Shumskaya

WURTZEL LAB
CLONE INFORMATION

Date Today: _10/__05_/_09 Entered into database yes 486_/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: A4_2236 (Bimolecular fluorescence complementation vector with CYP97A4)

Lab Clone Number/Name: _ A4_2236__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession____Original GenBank Accession: AK068163 (CYP97A4), DQ169003 (pSAT4(A)-nEYFP-N1)_

Clone Description: Two primers (forward 2455 and reverse 2456) were used to amplify the CYP97A4 ORF. XhoI and EcoRI restriction sites were added to primers 2455 and 2456, respectively. The amplified CYP97A4 ORF and pSAT4(A)-nEYFP-N1vector were cut with both restriction enzymes (XhoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT4(A)-nEYFP-N1contains the N-terminus of the Yellow Fluorescent Protein (EYFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of EYFP.

CYP97A4 **nYFP**

2455 Forward: ACCG CTC GAG ATG AGC TCA GCG ACG TCA GTG AG
XhoI
2456 Reverse: GAGA GAA TTC GAT TCG AGT TGC TGA GAC TTG CC
EcoRI
Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: _____Rena Quinlan_____Date rec′d _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I2____Conc. _0.6 µg/µl
Tube labeled as: __A4_2236_____
Strain Location (-80°C) Box Number: _18_____ Position: ____G2_____
Tube labeled as: __A4_2236_____ Strain: __TOP10F′_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __7/30/09_____

Original clone name (if different): _____

Organism source of gene: _Oryza sativa_____ Variety:__Nipponbare_____

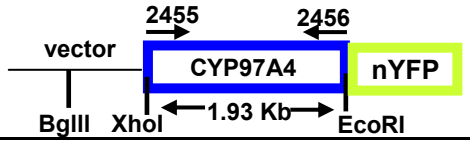
Cloning vector used: pSAT4(A)-nEYFP-N1__Vector size: 4378 bp__Insert size: _1932 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _XhoI and EcoRI__

Sequence verified: yes___x__no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT4(A)-nEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _7/__05_/_10 (MONTH)/(DAY)/(YEAR) Entered into database yes: 574 / no ___

CLONE NAME: C2_2236 (Bimolecular fluorescence complementation vector with CYP97C2)

Lab Clone Number/Name: _C2_2236__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK065689 (CYP97C2)

Clone Description: Two primers (forward 3025 and reverse 3026) were used to amplify the CYP97C2 ORF. XhoI and EcoRI restriction sites were added to primers 3025 and 3026, respectively. The amplified CYP97C2 ORF and pSAT4(A) -nEYFP-N1 vector were cut with both restriction enzymes (XhoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT4(A) -nEYFP-N1 contains the N-terminus of the Yellow Fluorescent Protein (EYFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of YFP.

3025 Forward: ACCG CTC GAG ATG GCC GCC GCC GCC GCC GCC GCC
XhoI
3026 Reverse: GAGA GAA TTC TCTGGACCCACTGAGTGCAAAATCAG
EcoRI
Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: Rena Quinlan Date rec′d _____
Purified by: _ Rena Quinlan
DNA Location (-20°C) Box Number: __6_____Position: _J1____Conc. _0.1 µg/µl
Tube labeled as: __C2_2236
Strain Location (-80°C) Box Number: _17_____ Position: ____E6
Tube labeled as: __C2_2236._____ Strain: __TOP10F′

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 10)_____Date: __6/27/10_____

Original clone name (if different): _____

Organism source of gene: _Oryza sativa_____ Variety:__Nipponbare_____

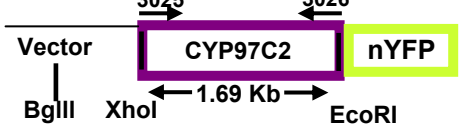
Cloning vector used: pSAT4(A) -nEYFP-N1 __Vector size: 4378 bp_Insert size: _1686bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _XhoI and EcoRI__

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT4(A)-nEYFP-N1 can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
(attach additional information here i.e. cartoon, VectorNTI details, sequence)
For 5′ Junction sequencing use the TI primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _10/__05_/_09 Entered into database yes_489_/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: H3_2236 (Bimolecular fluorescence complementation vector with diiron HYD3 beta-hydroxylase)

Lab Clone Number/Name: _ H3_2236__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: Original GenBank Accession: BM382572 (maize HYD3), DQ169003 (pSAT4(A)-nEYFP-N1)

Clone Description: Two primers (forward 2467 and reverse 2468) were used to amplify the diiron HYD3 ORF. XhoI and EcoRI restriction sites were added to primers 2467 and 2468, respectively. The amplified HYD3 ORF and pSAT4(A)-nEYFP-N1vector were cut with both restriction enzymes (XhoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT4(A)-nEYFP-N1contains the N-terminus of the Yellow Fluorescent Protein (EYFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of EYFP.

2467 Forward : ACCG CTC GAG ATG GCC GCC GCG ATG ACC
XhoI
2468 Reverse: GAGA GAA TTC GAA CTC ATT TGG CAC ACT CT
EcoRI
Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: _____Rena Quinlan_____Date rec′d _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I4____Conc. _0.1 µg/µl
Tube labeled as: __H3_2236_____
Strain Location (-80°C) Box Number: _18_____ Position: ____G4_____
Tube labeled as: __H3_2236_____ Strain: __TOP10F′_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __7/30/09_____

Original clone name (if different): _____

Organism source of gene: _Zea mays_____ Variety:__B73____

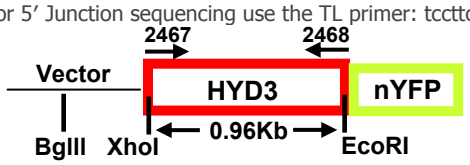
Cloning vector used: pSAT4(A)-nEYFP-N1__Vector size: 4378 bp_Insert size: _960 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _XhoI and EcoRI__

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT4(A)-nEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _10/__05_/_09 (MONTH)/(DAY)/(YEAR) Entered into database yes_488_/ no ___

CLONE NAME: H4_2236 (Bimolecular fluorescence complementation vector with diiron HYD4 beta-hydroxylase)

Lab Clone Number/Name: _ H4_2236__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: Original GenBank Accession: BG320875 (maize HYD4), DQ169003 (pSAT4(A)-nEYFP-N1)

Clone Description: Two primers (forward 2469 and reverse 2470) were used to amplify the diiron HYD4 ORF. XhoI and EcoRI restriction sites were added to primers 2469 and 2470, respectively. The amplified HYD4 ORF and pSAT4(A)-nEYFP-N1vector were cut with both restriction enzymes (XhoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT4(A)-nEYFP-N1contains the N-terminus of the Yellow Fluorescent Protein (EYFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of EYFP.

2469 Forward: ACCG CTC GAG ATG GCC GCC GGT CTG TCC
XhoI
2470 Reverse: GAGA GAA TTC GAT GGT CCG GCC GAT TCG
EcoRI

Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: _____Rena Quinlan_____Date rec′d _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I5____Conc. _0.1 µg/µl
Tube labeled as: __H4_2236_____
Strain Location (-80°C) Box Number: _18_____ Position: ____G5_____
Tube labeled as: __H4_2236_____ Strain: __TOP10F′_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __7/30/09_____

Original clone name (if different): _____

Organism source of gene: _Zea mays_____ Variety:__B73____

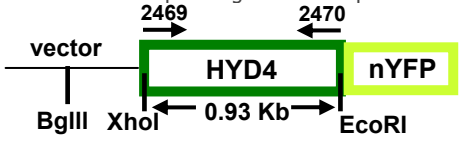
Cloning vector used: pSAT4(A)-nEYFP-N1__Vector size: 4378 bp_Insert size: _930 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _XhoI and EcoRI__

Sequence verified: yes___x__no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT4(A)-nEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _7/__05_/_10 (MONTH)/(DAY)/(YEAR) Entered into database yes: _573_/ no ___

CLONE NAME: A4_1476 (Bimolecular fluorescence complementation vector with CYP97A4)

Lab Clone Number/Name: _ A4_1476__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK068163 (CYP97A4)

Clone Description: Two primers (forward 3023 and reverse 3024) were used to amplify the CYP97A4 ORF. XhoI and EcoRI restriction sites were added to primers 3023 and 3024, respectively. The amplified CYP97A4 ORF and pSAT6-cEYFP-N1vector were cut with both restriction enzymes (XhoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT6-cEYFP-N1produces an in-frame fusion of the C-terminus of the protein of interest to the N-terminus of C-terminal half of YFP.

3023 Forward: ACCG C TCG AG ATG AGC TCA GCG ACG TCA GTG AG
XhoI
3024 Reverse: GAGA GAA TTC GAT TCG AGT TGC TGA GAC TTG CC
EcoRI

Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: _____Rena Quinlan_____Date rec′d _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I9____Conc. _0.2 µg/µl
Tube labeled as: __A4_1476_____
Strain Location (-80°C) Box Number: _17_____ Position: ____E5_____
Tube labeled as: __A4_1476_____ Strain: __TOP10F′_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 10)_____Date: __6/27/10_____

Original clone name (if different): _____

Organism source of gene: _Oryza sativa_____ Variety:__Nipponbare_____

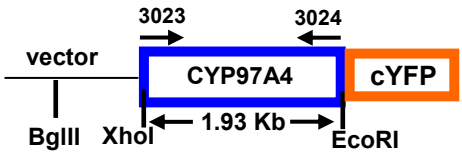
Cloning vector used: pSAT6-cEYFP-N1__Vector size: 4378 bp__Insert size: _1932 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _XhoI and EcoRI__

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT6-cEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _10/__05_/_09 (MONTH)/(DAY)/(YEAR) Entered into database yes: 487_/ no __

CLONE NAME: C2_1476 (Bimolecular fluorescence complementation vector with CYP97C2)

Lab Clone Number/Name: _C2_1476__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: AK065689 (CYP97C2)

Clone Description: Two primers (forward 2459 and reverse 2460) were used to amplify the CYP97C2 ORF. NcoI and EcoRI restriction sites were added to primers 2459 and 2460, respectively. The amplified CYP97C2 ORF and pSAT6 -cEYFP-N1 vector were cut with both restriction enzymes (NcoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT6-cEYFP-N1 generates fusion of C-terminus of tested protein to N-terminus of the C-terminal half of Yellow Fluorescent Protein (EYFP).

2459 Forward: ACCG CC ATG G CC GCC GCC GCC GCC
NcoI
2460 Reverse: GAGA GAA TTC TCTGGACCCACTGAGTGC
EcoRI
Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: _____Rena Quinlan_____Date rec′d _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I3____Conc. _2.4 µg/µl
Tube labeled as: __C2_1476_____
Strain Location (-80°C) Box Number: _18_____ Position: ____G3_____
Tube labeled as: __C2_1476_____ Strain: __TOP10F′_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __7/30/09_____

Original clone name (if different): _____

Organism source of gene: _Oryza sativa_____ Variety:__Nipponbare_____

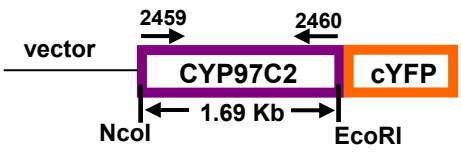
Cloning vector used: pSAT6 -cEYFP-N1 __Vector size: 4378 bp_Insert size: _1686bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _NcoI and EcoRI__

Sequence verified: yes___x__no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT6 -cEYFP-N1 can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _7/__05_/_10 (MONTH)/(DAY)/(YEAR) Entered into database yes: 575 / no ___

CLONE NAME: H3_1476 (Bimolecular fluorescence complementation vector with diiron HYD3 beta-hydroxylase)

Lab Clone Number/Name: _ H3_1476__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BM382572 (maize HYD3)_

Clone Description: Two primers (forward 2846 and reverse 2847) were used to amplify the diiron HYD3 ORF. BspHI and EcoRI restriction sites were added to primers 2846 and 2847, respectively. The amplified HYD3 ORF and pSAT6-cEYFP-N1vector were cut with both restriction enzymes (BspHI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT6)-cEYFP-N1produces an in-frame fusion of the C-terminus of the protein of interest to the N-terminus of the C-terminal half of YFP.

2846 Forward : ACCG TCA TG A TGGCCGCCGCGATGACCAG
BspHI
2847 Reverse: GAGA GAA TTC GAACTCATTTGGCACACTCTGGC
EcoRI
Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: _____Rena Quinlan_____Date rec′d _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _J2____Conc. _0.15 µg/µl
Tube labeled as: __H3_1476_____
Strain Location (-80°C) Box Number: _17____ Position: ____E7_____
Tube labeled as: __H3_1476_____ Strain: __TOP10F′_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 10)_____Date: __6/27/10_____

Original clone name (if different): _____

Organism source of gene: _Zea mays_____ Variety:__B73____

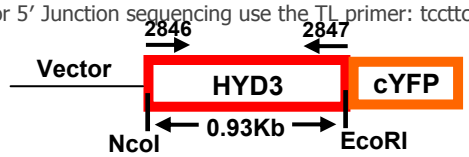
Cloning vector used: pSAT6-cEYFP-N1__Vector size: 4378 bp__Insert size: _960 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _BspHI and EcoRI__

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT6-cEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB
CLONE INFORMATION

Date Today: _7/__05_/_10 (MONTH)/(DAY)/(YEAR) Entered into database yes: 576 / no __

CLONE NAME: H4_1476 (Bimolecular fluorescence complementation vector with diiron HYD4 beta-hydroxylase)

Lab Clone Number/Name: _ H4_1476__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: BG320875 (maize HYD4), _

Clone Description: Two primers (forward 2848 and reverse 2849) were used to amplify the diiron HYD4 ORF. BspHI and EcoRI restriction sites were added to primers 2848 and 2849, respectively. The amplified HYD4 ORF and pSAT6-cEYFP-N1vector were cut with both restriction enzymes (BspHI and EcoRI). Resulting fragments were ligated and transformed into TOP10F′ cells. pSAT6-cEYFP-N1 produces an in-frame fusion of the C-terminus of the protein of interest to the N-terminus of the C-terminal half of YFP.

2848 Forward : ACCG TCA TG A TGGCCGCCGGTCTGTCCGG
BspHI
2849 Reverse: GAGA GAA TTC GATGGTCCGGCCGATTCGCG
EcoRI

Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Constructed by: Rena Quinlan Date rec′d _____
Purified by: _ Rena Quinlan
DNA Location (-20°C) Box Number: __6_____Position: _J3____Conc. _0.2 µg/µl
Tube labeled as: __H4_1476
Strain Location (-80°C) Box Number: _17_____ Position: ____E8
Tube labeled as: __H4_1476 Strain: __TOP10F′

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 10)_____Date: __5/14/10_____

Original clone name (if different): _____

Organism source of gene: _Zea mays Variety:__B73____

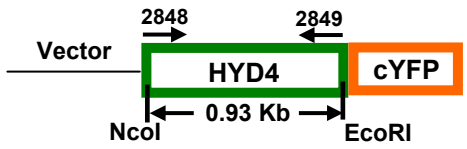
Cloning vector used: pSAT6-cEYFP-N1__Vector size: 4378 bp__Insert size: _930 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _BspHI and EcoRI__

Sequence verified: yes__x___no_____ Junction verified: yes__x____no_____

GenBank sheet attached: Yes_____No_____Hybridization results attached: Yes____ No____
Other: **Further details on pSAT4(A)-nEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB CLONE INFORMATION

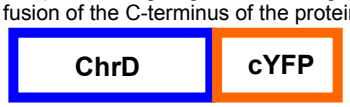
Date Today: _2/__02_/_10 Entered into database yes: _662_/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: ChrD_1476 (Bimolecular fluorescence complementation vector with ChrD)

Lab Clone Number/Name: _ ChrD_1476__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: DQ117526 (ChrD)_

Clone Description: Two primers (forward 2758 and reverse 2757) were used to amplify the ChrD ORF from ChrD_2485 vector. The ChrD protein from Cucumber is known to form dimers and localize to plastids. NcoI and EcoRI restriction sites were added to primers 2758 and 2757, respectively. The amplified ChrD ORF and pSAT6-cEYFP-N1vector were cut with both restriction enzymes (NcoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F' cells. pSAT6-cEYFP-N1 produces an in-frame fusion of the C-terminus of the protein of interest to the N-terminus of C-terminal half of YFP.



2758 Forward: ACCG CC ATG G CTATGTCCGCAGCTCAC
NcoI
2757 Reverse: GAGA GAA TTC AAGCACAGCTATTGCATCAATTTC
EcoRI

Refer to Libal-Weksler et al., (1997) Plant Physiology, **113:** 59-63

Constructed by: _____Rena Quinlan_____Date rec'd _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __7_____Position: _I4____Conc. _0.4 µg/µl
Tube labeled as: __ChrD_1476_____
Strain Location (-80°C) Box Number: _17_____ Position: ____I1_____
Tube labeled as: __ChrD_1476_____ Strain: __TOP10F'_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131, and Plant Physiology (1997), **113:** 59-63

Lab Notebook to reference: __Rena (Book 9)_____Date: __2/2/10_____

Organism source of gene: _*Cucumis sativus*_____ Variety:_____

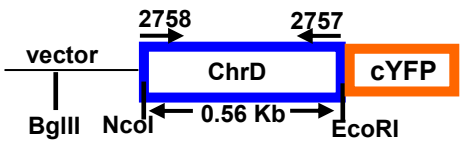
Cloning vector used: pSAT6-cEYFP-N1__Vector size: 4378 bp_Insert size: _564 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _NcoI and EcoRI__

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other: **Further details on pSAT6-cEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5' Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB CLONE INFORMATION

Date Today: _02/__02_/_10 Entered into database yes: 663 _/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: ChrD_2236 (Bimolecular fluorescence complementation vector with ChrD)

Lab Clone Number/Name: _ ChrD_2236__Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession___Original GenBank Accession: DQ117526 (ChrD), DQ169003 (pSAT4(A)-nEYFP-N1)_

Clone Description: Two primers (forward 2756 and reverse 2757) were used to amplify the ChrD ORF from ChrD_2485 vector. The ChrD protein from Cucumber is known to form dimers and localize to plastids. XhoI and EcoRI restriction sites were added to primers 2756 and 2757, respectively. The amplified ChrD ORF and pSAT4(A)-nEYFP-N1vector were cut with both restriction enzymes (XhoI and EcoRI). Resulting fragments were ligated and transformed into TOP10F' cells. pSAT4(A)-nEYFP-N1contains the N-terminus of the Yellow Fluorescent Protein (EYFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of EYFP.

2756 Forward: ACCG CTC GAG ATGGCTATGTCCGCAGCTCAC
XhoI
2757 Reverse: GAGA GAA TTC AAGCACAGCTATTGCATCAATTTC
EcoRI
Refer to Libal-Weksler et al., (1997) Plant Physiology, **113:** 59-63

Constructed by: _____Rena Quinlan_____Date rec'd _____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __7_____Position: _I5____Conc. _0.3 µg/µl
Tube labeled as: __ChrD_2236_____
Strain Location (-80°C) Box Number: _17_____ Position: ____J1_____
Tube labeled as: __ChrD_2236_____ Strain: __TOP10F'_____

Cited in journal: Plant Physiology, (1997) **113:** 59-63 and J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __2/2/10_____

Original clone name (if different): _____

Organism source of gene: _*Cucumis sativus*_____ Variety:_____

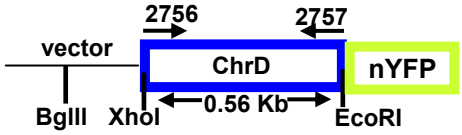
Cloning vector used: pSAT4(A)-nEYFP-N1__Vector size: 4378 bp__Insert size: _564 bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _XhoI and EcoRI__

Sequence verified: yes___x__no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other: **Further details on pSAT4(A)-nEYFP-N1clone can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc



Depositor's Name: Rena Quinlan

WURTZEL LAB CLONE INFORMATION

Date Today: _10/__05_/_09 Entered into database yes_490_/ no _
(MONTH)/(DAY)/(YEAR)

CLONE NAME: ChrD_2485 (Bimolecular fluorescence complementation vector with chrD)

Lab Clone Number/Name: ChrD_2485__Alternate name(s) pSAT6(A)-RFP-N1-pchrD-cuc: _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession:_____

Clone Description: The cucumber chrD gene (pchrD-cuc) was PCR-amplified and cloned into the Sal1-BamH1 sites of pSAT6(A)-mRFP-N1 producing pSAT6(A)-RFP-N1-pchrD-cuc. This construct can be used for RFP (Red Fluorescence Protein) expression of unprocessed chrD.

For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131 and Libal-Weksler et al., (1997) Plant Physiology, **113:** 59-63

Further details on pSAT6(A)-RFP-N1-pchrD-cuc can be found at:
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls

Constructed by: _____Dr. Citovsky_____Date rec'd _9/8/09_____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I6____Conc. _0.1 µg/µl
Tube labeled as: __ChrD_2485_____
Strain Location (-80°C) Box Number: _18_____ Position: ____G6_____
Tube labeled as: __ChrD_2485_____ Strain: __TOP10F'_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __9/17/09_____

Original clone name (if different): _____

Organism source of gene: _*Cucumis sativus*_____ Variety:_____

Cloning vector used:__Vector size: 4378 bp__Insert size: ___564bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _Sal1 and BamH1__

Sequence verified: yes_____no___x___ Junction verified: yes_____no___x___

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other: **Further details on pSAT6(A)-RFP-N1-pchrD-cuc can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Depositor's Name: _Rena Quinlan_____

WURTZEL LAB CLONE INFORMATION

Date Today: _10/__05_/_09 Entered into database yes_491_/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: ChrD_3038 (Bimolecular fluorescence complementation vector with chrD)

Lab Clone Number/Name: __ ChrD_3038 __Alternate name(s): pSAT6(A)-RFP-N1-pchrD-cuc _
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession:_____

Clone Description: The cucumber chrD gene (pchrD-cuc) was PCR-amplified and cloned into the Sal1-BamH1 sites of pSAT6(A)-mRFP-N1 producing pSAT6(A)-RFP-N1-pchrD-cuc. This construct can be used for RFP (Red Fluorescence Protein) expression of unprocessed chrD.
Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Further details on pSAT6(A)-RFP-N1-pchrD-cuc can be found at:
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls

Constructed by: _____Dr. Citovsky_____Date rec'd _9/8/09_____
Purified by: _ Rena Quinlan_____
DNA Location (-20°C) Box Number: __6_____Position: _I7____Conc. _0.77 µg/µl
Tube labeled as: __ChrD_3038_____
Strain Location (-80°C) Box Number: _18_____ Position: ____G7_____
Tube labeled as: _ChrD_3038_____ Strain: __TOP10F'

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 9)_____Date: __9/17/09_____

Original clone name (if different): _____

Organism source of gene: _*Cucumis sativus*_____ Variety:_____

Cloning vector used:__Vector size: 4378 bp__Insert size: __564bp____

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _Sal1 and BamH1__

Sequence verified: yes_____no_____ Junction verified: yes_____no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other: **Further details on pSAT6(A)-RFP-N1-pchrD-cuc can be found at:**
http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Depositor's Name: _Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today: 11 / 16 / 05 (MONTH) / (DAY)/ (YEAR) Entered into database yes 395 / no __

CLONE NAME: pCOLADuet-1

Lab Clone Number/Name: ____pCOLADuet-1_____Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _____

Clone Description: pCOLADuet-1(purchased from Novagen) is designed for the coexpression of two genes. This vector contains two MCS (each with a T7 promoter, lac operator, and rbs).

Constructed by: _____Date rec'd_2/26/05__
Purified by: _____
DNA Location (-20°C) Box Number: __6___Position: _A8____Conc. _0.5 µg/µl
Tube labeled as: _pCOLADuet-1_____
Strain Location (-80°C) Box Number: __15_____Position____A8 & A9_____
Tube labeled as: pColaDuet-1/DH5α__ Strain: __DH5α_____

Cited in journal:

Lab Notebook to reference: ___Rena (Book 3)_____Date: ___3/29/05_____

Original clone name (if different): _____

Organism source of gene: _____Variety:_____

Cloning vector used:_____Vector size: _3719 bp__Insert size: _____

Antibiotic markers: ____ amp; _____tet; _____chloramphenicol;__x__other: (_Kan_)

Restriction enzyme(s) to release insert: _____

Sequence verified: yes_____no___x___ Junction verified: _____no___x___

GenBank sheet attached: Yes_____No__x___
Hybridization results attached: Yes____ No__x__
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: __Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today: 6 / 16 / 08 (MONTH) / (DAY)/ (YEAR) Entered into database yes 431 / no __

CLONE NAME: pCDFDuet-1

Lab Clone Number/Name: ____pCDFDuet-1_____Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _____

Clone Description: pCDFDuet-1(purchased from EMD (Novagen) is designed for the coexpression of two genes. This vector contains two MCS (each with a T7 promoter, lac operator, and rbs). The vector also carries the CloDF13-derived CDF replicon, lacI gene and streptomycin/ Spectinomycin resistence gene.

Constructed by: _____Date rec'd_6/10/08__
Purified by: _____
DNA Location (-20°C) Box Number: __6___Position: _H2____Conc. _0.1 µg/µl
Tube labeled as: _pCDFDuet-1_____
Strain Location (-80°C) Box Number: __17_____Position____B3 & B4_____
Tube labeled as: pCDFDuet in TOP10F'__ Strain: __TOP10F'_____

Cited in journal:

Lab Notebook to reference: ___Rena (Book 6)_____Date: ___6/10/08_____

Original clone name (if different): _____

Organism source of gene: _____Variety:_____

Cloning vector used:_____Vector size: _3781 bp__Insert size: _____

Antibiotic markers: ____ amp; _____tet; _____chloramphenicol;__x__other: (Streptomycin)

Restriction enzyme(s) to release insert: _____

Sequence verified: yes_____no___x___ Junction verified: _____no___x___

GenBank sheet attached: Yes_____No__x___
Hybridization results attached: Yes____ No__x__
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today:__01/12_/__ 04_ Entered into database yes_340__ / no ___
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pTnT Vector
Lab Clone Number/Name: _____Alternate name(s): _____ Clone type: Genomic: _____ cDNA: Expression_____x_____Non-expression_____
New GenBank Accession: _____Original GenBank Accession: _AF479322__
Clone Description: The pTnT vector is designed for the expression of cloned genes using *in vitro* expression systems. pTnT contains both SP6 and T7 polymerase promoters which are adjacent to the MCS. Protein from cloned gene of interest can be expressed *in vitro* using an SP6- or T7-based, *in vitro* coupled transcription/translation system. pTnT contains a 5' Beta-globin leader sequence and a poly $(A)_{30}$ tail which should enhance expression. The vector also has a T7 terminator sequence.
Constructed by: _____ Date rec'd_____ Purified by: _____C. Zhu_____ DNA Location (-20°C) Box Number: ___7_____Position: ___D1___Conc. ___0.1__µg/µl Tube labeled as: ____pTnT Vector_____ Strain Location (-80°C) Box Number: _____15_____Position_____E1____ Tube labeled as: ___pTnT_____ Strain: _____Top10F'_____

Cited in journal:

Lab Notebook to reference: _____Zhu_____Date: _____10/16/3_____
Original clone name (if different): _____
Organism source of gene: _____Variety_____
Cloning vector used: _____Vector size: _2871bp__ Insert size: _____
Antibiotic markers: __X__amp; _____tet; ____chloramphenicol;_____other (__)
Restriction enzyme(s) to release insert: _____ _____
Sequence verified: yes__X__no_____ Junction verified: yes___X__no_____
GenBank sheet attached: Yes_____No__X__ Hybridization results attached: Yes____ No____ Other: (attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: C. Zhu

WURTZEL LAB
CLONE INFORMATION

Date Today:__03/03/ 04_ Entered into database yes_326__ / no ___
(MONTH) / (DAY)/ (YEAR)

CLONE NAME: pUC35S-sGFP-Nos

Lab Clone Number/Name: _____Alternate name(s): _____
Clone type: Genomic: _____
cDNA: Expression_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _____

Clone Description: Using BamHI and EcoRI cut pUC35S-GUS-Nos and pBIG121 (Okada et al., 2000, Plant Physiol, 122:1045-1056), then ligation and transformation.

Constructed by: _____C. Zhu_____ Date rec'd_____
Purified by: _____C. Zhu_____
DNA Location (-20^{o}C) Box Number: ___7_____Position: ___D4___Conc. ___0.3__µg/µl
Tube labeled as: ____pUC35S-sGFP-Nos_____
Strain Location (-80^{o}C) Box Number: _____15____Position_____B6_____
Tube labeled as: ___ pUC35S-sGFP-Nos _____ Strain: ____Top 10____

Cited in journal: (Okada et al., 2000, Plant Physiol, 122:1045-1056)

Lab Notebook to reference: _____Zhu_____Date: _____1/4/3_____

Original clone name (if different): _____

Organism source of gene: _____Variety_____

Cloning vector used: _____Vector size: _3500bp__ Insert size: __1000bp___

Antibiotic markers: __X__amp; _____tet; ____chloramphenicol;_____other (__)

Restriction enzyme(s) to release insert: ____BamHI/EcoRI____

Sequence verified: yes__X__no_____ Junction verified: yes___X__no_____

GenBank sheet attached: Yes_____No__X__
Hybridization results attached: Yes____ No____
Other:
(attach additional information here i.e. cartoon, VectorNTI details, sequence)

Depositor's Name: C. Zhu_____

WURTZEL LAB
CLONE INFORMATION

Date Today: _05/__01_/_08 Entered into database yes: 572_/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: pSAT 1476 (pSAT6-cEYFP –N1) (Bimolecular fluorescence complementation vector)

Lab Clone Number/Name: _ pSAT6-cEYFP –N1__Alternate name(s): _1476_____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession:___

Clone Description: This plasmid produces an in-frame fusion of the C-terminus of the protein of interest to the N-terminus of of the C-terminal half of YFP. When co-expressed with fusion of another tested protein to the N-terminal half of YFP produces YFP fluorescence if both proteins interact. Expression cassette can be transferred into PIPspI of ppzp-RCS-based binary vectors. Using these rare cutting nucleases, different combinations of the BiFC expression cassettes can be transferred from the pSATN BiFC vectors into the T-DNA region of Agrobacterium binary plasmids.

For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Further details on clone can be found at: http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls

Constructed by: _____Dr. Citovsky_____Date rec'd _4/30/08_____
Purified by: _____
DNA Location (-20°C) Box Number: __6_____Position: _I8____Conc. _0.2 µg/µl
Tube labeled as: __1476_____
Strain Location (-80°C) Box Number: __17____ Position: ____E4_____
Tube labeled as: ___1476_____ Strain: ___TOP10F'_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 6)_____Date: __4/30/08_____

Original clone name (if different): _____

Organism source of gene: _____ Variety_____

Cloning vector used: pUC18 derivative__Vector size: 4378 bp__Insert size: _792 bp____

Note: The insert includes N-terminus of EYFP (524bp) + MCS (58bp) + CaMV 35S terminator (210bp) (refer to Methods in Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _NcoI/BamHi___

Sequence verified: yes__x___no_____ Junction verified: yes__x____no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other: **Further details on clone can be found at:** http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
(attach additional information here i.e. cartoon, VectorNTI details, sequence)
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Depositor's Name: _Rena Quinlan_____

WURTZEL LAB
CLONE INFORMATION

Date Today: _05/__01_/_08 Entered into database yes_423_/ no ___
(MONTH)/(DAY)/(YEAR)

CLONE NAME: pSAT-2236 (pSAT4(A)-nEYFP–N1) (Bimolecular fluorescence complementation vector)

Lab Clone Number/Name: _ pSAT4(A)-nEYFP –N1__Alternate name(s): _2236_____
Clone type: Genomic: _____
cDNA: Expression_____x_____Non-expression_____

New GenBank Accession: _____Original GenBank Accession: _DQ169003_

Clone Description: Contains the N-terminus of the Yellow Fluorescent Protein (EYFP). This plasmid produces an in-frame fusion of the protein of interest to the N-terminus of EYFP. In the pSATN(A) vectors the N1 fusion MCS lacks the ATG codon, allowing the user to use the tested gene's own ATG as a start codon. Expression cassette can be inserted as an AgeI-NotI fragment into the pSAT4 series of vectors in which the expression cassettes are flanked with a rare-cutting endonuclease (I-SceI). Using these rare cutting nucleases, different combinations of the BiFC expression cassettes can be transferred from the pSATN BiFC vectors into the T-DNA region of Agrobacterium binary plasmids.

For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Further details on clone can be found at: http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls

Constructed by: _____Dr. Citovsky_____Date rec'd _4/30/08_____
Purified by: _____
DNA Location (-20°C) Box Number: __6_____Position: _F3____Conc. _0.2 µg/µl
Tube labeled as: __2236_____
Strain Location (-80°C) Box Number: __16____ Position: ____I5,I6_____
Tube labeled as: ___2236_____ Strain: ___TOP10F'_____

Cited in journal: J. Mol. Biol. (2006) **362**, 1120-1131.

Lab Notebook to reference: __Rena (Book 6)_____Date: __4/30/08_____

Organism source of gene: _____ Variety_____

Cloning vector used: pUC18 derivative__Vector size: 4378 bp__Insert size: _792 bp____

Note: The insert includes N-terminus of EYFP (524bp) + MCS (58bp) + CaMV 35S terminator (210bp) (refer to Methods in Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131

Antibiotic markers: __x__amp; _____tet; _____chloramphenicol;_____other (_____)

Restriction enzyme(s) to release insert: _see note below___

Note: The pSATN(A) vectors do not have the NcoI site (to release insert) and its resident ATG codon at the beginning of the MCS. Refer to Citovsky et al., (2006) J. Mol. Biol. **362**, 1120-1131.

Sequence verified: yes__x___no_____ Junction verified: yes___x___no_____

GenBank sheet attached: Yes_____No_____
Hybridization results attached: Yes____ No____
Other: **Further details on clone can be found at:** http://maize.lehman.cuny.edu/LabResources/CloneBook/pSATvectors/list of pSAT Plasmids ver5.xls
(attach additional information here i.e. cartoon, VectorNTI details, sequence)
For 5′ Junction sequencing use the TL primer: tccttcgcaagacccttcctc

Depositer's Name: Rena Quinlan

Bibliography

Akiyama K, Matsuzaki K-i, Hayashi H (2005) Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. Nature **435:** 824-827

Al-Babili S, von Lintig J, Haubruck H, Beyer P (1996) A novel, soluble form of phytoene desaturase from *Narcissus pseudonarcissus* chromoplasts is Hsp70-complexed and competent for flavinylation, membrane association and enzymatic activation. Plant J **9:** 601-612

Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, Ghisla S, Bouwmeester H, Beyer P, Al-Babili S (2012) The path from beta-carotene to carlactone, a strigolactone-like plant hormone. Science **335:** 1348-1351

Alvarez V, Rodriguez-Saiz M, de la Fuente JL, Gudina EJ, Godio RP, Martin JF, Barredo JL (2006) The crtS gene of *Xanthophyllomyces dendrorhous* encodes a novel cytochrome-P450 hydroxylase involved in the conversion of [beta]-carotene into astaxanthin and other xanthophylls. Fungal Genetics and Biology **43:** 261-272

Arnoux P, Morosinotto T, Saga G, Bassi R, Pignol D (2009) A structural basis for the pH-dependent xanthophyll cycle in Arabidopsis thaliana. Plant Cell **21:** 2036-2044

Auldridge ME, McCarty DR, Klee HJ (2006) Plant carotenoid cleavage oxygenases and their apocarotenoid products. Current Opinion in Plant Biology **9:** 315-321

Axarli I, Prigipaki A, Labrou NE (2005) Engineering the substrate specificity of cytochrome P450 CYP102A2 by directed evolution: production of an efficient enzyme for bioconversion of fine chemicals. Biomolecular Engineering **22:** 81-88

Bak S, Kahn RA, Nielsen HL, Møller BL, Halkier BA (1998) Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor* (L.) Moench by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside dhurrin. Plant Molecular Biology **36:** 393-405

Baltrusch M, Fulda M, Wolter F-P, Heinz E (1997) Cloning and sequencing of a cytochrome P450 from *Pisum sativum* L. (Accession No. Z49263). Plant Physiol. **114:** 1568

Bartley GE, Scolnik PA, Beyer P (1999) Two Arabidopsis thaliana carotene desaturases, phytoene desaturase and zeta-carotene desaturase, expressed in Escherichia coli, catalyze a poly-cis pathway to yield pro-lycopene. Eur J Biochem **259:** 396-403

Bartley GE, Viitanen PV, Pecker I, Chamovitz D, Hirschberg J, Scolnik PA (1991) Molecular cloning and expression in photosynthetic bacteria of a soybean cDNA coding for phytoene desaturase, an enzyme of the carotenoid biosynthesis pathway. Proc Natl Acad Sci U S A **88:** 6532-6536

Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts. Journal of Biotechnology **124:** 128-145

Bertrand E, Sakai R, Rozhkova-Novosad E, Moe L, Fox BG, Groves JT, Austin RN (2005) Reaction mechanisms of non-heme diiron hydroxylases characterized in whole cells. Journal of Inorganic Biochemistry **99:** 1998-2006

Beyer P, Mayer M, Kleinig H (1989) Molecular oxygen and the state of geometric isomerism of intermediates are essential in the carotene desaturation and cyclization reactions in daffodil chromoplasts. Eur J Biochem **184:** 141-150

Blasco F, Kauffmann I, Schmid RD (2004) CYP175A1 from Thermus thermophilus HB27, the first beta-carotene hydroxylase of the P450 superfamily. Appl Microbiol Biotechnol **64:** 671-674

Bonk M, Hoffmann B, Von Lintig J, Schledz M, Al-Babili S, Hobeika E, Kleinig H, Beyer P (1997) Chloroplast import of four carotenoid biosynthetic enzymes in vitro reveals differential fates prior to membrane binding and oligomeric assembly. Eur J Biochem **247:** 942-950

Booker J, Auldridge M, Wills S, McCarty D, Klee H, Leyser O (2004) MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. Curr Biol **14:** 1232-1238

Bouvier F, Backhaus RA, Camara B (1998) Induction and control of chromoplast-specific carotenoid genes by oxidative stress. J Biol Chem **273:** 30651-30659

Bouvier F, Keller Y, d'Harlingue A, Camara B (1998) Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum* L.). Biochim Biophys Acta **1391:** 320-328

Bouvier F, Suire C, Mutterer J, Camara B (2003) Oxidative remodeling of chromoplast carotenoids: Identification of the carotenoid dioxygenase CsCCD and CsZCD genes involved in Crocus secondary metabolite biogenesis. Plant Cell **15:** 47-62

Britton G, Liaaen-Jensen S, Pfander H (1995) Carotenoids- Volume 1A: Isolation and Analysis. Birkhäuser, Basel

Britton G, Liaaen-Jensen S, Pfander H, eds (2004) Carotenoids Handbook. Birkhäuser Verlag, Basel

Britton G, Liaaen-Jensen S, Pfander H (2004) Carotenoids Handbook. Birkhauser Verlag, Basel, Switzerland

Bruce BD, Perry S, Froehlich J, Keegstra K (1994) In vitro import of protein into chloroplasts. *In* Plant Molecular Biology Manual. Kluwer Academic Publishers, Boston, pp 1-15

Bugos RC, Yamamoto HY (1996) Molecular cloning of violaxanthin de-epoxidase from romaine lettuce and expression in Escherichia coli. PNAS **93:** 6320-6325

Camara B, Bardat F, Moneger R (1982) Sites of biosynthesis of carotenoids in *Capsicum* chromoplasts. Eur J Biochem **127:** 255-258

Cao FY, Yoshioka K, Desveaux D (2011) The roles of ABA in plant-pathogen interactions. J Plant Res **124:** 489-499

Carde JP, Joyard J, Douce R (1982) Electron microscopic studies of envelope membranes from spinach plastids. Biol Cell **44:** 315-324

Carretero-Paulet L, Cairo A, Botella-Pavia P, Besumbes O, Campos N, Boronat A, Rodriguez-Concepcion M (2006) Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. Plant Mol Biol **62:** 683-695

Carrillo N, Ceccarelli EA (2003) Open questions in ferredoxin-NADP+ reductase catalytic mechanism. Eur. J. Biochem. **270:** 1900-1915

Castillo R, Fernandez J-A, Gomez-Gomez L (2005) Implications of carotenoid biosynthetic genes in apocarotenoid formation during the stigma development of *Crocus sativus a*nd Its closer relatives. Plant Physiol. **139:** 674-689

Castillo R, Fernandez J-A, Gomez-Gomez L (2005) Implications of carotenoid biosynthetic genes in apocarotenoid formation during the stigma development of *Crocus sativus* and its closer relatives. Plant Physiol. **139:** 674-689

Ceccarelli EA, Arakaki AK, Cortez N, Carrillo N (2004) Functional plasticity and catalytic efficiency in plant and bacterial ferredoxin-NADP(H) reductases. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics **1698:** 155-165

Chapple C (1998) Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. Annu Rev Plant Physiol Plant Mol Biol **49:** 311-343

Chen D, Schnell DJ (1997) Insertion of the 34-kDa chloroplast protein import component, IAP34, into the chloroplast outer membrane is dependent on its intrinsic GTP-binding capacity. J Biol Chem **272:** 6614-6620

Chen Y, Li F, Wurtzel ET (2010) Isolation and characterization of the Z-ISO gene encoding a missing component of carotenoid biosynthesis in plants. Plant Physiol **153:** 66-79

Citovsky V, Lee LY, Vyas S, Glick E, Chen MH, Vainstein A, Gafni Y, Gelvin SB, Tzfira T (2006) Subcellular localization of interacting proteins by bimolecular fluorescence complementation in planta. J Mol Biol **362:** 1120-1131

Cunningham FX, Gantt E (1998) Genes and Enzymes of Carotenoid Biosynthesis in Plants. Annu Rev Plant Physiol Plant Mol Biol **49:** 557-583

Cunningham FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **49:** 557

Cunningham FX, Jr., Chamovitz D, Misawa N, Gantt E, Hirschberg J (1993) Cloning and functional expression in Escherichia coli of a cyanobacterial gene for lycopene cyclase, the enzyme that catalyzes the biosynthesis of beta-carotene. FEBS Lett **328:** 130-138

Cunningham FX, Jr., Lee H, Gantt E (2007) Carotenoid biosynthesis in the primitive red alga Cyanidioschyzon merolae. Eukaryot Cell **6:** 533-545

Cunningham FX, Jr., Pogson B, Sun Z, McDonald KA, DellaPenna D, Gantt E (1996) Functional analysis of the beta and epsilon lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. Plant Cell **8:** 1613-1626

Cunningham Jr. FX, Chamovitz D, Misawa N, Gantt E, Hirschberg J (1993) Cloning and functional expression in *Escherichia coli* of a cyanobacterial gene for lycopene cyclase, the enzyme that catalyzes the biosynthesis of beta-carotene. FEBS Lett **328:** 130-138

Cunningham Jr. FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. *In* Annu. Rev. Plant Physiol. Plant Mol. Biol., Vol 49, pp 557-583

Cunningham Jr. FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol **49:** 557-583

Cunningham Jr. FX, Gantt E (2001) One ring or two? Determination of ring number in carotenoids by lycopene epsilon-cyclases. Proc Natl Acad Sci USA **98:** 2905-2910

Cunningham Jr. FX, Pogson B, Sun Z, McDonald KA, DellaPenna D, Gantt E (1996) Functional analysis of the beta and epsilon lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. Plant Cell **8:** 1613-1626

Cutler AJ, Krochko JE (1999) Formation and breakdown of ABA. Trends Plant Sci **4:** 472-478

Cuttriss AJ, Cazzonelli CI, Wurtzel ET, Pogson BJ (2011) "Carotenoids". *In* Biosynthesis of Vitamins in Plants. eds. F. Rebeille &
R. Douce. Elsevier.

Cuttriss AJ, Cazzonelli, C.I., Wurtzel, E.T. and Pogson, B.J. (2011) Carotenoids. *In* F Rébeillé, R Douce, eds, Biosynthesis of Vitamins in Plants. Elsevier

Dall'Osto L, Fiore A, Cazzaniga S, Giuliano G, Bassi R (2007) Different roles of α- and β-branch xanthophylls in photosystem assembly and photoprotection. J. Biol. Chem. **282:** 35056-35068

Dall'Osto L, Fiore A, Cazzaniga S, Giuliano G, Bassi R (2007) Different roles of alpha- and beta-branch xanthophylls in photosystem assembly and photoprotection. J Biol Chem **282:** 35056-35068

Dall'Osto L, Lico C, Alric J, Giuliano G, Havaux M, Bassi R (2006) Lutein is needed for efficient chlorophyll triplet quenching in the major LHCII antenna complex of higher plants and effective photoprotection in vivo under strong light. BMC Plant Biol **6:** 32

Danielson PB (2002) The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. Curr Drug Metab **3:** 561-597

Davison PA (2002) Overexpression of beta-carotene hydroxylase enhances stress tolerance in *Arabidopsis.* Nature **418:** 203-206

DellaPenna D, Pogson BJ (2006) Vitamin Synthesis in Plants: Tocopherols and Carotenoids. Annual Review of Plant Biology **57:** 711-738

Douce R (1974) Site of galactolipid synthesis in spinach chloroplasts. Science **183:** 852-853

Ducreux LJ, Morris WL, Hedley PE, Shepherd T, Davies HV, Millam S, Taylor MA (2005) Metabolic engineering of high carotenoid potato tubers containing enhanced levels of beta-carotene and lutein. J Exp Bot **56:** 81-89

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci **8:** 978-984

Fiore A, Dall'osto L, Fraser PD, Bassi R, Giuliano G (2006) Elucidation of the beta-carotene hydroxylation pathway in Arabidopsis thaliana. FEBS Lett **580:** 4718-4722

Fiore A, Dall'Osto L, Fraser PD, Bassi R, Giuliano G (2006) Elucidation of the [beta]-carotene hydroxylation pathway in Arabidopsis thaliana. FEBS Letts. **580:** 4718-4722

Floss DS, Walter MH (2009) Role of carotenoid cleavage dioxygenase 1 (CCD1) in apocarotenoid biogenesis revisited. Plant Signal Behav **4:** 172-175

Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. Progress in Lipid Research **43:** 228-265

Gallagher CE, Matthews PD, Li F, Wurtzel ET (2004) Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses (Poaceae). Plant Physiol. **135:** 1776-1783

Galpaz N, Ronen G, Khalfa Z, Zamir D, Hirschberg J (2006) A chromoplast-specific carotenoid biosynthesis pathway is revealed by cloning of the tomato white-flower locus. Plant Cell **18:** 1947-1960

Garcia-Limones C, Schnabele K, Blanco-Portales R, Luz Bellido M, Caballero JL, Schwab W, Munoz-Blanco J (2008) Functional characterization of FaCCD1: a carotenoid cleavage dioxygenase from strawberry involved in lutein degradation during fruit ripening. J Agric Food Chem **56:** 9277-9285

Goss R, Opitz C, Lepetit B, Wilhelm C (2008) The synthesis of NPQ-effective zeaxanthin depends on the presence of a transmembrane proton gradient and a slightly basic stromal side of the thylakoid membrane. Planta **228:** 999-1009

Greenwood AD, Leech RM, Williams JP (1963) The osmiophilic globules of chloroplasts. I. Osmiophilic globules as a normal component of chloroplasts and their isolation and composition in *Vicia faba* L. Biochim Biophys Acta **78:** 148-162

Hanley BA, Schuler MA (1988) Plant intron sequences: evidence for distinct groups of introns. Nucleic Acids Res **16:** 7159-7176

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science **319:** 330-333

Henne A, Bruggemann H, Raasch C, Wiezer A, Hartsch T, Liesegang H, Johann A, Lienard T, Gohl O, Martinez-Arias R, Jacobi C, Starkuviene V, Schlenczeck S, Dencker S, Huber R, Klenk H-P, Kramer W, Merkl R, Gottschalk G, Fritz H-J (2004) The genome sequence of the extreme thermophile Thermus thermophilus. Nat Biotech **22:** 547-553

Hentschel V, Kranl K, Hollmann J, Lindhauer MG, Bohm V, Bitsch R (2002) Spectrophotometric determination of yellow pigment content and evaluation

of carotenoids by high-performance liquid chromatography in durum wheat grain. J Agric Food Chem **50:** 6663-6668

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol **4:** 210-218

Hoober JK, Boyd CO, Paavola LG (1991) Origin of Thylakoid Membranes in Chlamydomonas reinhardtii y-1 at 38 degrees C. Plant Physiol **96:** 1321-1328

Howitt CA, Cavanagh CR, Bowerman AF, Cazzonelli C, Rampling L, Mimica JL, Pogson BJ (2009) Alternative splicing, activation of cryptic exons and amino acid substitutions in carotenoid biosynthetic genes are associated with lutein accumulation in wheat endosperm. Funct Integr Genomics **9:** 363-376

Ilg A, Yu Q, Schaub P, Beyer P, Al-Babili S (2010) Overexpression of the rice carotenoid cleavage dioxygenase 1 gene in Golden Rice endosperm suggests apocarotenoids as substrates in planta. Planta **232:** 691-699

Inoue K (2004) Carotenoid hydroxylation - P450 finally! Trends in Plant Science **9:** 515-517

Inoue K, Furbee K.J., Uratsu, S., Masaya K., Dandekar A.M., Ikoma Y. (2006) Catalytic activities and chloroplast import of carotenogenic enzymes from citrus. Physiologia Plantarum **127:** 561-570

Inoue K, Furbee KJ, Uratsu S, Kato M, Dandekar AM, Ikoma Y (2006) Catalytic activities and chloroplast import of carotenogenic enzymes from citrus. Physiologia Plantarum **127:** 561-570

Isaacson T, Ohad I, Beyer P, Hirschberg J (2004) Analysis in vitro of the enzyme CRTISO establishes a poly-cis-carotenoid biosynthesis pathway in plants. Plant Physiol **136:** 4246-4255

Isaacson T, Ronen G, Zamir D, Hirschberg J (2002) Cloning of *tangerine* from tomato reveals a carotenoid isomerase essential for the production of β–carotene and xanthophylls in plants. Plant Cell **14:** 333-342

Isaacson T, Ronen G, Zamir D, Hirschberg J (2002) Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants. Plant Cell **14:** 333-342

Islam SN (2004) Survey of carotenoid variation and quantitative trait loci mapping for carotenoid and tocopherol variation in maize. Masters. University of Illinois at Urbana-Champaign, Urbana-Champaign

Jahns P, Holzwarth AR (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. Biochim Biophys Acta **1817:** 182-193

Jahns P, Holzwarth AR (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. Biochimica et Biophysica Acta (BBA) - Bioenergetics **1817:** 182-193

Jenkins C, Waterman M (1994) Flavodoxin and NADPH-flavodoxin reductase from *Escherichia coli* support bovine cytochrome P450c17 hydroxylase activities. J. Biol. Chem. **269:** 27401-27408

Joyard J, Douce R, Siebertz HP, Heinz E (1980) Distribution of radioactive lipids between envelopes and thylakoids from chloroplasts labelled in vivo. Eur J Biochem **108:** 171-176

Joyard J, Ferro M, Masselon C, Seigneurin-Berny D, Salvi D, Garin J, Rolland N (2009) Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways. Mol Plant **2:** 1154-1180

Kato M, Ikoma Y, Matsumoto H, Sugiura M, Hyodo H, Yano M (2004) Accumulation of carotenoids and expression of carotenoid biosynthetic genes during maturation in citrus fruit. Plant Physiol **134:** 824-837

Kessler F, Schnell D, Blobel G (1999) Identification of proteins associated with plastoglobules isolated from pea (*Pisum sativum* L.) chloroplasts. Planta **208:** 107-113

Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, Yamada H, Ooka H, Hotta I, Kojima K, Namiki T, Ohneda E, Yahagi W, Suzuki K, Li CJ, Ohtsuki K, Shishiki T, Otomo Y, Murakami K, Iida Y, Sugano S, Fujimura T, Suzuki Y, Tsunoda Y, Kurosaki T, Kodama T, Masuda H, Kobayashi M, Xie Q, Lu M, Narikawa R, Sugiyama A, Mizuno K, Yokomizo S, Niikura J, Ikeda R, Ishibiki J, Kawamata M, Yoshimura A, Miura J, Kusumegi T, Oka M, Ryu R, Ueda M, Matsubara K, Kawai J, Carninci P, Adachi J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Hayatsu N, Imotani K, Ishii Y, Itoh M, Kagawa I, Kondo S, Konno H, Miyazaki A, Osato N, Ota Y, Saito R, Sasaki D, Sato K, Shibata K, Shinagawa A, Shiraki T, Yoshino M, Hayashizaki Y, Yasunishi A (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. Science **301:** 376-379

Kim J, DellaPenna D (2006) Defining the primary route for lutein synthesis in plants: the role of Arabidopsis carotenoid beta-ring hydroxylase CYP97A3. Proc Natl Acad Sci U S A **103:** 3474-3479

Kim J, DellaPenna D (2006) Defining the primary route for lutein synthesis in plants: The role of Arabidopsis carotenoid beta-ring hydroxylase CYP97A3. Proc Natl Acad Sci USA **103:** 3474-3479

Kim J, Smith JJ, Tian L, Dellapenna D (2009) The evolution and function of carotenoid hydroxylases in Arabidopsis. Plant Cell Physiol **50:** 463-479

Kopsell DA, Kopsell DE (2006) Accumulation and bioavailability of dietary carotenoids in vegetable crops. Trends in Plant Science **11:** 499-507

Kozak M (1987) An analysis of 5--noncoding sequence from 699 vertebrate messenger RNAs. Nucleic Acids Research **15:** 8125-8148

Krapp AR, Rodriguez RE, Poli HO, Paladini DH, Palatnik JF, Carrillo N (2002) The flavoenzyme ferredoxin (flavodoxin)-NADP(H) reductase modulates NADP(H) homeostasis during the *soxRS* response of *Escherichia coli*. J. Bacteriol. **184:** 1474-1480

Kreuz K, Beyer P, Kleinig H (1982) The site of carotenogenic enzymes in chromoplasts from *Narcissus pseudonarcissus* L. Planta **154:** 66-69

Kroll D, Meierhoff K, Bechtold N, Kinoshita M, Westphal S, Vothknecht UC, Soll J, Westhoff P (2001) VIPP1, a nuclear gene of Arabidopsis thaliana

essential for thylakoid membrane formation. Proc Natl Acad Sci U S A **98:** 4238-4242

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *In* Bioinformatics, Vol 17, pp 1244-1245

Kurilich A, Juvik J (1999) Quantification of carotenoid and tocopherol antioxidants in *Zea mays*. J. Agric. Food Chem. **47:** 1948 -1955

Li F, Murillo C, Wurtzel ET (2007) Maize Y9 encodes a product essential for 15-cis-zeta-carotene isomerization. Plant Physiol **144:** 1181-1189

Li F, Vallabhaneni R, Wurtzel ET (2008a) PSY3, a new member of the phytoene synthase gene family conserved in the Poaceae and regulator of abiotic stress-induced root carotenogenesis. Plant Physiol **146:** 1333-1345

Li F, Vallabhaneni R, Yu J, Rocheford T, Wurtzel ET (2008b) The maize phytoene synthase gene family: overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. Plant Physiol **147:** 1334-1346

Li L, Paolillo DJ, Parthasarathy MV, Dimuzio EM, Garvin DF (2001) A novel gene mutation that confers abnormal patterns of beta-carotene accumulation in cauliflower (Brassica oleracea var. botrytis). Plant J **26:** 59-67

Li Z, Wakao S, Fischer BB, Niyogi KK (2009) Sensing and responding to excess light. Annu Rev Plant Biol **60:** 239-260

Liang C, Zhao F, Wei W, Wen Z, Qin S (2006) Carotenoid biosynthesis in cyanobacteria: structural and evolutionary scenarios based on comparative genomics. Int J Biol Sci **2:** 197-207

Libal-Weksler Y, Vishnevetsky M, Ovadis M, Vainstein A (1997) Isolation and regulation of accumulation of a minor chromoplast-specific protein from cucumber corollas

Plant Physiology **113:** 59-63

Lichtenthaler HK, Rohmer M, Schwender J (1997) Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants Physiologia Plantarum **101:** 643-652

Linden H, Misawa N, Chamovitz D, Pecker I, Hirschberg J, Sandmann G (1991) Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenoids. Z. Naturforsch. (C) **46c:** 1045-1051

Liochev S, Hausladen A, Beyer W, Jr, Fridovich I (1994) NADPH:ferredoxin oxidoreductase acts as a paraquat diaphorase and is a member of the *soxRS* regulon. PNAS **91:** 1328-1331

Lokstein H, Tian L, Polle JE, DellaPenna D (2002) Xanthophyll biosynthetic mutants of Arabidopsis thaliana: altered nonphotochemical quenching of chlorophyll fluorescence is due to changes in Photosystem II antenna size and stability. Biochim Biophys Acta **1553:** 309-319

Lokstein H, Tian L, Polle JEW, DellaPenna D (2002) Xanthophyll biosynthetic mutants of Arabidopsis thaliana: Altered nonphotochemical quenching of

chlorophyll fluorescence is due to changes in Photosystem II antenna size and stability. Biochim. Biophys. Acta **1553:** 309

Lopez AB, Yang Y, Thannhauser TW, Li L (2008) Phytoene desaturase is present in a large protein complex in the plastid membrane. Physiologia Plantarum **133:** 190-198

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J, Kochian LV, Kupper H, Earle ED, Cao J, Li L (2006) The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. Plant Cell **18:** 3594-3605

Lutzow M, Beyer P (1988) the isopentyl diphosphate delta-isomerase and its relation ot the phytoene synthase complex of daffodil chromoplasts. Biochim. Biophys. Acta **959:** 118-126

Mann V, Harker M, Pecker I, Hirschberg J (2000) Metabolic engineering of astaxanthin production in tobacco flowers. Nat Biotechnol **18:** 888-892

Mann V, Harker M, Pecker I, Hirschberg J (2000) Metabolic engineering of astaxanthin production in tobacco flowers. Nat Biotechnol **18:** 888-892

Marasco EK, Vay K, Schmidt-Dannert C (2006) Identification of carotenoid cleavage dioxygenases from Nostoc sp. PCC 7120 with different cleavage activities. J Biol Chem **281:** 31583-31593

Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH (2005) CDD: a Conserved Domain Database for protein classification. Nucl. Acids Res. **33:** D192-196

Marques JP, Schattat MH, Hause G, Dudeck I, Klosgen RB (2003) *In vivo* transport of folded EGFP by the ΔpH/TAT-dependent pathway in chloroplasts of *Arabidopsis thaliana*

Journal of Experimental Botany **55:** 1697-1706

Matthews PD, Luo R, Wurtzel ET (2003) Maize phytoene desaturase and zeta-carotene desaturase catalyse a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. J Exp Bot **54:** 2215-2230

Matthews PD, Wurtzel ET (2000) Metabolic engineering of carotenoid accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase. Appl Microbiol Biotechnol **53:** 396-400

Matthews PD, Wurtzel ET (2007) Biotechnology of Food Colorant Production. *In* C Socaciu, ed, Food Colorants: Chemical and Functional Properties. CRC Press, Boca Raton

Matusova R, Rani K, Verstappen FWA, Franssen MCR, Beale MH, Bouwmeester HJ (2005) The strigolactone germination stimulants of the

plant-parasitic *Striga* and *Orobanche* spp. are derived from the carotenoid pathway. Plant Physiol. **139:** 920-934

Maudinas B, Bucholtz ML, Papastephanou C, Katiyar SS, Briedis AV, Porter JW (1977) The partial purification and properties of a phytoene synthesizing enzyme system. Arch Biochem Biophys **180:** 354-362

Meier S, Tzfadia O, Vallabhaneni R, Gehring C, Wurtzel ET (2011) A transcriptional analysis of carotenoid, chlorophyll and plastidial isoprenoid biosynthesis genes during development and osmotic stress responses in Arabidopsis thaliana. BMC Syst Biol **5:** 77

Merzlyak MN, Solovchenko AE (2002) Photostability of pigments in ripening apple fruit: A possible photoprotective role of carotenoids during plant senescence. Plant Science **163:** 881-888

Mikkelsen MD, Hansen CH, Wittstock U, Halkier BA (2000) Cytochrome P450 CYP79B2 from Arabidopsis catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. J. Biol. Chem. **275:** 33712-33717

Milborrow BV (2001) The pathway of biosynthesis of abscisic acid in vascular plants: a review of the present state of knowledge of ABA biosynthesis. J Exp Bot **52:** 1145-1164.

Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, Nakamura K, Harashima K (1990) Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli. J Bacteriol **172:** 6704-6712

Moise AR, von Lintig J, Palczewski K (2005) Related enzymes solve evolutionarily recurrent problems in the metabolism of carotenoids. Trends in Plant Science **10:** 178-186

Morre DJ, Sellden G, Sundqvist C, Sandelius AS (1991) Stromal low temperature compartment derived from the inner membrane of the chloroplast envelope. Plant Physiol **97:** 1558-1564

Muhlethaler K, Frey-Wyssling A (1959) Entwicklung und Struktur der Propastiden. J. Biophys. Biochem. Cytol. **6:** 507-512

Naur P, Hansen CH, Bak S, Hansen BG, Jensen NB, Nielsen HL, Halkier BA (2003) CYP79B1 from *Sinapis alba* converts tryptophan to indole-3-acetaldoxime. Archives of Biochemistry and Biophysics **409:** 235-241

Nelson DR, Schuler MA, Paquette SM, Werck-Reichhart D, Bak S (2004) Comparative genomics of Rice and *Arabidopsis*. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. Plant Physiol. **135:** 756-772

Okada K, Saito T, Nakagawa T, Kawamukai M, Kamiya Y (2000) Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in Arabidopsis. Plant Physiol **122:** 1045-1056

Orihara N, Furihata K, Seto H (1997) Studies on the biosynthesis of terpenoidal compounds produced by actinomycetes. 2. Biosynthesis of carquinostatin B

via the non-mevalonate pathway in Streptomyces exfoliatus. J Antibiot (Tokyo) **50:** 979-981

Park H, Kreunen SS, Cuttriss AJ, DellaPenna D, Pogson BJ (2002) Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis. Plant Cell **14:** 321-332

Parker GD, Chalmers KJ, Rathjen AJ, Langridge P (1998) Mapping loci associated with flour colour in wheat (*Triticum aestivum* L.). Theor. Appl. Genetics **97:** 238-245

Parry AD, Horgan R (1992) Abscisic acid biosynthesis in roots I. The identification of potential abscisic acid precursors, and other carotenoids. Planta **187:** 185-191

Pogson B, McDonald KA, Truong M, Britton G, DellaPenna D (1996) Arabidopsis carotenoid mutants demonstrate that lutein is not essential for photosynthesis in higher plants. Plant Cell **8:** 1627-1639

Pogson BJ, Niyogi KK, Bjorkman O, DellaPenna D (1998) Altered xanthophyll compositions adversely affect chlorophyll accumulation and nonphotochemical quenching in Arabidopsis mutants. Proc. Nat. Acad. Sci. USA **95:** 13324

Pogson BJ, Niyogi KK, Bjorkman O, DellaPenna D (1998) Altered xanthophyll compositions adversely affect chlorophyll accumulation and nonphotochemical quenching in Arabidopsis mutants. Proc Natl Acad Sci U S A **95:** 13324-13329

Pogson BJ, Rissler HM (2000) Genetic manipulation of carotenoid biosynthesis and photoprotection. Philos Trans R Soc Lond B Biol Sci **355:** 1395-1403.

Qin X, Zeevaart JA (1999) The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. Proc Natl Acad Sci USA **96:** 15354-15361

Quinlan R, Jaradat T, Wurtzel ET (2007) *Escherichia coli* as a platform for functional expression of plant P450 carotene hydroxylases. Arch. Biochem. Biophysics **458:** 146-157

Quinlan RF, Jaradat TT, Wurtzel ET (2007) Escherichia coli as a platform for functional expression of plant P450 carotene hydroxylases. Arch Biochem Biophys **458:** 146-157

Quinlan RF, Jaradat TT, Wurtzel ET (2007) Escherichia coli as a platform for functional expression of plant P450 carotene hydroxylases. Arch. Biochem. Biophys. **458:** 146-157

Rodriguez-Concepcion M (2010) Supply of precursors for carotenoid biosynthesis in plants. Arch Biochem Biophys **504:** 118-122

Rodriguez-Concepcion M, Boronat A (2002) Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiol **130:** 1079-1089

Rosati C, Aquilani R, Dharmapuri S, Pallara P, Marusic C, Tavazza R, Bouvier F, Camara B, Giuliano G (2000) Metabolic engineering of beta-carotene and lycopene content in tomato fruit. Plant J **24:** 413-419.

Rothnie HM (1996) Plant mRNA 3'-end formation. Plant Mol Biol **32:** 43-61

Ryle MJ, Hausinger RP (2002) Non-heme iron oxygenases. Curr Opin Chem Biol **6:** 193-201

Sandmann G, Albrecht M, Schnurr G, Knorzer O, Boger P (1999) The biotechnological potential and design of novel carotenoids by gene combination in *Escherichia coli*. Trends Biotechnol **17:** 233-237

Schledz M, Al-Babili S, von Lintig J, Haubruck H, Rabbani S, Kleinig H, Beyer P (1996) Phytoene synthase from *Narcissus pseudonarcissus*: functional expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering. Plant J **10:** 781-792

Schmidt H, Kurtzer R, Eisenreich W, Schwab W (2006) The Carotenase AtCCD1 from Arabidopsis thaliana Is a Dioxygenase
10.1074/jbc.M511668200. J. Biol. Chem. **281:** 9845-9851

Schuler MA, Werck-Reichhart D (2003) Functional genomics of P450s. Annual Review of Plant Biology **54:** 629-667

Schwartz SH, Qin X, Loewen MC (2004) The biochemical characterization of two carotenoid cleavage enzymes from Arabidopsis indicates that a carotenoid-derived compound Inhibits lateral branching. J Biol Chem **279:** 46940-46945

Schwartz SH, Tan BC, Cage DA, Zeevaart JAD, McCarty DR (1997) Specific Oxidative Cleavage of Carotenoids by VP14 of maize. Science **276:** 1872-1874

Schwender J, Seemann M, Lichtenthaler HK, Rohmer M (1996) Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga Scenedesmus obliquus. Biochem J **316 (Pt 1):** 73-80

Seemann M, Tse Sum Bui B, Wolff M, Miginiac-Maslow M, Rohmer M (2006) Isoprenoid biosynthesis in plant chloroplasts via the MEP pathway: direct thylakoid/ferredoxin-dependent photoreduction of GcpE/IspG. FEBS Lett **580:** 1547-1552

Sheen J (1991) Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. Plant Cell **3:** 225-245

Shewmaker CK, Sheehy JA, Daley M, Colburn S, Ke DY (1999) Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. Plant J **20:** 401-412

Siefermann-Harms D, Hertzberg S, Borch G, Liaaen-Jensen S (1981) Lactucaxanthin, an ε,ε-carotene-3,3'-diol from *Lactuca sativa*. Phytochemistry **20:** 85-88

Siminszky B, Corbin FT, Ward ER, Fleischmann TJ, Dewey RE (1999) Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast

and tobacco enhances the metabolism of phenylurea herbicides. *In* Proc Natl Acad Sci U S A, Vol 96, pp 1750-1755

Simkin AJ, Schwartz SH, Auldridge M, Taylor MG, Klee HJ (2004) The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles -ionone, pseudoionone, and geranylacetone. The Plant Journal **40:** 882-892

Sommer A, Davidson FR (2002) Assessment and control of vitamin A deficiency: the Annecy Accords. J Nutr **132:** 2845S-2850S

Sun Z, Gantt E, Cunningham FX, Jr. (1996) Cloning and functional analysis of the beta-carotene hydroxylase of Arabidopsis thaliana. J Biol Chem **271:** 24349-24352

Sun Z, Gantt E, Cunningham FX, Jr. (1996) Cloning and functional analysis of the beta -carotene hydroxylase of *Arabidopsis thaliana*. J. Biol. Chem. **271:** 24349-24352

Sun Z, Hans J, Walter MH, Matusova R, Beekwilder J, Verstappen FW, Ming Z, van Echtelt E, Strack D, Bisseling T, Bouwmeester HJ (2008) Cloning and characterisation of a maize carotenoid cleavage dioxygenase (ZmCCD1) and its involvement in the biosynthesis of apocarotenoids with various roles in mutualistic and parasitic interactions. Planta **228:** 789-801

Tan BC, Cline K, McCarty DR (2001) Localization and targeting of the VP14 epoxy-carotenoid dioxygenase to chloroplast membranes. Plant J **27:** 373-382

Tan BC, Schwartz SH, Zeevaart JA, McCarty DR (1997) Genetic control of abscisic acid biosynthesis in maize. Proc Natl Acad Sci USA **94:** 12235-12240

Tevini M, Steinmuller D (1985) Composition and function of plastoglobuli. II. Lipid composition of leaves and plastoglobuli during senescence. Planta **163:** 91-96

Tian L, DellaPenna D (2001) Characterization of a second carotenoid β-hydroxylase gene from Arabidopsis and its relationship to the *LUT1* locus. Plant Molecular Biology **47:** 379-388

Tian L, DellaPenna D (2001) Characterization of a second carotenoid beta-hydroxylase gene from Arabidopsis and its relationship to the LUT1 locus. Plant Mol Biol **47:** 379-388

Tian L, DellaPenna D (2004) Progress in understanding the origin and functions of carotenoid hydroxylases in plants. Arch Biochem Biophys **430:** 22-29

Tian L, DellaPenna D, J.A.D. Z (2004) Effect of hydroxylated carotenoid deficiency on ABA accumulation in Arabidopsis. Physiologia Plantarum **122:** 314-320

Tian L, Magallanes-Lundback M, Musetti V, DellaPenna D (2003) Functional analysis of beta- and epsilon-ring carotenoid hydroxylases in Arabidopsis. Plant Cell **15:** 1320

Tian L, Magallanes-Lundback M, Musetti V, DellaPenna D (2003) Functional analysis of beta- and epsilon-ring carotenoid hydroxylases in Arabidopsis. Plant Cell **15:** 1320-1332

Tian L, Magallanes-Lundback M, Musetti V, DellaPenna D (2003) Functional analysis of beta- and epsilon-ring carotenoid hydroxylases in Arabidopsis. Plant Cell **15:** 1320-1332

Tian L, Musetti V, Kim J, Magallanes-Lundback M, DellaPenna D (2004) The Arabidopsis *LUT1* locus encodes a member of the cytochrome P450 family that is required for carotenoid ε-ring hydroxylation activity. PNAS **101:** 402-407

Tian L, Musetti V, Kim J, Magallanes-Lundback M, DellaPenna D (2004) The Arabidopsis LUT1 locus encodes a member of the cytochrome p450 family that is required for carotenoid epsilon-ring hydroxylation activity. Proc Natl Acad Sci U S A **101:** 402-407

Toledo-Ortiz G, Huq E, Rodriguez-Concepcion M (2010) Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. Proc Natl Acad Sci U S A **107:** 11626-11631

Tusnady GE, Simon I (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol **283:** 489-506

Vallabhaneni R, Bradbury LM, Wurtzel ET (2010) The carotenoid dioxygenase gene family in maize, sorghum, and rice. Arch Biochem Biophys **504:** 104-111

Vallabhaneni R, Gallagher CE, Licciardello N, Cuttriss AJ, Quinlan RF, Wurtzel ET (2009) Metabolite sorting of a germplasm collection reveals the hydroxylase3 locus as a new target for maize provitamin A biofortification. Plant Physiol **151:** 1635-1645

Vallabhaneni R, Gallagher CE, Licciardello N, Cuttriss AJ, Quinlan RF, Wurtzel ET (2009) Metabolite sorting of a germplasm collection reveals the *Hydroxylase3* locus as a new target for maize provitamin A biofortification. Plant Physiol. **151:** 1635-1645

Vallabhaneni R, Wurtzel ET (2009) Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. Plant Physiol **150:** 562-572

van Bokhoven H, Verver J, Wellink J, van Kammen A (1993) Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA. J Gen Virol **74 (Pt 10):** 2233-2241

van den Berg H, Faulks R, Fernando Granado H, Hirschberg J, Olmedilla B, Sandmann G, Southon S, Stahl W (2000) The potential for the improvment of carotenoid levels in foods and likely systemic effects. Journal of the Science of Food and Agriculture **80:** 880-912

Vidi PA, Kanwischer M, Baginsky S, Austin JR, Csucs G, Dormann P, Kessler F, Brehelin C (2006) Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. J Biol Chem **281:** 11225-11234

Vishnevetsky M, Ovadis M, Vainstein A (1999) Carotenoid sequestration in plants: the role of carotenoid-associated proteins. Trends Plant Sci **4:** 232-235

Vogel JT, Tan BC, McCarty DR, Klee HJ (2008) The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. J Biol Chem **283:** 11364-11373

von Lintig J (2010) Colors with functions: elucidating the biochemical and molecular basis of carotenoid metabolism. Annu Rev Nutr **30:** 35-56

von Lintig J, Vogt K (2004) Vitamin A formation in animals: molecular identification and functional characterization of carotene cleaving enzymes. J. Nutr. **134:** 251S-256

von Wettstein D (1958) The formation of plastid structure. Brookhaven Symp. Biol. **11:** 138

Walter MH, Floss DS, Strack D (2010) Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles. Planta **232:** 1-17

Welsch R, Beyer P, Hugueney P, Kleinig H, von Lintig J (2000) Regulation and activation of phytoene synthase, a key enzyme in carotenoid biosynthesis, during photomorphogenesis. Planta **211:** 846-854

Welsch R, Wust F, Bar C, Al-Babili S, Beyer P (2008) A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. Plant Physiol **147:** 367-380

Westphal S, Soll J, Vothknecht UC (2001) A vesicle transport system inside chloroplasts. FEBS Lett **506:** 257-261

Woggon WD (2005) Metalloporphyrines as active site analogues--lessons from enzymes and enzyme models. Acc Chem Res **38:** 127-136

Wong JC, Lambert RJ, Rocheford TR (2002) Comparing QTL and candidate genes for carotenoids and tocopherols in two maize populations. Proc. 38th Ann. Illinois Corn Breeders School**:** 145 -170

Wong JC, Lambert RJ, Wurtzel ET, Rocheford TR (2004) QTL and candidate genes phytoene synthase and zetacarotene desaturase associated with the accumulation of carotenoids in maize. Theor. Appl. Genetics **108:** 349-359

Wurtzel ET (2004) Genomics, genetics, and biochemistry of maize carotenoid biosynthesis. *In* Recent Advances in Phytochemistry (J. Romero, ed.), Vol 38. Elsevier Ltd, Oxford, UK, pp 85-110

Wurtzel ET (2004) Genomics, genetics, and biochemistry of maize carotenoid biosynthesis. *In* J Romeo, ed, Recent Advances in Phytochemistry, Vol 38. Elsevier Ltd., pp 85-110

Wurtzel ET, Cuttriss AJ, Vallabhaneni R (2012) Maize provitamin A carotenoids, current resources, and future metabolic engineering challenges. Frontiers in Plant Science **3:** 1-12

Yan J, Kandianis CB, Harjes CE, Bai L, Kim EH, Yang X, Skinner DJ, Fu Z, Mitchell S, Li Q, Fernandez MG, Zaharieva M, Babu R, Fu Y, Palacios N, Li J, Dellapenna D, Brutnell T, Buckler ES, Warburton ML, Rocheford T (2010) Rare genetic variation at Zea mays crtRB1 increases beta-carotene in maize grain. Nat Genet **42:** 322-327

Yang S, Wu RSS, Mok HOL, Zhang ZP, Kong RYC (2003) Identification of a novel cytochrome P450 cDNA, CYP97E1, from the marine diatom *Skeletonema costatum* Bacillariophyceae. Journal of Phycology **39:** 555-560

Yano JK, Blasco F, Li H, Schmid RD, Henne A, Poulos TL (2003) Preliminary characterization and crystal structure of a thermostable cytochrome P450 from *Thermus thermophilus.* J. Biol. Chem. **278:** 608-616

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

Zhang W, Lukaszewski AJ, Kolmer J, Soria MA, Goyal S, Dubcovsky J (2005) Molecular characterization of durum and common wheat recombinant lines carrying leaf rust resistance (*Lr19*) and yellow pigment (*Y*) genes from *Lophopyrum ponticum*. TAG Theoretical and Applied Genetics **111:** 573-582

Zhao J, Hyman L, Moore C (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev. **63:** 405–445