Cloning and comparative analysis of carotenoid β -hydroxylase genes provides new insights into carotenoid metabolism in tetraploid (*Triticum turgidum* ssp. *durum*) and hexaploid (*Triticum aestivum*) wheat grains

Xiaoqiong Qin, Wenjun Zhang, Jorge Dubcovsky, and Li Tian*

Department of Plant Sciences, University of California, Davis

Abstract

Carotenoid β -hydroxylases attach hydroxyl groups to the β -ionone rings (β -rings) of carotenoid substrates, resulting in modified structures and functions of carotenoid molecules. We cloned and characterized two genes (each with three homeologs), *HYD1* and *HYD2*, which encode β -hydroxylases in wheat. The results from bioinformatic and nested degenerate PCR analyses collectively suggest that *HYD1* and *HYD2* may represent the entire complement of non-heme di-iron β -hydroxylases in wheat. The homeologs of wheat *HYDs* exhibited major β -ring and minor ϵ -ring hydroxylation activities in carotenoid-accumulating *E. coli* strains. Distinct expression patterns were observed for different *HYD* genes and homeologs in vegetative tissues and developing grains of tetraploid and hexaploid wheat, suggesting their functional divergence and differential regulatory control in tissue-, grain development-, and ploidy-specific manners. An intriguing observation was that the expression of *HYD1*, particularly *HYD-B1*, reached highest levels at the last stage of tetraploid and hexaploid grain development, suggesting that carotenoids (at least xanthophylls) were still actively synthesized in mature grains. This result challenges the common perception that carotenoids are simply being turned over during wheat grain development after their initial biosynthesis at the early grain development stages. Overall, this improved understanding of carotenoid biosynthetic gene expression and carotenoid metabolism in wheat grains will contribute to the improvement of the nutritional value of wheat grains for human consumption.

Keywords

β -hydroxylase; carotenoid; homeolog; lutein; provitamin A; wheat

Introduction

Vitamin A nutrition is essential to human health and survival. However, humans are incapable of *de novo* synthesis of vitamin A and have to obtain this important nutrient from dietary sources, such as provitamin A carotenoids in plant-based foods. Provitamin A

*Corresponding author: Department of Plant Sciences, Mail Stop 3, University of California, Davis, Davis, CA 95616, Phone: (530) 7520940, Fax: (530) 7529659, ltian@ucdavis.edu.

carotenoids contain at least one unmodified β -ionone ring (β -ring) and include γ -carotene (β,ψ -carotene), α -carotene (β,ϵ -carotene), β -carotene (β,β -carotene), and β -cryptoxanthin (a mono-hydroxylated β -carotene derivative). Hydroxylation of β -rings by carotenoid hydroxylases constitutes a key step that depletes the provitamin A activities of the precursor carotenoid molecules (Fig. 1a).

Two classes of carotenoid hydroxylases with overlapping activities have been identified in plants: the cytochrome P450 hydroxylases (ϵ -hydroxylases) that act primarily on α -carotene and its derivatives and the non-heme di-iron hydroxylases (β -hydroxylases) that prefer β -carotene and β -cryptoxanthin as substrates (Fiore et al., 2006; Kim et al., 2009; Tian et al., 2003). In dicotyledonous plants, duplicated β -hydroxylase genes have been cloned and characterized from Arabidopsis, tomato, and pepper (Bouvier et al., 1998; Galpaz et al., 2006; Sun et al., 1996; Tian and DellaPenna, 2001). Among monocotyledonous plants, between two to four β -hydroxylase genes have been tentatively identified from sorghum, rice, and maize based on sequence homology to known β -hydroxylases. However, only two maize β -hydroxylases, ZmHYD3 and ZmHYD4, have been functionally characterized thus far (Vallabhaneni et al., 2009; Yan et al., 2010).

β -hydroxylase control of β -carotene and β -carotene derived xanthophyll (oxygenated carotenoid) accumulation was previously demonstrated in different plant tissues, such as potato tubers and maize grains (Brown et al., 2006; Yan et al., 2010). In potato tubers, a polymorphism at a β -hydroxylase locus was shown to at least partially account for variation in β -carotene accumulation (Brown et al., 2006). In maize, ZmHYD3 displayed expression patterns that correlated with hydroxylated β -carotene derivative (β -xanthophyll) accumulation during endosperm development (Vallabhaneni et al., 2009). In addition, polymorphisms within the β -hydroxylase gene (*CrtR-B1*) are associated with a major QTL that controls β -carotene content in maize grains (Yan et al., 2010). Furthermore, maize plants that contain *CrtR-B1* alleles with much reduced *CrtR-B1* transcripts showed increased β -carotene accumulation in endosperm (Yan et al., 2010). Because β -hydroxylases contribute to the turnover of β -carotene to non-provitamin A carotenoids, these genes/enzymes have become a target for improving the provitamin A content of food, particularly in storage organs of crops that serve as major energy sources for humans. For instance, potato tubers possessing an antisense silenced β -hydroxylase showed up to 14 fold increases in β -carotene production (Diretto et al., 2006; Diretto et al., 2007).

Wheat grains play a major role in sustaining the caloric needs of the growing world population. However, mature wheat grain endosperm (the grain tissue that is used to make pasta and bread) is devoid of provitamin A carotenoids. Wheat occurs at different ploidy levels due to the hybridization of diploid and tetraploid genomes during domestication. The mature tetraploid wheat (*Triticum turgidum* ssp. *durum*) grains appear yellow and accumulate mainly lutein (a non-provitamin A carotenoid). High levels of yellow pigments in pasta/tetraploid wheat are favored by durum wheat breeders due to consumer preferences. In contrast, consumers desire white flour for bread baking, and consequently hexaploid/bread wheat (*Triticum aestivum*) varieties have been selected for very low levels of carotenoid (yellow) pigments (Hentschel et al., 2002). Carotenoid accumulation during wheat endosperm development was previously examined using a doubled haploid bread

wheat population (Howitt et al., 2009). The highest amount of lutein (derived from the β,ϵ -carotene branch) and zeaxanthin (derived from the β,β -carotene branch) were detected at 10 days after pollination (DAP). Although the level of lutein did not change significantly during endosperm development, zeaxanthin and two other β -carotene derived xanthophylls, violaxanthin and antheraxanthin, declined gradually through grain development and were undetectable in mature endosperms (Howitt et al., 2009). These data collectively suggest that genes for β -carotene biosynthesis and turnover/degradation are expressed in wheat grains and are possibly developmentally regulated.

β -carotene is the most efficient form of provitamin A (Britton, 2009). Based on the above-mentioned observations on carotenoid accumulation in wheat grains and the successful engineering of β -carotene production in potato tubers, it is conceivable that β -carotene content in wheat grains can potentially be increased by blocking the competing reactions that lead to the biosynthesis of β,ϵ -carotene branch carotenoids as well as the turnover of β -carotene by manipulating the expression of the respective carotenogenic genes. However, most of the carotenoid biosynthetic genes, except for phytoene synthase (*PSY*) and lycopene ϵ -cyclase (*LCY-E*) (Howitt et al., 2009), have not been cloned and characterized in wheat, possibly due to the complex polyploid nature of wheat. In addition to the lack of molecular information about the structural genes, regulation of carotenoid accumulation in wheat grains is also not well understood.

We report the isolation of the three homeologs for each of the two β -hydroxylases present in wheat and the characterization of their functions in a bacterial system. In addition, the expression profiles of the different β -hydroxylase genes and their specific homeologs in vegetative tissues and developing grains of tetraploid and hexaploid wheat were also determined and compared. The carotenoid levels in developing tetraploid and hexaploid wheat grains were analyzed in parallel to the gene expression analysis. Overall, the comparative metabolite and gene expression analyses provided new insights into carotenoid metabolism as well as the function and regulation of β -hydroxylase genes and homeologs in tetraploid and hexaploid wheat grains.

Materials and methods

Plant materials

Tetraploid wheat (*Triticum turgidum* ssp. *durum*) var. Kronos and hexaploid wheat (*Triticum aestivum*) breeding line UC1041 (Tadinia/Yecora Rojo) seedlings were grown in a temperature controlled greenhouse under natural light conditions. Kronos and UC1041 were selected for the gene expression and carotenoid analyses because they were the parental lines used for the generation of tetraploid and hexaploid wheat TILLING (Targeting Induced Local Lesions IN Genomes) mutant populations (Uauy et al., 2009), which will be the source for future screening of down-regulation/knockout mutant of wheat carotenoid biosynthetic genes. After the ear emerged from the leaf sheath, plants were checked daily for anthesis and allowed to self-pollinate. Grains were collected between 5 and 30 days after pollination (DAP) at 5-day intervals and mature grains were collected at around 40 DAP. Leaf, stem, and root tissues were collected from 3-week-old tetraploid and hexaploid wheat seedlings grown in vermiculite in a growth chamber with a 16-h photoperiod. All tissue

samples were immediately frozen in liquid nitrogen upon collection and stored at -80°C until analysis.

RNA and genomic DNA isolation

Total RNA was extracted from different wheat tissues using Trizol reagent (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. RNA concentration and purity (A260/A280 and A260/A230 ratios) were determined using a Nanodrop[®] ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE). An aliquot of the RNA sample was separated on a non-denaturing agarose gel to assess its integrity. Genomic DNA was isolated from leaves of nullisomic-tetrasomic and ditelosomic wheat lines as previously described (Dvorak et al., 1988).

Cloning of wheat β -hydroxylase genes

The Arabidopsis and maize β -hydroxylase genes were used as queries for identification of homologous sequences in three public databases that contain wheat ESTs, including TIGR Gene Indices, HarvEST, and NCBI. Multiple wheat contigs and singletons annotated as β -hydroxylases were identified and assembly into two unigenes (each with three homeologs). One homeolog of a unigene did not appear to be full length and the missing sequence was obtained using rapid amplification of cDNA ends (RACE) PCR (Clontech, Mountain View, CA).

To clone full-length wheat β -hydroxylases, primers were designed based on the homeologous sequences of each gene. First strand cDNA was synthesized from 1 μg total RNA using a Superscript[®] III reverse transcription system (Invitrogen). The high fidelity Platinum[®] Pfx DNA polymerase (Invitrogen) was used to amplify GC-rich wheat cDNA templates. The PCR reaction (20 μL) contained 2 $\times$ PCR buffer, 3 $\times$ PCRx enhancer, 0.3 mM dNTPs, 1 mM MgSO_4 , 0.3 μM each primer, 1 μL cDNA, and 0.5 unit Pfx DNA polymerase. The PCR parameters were 94°C for 5 min, 35 cycles of 94°C for 15 sec, 58°C for 30 sec, and 68°C for 1 min, followed by 68°C for 5 min. PCR products of expected sizes were gel-purified (Qiagen, Valencia, CA), cloned into the pENTR/D-TOPO vector (Invitrogen), and transformed into *E. coli* Top10 competent cells. Several colonies were randomly picked and used for inoculation of liquid cultures. DNA plasmids were extracted and sequenced using M13 forward and reverse primers. Betaine was added to DNA sequencing reactions to relax secondary structures formed in the GC-rich DNA templates.

The high fidelity Phusion[®] DNA polymerase (New England Biolabs, Ipswich, MA) was used to amplify wheat β -hydroxylases from genomic DNA templates. The PCR reaction mix (20 μL) included 1 $\times$ Phusion[®] GC buffer, 0.2 mM dNTPs, 0.5 μM each primer, 100 ng genomic DNA, 5% DMSO, and 0.4 unit Phusion[®] DNA polymerase. The PCR parameters were 98°C for 1 min, 35 cycles of 98°C for 10 sec, 55°C for 30 sec, and 72°C for 75 sec, followed by 72°C for 10 min. The PCR products were gel-purified (Qiagen). After A-addition, the DNA fragment was cloned into the pGEM-T Easy vector (Promega) and then subjected to sequencing reactions. Sequence assembly as well as cDNA and genomic DNA sequence comparisons were performed using Vector NTI[™] (Invitrogen).

The wheat β -hydroxylase genes were tentatively assigned to different chromosomes based on the rice-wheat colinear relationships (Sorrells et al., 2003). To verify the chromosomal locations of wheat β -hydroxylases, a series of nullisomic-tetrasomic and ditelosomic lines of hexaploid wheat var. Chinese Spring were used as template and amplified with homeolog-specific primers for each gene (Fig. 1b). Primers used for RACE PCR, cDNA and genomic DNA cloning, and nested degenerate PCR are listed in Table S1. GenBank accession numbers of *HYD1* and *HYD2* homeologs are: *HYD-A1* (JX171670), *HYD-B1* (JX171671), *HYD-D1* (JX171672), *HYD-A2* (JX171673), *HYD-B2* (JX171674), and *HYD-D2* (JX171675).

Nested degenerate PCR analysis

To examine the possible presence of additional *HYD* genes in the wheat genomes, nested degenerate PCR primers were designed that span the most conserved regions of selected monocotyledonous and dicotyledonous *HYD* genes (Fig. S1). Genomic DNA extracted from hexaploid wheat breeding line UC1041 and diploid wheat *Ae. tauschii* were used as template for PCR amplifications using the high-fidelity Phusion[®] DNA polymerase (New England Biolabs). The degree of degeneracy and amplicon sizes were taken into consideration when designing the degenerate and nested degenerate primers (Lang and Orgogozo, 2011). The PCR mixture (25 μ L) included PCR buffer, 0.5 mM each primer, 200 μ M dNTPs, 100 ng genomic DNA, 1 μ L DMSO, and 0.4 unit Phusion[®] DNA polymerase. The PCR parameters were 98°C for 1 min, 10 cycles of 98°C for 10 sec, 45°C for 30 sec, and 72°C for 15 sec, 25 cycles of 98°C for 10 sec, 58°C for 30 sec, and 72°C for 15 sec, and 72°C for 10 min. The first round PCR products were diluted 100-fold and used as template for the nested degenerate PCR under the same PCR conditions. The nested degenerate PCR products were purified using a PCR purification kit (Qiagen) to remove primer dimers and then cloned into the pGEM-T Easy vector. The plasmids were transformed into chemically competent *E. coli* DH5 α cells which were spread on LB agar plates supplemented with 100 μ g/mL ampicillin. A total of 18 (*Ae. tauschii* template) and 64 (UC1041 template) colonies were randomly picked. The plasmid DNA was extracted and subjected to sequencing using the M13 forward primer.

Phylogenetic analysis

Protein sequences of selected monocotyledonous and dicotyledonous β -hydroxylases were obtained from GenBank, Gramene and Phytozome databases. The accession numbers are: AtBCH1 (NM_001036638), AtBCH2 (NM_124636), OsHYD1 (Gramene LOC_Os04g48880), OsHYD2 (Gramene LOC_Os10g38940), OsHYD3 (Gramene LOC_Os03g03370), SbHYD1 (Phytozome Sb06g026190), SbHYD2 (Phytozome Sb01g048860), SlCrR-b1 (Y14809), SlCrR-b2 (Y14810), wheat HYD1 (JX171670), wheat HYD2 (JX171673), ZmHYD3 (AY844958), ZmHYD4 (AY844956), ZmHYD5 (NM_001154613), ZmHYD6 (BQ619575). The protein sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar, 2004) and the sequence alignment of full-length proteins was used for constructing a neighbor-joining (NJ) tree with pairwise deletion option and a p-distance matrix in MEGA5 (Tamura et al., 2011). Bootstrap analysis of the NJ tree was performed using 1,000 replicates.

Functional characterization of wheat β -hydroxylases in *E. coli*

Open reading frames of wheat β -hydroxylase homeologs were subcloned into the Gateway pENTR-D vector and then transferred to pDEST17 via LR cloning reactions (Invitrogen). The recombinant plasmids were transformed into *E. coli* JM109(DE3) competent cells harboring either pAC-BETA, pAC-DELTA, or pAC-EPSILON, which contain biosynthetic genes for β -carotene, δ -carotene, and ϵ -carotene production, respectively (Cunningham and Gantt, 2007). At least three colonies were randomly picked from each transformation and used for inoculation of a 30 mL Luria Bertani (LB) culture. The bacterial cells were grown at 28°C in dark until cell density at 600 nm reached 0.8. Isopropyl- β -D-thiogalactopyranoside (IPTG) was added to the bacterial culture to a final concentration of 0.5 mM and the cells continued to grow at 28°C for another 4 h. The bacterial cells were then harvested by centrifugation and total carotenoids were extracted as described (Schwartz et al., 2001). The carotenoid extracts were separated on a reverse phase HPLC column (Agilent Zorbax SB-C18, 5 μ m, 4.6 $\times$ 150 mm) using a previously established gradient (Laur and Tian, 2011). The above-mentioned carotenoid accumulating *E. coli* JM109(DE3) cells were also transformed with pDEST17-*AtBCH1* and pDEST17-*GUS* for positive and negative controls of β -ring/ ϵ -ring hydroxylation activities, respectively.

Real-time qPCR analysis

Total RNA extracted from different wheat tissues was treated with RNase-free DNase I (Fermentas, Glen Burnie, MD) to remove any residual genomic DNA that might be carried through the extraction process. Reverse transcription was performed with 0.9 μ g total RNA using an iScript™ cDNA synthesis kit and random hexamers (BioRad, Hercules, CA). Real-time qPCR reactions for each target gene/homeolog were carried out using three biological replicates with three technical duplicates each. Two sets of primers were designed (Table S2). The first set of primers amplify all three homeologs of each gene and were used to examine gene-specific expression. The second set of primers recognize specific homeologs and were used to determine the relative expression of different homeologs. The gene- and homeolog-specific primers were verified via PCR amplifications using DNA extracted from nullisomic-tetrasomic and ditelosomic lines of hexaploid wheat var. Chinese Spring (Fig. S2). The amplicon sizes ranged from 103 bp to 341 bp (Table S2; Fig. S2). The wheat *LCY-E* gene-specific primers were designed based on the previously reported sequences (Howitt et al., 2009).

Real-time qPCR reactions were performed using 0.2 μ L cDNA, 200 nM each primer and iTaq™ SYBR® Green Supermix (BioRad) on an ABI Prism® 7300 Real-time qPCR system (Applied Biosystems, Foster City, CA, USA). The PCR cycling parameters were 1 cycle of 3 min at 95°C, followed by 40 cycles of 15 sec at 95°C and 45 sec at 60°C. No-template and no-reverse transcription controls were also assayed for each primer pair to verify the quality of the cDNA templates and PCR amplifications. Dissociation curve analysis was performed following qPCR and a single peak was observed for each primer pair. A portion of the qPCR products was separated on agarose gels and single products at expected sizes were detected. The efficiency of qPCR amplifications, based on the slope of the standard curve for each primer pair (between -3.25 and -3.52), was between 92% and 103%, except for *HYD-D2*, which had an efficiency of 84% (the slope of the standard curve was -3.78).

A relative standard curve method was used to compare the relative abundance of *HYD* genes and homeologs as well as *LCY-E* in various wheat tissues. Previous studies showed that Ta2291 and Ta54227 were most stably expressed in different wheat tissues and the normalization factor derived from these two reference genes further improved the reliability of reference expression levels as compared to the single reference genes (Paolacci et al., 2009). Therefore, the geometric mean of Ta2291 and Ta54227 was used for normalization of *HYD* gene and homeolog expression. All of the *HYD* homeologs, *LCY-E*, and the reference genes (Ta2291 and Ta54227), are expressed in leaves of tetraploid and hexaploid wheat (based on our preliminary studies and also shown in Figs. 4 and 6). A large volume of cDNA synthesis was carried out using leaf total RNA as template. Standard curves for the target and endogenous reference genes were constructed using the same batch of leaf cDNA as template. Standard-curve and sample reactions for each gene/homeolog were included in the same run (i.e. located on the same 96-well optical plate). The abundance (copies per ng cDNA) of *HYD* genes and homeologs, *LCY-E*, and the reference genes were interpolated from the corresponding standard curves. The transcript quantity of *HYD* homeologs was first normalized to the geometric means of the reference genes in each biological replicate. The relative expression (fold difference) of the *HYD* homeologs was then compared to the A genome homeolog (calibrator) of each *HYD* gene. The standard deviation of the quotient and the relative fold change were calculated as previously described (Applied Biosystems, 2004).

Carotenoid analysis of wheat grains

Wheat grains of the same developmental stage were pooled, weighed, and ground into fine powder in liquid nitrogen using mortar and pestle. To 200 mg ground grain tissue, 900 μ L acetone:ethyl acetate (3:2, v/v) was added and the extraction was carried out in dark at room temperature for 1 h with occasional mixing. An internal standard, β -apo-8'-carotenal, was also added to the extraction buffer. Following the incubation, 600 μ L H₂O was added to the mixture, which was then vortexed and centrifuged at $13,000 \times g$ for 10 min. A portion of the ethyl acetate phase (200 μ L) was transferred to an HPLC vial and 10 μ L was injected into the HPLC column. The HPLC separation was between (A) acetonitrile:H₂O:triethylamine (900:99:1, v/v/v) and (B) ethyl acetate with a gradient of 0–5 min, 100–75% A; 5–10 min, 75–30% A; 10–15 min, 30–0% A; 15–16 min, 0–100% A, and 16–17 min, 100% A. β -carotene, lutein and β -apo-8'-carotenal analytical standards were purchased from Sigma-Aldrich (St. Louis, MO). Neoxanthin, violaxanthin and zeaxanthin were isolated from spinach and the Arabidopsis *lut2* mutant leaves using thin layer chromatography (TLC) and HPLC following an established method (Britton, 1995; Schiedt and Liaaen-Jensen, 1995). Quantity of carotenoids was extrapolated from standard curves.

To determine the carotenoid content of mature embryos, 100 mature tetraploid (Kronos) and hexaploid (UC1041) wheat grains were soaked in water for 2 hrs at room temperature to facilitate dissection and the embryos were carefully removed with a scalpel to ensure that pericarp and endosperm tissues were not attached to the embryos. The dissected embryos were weighed and ground into fine powder in liquid nitrogen. Total carotenoids were extracted and analyzed by reverse phase HPLC using the above-mentioned methods for developing wheat grains.

Statistical analysis

The carotenoid content and the *HYD* gene-specific expression in different wheat tissues and developing grains were analyzed using Tukey's Honestly Significant Difference (HSD) test at a 95% confidence level using the JMP software (SAS Institute, Cary, NC). The *HYD* homeolog-specific expression and *LCY-E* expression were compared using a paired Student's *t* test (Microsoft® Excel, Redmond, WA).

Results

Two wheat β -hydroxylase genes were cloned and are closely related to the monocotyledonous homologs

A combination of keyword- and sequence homology-based database searches using known plant β -hydroxylases identified several tentative contig (TC) and expressed sequence tag (EST) sequences annotated as β -hydroxylases. These TCs and ESTs were further assembled into two paralogous genes, each with three homeologs, which encode proteins that are 72% identical to each other over 90% of the protein length. Several acronyms have been used in the literature for β -hydroxylases cloned from different plant species, including *BCH* in *Arabidopsis*, *CrtR-b* in tomato, *CHY* in potato, *BCH* and *HYD* in rice, *CrtR-B* and *HYD* in maize, and *HYD* in sorghum (Diretto et al., 2007; Du et al., 2010; Galpaz et al., 2006; Kim et al., 2009; Tian et al., 2003; Vallabhaneni et al., 2009; Yan et al., 2010). To avoid further complications in β -hydroxylase nomenclature, we designated the two newly cloned wheat β -hydroxylases as *HYD1* and *HYD2*, to be consistent with the majority of their monocotyledonous homologs (Vallabhaneni et al., 2009). While a two-letter prefix was generally placed before the β -hydroxylase gene name to denote the plant species (except for wheat; e.g. *At* for *Arabidopsis thaliana*, *Os* for *Oryza sativa*, *Sb* for *Sorghum bicolor*, and *Zm* for *Zea mays*), the *HYD* gene symbol without a prefix was used to collectively refer to the β -hydroxylases from wheat at different ploidy levels. The homeologs of *HYD* genes were named following the guidelines for gene symbols in wheat (<http://wheat.pw.usda.gov/ggpages/wgc/98/>), with capital letters indicating different genomes and numbers indicating different paralogs (e.g. *HYD-B2* is the B-genome homeolog of *HYD2*).

Based on the specific amplification from nullisomic-tetrasomic and ditelosomic wheat DNA using homeolog-specific primers, *HYD1* homeologs were located on the long arms of wheat chromosomes 2A, 2B, and 2D, while *HYD2* homeologs were mapped to the long arms of wheat chromosomes 5A, 4B, and 4D due to a translocation between 4AL and 5AL during wheat evolution (Fig. 1b) (Devos et al., 1995). The chromosomal locations of *HYD1* and *HYD2* in wheat correlate with the location of their closest homologs Os04g48880 and Os03g03370 on chromosomes 4 and 3 of rice, which are collinear with wheat homeologous groups 2 and 4, respectively (Sorrells et al., 2003). A splice variant of *HYD-B1* was identified during the PCR cloning of *HYD1* homeologs. The *HYD-B1* splice variant retains the last intron (intron 5), which leads to a frame shift in the last exon (exon 6) and a loss of the stop codon.

A comparison of the gene structures of the six wheat β -hydroxylase homeologs indicated that they all contain six exons and five introns and the exon sizes are highly conserved

(Table S3). This intron-exon organization is also consistent with the maize β -hydroxylase genes (Vallabhaneni et al., 2009). The Arabidopsis β -hydroxylases, on the other hand, contain seven exons and six introns (Tian et al., 2003). Similar to the preservation of gene structures, phylogenetic analysis revealed that HYD1 and HYD2 group with the monocotyledonous β -hydroxylases, and are more distantly related to the dicotyledonous β -hydroxylases (Fig. 1c). Among monocotyledonous plants, two β -hydroxylase homologs have been identified from diploid sorghum (Vallabhaneni et al., 2009), four from ancient tetraploid maize (Vallabhaneni et al., 2009; Yan et al., 2010), and six from hexaploid wheat (this study), suggesting that a single duplication of the ancestral β -hydroxylase gene occurred before the divergence of the grass subfamilies (Fig. 1c, bootstrap confidence 100%). In rice, a more recent duplication resulted in two *OsHYD* paralogs within the β -hydroxylase 2 cluster (Fig. 1c).

Nested degenerate PCR analysis did not identify additional wheat β -hydroxylases

In addition to the results from the EST database searches, *HYD1* and *HYD2* are also the only β -hydroxylase genes present in the draft genome assemblies of the hexaploid wheat var. Chinese Spring (estimated $5 \times$ coverage of the genome; www.cerealsdb.uk.net) and the diploid wheat *Ae. tauschii* (estimated $50 \times$ coverage of the genome; <http://www.cshl.edu/genome/wheat>). To verify the findings from the bioinformatic analysis, PCR reactions were carried out using degenerate and nested degenerate primers that amplify a region highly conserved among β -hydroxylases from different plant species (Fig. S1). A range of products with similar sizes were obtained from the nested degenerate PCR. Out of the 82 clones that were randomly picked and sequenced, 66 (80%) were identical to the homeologs of *HYD1* or *HYD2* (Table S4). The other 16 clones were non-specific PCR products and did not exhibit significant (<10%) sequence homology to β -hydroxylases. In summary, the results from extensive database searches and the nested degenerate PCR analysis suggest that *HYD1* and *HYD2* may represent the entire complement of non-heme di-iron β -hydroxylases in wheat.

Wheat β -hydroxylase homeologs demonstrated major β -ring and minor ϵ -ring hydroxylation activities in *E. coli*

To determine the hydroxylation activities of wheat β -hydroxylases, the open reading frames of *HYD1* and *HYD2* homeologs were cloned into the bacterial expression vector pDEST17. The recombinant plasmids were transformed into JM109(DE3) cells that contain either pAC-BETA (leads to accumulation of β -carotene that has two β -rings), pAC-DELTA (leads to accumulation of δ -carotene that has one ϵ -ring), or pAC-EPSILON (leads to accumulation of ϵ -carotene that has two ϵ -rings). pDEST17-*AtBCH1* (expresses the Arabidopsis β -hydroxylase 1) and pDEST17-*GUS* (expresses a β -glucuronidase) were also transformed into the carotenoid-accumulating *E. coli* cells and used as positive and negative controls, respectively, for β -ring and ϵ -ring hydroxylation activities. In *HYD1* and *HYD2* homeolog-expressing pAC-BETA cells, a majority of the β -carotene substrate was converted into zeaxanthin (di-hydroxylated β -carotene) and a low level of β -cryptoxanthin (mono-hydroxylated β -carotene) production was observed (Fig. 2). In contrast to their high level of hydroxylation activities towards β -rings, only minor mono- ϵ -ring hydroxylation products were detected for all *HYD* homeologs (Fig. S3), similar to those previously shown for

AtBCH1 (Sun et al., 1996). The HYD-B1 splice variant was not functional in any of the carotenoid accumulating *E. coli* strains examined (Fig. 2 and Fig. S3).

β -hydroxylase genes and homeologs exhibit distinct expression patterns in vegetative tissues and developing grains of tetraploid and hexaploid wheat

To examine the spatial and temporal expression patterns of the two wheat β -hydroxylase paralogs and their corresponding homeologs, real-time qPCR analysis was carried out using three vegetative tissues (leaf, stem, and root) and grains that encompass six developmental stages (Fig. 3). Grain 1 (4–10 DAP) and grain 2 (10–16 DAP) represent the early phase of grain development (watery and early milk stages), grain 3 (16–20 DAP) and grain 4 (20–25 DAP) correspond to the grain filling period (late milk and soft dough stages), and grain 5 (25–35 DAP) and grain 6 (35–45 DAP) reflect the late stage of grain maturation (hard dough and ripening stages) (Fig. 3c). *HYD1* and *HYD2* gene- and homeolog-specific primers were designed, verified, and used for the gene expression analysis (Fig. S2; Table S2). The abundance of *HYD* genes and their respective homeologs were quantified using a relative standard curve method (Applied Biosystems, 2004) and were normalized to two endogenous reference genes, Ta2291 (encoding an ADP-ribosylation factor) and Ta54227 (encoding a cell division control protein), which allow direct comparison of gene/homeolog expression within the same tissue and among different tissues for tetraploid or hexaploid wheat (see Materials and methods section).

HYD1 and *HYD2* are expressed in all tetraploid and hexaploid wheat tissues examined (Fig. 3). In tetraploid wheat, these two β -hydroxylase genes exhibited comparable transcript levels in the vegetative tissues, except for leaf where *HYD2* showed higher expression than *HYD1*. Both genes also showed similar and consistent expression in the first five grain development stages. Most significantly, however, at 35–45 DAP (grain 6), *HYD1* rose to the highest expression of all tissues while *HYD2* transcript levels plummeted.

In the vegetative tissues of hexaploid wheat, the highest expression of *HYD1* and *HYD2* were observed in leaf followed by stem and root (Fig. 3b). Similar to tetraploid wheat, generally comparable levels of *HYD1* and *HYD2* transcripts are present in grain 1–grain 5. *HYD1* expression also peaks at 35–45 DAP (grain 6) in hexaploid wheat, to a level that is indistinguishable from that in leaf. Overall, *HYD1* is the dominant β -hydroxylase transcript at 35–45 DAP (grain 6) in both tetraploid and hexaploid wheat (18- and 28-fold higher than *HYD2*, respectively).

To understand homeolog-specific contributions to *HYD* gene expression, the transcript levels of *HYD1* and *HYD2* homeologs in vegetative tissues and developing grains of tetraploid and hexaploid wheat were also determined and compared (Fig. 4). In tetraploid wheat, *HYD-A1* and *HYD-B1* contribute similarly to *HYD1* expression in the vegetative tissues examined. However, *HYD-B1* is the major *HYD1* transcript in all stages of grain development (2–6 fold higher expression than *HYD-A1*). It should be noted that the *HYD-B1*-specific primers only amplified the alternatively spliced *HYD-B1* isoform that does not contain a retained intron (i.e. the isoform that is functional in *E. coli*). Leaf, stem, and the last two stages of grain development (grains 5 and 6) contain similar levels of *HYD-A2* and

HYD-B2 transcripts, whereas *HYD-A2* contributes significantly more to *HYD2* expression than *HYD-B2* in root as well as early and mid-stages of grain development (grain 1-grain 4).

In hexaploid wheat, *HYD-A1* and *HYD-D1* displayed similar transcript levels in most tissues tested except for root and grain 6 where *HYD-D1* expression was undetectable (Fig. 4b). *HYD-B1* expression did not differ significantly from *HYD-A1* and *HYD-D1* in most tissues, but was more prominent than *HYD* homeologs from the other two genomes in stem as well as the first (grain 1) and last (grain 6) stage of grain development. Particularly, *HYD-B1* accounts for approximately 80% of *HYD1* transcripts in grain 6. *HYD-A2* represents the major *HYD2* transcript in leaf, stem and grain 1, while *HYD-B2* and *HYD-D2* have moderately higher expression than *HYD-A2* in root, and grain developmental stages 3, 4, and 5 (Fig. 4b).

Tetraploid and hexaploid wheat show decreasing trend in carotenoid accumulation in developing grains

In parallel to the gene expression analysis, carotenoid profiles of developing tetraploid and hexaploid wheat grains were also analyzed and compared. As indicated by the greenish appearance of the early and mid-developmental stages of wheat grains (grain 1-grain 5) (Fig. 3c), HPLC analysis verified that chlorophylls were still present in these grains (Fig. 5). Carotenoid pigments, including neoxanthin, violaxanthin, lutein, zeaxanthin, and β -carotene, were found in grain 1-grain 5 and their concentrations decreased progressively during grain development. At 35–45 DAP (grain 6), chlorophylls were absent and lutein was the only carotenoid molecule that was detectable in mature tetraploid and hexaploid wheat grains (Fig. 5). In addition, the level of lutein in grain 6 of tetraploid wheat (with yellow mature endosperm) was one-fold higher than that of hexaploid wheat (with white mature endosperm) (Tables 1 and 2). It is worth noting that lutein content of whole grains, not the separated endosperm sections, was measured in the above-mentioned carotenoid analysis. To determine the carotenoid composition in mature embryos, 100 mature wheat grains were used for dissection of embryo tissues, which were then pooled and subjected to HPLC analysis. Zeaxanthin and lutein are present in mature wheat embryos at very low levels (Fig. S4). Several other compounds also showed absorption at 440 nm, but the peaks were too small for integration and identification (Fig. S4).

LCY-E catalyzes the formation of an ϵ -ring in lutein and is a key enzyme involved in lutein biosynthesis, in addition to the β -hydroxylases. To understand whether changes in *LCY-E* expression contributes to the observed reduction in lutein accumulation, *LCY-E* transcripts in developing tetraploid and hexaploid wheat grains were determined (Fig. 6). Similar levels of *LCY-E* expression were observed in grain 4-grain 6 of tetraploid wheat, which were reduced from grain 1-grain 3. In contrast, grain 1 of hexaploid wheat exhibited the highest level of *LCY-E* expression of all grain developmental stages and a six-fold decrease in *LCY-E* expression was observed in grain 2. *LCY-E* transcript levels remained constant from grain 2 to grain 5 and it rose again in grain 6.

Discussion

Tissue- and grain developmental stage-specific expression of β -hydroxylase genes and homeologs suggest gene/homeolog subfunctionalization and differential regulation in these tissues

Numerous studies have shown that gene duplication facilitates functional divergence of the duplicated genes (Taylor and Raes, 2004). The different expression patterns of *HYD1* and *HYD2* in vegetative tissues and developing grains of tetraploid and hexaploid wheat imply that these two β -hydroxylase paralogs possibly play distinct roles in different tissues and grain developmental stages (Fig. 3). In addition to presenting signs of early subfunctionalization, the expression patterns of *HYD* paralogs also serve as an indication of differential regulation, possibly by the metabolic needs and status of various tissues and grain developmental stages.

Previous studies showed that homeologs of the same metabolic gene could also contribute differently to metabolite biosynthesis and accumulation in hexaploid wheat (Nomura et al., 2005). In several instances, significant differences in *HYD1*- or *HYD2*-specific homeolog expression among different genomes (A, B, and D genomes) were observed (Fig. 4). For example, *HYD-B1* is the dominant *HYD1* transcript in tetraploid wheat grains; *HYD-B1* expressed four-fold higher than *HYD-A1* in hexaploid wheat grains at 35–45 DAP where *HYD-D1* expression was undetectable; *HYD-A2* accounted for more than 50% of *HYD2* expression in hexaploid leaves (Fig. 4). This diverse expression of *HYD1/HYD2* homeologs (Fig. 4) supports the notion that they may render differential contribution to β -ring hydroxylation activities in different wheat tissues or grain developmental stages.

As shown in Fig. 1c, the duplication of *HYD1* and *HYD2* took place prior to the divergence of the grass subfamilies (at least 50 million years ago/MYA), which provides an adequate evolutionary time span for functional diversification of the *HYD* paralogs. On the other hand, the more recent separation of the A, B, and D genomes of wheat (3–5 MYA) affords a much shorter period of time for functional divergence among the homeologs of *HYD1* or *HYD2*. The coexistence of these homeologs in the polyploid wheat genomes is even more recent (A and B genomes in tetraploid wheat at ~300,000–500,000 years ago and D genome in hexaploid wheat at ~10,000 years ago) (Dubcovsky and Dvorak, 2007), allowing limited time for additional subfunctionalization during polyploid evolution. Therefore the observed differential expression of specific homeologs of *HYD1* or *HYD2* (Fig. 4) suggests that the relaxed selection of duplicated genes in polyploid wheat genomes may have facilitated the accumulation of changes among the A, B, and D genomes and led to functional diversification of different homeologs (Dvorak and Akhunov, 2005). Overall, different functions and regulation of *HYD1* and *HYD2* paralogs and their respective homeologs provide additional diversity that can be used by natural or human selection to generate different carotenoid profiles.

The predominant expression of *HYD1*, particularly *HYD-B1*, in mature grains may suggest its possible role in embryonic carotenoid biosynthesis

Previously the expression of β -hydroxylase genes was shown to positively correlate with xanthophyll accumulation in maize grains (Vallabhaneni et al., 2009; Yan et al., 2010). The xanthophyll (lutein, neoxanthin, violaxanthin, and zeaxanthin) content decreased during wheat grain development (Tables 1 and 2) and one may expect a parallel decrease in β -hydroxylase gene expression. However, a significant increase in *HYD1*, particularly *HYD-B1*, expression was observed at the last stage of grain development for both tetraploid and hexaploid wheat (Figs. 3 and 4), suggesting that grains at this stage may have increased capacity for xanthophyll biosynthesis as compared to the early grain stages. One possible explanation for this lack of correlation between increased β -ring hydroxylation capacity and decreased xanthophyll accumulation in mature wheat grains could be that the xanthophylls formed via increased synthesis are readily converted into other compounds and lead to a net decreased xanthophyll accumulation.

The whole grain samples analyzed in this study include a mixture of different tissues (i.e. different parts of the grains). Therefore, heterogeneity in the spatial distribution of xanthophyll accumulation and β -hydroxylase gene expression in developing wheat grains could also contribute to the discrepancy observed between these two attributes. Wheat grains can be divided into three structural components: the pericarp, the endosperm, and the embryo/germ. The pericarp of immature wheat grains contain chloroplasts and are photosynthetically active. Four carotenoid molecules, lutein, neoxanthin, violaxanthin, and β -carotene, are highly conserved in the chloroplasts of flowering plants due to their essential functions in light harvesting and photoprotection (Lokstein et al., 2002). Chloroplasts are degraded after desiccation (Fig. 5) and therefore the pericarp of mature wheat grains lacks carotenoids in the absence of intact chloroplasts. As to the endosperm tissue, previous carotenoid analysis indicated that while β -carotene derived xanthophylls in endosperms gradually declined during grain development, lutein levels remained constant in developing wheat endosperms (Howitt et al., 2009). Although the carotenoid content in immature wheat embryos has not been reported, lutein and zeaxanthin were found to be the major carotenoids in mature wheat embryos (Panfili et al., 2003), a result that was also confirmed by this study (Fig. S4). However, as shown in Fig. 5, zeaxanthin was undetectable in mature whole wheat grains (35–45 DAP). Zeaxanthin and lutein are present at very low levels in mature embryos when embryos dissected from 100 mature grains were used for the HPLC analysis (Fig. S4). On the other hand, several folds less whole grain tissues (< 20 grains) were used for the whole grain carotenoid analysis (Tables 1 and 2). The low embryonic zeaxanthin content and the relatively small amount of whole grains (embryos only account for a portion of the whole grains) being analyzed may explain the inability to detect zeaxanthin in mature whole wheat grains (Fig. 5).

Taken together, the trend of spatial carotenoid accumulation in different grain tissues and *HYD* paralog and homeolog expression patterns suggest that *HYD1*, specifically *HYD-B1*, may possibly be responsible for lutein and zeaxanthin biosynthesis in wheat embryos. Previously, the expression of 55,052 transcripts were studied in developing (ranges from 6–42 DPA) hexaploid wheat grains (Wan et al., 2008). Distinct clusters were formed with

genes that exhibited similar changes in expression during wheat grain development and were putatively assigned to different grain tissue locations based on the genes with known locations within each cluster (Wan et al., 2008). Interestingly, when compared with these wheat grain-expressed genes (Wan et al., 2008), *HYD1* exhibited an embryo-like expression pattern that entails increased transcript accumulation through development, while *HYD2* showed similar decreasing trends as genes expressed in endosperm and pericarp tissues (Figs. 3 and 4). The enhanced expression of a β -hydroxylase gene in the embryo tissue may provide the required intermediate for synthesis of abscisic acid (ABA), which increases towards maturation of wheat grains (Walker-Simmons, 1987). Further studies are required to understand the spatial expression of carotenogenic genes and homeologs in different sections of wheat grains and how they contribute to β -carotene/ β -carotene-derived xanthophyll accumulation in these tissues.

The notion that *HYD1* may contribute to lutein and zeaxanthin (xanthophylls derived from the β,ϵ - and β,β -branch of the carotenoid pathway, respectively) biosynthesis in mature embryos prompted us to examine the possible ϵ -ring hydroxylation activity of *HYD1* for lutein formation. Only minor ϵ -ring hydroxylation activities were observed for all *HYD1* and *HYD2* homeologs in *E. coli*, suggesting that *HYD1* and *HYD2* may not significantly contribute to ϵ -ring hydroxylation in wheat. This result resonates with previous observations in *Arabidopsis*, where, though overlapping activities exist for β - and ϵ -hydroxylases, ϵ -hydroxylase is still the major activity for ϵ -ring hydroxylation (Tian et al., 2003; Tian et al., 2004). However, one should also bear in mind that the *in planta* functions of β -hydroxylases may deviate from their *in vitro*/*in E. coli* activities due to various factors such as substrate availability and cellular environment. For instance, a recent report showed that in a quadruple *Arabidopsis* mutant that is deprived of α -carotene (due to a mutation in *LCY-E*) and contains *LUT1* (an *Arabidopsis* ϵ -hydroxylase) as the only functional carotenoid hydroxylase, significant accumulation of β,β -xanthophylls was observed, suggesting that *LUT1* could function towards β -rings in the absence of α -carotene, its preferred substrate (Fiore et al., 2012). Therefore, even though low ϵ -ring hydroxylation activities were evident for the wheat β -hydroxylases in *E. coli*, they could still be involved in lutein biosynthesis *in planta*. Our preliminary database searches identified putative ϵ -hydroxylase homologs in wheat (data not shown). Future cloning and functional characterization of these ϵ -hydroxylase genes are expected to facilitate the delineation of β - and ϵ -hydroxylase activities towards lutein production in wheat grains.

New insights have emerged on carotenoid metabolism in tetraploid and hexaploid wheat grains

In addition to the overall trend of reduction in lutein content, our data also revealed two new insights on carotenoid metabolism for developing grains of tetraploid and hexaploid wheat. First, the ratio between the β,ϵ - and β,β -branch carotenoids remains relatively constant in developing grains of tetraploid and hexaploid wheat (Tables 1 and 2). Second, although similar decreases in lutein content were observed at early and late developmental stages of tetraploid and hexaploid wheat grains, hexaploid wheat showed more rapid reduction of lutein during grain filling than tetraploid wheat (Fig. 6c).

The relatively constant ratio between β,ϵ - and β,β -branch carotenoids in developing tetraploid and hexaploid wheat grains suggests that the partition of carbon flow between these two branches is tightly controlled in wheat grains. A key step for lutein biosynthesis, which also impacts the division between β,ϵ - and β,β -branch carotenoid formation, is the introduction of an ϵ -ring, catalyzed by *LCY-E* (Cunningham et al., 1996). It was shown that naturally occurring polymorphisms of *ZmLCY-E* account for 58% of variations in the β,ϵ -/ β,β -carotene branch ratios of the carotenoid pathway in maize grains (Harjes et al., 2008). Interestingly, while *LCY-E* expression was generally decreased in developing tetraploid wheat grains (similar to the reducing trend of lutein accumulation), it showed a steady increase during hexaploid wheat grain development after an initial drop at grain 2, a pattern that somewhat resembles *HYD1* expression (Fig. 3b and Fig. 6c). In view of the changing *LCY-E* expression levels that did not directly impact the β,ϵ -/ β,β -branch carotenoid ratios, it suggests that a coordination between *LCY-E* and lycopene β -cyclase/*LCY-B* (catalyzes the formation of a β -ring; Fig. 1a) expression may be necessary for regulation of carbon fluxes through different branches of the carotenoid pathway. Alternatively, this key branch point could be regulated at post-transcriptional, translational, or post-translational levels.

Previously *LCY-E* expression was shown to correlate well with lutein accumulation in maize grains (Naqvi et al., 2011). However, the *LCY-E* expression pattern cannot explain the (differential) decrease in lutein accumulation in tetraploid and hexaploid wheat (Fig. 6). On the contrary, the results from gene expression analysis suggest that, at least for hexaploid wheat, there may even be a rise of lutein biosynthesis at the last stage of grain development (Fig. 6b). It could be that the observed decreases in lutein content in developing grains of hexaploid wheat are the net results of a larger increase in turnover relative to biosynthesis processes. On the other hand, the gene expression data also suggest that *LCY-E* is not rate-limiting for lutein formation in wheat grains. Other enzymes may be involved in controlling lutein accumulation in wheat grains. A better understanding of lutein biosynthesis and turnover during grain development is expected to shine light on this (differential) decreases in tetraploid and hexaploid wheat grains.

Future perspectives

Cloning and functional characterization of wheat β -hydroxylase genes provides the knowledge base required for future manipulation of β -carotene content in wheat grains. Grains with elevated β -carotene content are highly desirable for people in developing countries as it provides an affordable dietary source of vitamin A. As shown in the potato tubers, increased β -carotene accumulation in a storage tissue can be achieved by blocking the expression of *LCY-E* and β -hydroxylases (Diretto et al., 2006; Diretto et al., 2007). The currently available TILLING mutant populations of tetraploid (var. Kronos) and hexaploid wheat (breeding line UC1041) (Uauy et al., 2009) will be used to select reduced- or loss-of-function mutations in wheat *LCY-E* and *HYDs* with the objective of increasing β -carotene content in wheat grains. The induced TILLING mutants are not subject to the expensive and time-consuming regulatory processes required for transgenic crops, which is expected to accelerate their incorporation in wheat breeding programs, and hopefully expedite the development of high β -carotene wheat.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Nadia Ono for critical reading of the manuscript, Dr. Diane Beckles for helpful discussions, Dr. Francis Cunningham for providing us the pAC-BETA, pAC-DELTA, and pAC-EPSILON plasmids, and Drs. Jan Dvorak, W. Richard McCombie, and Doreen Ware for early access to the assembly of the *Ae. tauschii* genome. This work was supported by the UC Davis new faculty startup fund to LT and by the Howard Hughes Medical Institute and Betty and Gordon Moore Foundation and USDA-AFRI grant 2011-68002-30029 to JD.

Abbreviations

ABA	Absciscic acid
β-ring	β-ionone ring
CRTISO	Carotenoid isomerase
CYP	Carotenoid ε-hydroxylase (cytochrome P450 type)
DAP	Days after pollination
EST	Expressed sequence tags
GGPP	Geranylgeranyl diphosphate
HYD	Carotenoid β-hydroxylase (non-heme di-iron type)
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Luria Bertani
LCY-B	Lycopene β-cyclase
LCY-E	Lycopene ε-cyclase
MUSCLE	Multiple Sequence Comparison by Log-Expectation
MYA	Million years ago
NJ	Neighbor-joining
NXS	Neoxanthin synthase
PDS	Phytoene desaturase
PSY	Phytoene synthase
RACE	Rapid amplification of cDNA ends
TC	Tentative contig
TLC	Thin layer chromatography
TILLING	Targeting Induced Local Lesions IN Genomes
VDE	Violaxanthin de-epoxidase
ZDS	ζ-carotene desaturase

ZEP	Zeaxanthin epoxidase
Z-ISO	ζ-carotene isomerase

References

- Applied Biosystems. Guide to performing relative quantitation of gene expression using Real-Time quantitative PCR. 2004
- Bouvier F, Keller Y, d'Harlingue A, Camara B. Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum* L.). *Biochim Biophys Acta*. 1998; 1391:320–328. [PubMed: 9555077]
- Britton, G. UV/visible spectroscopy. In: Britton, G.; Jensen, Liaaen-S.; Pfander, H., editors. *Carotenoids, Spectroscopy*. Vol. 1B. Birkhauser, Basel, Switzerland: 1995. p. 13-62.
- Britton, G. Vitamin A and vitamin A deficiency Carotenoids. In: Britton, G.; Liaaen-Jensen, S.; Pfander, H., editors. *Nutrition and Health*. Vol. 5. Birkhäuser Verlag Basel: 2009.
- Brown C, Kim T, Ganga Z, Haynes K, De Jong D, Jahn M, Paran I, De Jong W. Segregation of total carotenoid in high level potato germplasm and its relationship to beta-carotene hydroxylase polymorphism. *Am J Potato Res*. 2006; 83:365–372.
- Carpentier S, Knaus M, Suh M. Associations between lutein, zeaxanthin, and age-related macular degeneration: an overview. *Crit Rev Food Sci Nutr*. 2009; 49:313–326. [PubMed: 19234943]
- Cunningham F, Pogson B, Sun Z, McDonald K, DellaPenna D, Gantt E. Functional analysis of the beta and epsilon lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell*. 1996; 8:1613–1626. [PubMed: 8837512]
- Cunningham FJ, Gantt E. A portfolio of plasmids for identification and analysis of carotenoid pathway enzymes: *Adonis aestivalis* as a case study. *Photosynth Res*. 2007; 92:245–259. [PubMed: 17634749]
- Devos K, Dubcovsky J, Dvorak J, Chinoy C, Gale M. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor Appl Genet*. 1995; 91:282–288. [PubMed: 24169776]
- Diretto G, Tavazza R, Welsch R, Pizzichini D, Mourgues F, Papacchioli V, Beyer P, Giuliano G. Metabolic engineering of potato tuber carotenoids through tuber-specific silencing of lycopene epsilon cyclase. *BMC Plant Biol*. 2006; 6:13. [PubMed: 16800876]
- Diretto G, Welsch R, Tavazza R, Mourgues F, Pizzichini D, Beyer P, Giuliano G. Silencing of beta-carotene hydroxylase increases total carotenoid and beta-carotene levels in potato tubers. *BMC Plant Biol*. 2007; 7:11. [PubMed: 17335571]
- Du H, Wang N, Cui F, Li X, Xiao J, Xiong L. Characterization of the β-carotene hydroxylase gene *DSM2* conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice. *Plant Physiol*. 2010; 154:1304–1318. [PubMed: 20852032]
- Dubcovsky J, Dvorak J. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*. 2007; 316:1862–1866. [PubMed: 17600208]
- Dvorak J, Akhunov E. Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the Aegilops-Triticum alliance. *Genetics*. 2005; 171:323–332. [PubMed: 15996988]
- Dvorak J, McGuire P, Cassidy B. Apparent sources of the A genomes of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. *Genome*. 1988; 30:680–689.
- Edgar R. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res*. 2004; 32:1792–1797. [PubMed: 15034147]
- Fiore A, Dall'osto L, Cazzaniga S, Diretto G, Giuliano G, Bassi R. A quadruple mutant of *Arabidopsis* reveals a β-carotene hydroxylation activity for LUT1/CYP97C1 and a regulatory role of xanthophylls on determination of the PSI/PSII ratio. *BMC Plant Biol*. 2012; 12:50. [PubMed: 22513258]

- Fiore A, Dall'osto L, Fraser PD, Bassi R, Giuliano G. Elucidation of the β -carotene hydroxylation pathway in *Arabidopsis thaliana*. FEBS Lett. 2006; 580:4718–4722. [PubMed: 16890225]
- Galpaz N, Ronen G, Khalfa Z, Zamir D, Hirschberg J. A chromoplast-specific carotenoid biosynthesis pathway is revealed by cloning of the tomato white-flower locus. Plant Cell. 2006; 18:1947–1960. [PubMed: 16816137]
- Harjes C, Rocheford T, Bai L, Brutnell T, Kandianis C, Sowinski S, Stapleton A, Vallabhaneni R, Williams M, Wurtzel E, Yan J, Buckler E. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science. 2008; 319:330–333. [PubMed: 18202289]
- Hentschel V, Kranl K, Hollmann J, Lindhauer M, Bohm V, Bitsch R. Spectrophotometric determination of yellow pigment content and evaluation of carotenoids by high-performance liquid chromatography in durum wheat grain. J Agric Food Chem. 2002; 50:6663–6668. [PubMed: 12405758]
- Howitt C, Cavanagh C, Bowerman A, Cazzonelli C, Rampling L, Mimica J, Pogson B. 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. 2009; 9:363–376. [PubMed: 19330366]
- Kim J, Smith J, Tian L, DellaPenna D. The evolution and function of carotenoid hydroxylases in Arabidopsis. Plant Cell Physiol. 2009; 50:463–479. [PubMed: 19147649]
- Lang M, Orgogozo V. Identification of homologous gene sequences by PCR with degenerate primers. Methods Mol Biol. 2011; 772:245–256. [PubMed: 22065442]
- Laur L, Tian L. Provitamin A and vitamin C content in selected California-grown cantaloupe and honeydew melons and imported melons. J Food Comp Anal. 2011; 24:194–201.
- Lokstein H, Tian L, Polle JE, DellaPenna D. 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. 2002; 1553:309–319. [PubMed: 11997140]
- Naqvi S, Zhu C, Farre G, Sandmann G, Capell T, Christou P. Synergistic metabolism in hybrid corn indicates bottlenecks in the carotenoid pathway and leads to the accumulation of extraordinary levels of the nutritionally important carotenoid zeaxanthin. Plant Biotechnol J. 2011; 9:384–393. [PubMed: 20807370]
- Nomura T, Ishihara A, Yanagita R, Endo T, Iwamura H. Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. Proc Natl Acad Sci USA. 2005; 102:16490–16495. [PubMed: 16260753]
- Panfil G, Cinquanta L, Fratianni A, Cubadda R. Extraction of wheat germ oil by supercritical CO₂: Oil and defatted cake characterization. J Am Oil Chem Soc. 2003; 80:157–161.
- Paolacci A, Tanzarella O, Porceddu E, Ciaffi M. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol Biol. 2009; 10:11. [PubMed: 19232096]
- Schiedt, K.; Liaaen-Jensen, S. Isolation and analysis. In: Britton, G.; Jensen, Liaaen-S.; Pfander, H., editors. Carotenoids, Isolation and Analysis. Vol. 1A. Birkhauser, Basel, Switzerland: 1995. p. 81-108.
- Schwartz S, Qin X, Zeevaert J. Characterization of a novel carotenoid cleavage dioxygenase from plants. J Biol Chem. 2001; 276:25208–25211. [PubMed: 11316814]
- Sorrells M, La Rota M, Bermudez-Kandianis C, Greene R, Kantety R, Munkvold J, Miftahudin, Mahmoud A, Ma X, Gustafson P, Qi L, Echalié B, Gill B, Matthews D, Lazo G, Chao S, Anderson O, Edwards H, Linkiewicz A, Dubcovsky J, Akhunov E, Dvorak J, Zhang D, Nguyen H, Peng J, Lapitan N, Gonzalez-Hernandez J, Anderson J, Hossain K, Kalavacharla V, Kianian S, Choi D, Close T, Dilbirli M, Gill K, Steber C, Walker-Simmons M, McGuire P, Qualset C. Comparative DNA sequence analysis of wheat and rice genomes. Genome Res. 2003; 13:1818–1827. [PubMed: 12902377]
- Sun Z, Gantt E, Cunningham FX Jr. Cloning and functional analysis of the β -carotene hydroxylase of *Arabidopsis thaliana*. J Biol Chem. 1996; 271:24349–24352. [PubMed: 8798688]
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28:2731–2739. [PubMed: 21546353]

- Tanaka T, Shnimizu M, Moriwaki H. Cancer chemoprevention by carotenoids. *Molecules*. 2012; 17:3202–3242. [PubMed: 22418926]
- Taylor J, Raes J. Duplication and divergence: the evolution of new genes and old ideas. *Annu Rev Genet*. 2004; 38:615–643. [PubMed: 15568988]
- Tian L, DellaPenna D. Characterization of a second carotenoid β -hydroxylase gene from *Arabidopsis* and its relationships to the *LUT1* locus. *Plant Mol Biol*. 2001; 45:379–388. [PubMed: 11587509]
- Tian L, Magallanes-Lundback M, Musetti V, DellaPenna D. Functional analysis of β - and ϵ -ring carotenoid hydroxylases in *Arabidopsis thaliana*. *Plant Cell*. 2003; 15:1320–1332. [PubMed: 12782726]
- Tian L, Musetti V, Kim J, Magallanes-Lundback M, DellaPenna D. The *Arabidopsis LUT1* locus encodes a member of the cytochrome P450 family that is required for carotenoid ϵ -ring hydroxylation activity. *Proc Natl Acad Sci USA*. 2004; 101:402–407. [PubMed: 14709673]
- Uauy C, Paraiso F, Colasuonno P, Tran R, Tsai H, Berardi S, Comai L, Dubcovsky J. A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biol*. 2009; 9:115. [PubMed: 19712486]
- Vallabhaneni R, Gallagher CE, Licciardello N, Cuttriss AJ, Quinlan RF, Wurtzel ET. Metabolite sorting of a germplasm collection reveals the Hydroxylase3 locus as a new target for maize provitamin A biofortification. *Plant Physiol*. 2009; 151:1635–1645. [PubMed: 19767386]
- Walker-Simmons M. ABA levels and sensitivity in developing wheat embryos of sprouting resistant and susceptible cultivars. *Plant Physiol*. 1987; 84:61–66. [PubMed: 16665406]
- Wan Y, Poole R, Huttly A, Toscano-Underwood C, Feeney K, Welham S, Gooding M, Mills C, Edwards K, Shewry P, Mitchell R. Transcriptome analysis of grain development in hexaploid wheat. *BMC Genomics*. 2008; 9:121. [PubMed: 18325108]
- Yan J, Kandianis C, Harjes C, Bai L, Kim E, Yang X, Skinner D, Fu Z, Mitchell S, Li Q, Fernandez M, Zaharieva M, Babu R, Fu Y, Palacios N, Li J, DellaPenna D, Brutnell T, Buckler E, Warburton M, Rocheford T. Rare genetic variation at *Zea mays crtRB1* increases beta-carotene in maize grain. *Nat Genet*. 2010; 42:322–327. [PubMed: 20305664]

Fig. 1.

Cloning of two carotenoid β-hydroxylases, *HYD1* and *HYD2*, from wheat. **a** Biochemical pathways leading to the formation of α-/β,ε- and β-/β,β-carotene derived xanthophylls. Enzymes that catalyze the reactions are indicated. GGPP, geranylgeranyl diphosphate; PSY, phytoene synthase; PDS, phytoene desaturase; Z-ISO, ζ-carotene isomerase; ZDS, ζ-carotene desaturase; CRTISO, carotenoid isomerase; LCY-B, lycopene β-cyclase; LCY-E, lycopene ε-cyclase; HYD, carotenoid β-hydroxylase (non-heme di-iron type); CYP, carotenoid ε-hydroxylase (cytochrome P450 type); ZEP, zeaxanthin epoxidase; VDE, violaxanthin de-epoxidase; NXS, neoxanthin synthase. **b** Chromosomal locations of the wheat *HYD* homeologs were verified using homeolog-specific primers and nullisomic-tetrasomic and ditelosomic lines of hexaploid wheat var. Chinese Spring. N2BT2D (nullisomic for 2B and tetrasomic for 2D), N2DT2A (nullisomic for 2D and tetrasomic for 2A), N5AT5D (nullisomic for 5A and tetrasomic for 5D), N4DT4A (nullisomic for 4D and tetrasomic for 4A), DT2AS (the long arm of chromosome 2A is missing), DT4BS (the long arm of chromosome 4B is missing). Kronos and UC1041 are wild type tetraploid and hexaploid wheat, respectively. **(C)** A neighbor-joining tree of carotenoid β-hydroxylases from representative monocotyledonous and dicotyledonous plants. HYD, BCH, and CrR-b have all been used to refer to β-hydroxylases in the literature. Bootstrap values (1,000 replicates) are shown next to the branches. Wheat HYD1 and HYD2 are highlighted in bold. At, *Arabidopsis thaliana*; Os, *Oryza sativa*; Sb, *Sorghum bicolor*; Sl, *Solanum lycopersicum*; Zm, *Zea mays*.

Fig. 2.

Functional characterization of wheat HYD1 and HYD2 homeologs in β -carotene-accumulating *E. coli*. The plasmid pAC-BETA contains all the genes necessary for β -carotene (β -car) production. pAC-BETA-expressing *E. coli* cells were transformed with wheat *HYD1* and *HYD2* homeologs cloned in the pDEST17 vector. Both β -cryptoxanthin (β -cry; mono-hydroxylated β -carotene derivative) and zeaxanthin (zea; di-hydroxylated β -carotene derivative) were produced. HPLC elution profiles of **a** pAC-BETA, **b** pAC-BETA + pDEST17-AtBCH1, **c** pAC-BETA + pDEST17-GUS, **d** pAC-BETA + pDEST17-HYD-A1, **e** pAC-BETA + pDEST17-HYD-B1, **f** pAC-BETA + pDEST17-HYD-B1 splice variant, **g** pAC-BETA + pDEST17-HYD-D1, **h** pAC-BETA + pDEST17-HYD-A2, **i** pAC-BETA + pDEST17-HYD-B2, and **j** pAC-BETA + pDEST17-HYD-D2 are shown. pAC-BETA transformed with pDEST17-AtBCH1 and pDEST17-GUS were used for positive and negative controls of β -ring hydroxylation activities, respectively. **k-n** Absorption spectra of

β -carotene, β -cryptoxanthin, *trans*-zeaxanthin (*trans*-zea), and *cis*-zeaxanthin (*cis*-zea), respectively

(A) Tetraploid wheat var. Kronos **(B) Hexaploid wheat breeding line UC1041** **(C)** **Fig. 3.**

Expression of *HYD1* and *HYD2* in different wheat tissues and during grain development determined by real-time qPCR analysis. Relative transcript abundance of *HYD1* and *HYD2* in tetraploid wheat var. Kronos (**a**) and hexaploid wheat breeding line UC1041 (**b**) are shown. Gene expression was normalized to the geometric mean of two reference genes, *Ta2291* and *Ta54227*. Data presented are mean $\pm$ SD ($n=9$). Different letters denote significant difference ($P < 0.05$) in transcript abundance with Tukey's HSD test. **c** Wheat grains were collected at six developmental stages. DAP, days after pollination. Grain 1, 4–

10 DAP; Grain 2, 10–16 DAP; Grain 3, 16–20 DAP; Grain 4, 20–25 DAP; Grain 5, 25–35 DAP; Grain 6, 35–45 DAP

(A) Tetraploid wheat var. Kronos **(B) Hexaploid wheat breeding line UC1041** **Fig. 4.**

Expression of *HYD1* and *HYD2* homeologs in different wheat tissues and during grain development determined by real-time qPCR analysis. Relative transcript abundance of *HYD1* and *HYD2* homeologs in tetraploid wheat var. Kronos (a) and hexaploid wheat breeding line UC1041 (b) are shown. Gene expression was normalized to the geometric mean of two reference genes, *Ta2291* and *Ta54227*. Data presented are mean $\pm$ SD ($n=9$). For tetraploid wheat, significant differences ($P < 0.05$) between A and B homeologs in each tissue, examined by a paired Student's *t* test, are indicated by asterisks. For hexaploid wheat, different letters indicate significant differences ($P < 0.05$) in relative transcript abundance between different homeologs in each tissue according to a paired Student's *t* test.

Fig. 5.

Carotenoid profiles during wheat grain development. HPLC elution profiles of tetraploid wheat var. Kronos (a) and hexaploid wheat breeding line UC1041 (b) grains at six different developmental stages (grain 1-grain 6) are shown. Two hundred mg of ground grain tissue was used for carotenoid extraction and a portion of the extract was injected on HPLC. The HPLC traces were drawn to the same scale. c Absorption spectra of peaks 1–8. Peak 1, neoxanthin; Peak 2, *trans*-violaxanthin; Peak 3, *cis*-violaxanthin; Peak 4, lutein; Peak 5, *cis*-zeaxanthin; Peak 6, chlorophyll *b*; Peak 7, chlorophyll *a*; Peak 8, β-carotene

(A) Tetraploid wheat var. Kronos **(B) Hexaploid wheat breeding line UC1041** **Fig. 6.**

Lycopene ϵ -cyclase (*LCY-E*) expression and lutein accumulation during wheat grain development. *LCY-E* expression in tetraploid wheat var. Kronos (a) and hexaploid wheat breeding line UC1041 (b) grains was determined by real-time qPCR. Normalized gene expression to the geometric mean of two reference genes, *Ta2291* and *Ta54227*, is shown. Data presented are mean $\pm$ SD (n=9). Different letters indicate significant differences ($P < 0.05$) in relative transcript abundance according to a paired Student's *t* test. c Lutein content in tetraploid (var. Kronos) and hexaploid (breeding line UC1041) wheat grains decreases

during grain maturation. Lutein content in grain 1 of tetraploid and hexaploid wheat grains is not statistically significant ($P < 0.05$)

Table 1

Carotenoid composition in developing grains of tetraploid wheat var. Kronos.

Grain	μg carotenoid pigment/g grain					
	Neoxanthin	Violaxanthin	Lutein	Zeaxanthin	β -carotene	Total
1	1.86 $\pm$ 0.36 ^a	3.86 $\pm$ 0.29 ^a	8.92 $\pm$ 0.84 ^a	0.66 $\pm$ 0.14 ^a	3.87 $\pm$ 0.38 ^a	19.17 $\pm$ 1.78 ^a
2	1.61 $\pm$ 0.23 ^a	4.08 $\pm$ 0.19 ^a	7.36 $\pm$ 0.4 ^b	0.63 $\pm$ 0.1 ^{a,b}	3.03 $\pm$ 0.11 ^b	16.7 $\pm$ 0.02 ^b
3	1.53 $\pm$ 0.18 ^a	4.29 $\pm$ 0.14 ^a	7.09 $\pm$ 0.67 ^b	0.57 $\pm$ 0.07 ^{a,b}	2.98 $\pm$ 0.23 ^b	16.45 $\pm$ 0.6 ^{b,c}
4	1 $\pm$ 0.04 ^b	3.77 $\pm$ 0.08 ^a	6.44 $\pm$ 0.41 ^b	0.43 $\pm$ 0.03 ^b	2.57 $\pm$ 0.07 ^b	14.21 $\pm$ 0.58 ^c
5	0.18 $\pm$ 0.04 ^c	1.18 $\pm$ 0.28 ^b	2.98 $\pm$ 0.25 ^c	0.2 $\pm$ 0.02 ^c	0.99 $\pm$ 0.07 ^c	5.54 $\pm$ 0.66 ^d
6	ND	ND	1.61 $\pm$ 0.14 ^c	ND	ND	1.61 $\pm$ 0.14 ^e

Data presented are mean $\pm$ SD of three biological replicates of pooled grains. The ratios between β , ϵ - and β , β -branch carotenoids (β , ϵ / β , β) are also shown. Different letters indicate significant differences ($P < 0.05$) in carotenoid pigment content or β , ϵ / β , β ratios within a column determined by Tukey's HSD test. ND, not detectable

Table 2

Carotenoid composition in developing grains of hexaploid wheat breeding line UC1041.

Grain	μg carotenoid pigment/g grain						
	Neoxanthin	Violaxanthin	Lutein	Zeaxanthin	β -carotene	Total	β , ϵ / β , β
1	2.42 $\pm$ 0.09 ^a	5.3 $\pm$ 0.28 ^a	10.09 $\pm$ 0.3 ^a	0.68 $\pm$ 0.03 ^a	5.71 $\pm$ 0.23 ^a	24.19 $\pm$ 1 ^a	0.72 $\pm$ 0.01 ^a
2	2.06 $\pm$ 0.13 ^b	5.23 $\pm$ 0.25 ^a	8.26 $\pm$ 0.98 ^b	0.62 $\pm$ 0.09 ^a	4.82 $\pm$ 0.58 ^b	20.99 $\pm$ 1.76 ^b	0.65 $\pm$ 0.04 ^b
3	1.59 $\pm$ 0.09 ^c	4.55 $\pm$ 0.22 ^b	6.98 $\pm$ 0.32 ^b	0.46 $\pm$ 0.02 ^b	4.16 $\pm$ 0.17 ^b	17.75 $\pm$ 0.82 ^c	0.65 $\pm$ 0.01 ^b
4	1.14 $\pm$ 0.08 ^d	3.76 $\pm$ 0.21 ^c	5.58 $\pm$ 0.35 ^c	0.37 $\pm$ 0.02 ^b	3.25 $\pm$ 0.18 ^c	14.1 $\pm$ 0.83 ^d	0.65 $\pm$ 0.01 ^b
5	0.21 $\pm$ 0.04 ^e	1.11 $\pm$ 0.14 ^d	1.61 $\pm$ 0.16 ^d	0.11 $\pm$ 0.01 ^c	0.99 $\pm$ 0.12 ^d	4.04 $\pm$ 0.45 ^e	0.67 $\pm$ 0.02 ^{a,b}
6	ND	ND	0.76 $\pm$ 0.08 ^d	ND	ND	0.76 $\pm$ 0.08 ^f	-

Data presented are mean $\pm$ SD of three biological replicates of pooled grains. The ratios between β , ϵ - and β , β -branch carotenoids (β , ϵ / β , β) are also shown. Different letters indicate significant differences ($P < 0.05$) in carotenoid pigment content or β , ϵ / β , β ratios within a column determined by Tukey's HSD test. ND, not detectable