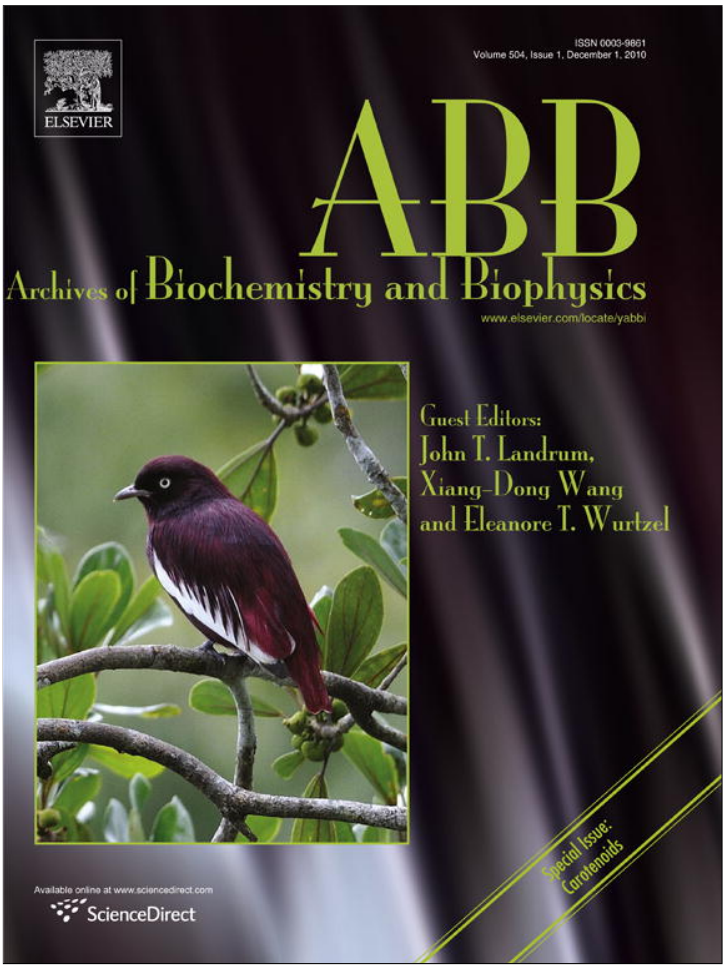

Archives of Biochemistry and Biophysics 504 (2010) 104–111

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

The carotenoid dioxygenase gene family in maize, sorghum, and rice

Ratnakar Vallabhaneni [a,b], Louis M.T. Bradbury [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, United States
[b] The Graduate School and University Center-CUNY, 365 Fifth Ave., New York, NY 10016-4309, United States

ARTICLE INFO

Article history:
Available online 27 July 2010

Keywords:
Carotenoid cleavage dioxygenases
Zea mays
Oryza sativa
Sorghum bicolor
Apocarotenoids
Provitamin A
Abscisic acid

ABSTRACT

Carotenoids and their apocarotenoid derivatives play essential physiological and developmental roles and provide plants tolerance to a variety of stresses. Carotenoid cleavage dioxygenases mediate the degradation of carotenoids to apocarotenoids. A better understanding of biosynthesis vs. degradation could be useful for controlling carotenoid levels leading to improved plant fitness and/or enhanced content of nutritionally valuable carotenoids. The Poaceae (grass) plant family contains many crops of agronomic value. Therefore this study focused on characterizing the carotenoid dioxygenase gene family in the grass species maize, rice, and sorghum with comparison made to newly identified gene families in two non-seed plants as well as an alga and previously identified eudicot genes. Genome analysis was used to map grass genes encoding the carotenoid dioxygenases to chromosome locations. Sequences of encoded proteins were phylogenetically compared. CCD8b was identified as a new class of cleavage dioxygenases that may play a specialized role in apocarotenoid biogenesis. A simple PCR assay was developed to measure *CCD1* gene copy number which is known to vary in maize. Using a panel of maize inbred lines varying in carotenoid content, linear regression analysis revealed a statistically significant negative correlation between copy number of *CCD1* and carotenoid content, an effect likely mediated through the resulting elevated levels of endosperm *CCD1* transcripts in high copy number lines. The PCR assay adds to a growing toolbox for metabolic engineering of maize endosperm carotenoids. This new tool can be used to select maize lines that are less likely to promote endosperm carotenoid degradation, thus predicting optimal results in metabolic engineering of endosperm provitamin A and/or nonprovitamin A carotenoids.

Introduction

In nature, carotenoids have diverse functions in both plants and animals. In plants, algae and cyanobacteria, carotenoids serve as structural and accessory pigments within the light harvesting complex to mediate photosynthesis and photoprotection. Carotenoids are also enzymatically cleaved to produce apocarotenoids, such as strigolactones, abscisic acid (ABA)[1] and other volatile compounds that contribute to the aroma of fruits and flowers [reviewed in 1–3]. Apocarotenoids play numerous physiological roles, including control of plant architecture, dormancy and stress responses and signals attracting beneficial mycorrhizal fungi that aid in nutrient uptake and promote plant growth. These same signals are utilized by harmful parasitic plants (e.g. *Striga*) that compete for plant resources [reviewed in 4].

The primary route in the formation of biologically active apocarotenoids is oxidative cleavage of carotenoids mediated by carotenoid dioxygenases [5,6]. Two types of carotenoid dioxygenases have been identified in plants, 9-cis carotenoid cleavage dioxygenases (NCEDs) and carotenoid cleavage dioxygenases (CCDs) (Fig. 1). These enzymes were identified by analyzing *viviparous* (*vp*) mutants of maize, which led to cloning of maize *NCED1* (*Vp14*) [7]. NCED enzymes cleave the 11,12 (11′,12′) double bond of 9-cis-violaxanthin or 9-cis-neoxanthin, catalyzing the first step in ABA biosynthesis [7,8]. In contrast, CCD enzymes do not share cleavage specificity with NCED enzymes. Some CCD enzymes recognize specific carotenoid or apocarotenoid substrates while others show promiscuity in choice of substrate as evidenced in heterologous systems [reviewed in 6,9–11]. CCD4 enzymes cleave β-carotene and/or 8′-apo-β-caroten-8-al, although there is no consistent CCD4 substrate among taxa, except for regioselectivity of cleavage at the 9,10 (9′,10′) positions [11,12]. In comparison, maize CCD1 has been shown *in vitro* to cleave a wide range of substrates (e.g. lycopene, β-carotene, zeaxanthin, etc.) and like other CCD1 enzymes, it is not regioselective in cleavage site [9,10,13]. CCD7 cleaves β-carotene asymmetrically, producing one β-ionone and C_{27} 10′-apo-β-carotenal; the latter product can be further cleaved by CCD8 generating C_{18} 13′-apo-β-carotenal [14] or CCD1 generating another β-ionone and one apo-10,10′-carotendial [13].

* Corresponding author at: Department of Biological Sciences, Lehman College, The City University of New York, 250 Bedford Park Boulevard West, Bronx, NY 10468, United States. Fax: +1 718 960 8236.
E-mail address: wurtzel@lehman.cuny.edu (E.T. Wurtzel).

[1] *Abbreviations used:* ABA, abscisic acid; CCDs, carotenoid cleavage dioxygenases; ESTs, expressed sequence tags; CT, threshold cycle; DAP, days after pollination.

0003-9861/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2010.07.019

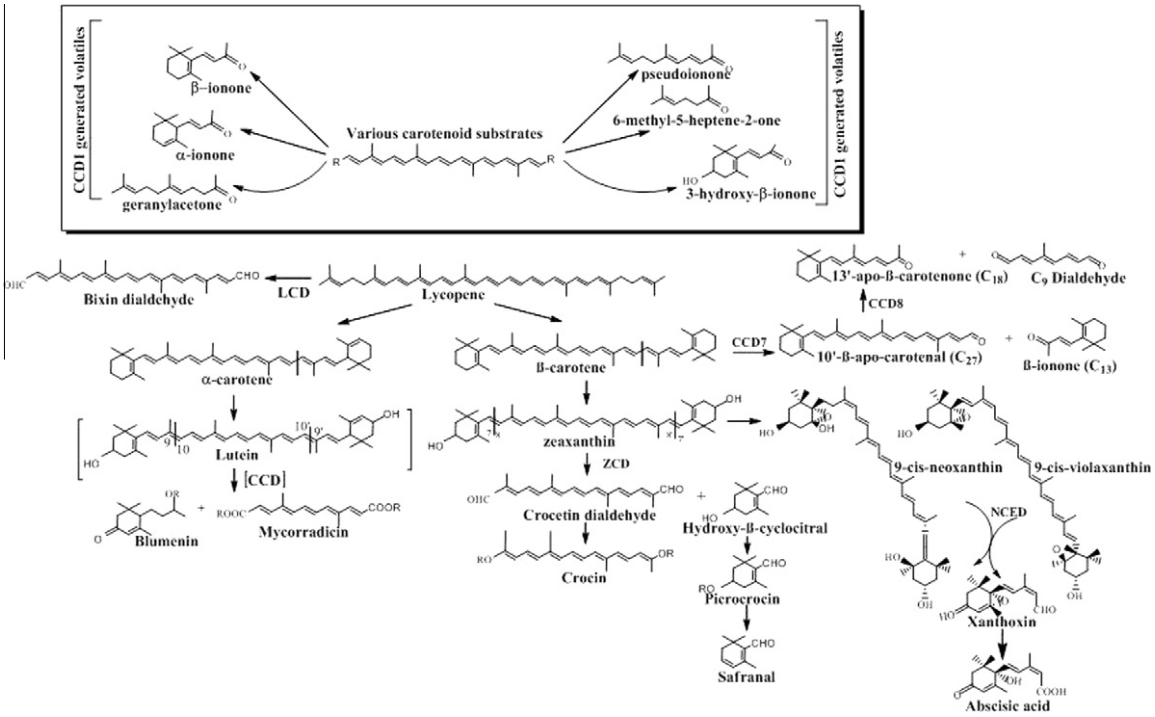

Fig. 1. Carotenoid cleavage dioxygenase activities in various species. CCD1 is promiscuous in the choice of substrate, thus producing an array of volatiles as shown inside the box [47]. Zeaxanthin cleavage dioxygenase, ZCD (*Crocus sativus*) [50]; lycopene cleavage dioxygenase, LCD (*Bixa orellana*) [5]. Formation of abscisic acid is mediated by 9-cis-epoxy carotenoid dioxygenases (NCED) [8]. CCD activity in mycorrhizal maize roots leads to formation of apocarotenoids mycorradicin and blumenin [22,51]; the hypothetical lutein substrate is indicated in square parentheses. CCD7 and CCD8 participate sequentially as shown [5].

Plant CCD enzymes generate various phytohormones and aroma compounds, although the cleavage pathways in plants are poorly understood. For example, *CCD1* expression causes the notable emission of diverse volatiles characteristic of fruit and flowers [15,16]. From studies in Arabidopsis, peas and rice, CCD7 and CCD8 were shown to be involved in generating the branching inhibition strigolactone hormones formed in roots [17–20]. These hormones signal symbiotic mycorrhizal fungal hyphae to branch and associate with plant roots or can be transported to aerial parts of the plant to inhibit branching or tillering of the plant, an important agronomic trait in the grasses [21].

Carotenoid dioxygenase genes have been identified in various plant species [22]. In the complete genome of *Arabidopsis*, nine genes similar to *VP14* were identified and characterized with regard to their tissue specific expression pattern and predicted subcellular localization [23–25]. Five genes encoding NCEDs are involved in ABA biosynthesis, while the remaining four genes encode CCDs.

Less is known about carotenoid cleavage enzymes in crops of agronomic importance. This enzyme class may potentially influence nutritional content (e.g. provitamin A carotenoids) and plant yield. The Poaceae includes the major food crops, for which carotenoid content is a target for improvement [reviewed in 26,27]. Efforts to engineer carotenoid accumulation in the grasses involve control of factors that influence net carotenoid accumulation, a balance of synthesis and degradation. Recent studies have shed light on the timing of expression and control points affecting pathway flux for carotenoid biosynthesis and accumulation in endosperm of maize [27–36]. Although there are some studies on a few individual cleavage enzymes, carotenoid degradation in the grasses is not well understood. This deficiency in knowledge will undoubtedly stymie attempts to breed high levels of carotenoids. Therefore this study focused on identifying and characterizing the entire carotenoid dioxygenase gene family in maize, rice, and sorghum.

Materials and methods

Sequence analysis and chromosome map localization

Rice genes (www.gramene.org) were used as a query to identify orthologs from *Zea mays* (www.tigr.org; www.plantgdb.org) and to decipher gene families. The genomic sequence, obtained from the respective BAC clones (Table 1), was used to deduce the full length cDNA sequence. The exon–intron boundaries were confirmed through selected ESTs, where available. Sequence analysis was performed using Vector NTI Suite 9.0 (Invitrogen, Carlsbad, CA). Translation of expressed sequence tags (ESTs) was used to distinguish gene paralogs. Chromosomal positions of genes in the *Z. mays* B73 inbred line were localized either by utilizing tools available at WebAGCoL package (http://www.agcol.arizona.edu/software/webagcol/) [37] or Maize GDB. Prediction of chloroplast targeting site was made using the ChloroP software (http://www.cbs.dtu.dk/services/ChloroP/) [38] and pSORT (http://wolfpsort.org/). Phytozome (http://www.phytozome.org) was used to retrieve sequences from *Sorghum bicolor* and all nonflowering plant and algal sequences. Phylogenetic analysis was performed using MEGA2 [39]. Deduced open reading frames and predicted proteins are presented in Tables S1 and S2, respectively.

Quantitative RT-PCR

Plants and tissues of different maize inbreds in a core subset of a maize germplasm collection were collected as described [29]. Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen Sciences, Maryland), and DNase I-treated (Invitrogen, Carlsbad, CA) prior to first strand cDNA synthesis using oligo (dT) as a primer and SuperscriptTM III RT (Invitrogen, Carlsbad, CA). One microliters of 50 μM oligo $(dT)_{20}$ and 1 μl of 10 mM dNTP mix were mixed with 8 μl of DNase I treated total RNA (~1 μg) and incubated at 65 °C for

Table 1
Summary of carotenoid cleavage dioxygenase gene families. For maize genes, information is provided for chromosome location (Map); sizes of genomic DNA, cDNA, and open reading frame (ORF) are given in base pairs. Genomic DNA sequence covers from ATG to stop codon. The predicted cDNA sequences starts from predicted transcription start site to poly-A signal. For maize protein information, number of amino acid residues is given for the mature protein (M) and TP (transit peptide). Loc, predicted cellular location; P, plastid and C, cytosolic. Pre/pI, preprotein (MW in kD)/isolectric point; M/pI, predicted mature protein (MW in kD)/isoelectric point. Rice genes (www.gramene.org) were used as a query to identify orthologs from *Zea mays* (www.tigr.org; www.plantgdb.org; www.maizesequence.org) and to decipher gene families. Phytozome (http://www.phytozome.org) was used to retrieve sequences from Sorghum bicolor and all nonflowering plant and algal sequences.

	Oryza sativa	*Zea mays*												*Sorghum bicolor*
		Accessions		Gene					Protein					
		Genomic	ESTs	Map	Introns	Genomic	cDNA[a]	ORF	M	TP	Loc	Pre/pI	M/pI	
NCED1		AC201886	U95953	1.08	0	1815	2498	1815	605	23	P	65/5.5	63/5.3	
NCED2	LOC_Os12g42280	AC199036	DV032362	3.04	0	NA	NA	1008[b]	403[b]	–	–	–	–	Sbi_0.36552[b]
NCED3a	LOC_Os07g05940	AC212820	DY541318	NA	0	1722	2208	1722	574	46	P	65/6.1	57/5.8	Sb02g003230
NCED3b		AC205109	NA	2.09	0	1728	2843	1728	576	52	P	62/6.6	56/5.8	
NCED9	LOC_Os03g44380	AC190614	NA	5.03	0	1806	2157	1806	602	23	P	65/6.4	62/6.1	Sb01g013520
CCD1	LOC_Os12g44310	DQ100347	DQ100346	9.07	13	5414	2388	1653	551	–	C	62/5.9	55/5.8	Sb01g047540
CCD4a	LOC_Os02g47510	AC190588	EE172116	4.06	0	1917	2262	1917	639	50	P	68/6.3	63/5.7	Sb04g030640
CCD4b	LOC_Os12g24800	AC194862	NA	5.06	0	1920	2382	1920	640	52	P	69/6.8	63/5.8	
CCD7	LOC_Os04g46470	AC211432	NA	2.03	6	2607	3140	1860	620	29	P	68/8.7	65/7.8	Sb06g024560
CCD8a	LOC_Os01g54270 (CCD8a)	AC185113	CO526902	3.07	3	3007	2672	1662	554	36	P	61/6.5	57/6.0	Sb05g009950 (CCD8a)
CCD8b	LOC_Os09g15240 (CCD8b)	AC198395	DW954641	7.02	11	3771	2402	1632	544	79	P	60/5.6	51/5.5	Sb03g034400 (CCD8b)
CCD8c	LOC_Os01g38580 (CCD8c)													Sb10g004360 (CCD8c) Sb10g004370 (CCD8c-like)
CCD8d	LOC_Os08g28240 (CCD8d)													Sb02g021490 (CCD8d) Sb07g024250 (CCD8d-1ike)

[a] Predicted.
[b] Partial; NA = no specific EST to a paralog.

5 min, and left on ice for at least 1 min. Ten microliters of cDNA synthesis mix (2 μl of 10× RT buffer, 4 μl of 25 mM $MgCl_2$, 2 μl of 0.1 M DTT, 1 μl of RNaseOUT™ (40 U/μl), 1 μl of Superscript™ III RT (200 U/μl) were added and incubated for 50 min at 50 °C and reactions terminated at 85 °C for 5 min. Samples were collected by brief centrifugation and 1 μl of RNase H added and incubated for 20 min at 37 °C. cDNA samples were amplified using gene specific primers (Table S3) on the MyIQ Single-Color Real-Time PCR detection system (Bio-Rad, Hercules, CA), using iQTM SYBR Green Supermix (Bio-Rad, Hercules, CA). Two microliters (5 ng/μl) of cDNA; 15 μl of 2× iQTM SYBR Green Supermix; 11 μl of nuclease-free water; 1 μl (20 ρm/μl) of each primer were used in a 30 μl reaction volume. Thermal cycling conditions included an initial incubation at 94 °C for 10 s, followed by 35 cycles of 95 °C for 10 s, 58 °C for 35 s, and 72 °C 10 s. Melt curve analysis was performed to verify primer specificity, and PCR products were confirmed by sequencing. The relative quantity of the transcripts was calculated by using the comparative threshold cycle (CT) method. Actin mRNA was amplified simultaneously for normalization between samples. Values are expressed as the mean of three RT-PCR replicates ±standard deviation.

Relative quantification of gene copy number

Genomic DNA was quantified using the NanoDrop 1000 (Thermo-Scientific) and diluted to 10 ng/μl. Reactions for quantitative real-time PCR were prepared as follows: the 30 μl reaction volume contained 15 μl Sybr Green master mix (Bio-Rad, CA), 1 μl (20 ρmol/μl) of each primer (Table S1), 11 μl of water and 2 μl (10 ng/μl) of template. Amplifications were performed in a Bio-Rad-iCycler real-time PCR machine. Thermal cycling conditions included an initial incubation at 94 °C for 10 s, followed by 35 cycles of 95 °C for 10 s, 58 °C for 35 s, and 72 °C 10 s. Both control *phytoene desaturase* (*PDS*) and target *carotenoid cleavage1* (*CCD1*) genes were amplified from each sample to calculate ΔCt (ΔCt = Control Ct − Target Ct) values. *PDS*, a single copy gene in the maize genome, was used to normalize values for copy number. All samples were run in triplicate and standard deviation was calculated. Relative quantification of gene copy number ($2^{\Delta Ct}$) was used to plot against total carotenoid content of each sample using available data on carotenoids from different maize inbred lines [33].

Statistical analyses

Pearson correlation analysis of transcript and carotenoid composition of different maize inbred lines in a core subset of a maize germplasm collection [29] and linear regression plot for gene copy number vs. total carotenoid content was performed using JMP v. 5.1.2 (SAS Institute, Cary, NC) to test the statistical significance ($p \leqslant 0.05$) of the relationship.

Results and discussion

Carotenoid dioxygenase gene family in the grasses

Using available sequence data, genes encoding carotenoid dioxygenases were identified from the grass species maize, rice and

sorghum, and from an alga (*Chlamydomonas reinhardtii*) and two non-seed plants (*Selaginella moellendorffii*, a lycophyte and *Physcomitrella patens*, a bryophyte) (Tables 1 and 2; Fig. 2). Genes were characterized and named according to homology with genes from *Arabidopsis* [23]. For the three species spanning two subfamilies of the Poaceae, carotenoid dioxygenases were found to be encoded by a large gene family. This family was divided into seven groups encompassing two clades having homology to genes encoding 9-cis-epoxy carotenoid cleavage dioxygenases (NCEDs) and five clades having homology to carotenoid cleavage dioxygneases (CCDs).

Table 1 summarizes the structural properties of the genes and encoded proteins from maize, sorghum and rice. The alignment of the predicted proteins from maize is shown in Fig. S1. The amino acid sequences of carotenoid dioxygenases reveal the presence of four conserved histidines that coordinate iron atoms and are required for catalytic activity [8,40,41]. All maize cleavage dioxygenases are predicted to be plastid-localized except for CCD1, as previously reported [42]. Chromosome location was determined (Table 1) and no apparent linkage was seen among the gene family members of maize.

As shown in Fig. 2, the *NCED* genes and the *CCD4*-related genes of grasses lack introns, while all other *CCD* genes contained multiple introns. There are reports of *CCD4* genes from a few plant species containing one or two introns, these include *Chrysanthemum morifolium*, [43], *Malus domestica* and *Osmanthus fragrans* [11], while others from *Arabidopsis thaliana* and *Rosa damascena* are intronless [11].

Fig. 3 shows analysis of the grass enzymes in a broader phylogenetic context including eudicots and lower plants. In the *NCED* clade, the phylogenetic tree shows that gene duplications arose after the monocot–eudicot evolutionary split. Based on this evidence, expression profiles in eudicots would not be predictive of *NCED* transcript profiles in the grasses, thus warranting further analysis in the grasses. The *NCED* genes identified using the recently published *S. moellendorffii* genome (labeled as *Sm*) showed sequence placement basal to all other NCEDs in the tree.

Table 2
Other carotenoid cleavage dioxygenase genes identified in this study. *Vv*, *Vitis vinifera*; *Sm*, *Selaginella moellendorfii*; *Ppa*, *Physcomitrella patens*; *Cre*, *Chlamydomonas reinhardtii*. Note that *Physcomitrella* and *Chlamydomonas* contain only one *CCD8* gene of the *CCD8a* type (see Fig. 3).

Species	# Genes	Gene	Accession numbers
Vitis vinifera	9	*VvNCED2*	GSVIVT00028310001
		VvNCED3	GSVIVT00000988001
		ViNCED6	GSVIVT00020467001
		VvCCD1	GSVIVT00028793001
		VvCCD4a	GSVIVT00001163001
		VvCCD4b	GSVIVT00013949001
		VvCCD7	GSVIVT0001914600
		VvCCD8a	GSVIVT0003242300
		VvCCD8b	GSVIVT00032436001
Selaginella moellendorfii	15	*SmNCEDa*	JGI11287, JGI11289, JGI11292, JGI11304, JGI79628, JGI80651, JGI91815, JGI94523
		SmNCEDb	JGI233638
		SmNCEDc	JGI75383
		SmCCD1	JGI165469, JGI272067
		SmCCD7	JGI11120
		SmCCD8a	JGI173016
		SmCCD8b	JGI132803
Physcomitrella patens	6	*PpaNCED*	JGI57876
		PpaCCD1	JGI159406, JGI158869, JGI108830
		PpaCCD7	JGI115720
		PpaCCD8	JGI115616
Chlamydomonas reinhardtii	2	*CreCCD7*	JGI144745
		CreCCD8	JGI141185

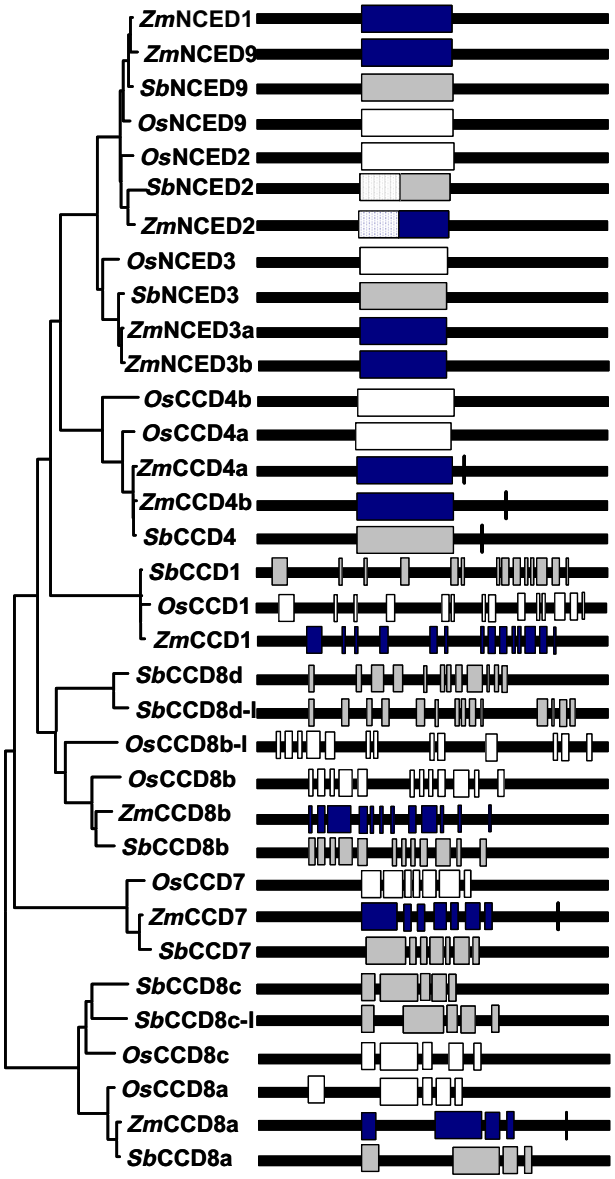

Fig. 2. Phylogenetic analysis of carotenoid cleavage dioxygenases in the Grasses. Un-rooted neighbor-joining tree was constructed using protein sequences shown at left. Corresponding gene structures are shown to right of each protein. *Zea mays* (*Zm*); *Oryza sativa* (*Os*); *Sorghum bicolor* (*Sb*). Exons for maize, sorghum and rice are indicated by boxes colored dark, grey, and white, respectively; thick horizontal lines are introns. *ZmNCED2* and *SbNCED2* are partial as indicated by dotted boxes. Exon structures starts with ATG, and vertical thin lines indicate poly-A signals.

Maize and rice CCD1 enzymes cleave a wide range of carotenoid substrates without any common mechanism of cleavage regioselectivity [9,10,13]. Here we identified *CCD1* gene homologs in the grasses along with previously unknown sequences in *Selaginella* (*Sm*) and *P. patens* (*Ppa*). *CCD4* homologs were only found in the flowering plants and not in *S. moellendorffii*, *P. patens*, or *C. reinhardtii*. Preferential expression of *CCD4* in floral organs of *Chrysanthemum* and *Crocus* [12,43] suggests that evolution of *CCD4* genes may have facilitated adaptation for specific physiological capacities unique to flowering plants. A new class of *CCD* genes, discovered in the phylogenetic analysis (Fig. 3) was the group containing maize *CCD8b*. The aligned protein sequences for CCD8-related sequences are shown in Fig. S2 (CCD8a) and Fig. S3 (CCD8b). *CCD8b* is phylogenetically closer to *CCD1* and has a unique exon–intron structure distinguishable from the *CCD8a* genes (Fig. 2). *CCD8b* homologs were not found in *Arabidopsis*. The *CCD8b*

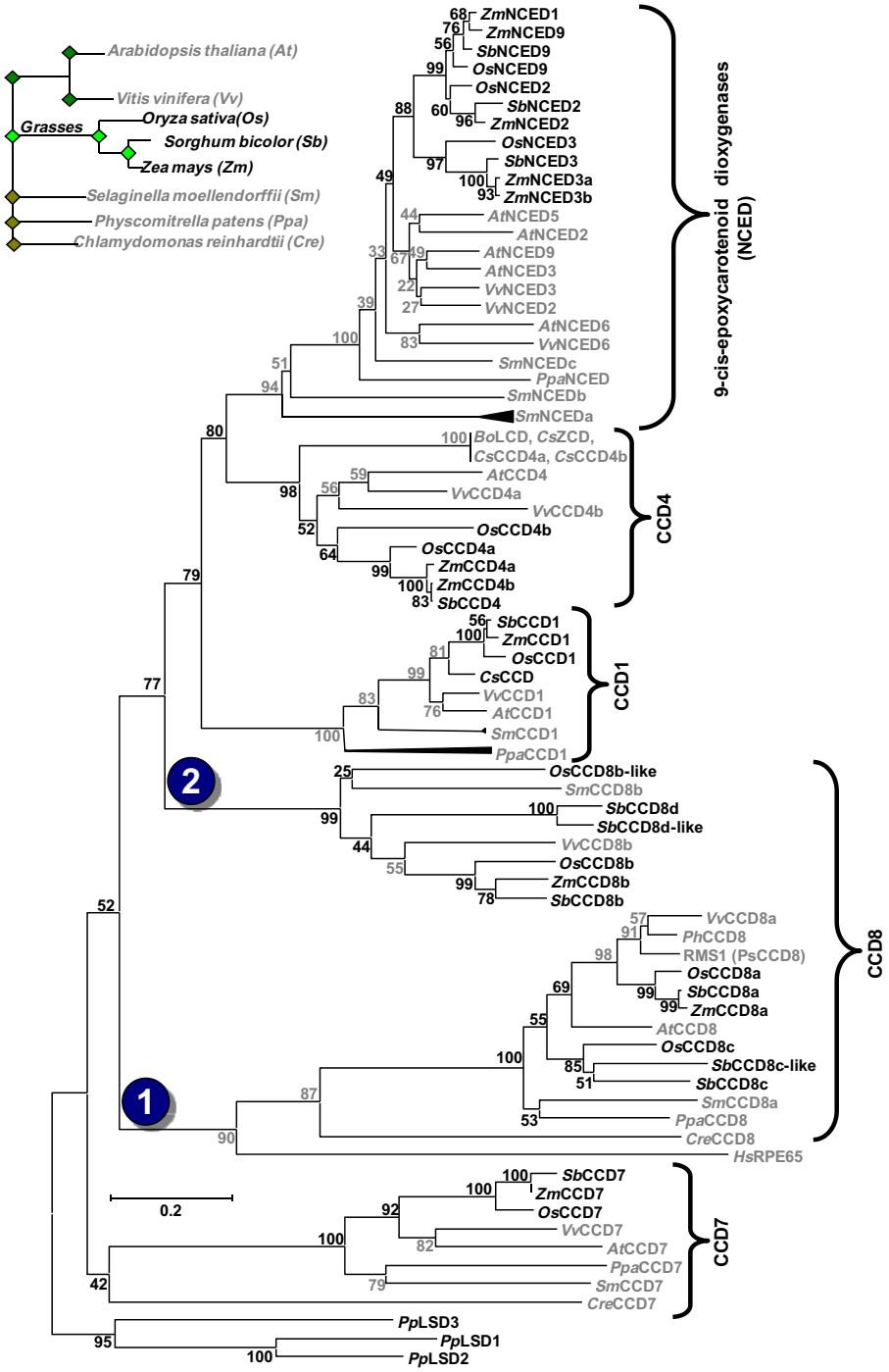

Fig. 3. Phylogenetic analysis of carotenoid cleavage dioxygenases from diverse taxa. Neighbor-joining tree was constructed using protein sequences. *Zea mays, Oryza sativa, Sorghum bicolor* (see Table 1 for accession numbers), *Arabidopsis thaliana* (AtNCED2, At4g18350; AtNCED3, At3g14440; AtNCED5, At1g30100; AtNCED6, At3g24220; AtNCED9, At1g78390; AtCCD1, At3g63520; AtCCD4, At4g19170; AtCCD7, At2g44990; AtCCD8, At4g32810), *Bixa orellana* (BoLCD, AJ489277); *Crocus sativus* (*Cs*ZCD, CAD33262, *Cs*CCD, AJ132927; CsCCD4a, EU523662; CsCCD4b, EU523663); *Sphingomonas paucimobilis* (*Pp*LSD1, S80637; *Pp*LSD2, S65040; *Pp*LSD3, AB073227); *Homo sapiens* (*Hs*RPE65, AF039855); *Pisum sativum* (RMS1, AY557341); *Petunia hybrida* (*Ph*CCD8, AY743219). Refer to Table 2 for genes from: *Chlamydomonas reinhardtii Selaginella moellendorfii, Physcomitrella patens* and *Vitis vinifera*. Numbers 1 and 2 indicate two distinct clusters of CCD8. Amino acid distances were corrected by Poisson correction. Numbers indicate bootstrap support for individual nodes. Bootstrap support higher than 40% is indicated at respective nodes ($n = 500$). Grass species enzymes are shown in black; except for *Pp* genes, others are grey. Black triangles represent additional gene copies shown in a compressed fashion (detailed in Table 2). The evolution of representative green plants is shown at the top left hand corner.

group is not unique to the Grasses since homologs were found for a eudicot, *Vitis vinifera* (*Vv*), and the nonflowering plants and alga. Homologs of the sister group, *CCD8a*, have been described in *Arabidopsis* and other species. Mutants in the *CCD8a* homolog of *Arabidopsis* and pea (e.g. *MAX4, RMS1*) are known to interfere with plant development [44]. CCD8a cleaves the apocarotenoid C_{27} 10′-apo-β-carotenal generating C_{18} 13′-apo-β-carotenal [14] in formation of strigolactones, the branching inhibition hormones [19,20]. The human CCD enzyme (*Hs*RPE65), phylogenetic sister of both CCD8a and CCD8b enzymes, cleaves β-carotene at the 15, 15′ dou-

ble bond to form vitamin A [45]. The CCD8b enzymes form a separate group from CCD8a suggesting that CCD8b enzymes could potentially serve a novel role unrelated to strigolactone biosynthesis. Alternatively, *CCD8b* genes may be duplicative in terms of enzyme activity but have alternative expression profiles or cell-type specificities. The specific enzymatic role of the *CCD8b* group remains elusive.

The CCD8a substrate for biosynthesis of strigolactones is produced by CCD7, represented by the remaining CCD clade in the phylogenetic tree. *CCD7* homologs were identified in maize (*Zm*), rice (*Os*), and sorghum (*Sb*), along with *S. moellendorffii* (*Sm*), *P. patens* (*Ppa*) and *C. reinhardtii* (*Cre*). To date, all *CCD7* genes identified are single copy. The significance of the single gene copy of the *CCD7* clade in comparison to the *CCD8* clade is unknown. Homologs for CCD7 and *CCD8* were also found in *C. reinhardtii* (*Cre*); interestingly this algal species lacked homologs for all other CCD and NCED genes. The CCD7 clade remained most closely related to lipoxygenase dioxygenases that are encoded by the soil bacterium *Sphingomonas paucimobilis* [46].

Tissue specific distribution of dioxygenase gene transcripts

Tissue specificity was examined in the B73 maize inbred line, where transcript levels for the various *NCED* and *CCD* genes were assessed in a range of tissues including leaves, roots, endosperm and embryo, the latter two staged at 20 days after pollination (DAP) as well as 25 DAP for endosperm (Figs. 4 and 5). This particular inbred line was chosen since it was the source of the first maize genome sequence and has been used extensively for analysis of genes that influence carotenoid content of maize [29,31,32,36].

NCED transcripts

In leaf tissue, where carotenoids are abundant, NCED cleavage of 9-*cis*-epoxycarotenoids represents the rate-limiting step in biosynthesis of ABA [3]. While *NCED1* transcript accumulation is induced in roots of abiotic-stressed maize plants, ABA formation in roots is limited by biosynthesis of carotenoid precursors [32]. To assess the potential contribution of *NCED* gene transcripts for carotenoid degradation in different tissues under non-stressed conditions, transcript levels were quantitatively measured for the various tissues (Fig. 4). In most tissues including leaves, *NCED1* transcripts represented the most prevalent NCED transcript. *NCED3a* and *NCED9* transcripts were also present in leaves, while roots contained equivalent levels of *NCED1* and *NCED3a* with 60% relative levels of *NCED9* transcripts. Embryos contained mostly *NCED1* and ~25% relative levels of *NCED9* transcripts. In endosperm, transcripts of *NCED1* followed by *NCED9* represented the most prevalent *NCED* classes. However, as the endosperm approached dormancy from 20 to 25 DAP, relative levels switched between the two genes as *NCED9* transcripts increase two to threefold.

CCD transcripts

CCD enzymes vary in choice of substrate and in potential apocarotenoid product [reviewed in 6,47]. Differential expression of CCD genes combined with tissue-specific carotenoid profiles will yield a suite of apocarotenoid products depending on combinations of genes expressed. Leaves showed the highest expression of *CCD* transcripts, with *CCD1* and *CCD4a* representing the most abundant transcript classes, respectively (Fig. 4); lower amounts of *CCD4b* and barely detectable levels of *CCD7* transcripts were also detected in leaves. Leaves did not contain any detectable levels of *CCD8a* or *CCD8b* transcripts. In contrast to the high levels seen in leaves, the endosperm of the maize B73 inbred line had *CCD1* mRNA levels that were three orders of magnitude lower. Under normal growth conditions, the prevalent CCD classes found in roots, in order of abundance, were *CCD1*, *CCD8b*, and *CCD7*. However, these levels were significantly lower than those found in leaf tissue. For example, *CCD1* was about 30-fold lower in roots as compared to leaves.

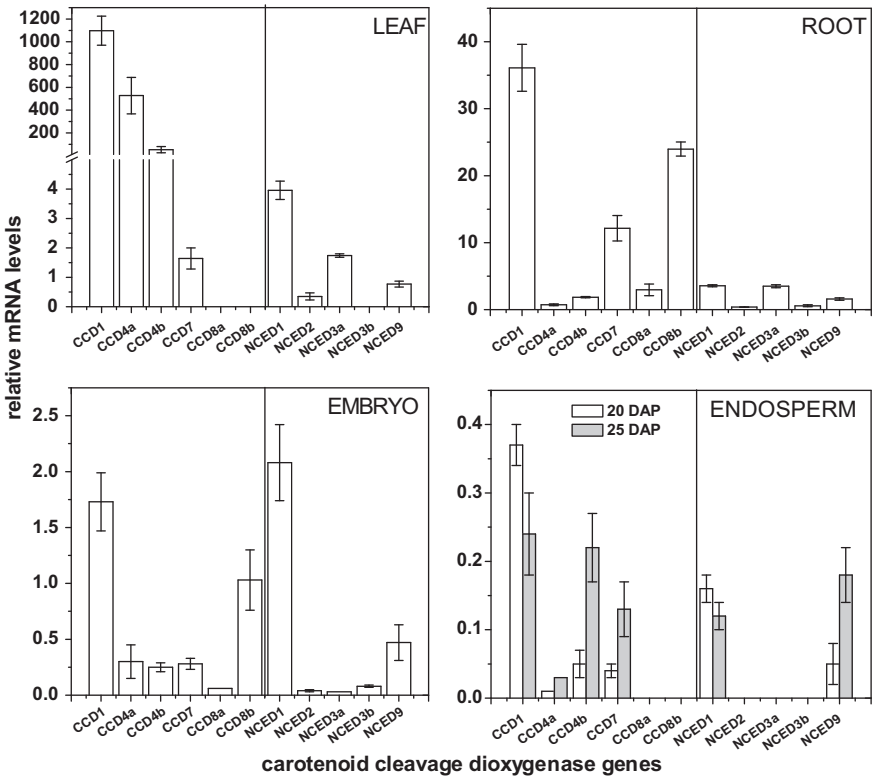

Fig. 4. Maize B73 inbred line transcript profiles of genes encoding (A) carotenoid cleavage dioxygenases (CCD) (B) 9-cis-epoxy carotenoid dioxygenases (NCED). Mean of three replicates ±standard deviation. Embryo collected at 20 DAP and endosperm collected at 20 and 25 DAP from field grown plants; leaf and root samples collected from seedlings at the six-leaf stage.

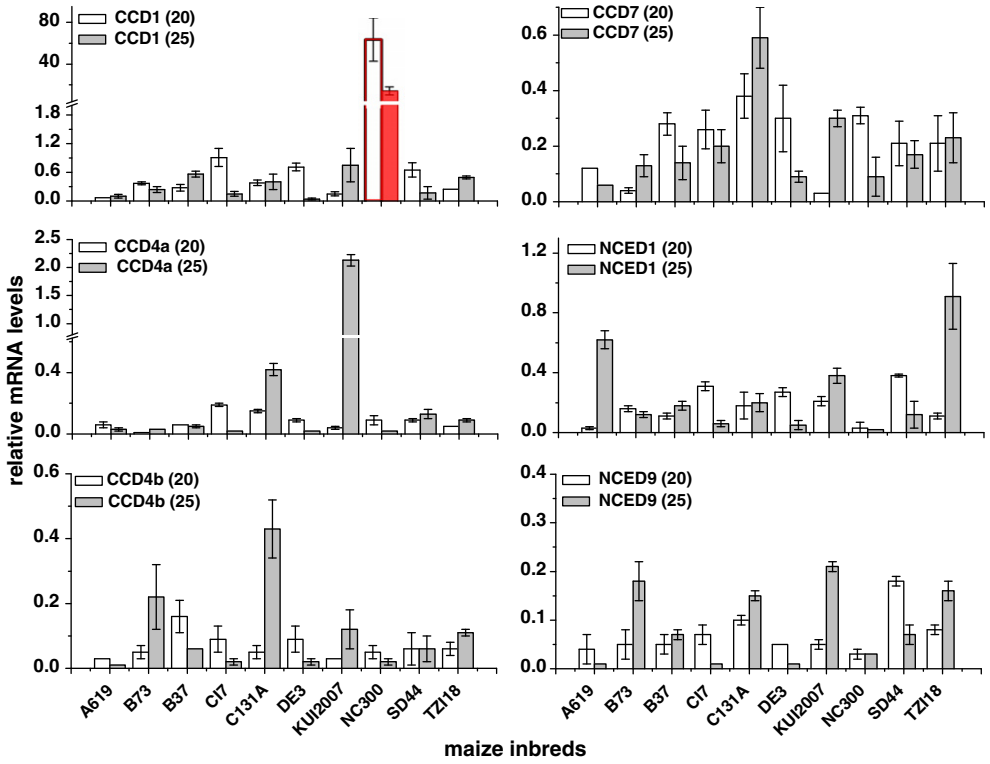

Fig. 5. Transcript profiles of genes showing detectable levels of expression in a core subset of maize inbred lines. Transcript levels were measured at both 20 and 25 DAP during endosperm development. Mean of three replicates ±standard deviation.

CCD1 expression in roots has been associated with formation of mycorradicin, a root apocarotenoid produced after mycorrhizal fungal colonization [48,9]. Roots proved to be the tissue where the highest levels of *CCD7, CCD8a* and *CCD8b* were detected. The expression of *CCD7* in roots was consistent with expression of orthologs in *Arabidopsis*, pea and petunia [17,42,44,49]. Maize roots preferentially accumulated CCD8b over transcripts of CCD8a. Among the tissues sampled, CCD transcripts were least prevalent in 20 DAP embryo.

Other maize inbred lines

While the dioxygenase gene family members showed specific patterns of expression in B73, there may be variation among genetically diverse maize cultivars. Therefore, transcript levels were assessed for the NCED and CCD family members in a 10-line core set of genetically diverse maize inbred lines. These particular lines have been used for identifying control points in the carotenoid biosynthetic pathway. We previously showed that transcript levels of genes encoding several key steps in carotenoid biosynthesis correlated with carotenoid content and/or altered composition in maize endosperm [28,29,31,33]. Expression in endosperm was given specific focus here since it was unknown whether there might also be variation in transcript levels of carotenoid cleavage genes that could counteract efforts to maximize levels of nutritionally targeted carotenoids in endosperm. Of all the cleavage genes analyzed, *CCD1* was the only gene family member showing unusually elevated transcripts in one maize line, NC300 (Figs. 5 and 6A). The peak of CCD1 transcript levels in NC300 was at 20 days after pollination (DAP), the stage found to be a critical time point of gene expression related to endosperm carotenoid accumulation [29,31]. In the other nine maize lines, *CCD1* mRNA level ranged from ~100 to 300-fold lower. Some genes lacked detectable transcripts in endosperm of the ten lines and were therefore not

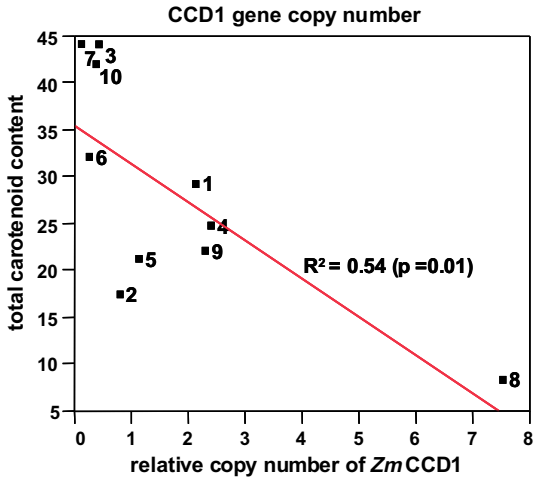

Fig. 6. Gene copy number of *ZmCCD1* relative to *ZmPDS*. Mean of three replicates ±standard deviation. Line indicates linear fit line. Inbred lines are: 1, A619; 2, B73; 3, B37; 4, CI.7; 5, C131A; 6, DE3; 7, KUI2007; 8, NC300; 9, SD44; 10, TZI18.

shown. For the five other cleavage genes shown, there was variation in expression, but not of such magnitude.

Maize *CCD1* is located on *Chromosome* 9.07 (see chromosome locations, Table 1), and is linked to the *white cap1* (*wc1*) locus for which dominant alleles exhibit reduced endosperm carotenoids (Maize Genetics and Genomics Database: http://www.maizegdb.org/). The Maize Genetics and Genomics Database provides unpublished data indicating that the *wc1* locus encoding *CCD1* is associated with variation in gene copy number at the locus (www.maizegdb.org). Therefore, correlation between gene copy number and carotenoid content of the 10 lines was tested. Quantitative real-time PCR was conducted using genomic DNA from each

line, to establish copy number at the *CCD1* locus by normalizing with *PDS* (phytoene desaturase), a single copy gene (as described in Materials and Methods). As shown in Fig. 6B, linear regression analysis revealed a statistically significant correlation ($R^2 = 0.54$ and $p = 0.01$) between copy number of *CCD1* and carotenoid content. Of note was line NC300 with ~7–8 copies of the *CCD1* gene which could explain the unusually high *CCD1* transcript level in that line compared to the other nine lines tested.*CCD1* of maize has been shown to cleave multiple carotenoid substrates at varying positions [9,10]. Together these results suggest that *CCD1* expression may limit maize seed carotenoid content regardless of which carotenoid is preferentially accumulated, thereby posing a challenge for metabolic engineering efforts.

Conclusion

The carotenoid cleavage dioxygenase gene family was defined for three grass species in the context of phylogenetic analysis for a wide evolutionary range of plants. Gene loci and transcript profiling for maize, together with identification of homologs in non-flowering plants, provide a foundation needed for basic research and future crop improvement. Apocarotenoids contribute to valuable agronomic traits such as tillering and abiotic and biotic stress responses. However, apocarotenoid formation occurs at the expense of carotenoid accumulation. The finding that *CCD1* copy number is associated with reduced endosperm carotenoids has important implications for metabolic engineering of endosperm. Since maize CCD1 is known to cleave multiple carotenoid substrates, efforts to enhance provitamin A or nonprovitamin A carotenoids could be negatively affected by presence of a high copy *CCD1* allele. The PCR assay described here is a valuable new tool for identifying optimal breeding lines that are more likely to exhibit carotenoid stability in endosperm because *CCD1* gene copy number will be predicted to be low. This new tool is an important component adding to other recently described tools and/or control points [28,29,31,33] in a growing toolbox needed to achieve high provitamin A carotenoid endosperm in maize for combating global vitamin A deficiency.

Acknowledgments

We thank Dr. Dwight Kincaid (Lehman College) for advice on the statistical analysis. This research was supported by Grants (to ETW) from NIH (S06-GM08225, 1SC1GM081160-01, and 5SC1G M081160-02).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2010.07.019.

References

[1] H.J. Klee, New Phytol. 187 (2010) 44–56.
[2] M.H. Walter, D.S. Floss, D. Strack, Planta 232 (2010) 1–17.
[3] E. Nambara, A. Marion-Poll, Annu. Rev. Plant Biol. 56 (2005) 165–185.
[4] K. Akiyama, H. Hayashi, New Phytol. 178 (2008) 695–698.
[5] F. Bouvier, J.-C. Isner, O. Dogbo, B. Camara, Trends Plant Sci. 10 (2005) 187–194.
[6] A. Ohmiya, Plant Biotechnol. (Japan) 26 (2009) 351–358.
[7] B.C. Tan, S.H. Schwartz, J.A. Zeevaart, D.R. McCarty, Proc. Natl. Acad. Sci. USA 94 (1997) 12235–12240.
[8] S.H. Schwartz, B.C. Tan, D.A. Gage, J.A. Zeevaart, D.R. McCarty, Science 276 (1997) 1872–1874.
[9] Z. Sun, J. Hans, M. Walter, R. Matusova, J. Beekwilder, F. Verstappen, Z. Ming, E. van Echtelt, D. Strack, T. Bisseling, H. Bouwmeester, Planta 228 (2008) 789–801.
[10] J.T. Vogel, B.-C. Tan, D.R. McCarty, H.J. Klee, J. Biol. Chem. 283 (2008) 11364–11373.
[11] F.C. Huang, P. Molnar, W. Schwab, J. Exp. Botany 60 (2009) 3011–3022.
[12] A. Rubio, J.L. Rambla, M. Santaella, M.D. Gomez, D. Orzaez, A. Granell, L. Gomez-Gomez, J. Biol. Chem. 283 (2008) 24816–24825.
[13] A. Ilg, P. Beyer, S. Al-Babili, Febs J. 276 (2009) 736–747.
[14] S.H. Schwartz, X. Qin, M.C. Loewen, J. Biol. Chem. 279 (2004) 46940–46945.
[15] A.J. Simkin, S.H. Schwartz, M. Auldridge, M.G. Taylor, H.J. Klee, Plant J. 40 (2004) 882–892.
[16] A.J. Simkin, B.A. Underwood, M. Auldridge, H.M. Loucas, K. Shibuya, E. Schmelz, D.G. Clark, H.J. Klee, Plant Physiol. 136 (2004) 3504–3514.
[17] J. Booker, M. Auldridge, S. Wills, D. McCarty, H. Klee, O. Leyser, Curr. Biol. 14 (2004) 1232–1238.
[18] K. Sorefan, J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, E. Foo, S. Chatfield, S. Ward, C. Beveridge, C. Rameau, O. Leyser, Genes Dev. 17 (2003) 1469–1474.
[19] V. Gomez-Roldan, S. Fermas, P.B. Brewer, V. Puech-Pages, E.A. Dun, J.P. Pillot, F. Letisse, R. Matusova, S. Danoun, J.C. Portais, H. Bouwmeester, G. Becard, C.A. Beveridge, C. Rameau, S.F. Rochange, Nature 455 (2008) 189–194.
[20] M. Umehara, A. Hanada, S. Yoshida, K. Akiyama, T. Arite, N. Takeda-Kamiya, H. Magome, Y. Kamiya, K. Shirasu, K. Yoneyama, J. Kyozuka, S. Yamaguchi, Nature 455 (2008) 195–200.
[21] M.D. Purugganan, D.Q. Fuller, Nature 457 (2009) 843–848.
[22] G. Giuliano, S. Al-Babili, J. von Lintig, Trends Plant Sci. 8 (2003) 145–149.
[23] B.C. Tan, L.M. Joseph, W.T. Deng, L. Liu, Q.B. Li, K. Cline, D.R. McCarty, Plant J. 35 (2003) 44–56.
[24] S. Iuchi, M. Kobayashi, T. Taji, M. Naramoto, M. Seki, T. Kato, S. Tabata, Y. Kakubari, K. Yamaguchi-Shinozaki, K. Shinozaki, Plant J. 27 (2001) 325–333.
[25] V. Lefebvre, H. North, A. Frey, B. Sotta, M. Seo, M. Okamoto, E. Nambara, A. Marion-Poll, Plant J. 45 (2006) 309–319.
[26] G. Giuliano, R. Tavazza, G. Diretto, P. Beyer, M.A. Taylor, Trends Biotech. 26 (2008) 139–145.
[27] P.D. Matthews, E.T. Wurtzel, Biotechnology of food colorant production, in: C. Socaciu (Ed.), Food Colorants: Chemical and Functional Properties, CRC Press, Boca Raton, 2007, pp. 347–398.
[28] R. Vallabhaneni, C.E. Gallagher, N. Licciardello, A.J. Cuttriss, R.F. Quinlan, E.T. Wurtzel, Plant Physiol. 151 (2009) 1635–1645.
[29] R. Vallabhaneni, E.T. Wurtzel, Plant Physiol. 150 (2009) 562–572.
[30] F. Li, O. Tzfadia, E.T. Wurtzel, Plant Signal. Behav. 4 (2009) 208–211.
[31] F. Li, R. Vallabhaneni, J. Yu, T. Rocheford, E.T. Wurtzel, Plant Physiol. 147 (2008) 1334–1346.
[32] F. Li, R. Vallabhaneni, E.T. Wurtzel, Plant Physiol. 146 (2008) 1333–1345.
[33] C.E. Harjes, T.R. Rocheford, L. Bai, T.P. Brutnell, C.B. Kandianis, S.G. Sowinski, A.E. Stapleton, R. Vallabhaneni, M. Williams, E.T. Wurtzel, J. Yan, E.S. Buckler, Science 319 (2008) 330–333.
[34] F. Li, C. Murillo, E.T. Wurtzel, Plant Physiol. 144 (2007) 1181–1189.
[35] E.T. Wurtzel, Genomics, genetics, and biochemistry of maize carotenoid biosynthesis, in: J. Romeo (Ed.), Recent Advances in Phytochemistry, Elsevier Ltd., 2004, pp. 85–110.
[36] C.E. Gallagher, P.D. Matthews, F. Li, E.T. Wurtzel, Plant Physiol. 135 (2004) 1776–1783.
[37] V. Pampanwar, F. Engler, J. Hatfield, S. Blundy, G. Gupta, C. Soderlund, Plant Physiol. 138 (2005) 116–126.
[38] O. Emanuelsson, H. Nielsen, G. von Heijne, Protein Sci. 8 (1999) 978–984.
[39] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, Bioinformatics (2001) 244–1245.
[40] E.K. Marasco, K. Vay, C. Schmidt-Dannert, J. Biol. Chem. 281 (2006) 31583–31593.
[41] C. Kiefer, S. Hessel, J.M. Lampert, K. Vogt, M.O. Lederer, D.E. Breithaupt, J. von Lintig, J. Biol. Chem. 276 (2001) 14110–14116.
[42] M.E. Auldridge, A. Block, J.T. Vogel, C. Dabney-Smith, I. Mila, M. Bouzayen, M. Magallanes-Lundback, D. DellaPenna, D.R. McCarty, H.J. Klee, Plant J. 45 (2006) 982–993.
[43] A. Ohmiya, S. Kishimoto, R. Aida, S. Yoshioka, K. Sumitomo, Plant Physiol. 142 (2006) 1193–1201.
[44] K. Sorefan, J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, Genes Dev. 17 (2003) 1469–1474.
[45] T. Redmond, S. Yu, E. Lee, D. Bok, D. Hamasaki, N. Chen, P. Goletz, J. Ma, R. Crouch, K. Pfeifer, Nat. Genet. 20 (1998) 344–351.
[46] T.M. Redmond, S. Gentleman, T. Duncan, S. Yu, B. Wiggert, E. Gantt, F.X. Cunningham, J. Biol. Chem. 276 (2001) 6560–6565.
[47] M.E. Auldridge, D.R. McCarty, H.J. Klee, Curr. Opin. Plant Biol. 9 (2006) 315–321.
[48] D.S. Floss, W. Schliemann, J. Schmidt, D. Strack, M.H. Walter, Plant Physiol. 148 (2008) 1267–1282.
[49] K.C. Snowden, A.J. Simkin, B.J. Janssen, K.R. Templeton, H.M. Loucas, Plant Cell 17 (2005) 746–759.
[50] F. Bouvier, C. Suire, J. Mutterer, B. Camara, Plant Cell 15 (2003) 47–62.
[51] M.H. Walter, T. Fester, D. Strack, Plant J. 21 (2000) 571–578.