ZEAXANTHIN EPOXIDASE Activity Potentiates Carotenoid Degradation in Maturing Seed[1][OPEN]

Sabrina Gonzalez-Jorge, Payam Mehrshahi[2], Maria Magallanes-Lundback[2], Alexander E. Lipka, Ruthie Angelovici, Michael A. Gore, and Dean DellaPenna*

Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824 (S.G.-J., P.M., M.M.-L., R.A., D.D.P.); Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA United Kingdom (S.G.-J., P.M.); Department of Crop Sciences, College of Agricultural, Consumer, and Environmental Sciences, University of Illinois, Urbana-Champaign, Illinois 61801 (A.E.L.); Division of Biological Sciences, University of Missouri, Columbia, Missouri 65201 (R.A.); and Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, New York 14853 (M.A.G.)

ORCID IDs: 0000-0002-0783-1929 (S.G.-J.); 0000-0002-7192-0942 (P.M.); 0000-0001-6896-8024 (M.A.G.).

Elucidation of the carotenoid biosynthetic pathway has enabled altering the composition and content of carotenoids in various plants, but to achieve desired nutritional impacts, the genetic components regulating carotenoid homeostasis in seed, the plant organ consumed in greatest abundance, must be elucidated. We used a combination of linkage mapping, genome-wide association studies (GWAS), and pathway-level analysis to identify nine loci that impact the natural variation of seed carotenoids in Arabidopsis (*Arabidopsis thaliana*). *ZEAXANTHIN EPOXIDASE* (*ZEP*) was the major contributor to carotenoid composition, with mutants lacking ZEP activity showing a remarkable 6-fold increase in total seed carotenoids relative to the wild type. Natural variation in *ZEP* gene expression during seed development was identified as the underlying mechanism for fine-tuning carotenoid composition, stability, and ultimately content in Arabidopsis seed. We previously showed that two CAROTENOID CLEAVAGE DIOXYGENASE enzymes, CCD1 and CCD4, are the primary mediators of seed carotenoid degradation, and here we demonstrate that *ZEP* acts as an upstream control point of carotenoid homeostasis, with *ZEP*-mediated epoxidation targeting carotenoids for degradation by CCD enzymes. Finally, four of the nine loci/enzymatic activities identified as underlying natural variation in Arabidopsis seed carotenoids also were identified in a recent GWAS of maize (*Zea mays*) kernel carotenoid variation. This first comparison of the natural variation in seed carotenoids in monocots and dicots suggests a surprising overlap in the genetic architecture of these traits between the two lineages and provides a list of likely candidates to target for selecting seed carotenoid variation in other species.

Carotenoids are a group of more than 700 lipid-soluble pigments synthesized by all photosynthetic organisms and some nonphotosynthetic bacteria and fungi (Ruiz-Sola and Rodríguez-Concepción, 2012). In leaves of higher plants, carotenoids are essential structural and functional components of the photosynthetic machinery with roles in light harvesting, in non-photochemical quenching, and in limiting membrane damage by reactive oxygen species and singlet oxygen species (DellaPenna and Pogson, 2006; Howitt and Pogson, 2006; Cuttriss et al., 2011). Carotenoids also are essential for seed maturation and dormancy as biosynthetic precursors for the synthesis of the plant hormones abscisic acid (ABA) and strigolactones (Koornneef et al., 2002; Frey et al., 2006; Gomez-Roldan et al., 2008; Alder et al., 2012).

Carotenoids also are essential for animal growth, development, and maintenance, primarily through their roles as biosynthetic precursors for vitamin A synthesis. As animals cannot de novo synthesize carotenoids, they must be obtained through dietary intake, most often from plant-derived foods (Krinsky and Johnson, 2005). Such dietary carotenoids are divided into two classes: the nonprovitamin A carotenoids, which include the majority of carotenoids; and the provitamin A carotenoids, which are limited in number and include α-, β-, and γ-carotenes and β-cryptoxanthin. Provitamin A carotenoids are distinguished by the presence of a β-ionone ring, which allows their enzymatic conversion to vitamin A in the body. Although nonprovitamin A carotenoids by definition cannot

[1] This work was supported by the National Science Foundation (grant no. DBI–4 0922493 to D.D.P.).

[2] These authors contributed equally to the article.

* Address correspondence to dellapen@msu.edu.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Dean DellaPenna (dellapen@msu.edu).

S.G.-J. designed and performed research and wrote the article; P.M. designed research and wrote the article; M.M.-L., A.E.L., and R.A. performed data analysis; M.A.G. designed and performed data analysis; D.D.P. designed research and wrote the article.

[OPEN] Articles can be viewed without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.16.00604

serve as substrates for vitamin A synthesis, they are important nonetheless for human health. For example, elevated dietary intake of the nonprovitamin A carotenoids lutein, zeaxanthin, and lycopene has been implicated in the prevention of macular degeneration and in reducing the risk of cardiovascular disease and some cancers (Landrum and Bone, 2004; Hobbs and Bernstein, 2014).

In plants, all carotenoids are synthesized in the plastid from geranylgeranyl diphosphate produced by the plastid-localized methylerythritol 4-phosphate (MEP) pathway (Supplemental Fig. S1). The committed step in the carotenoid pathway is the condensation of two geranylgeranyl diphosphate molecules to yield phytoene, which is further desaturated and isomerized to trans-lycopene. At this point, the pathway bifurcates to the α- and β-carotene branches, which serve as substrates for four carotenoid hydroxylases that collectively convert them to lutein and zeaxanthin, respectively (DellaPenna and Pogson, 2006). The hydroxylated β-rings of zeaxanthin are subject to sequential epoxidation by *ZEAXANTHIN EPOXIDASE* (*ZEP*) to yield antheraxanthin and then violaxanthin. Violaxanthin is the substrate for the synthesis of neoxanthin or can be converted back to zeaxanthin by *VIOLAXANTHIN DEEPOXIDASE*. This reversible epoxidation/deepoxidation is termed the xanthophyll cycle, in which deepoxidation to zeaxanthin is favored in high-light conditions while epoxidation to violaxanthin predominates in moderate-light conditions. The rapid formation of zeaxanthin by the xanthophyll cycle is integral for the dissipation of excess energy by nonphotochemical quenching (Jahns and Holzwarth, 2012), a process critical for plant fitness and stress tolerance (Fray et al., 1995; Niyogi, 1999; Busch et al., 2002; Davison et al., 2002; Dall'Osto et al., 2007; Qin et al., 2007; Bailey and Grossman, 2008; Alboresi et al., 2011). Finally, violaxanthin is converted to 9-cis-violaxanthin and 9-cis-neoxanthin, which, through oxidative cleavage performed by 9-cis-epoxy-carotenoid dioxygenase (NCED) enzymes, gives rise to xanthoxin. Upon export to the cytosol, xanthoxin ultimately is converted to ABA, an important plant hormone involved in the regulation of seed development, dormancy, germination, vegetative growth, and environmental stress responses (McCarty, 1995; Koornneef et al., 2002; Frey et al., 2006).

Given the numerous structural and functional roles of carotenoids in photosystems, it is not surprising that carotenoid composition and content are evolutionarily conserved and tightly regulated in photosynthetic tissues (in order of abundance: lutein, β-carotene, violaxanthin, and neoxanthin). However, in tissues that are not primarily photosynthetic (e.g. flowers, fruits, and seeds), such constraints are relaxed and tremendous variation in carotenoid composition and content is observed. For example, ripe tomato (*Solanum lycopersicum*) fruit are red due to lycopene, which accumulates due to the lack of cyclization activities that would otherwise form α- and β-carotenes (Paran and van der Knaap, 2007). While the absence of cyclization activities does not impact tomato fruit development, mutations that similarly disrupt lycopene cyclization in photosynthetic tissues are invariably lethal (Young, 1993).

In contrast to such early pathway mutations, both photosynthetic and nonphotosynthetic tissues can tolerate the mutation of postcyclization steps, although often with very different consequences. For example, lutein is the most abundant xanthophyll in Arabidopsis (*Arabidopsis thaliana*) leaves and seeds, but mutation of *EPSILON-RING HYDROXYLASE1* (*LUT1*), which encodes the ε-ring hydroxylase responsible for the last step in lutein synthesis, has very different effects in these two tissues (Kim et al., 2009). In *lut1* leaves, the total carotenoid level is unchanged, and zeinoxanthin, the monohydroxy substrate of the enzyme, accumulates to approximately 50% of the lutein level in the wild type (Columbia-0 [Col-0]), with the xanthophyll cycle pigments, violaxanthin, zeaxanthin, and antheraxanthin, collectively accounting for the remaining 50%. In contrast to leaves, in *lut1* seed, the total carotenoid level is decreased approximately 75% due to an absence of lutein, and zeinoxanthin and xanthophyll cycle pigment levels are unchanged (Kim et al., 2009). Similarly, divergent impacts in seeds and leaves are observed for other mutations in carotenoid biosynthetic enzymes in Arabidopsis (Tian et al., 2003; Kim et al., 2009) and carotenoid cleavage enzymes in Arabidopsis and potato (*Solanum tuberosum*; Auldridge et al., 2006; Campbell et al., 2010; Gonzalez-Jorge et al., 2013).

While the mechanisms controlling the synthesis and accumulation of specific carotenoids in leaves are well studied (Rock and Zeevaart, 1991; Giuliano et al., 1993; Marin et al., 1996; Fraser et al., 1999; Carol and Kuntz, 2001; Fang et al., 2008; Toledo-Ortiz et al., 2010; Avendaño-Vázquez et al., 2014), those operating in seed, the plant organ consumed in the greatest abundance by humans, have not (Nisar et al., 2015). Quantitative trait locus (QTL) analyses have demonstrated that significant levels of natural variation for seed carotenoid traits exist in several plant species (Wong et al., 2004; Abbo et al., 2005; Fernandez et al., 2008; Blanco et al., 2011; Kandianis et al., 2013), but these intervals typically contain hundreds or thousands of genes; as a result, few causal loci underlying QTLs have been defined (Pozniak et al., 2007; Chandler et al., 2013; Gonzalez-Jorge et al., 2013). Significant natural variation for seed carotenoids also is present in diversity panels of maize (*Zea mays*), wheat (*Triticum aestivum*), *Citrus* spp., chickpea (*Cicer arietinum*), carrot (*Daucus carota*), pea (*Pisum sativum*), cassava (*Manihot esculenta*), and Arabidopsis (Abbo et al., 2010; Arango et al., 2010; Welsch et al., 2010; Yan et al., 2010; Blanco et al., 2011; Cook et al., 2012; Gonzalez-Jorge et al., 2013; Kandianis et al., 2013; Owens et al., 2014; Jourdan et al., 2015; Suwarno et al., 2015). Advances in high-throughput phenotyping and genotyping are allowing some of the genes underlying this natural variation to be identified through genome-wide association studies (GWAS; Harjes et al., 2008; Zhou et al., 2012; Gonzalez-

Jorge et al., 2013; Owens et al., 2014; Wen et al., 2014; Sonah et al., 2015). In this study, we integrate GWAS, QTL analysis, and reverse genetics in the model plant Arabidopsis to functionally define the loci contributing to natural variation in seed carotenoid levels. In particular, we show *ZEP* as the major-effect locus at the genome-wide level for natural variation in β-carotene-derived xanthophylls in Arabidopsis seed, and through a pathway-level analysis, we demonstrate the contribution of several other carotenoid pathway-associated genes to these seed traits.

RESULTS

GWAS and Pathway-Level Analysis Identify ZEP as a Major Locus Affecting Seed Carotenoid Content and Composition

To characterize natural variation for carotenoids in dry mature Arabidopsis seed, the six major seed carotenoids (in order of abundance: lutein, violaxanthin, zeaxanthin, antheraxanthin, neoxanthin, and β-carotene) were analyzed in three biological outgrowths of a 315-member Arabidopsis diversity panel. Individual carotenoids showed large variation across the population: 4-fold for lutein; 9- to 11-fold for neoxanthin, zeaxanthin, and β-carotene; 13-fold for violaxanthin; and 14-fold for antheraxanthin, with broad sense heritabilities from 0.61 to 0.78 (Supplemental Table S1). A total of 33 carotenoid traits (six individual traits and 27 sums and ratios) were analyzed through a GWAS using a diversity panel previously genotyped with a 250K single-nucleotide polymorphism (SNP) array (Baxter et al., 2010; Platt et al., 2010; Horton et al., 2012). For each trait, a unified mixed linear model (Yu et al., 2006) accounting for population structure (principal component [PC]) and familial relatedness (kinship [K]) was fitted to a total of 170,344 SNPs with a minor allele frequency greater than 0.05. Of the 33 traits assessed, only the ratio of violaxanthin to antheraxanthin (V/A) had associations that were significant using the Benjamini and Hochberg (1995) procedure to control a genome-wide false discovery rate (FDR) at 5% (Fig. 1A). These six significant V/A SNP associations all were present in an 18.8-kb region of chromosome 5, with SNP213665 (at 26,753,237 bp) being the most strongly associated SNP (i.e. the peak SNP), with $P = 4.64E\text{-}10$. Interestingly, this genomic region contains nine genes, including the carotenoid biosynthetic gene *ZEP* (At5g67030). Given that *ZEP* converts zeaxanthin to violaxanthin via the intermediate antheraxanthin (Supplemental Fig. S1), it was an obvious candidate gene for this trait.

To explore the potential contributions of other loci also known to play roles in carotenoid synthesis and turnover, a pathway-level analysis was performed (Segura et al., 2012; Lipka et al., 2013; Owens et al., 2014) with 42 genes known from prior studies to be associated with carotenoid synthesis or degradation (DellaPenna and Pogson, 2006; Ruiz-Sola and Rodríguez-Concepción, 2012; Nisar et al., 2015). These included 16 genes for the synthesis of isopentenyl diphosphate from the MEP pathway, 15 carotenoid biosynthetic pathway genes (including *ZEP*), and 11 genes involved in carotenoid degradation and ABA synthesis (Supplemental Table S2). For this analysis, a total of 10,922 SNP markers located within ±100 kb of these 42 a priori candidates were utilized. For each trait, FDR adjustment was conducted using only this subset of SNPs from respective GWAS analysis results. A total of 29 unique SNPs in the vicinity of five candidate genes were significant in the pathway-level analysis at the 5% FDR level, with some of these SNPs associating with multiple traits (Table I; Supplemental Table S3). Three of these genes were in the carotenoid biosynthetic pathway: *ZEP* (At5g67030), identified for a total of six traits, including V/A (for information on additional traits, see Table I; Supplemental Table S3); *BETA-CAROTENE HYDROXYLASE2* (*βOHase2*; At5g52570), identified for three traits; and *LYCOPENE CYCLASE PARALOG* (*CruP*; At2g32640), identified for a single trait. Two additional loci were identified from the MEP pathway for the V/A trait: *1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE3* (*DXS3*; At5g11380) and *GERANYLGERANYL PYROPHOSPHATE SYNTHASE6* (*GGPS6*; At4g36810; Table I; Supplemental Table S3). When the FDR cutoff was relaxed to 10%, weaker associations also were observed for the carotenoid pathway enzymes *PHYTOENE SYNTHASE* (*PSY*; At5g17230) and *ZETA-CAROTENE DESATURASE* (*ZDS*; At3g04870), the carotenoid degradation enzyme *CAROTENOID CLEAVAGE DIOXYGENASE4* (*CCD4*; At4g19710), and the MEP pathway enzyme *1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE* (*DXR*; At5g62790; Supplemental Table S4).

Molecular Basis for Natural Variation in Arabidopsis Seed V/A Levels

Of the nine loci identified by GWAS and pathway-level analysis, associations with the V/A trait were strongest with six significant SNPs located in a chromosome 5 interval containing *ZEP*, with the most significant, SNP213665 ($P = 4.64E\text{-}10$), explaining 19% of the genetic variance for the V/A trait in the association panel (Fig. 1; Table I; Supplemental Table S3). To further assess the contribution of SNP213665 and the five other significant SNPs present in this region (Fig. 1, A and B) to V/A levels, a multilocus mixed model (MLMM) analysis (Segura et al., 2012) was conducted at the genome-wide, chromosome-wide, and local (±100-kb region centered on peak SNP213665) levels (Table II). At all levels, the MLMM model included only the peak SNP identified from GWAS (i.e. SNP213665). The contribution of SNP213665 was explored further by rerunning a GWAS with SNP213665 included as a covariate, which resulted in all other previously statistically significant associations being lost (Fig. 2). Thus, the

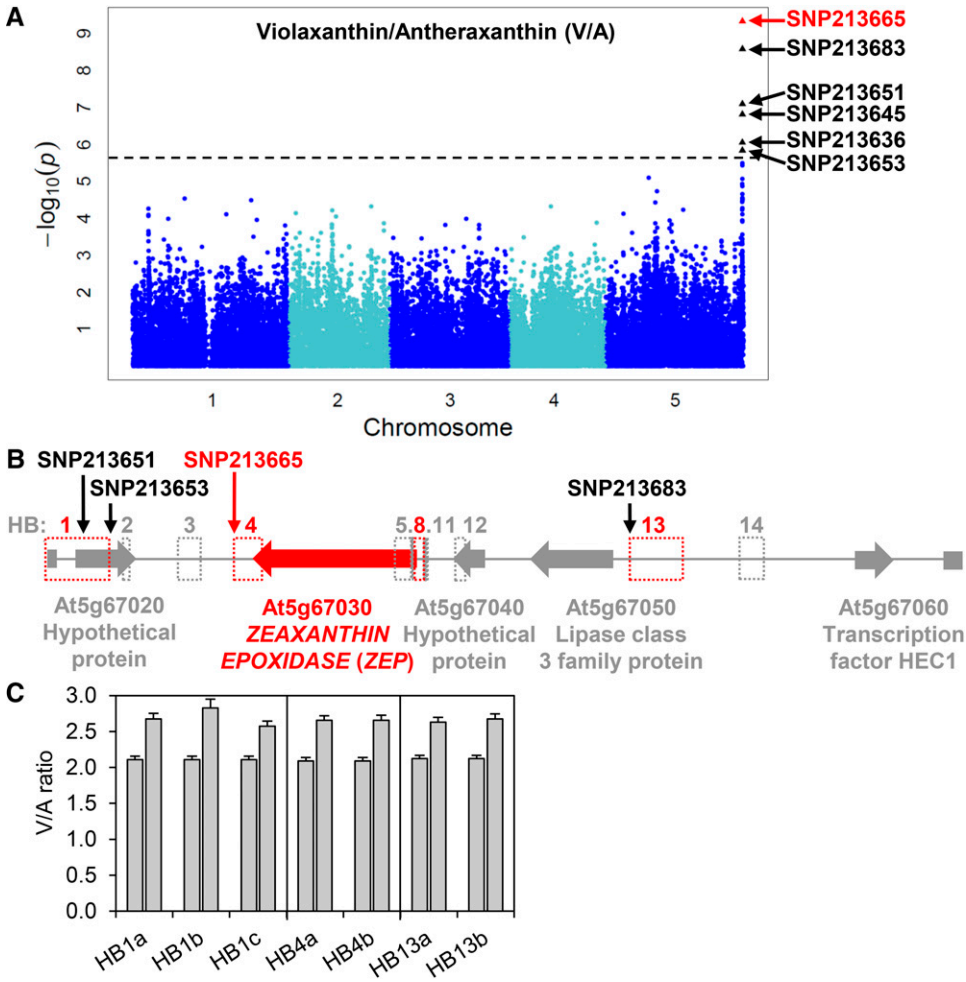

Figure 1. Genome-wide association identifies *ZEP* as a major contributor to seed carotenoid levels. A, Manhattan plot of genome-wide association for the V/A seed carotenoid trait using a unified mixed linear model. The *y* axis shows the negative log of *P* values for SNPs spanning the five Arabidopsis chromosomes, shown in alternating colors on the *x* axis. Black triangles denote SNPs that are statistically significant at a 5% FDR (indicated by the horizontal dashed line). The red triangle indicates the peak SNP, SNP213665, located on chromosome 5 at 26,753,237 bp. B, Haploblock analysis of V/A in the chromosome 5 region encompassing the most statistically significant SNPs identified in the GWAS. Linkage disequilibrium (LD) analysis centered on the peak SNP, SNP213665, defined as the 19.1-kb LD interval shown, which encompasses five genes. A total of 14 haploblocks (indicated by dashed boxes) were identified, of which only four were statistically significant (indicated by red dashed boxes): HB1, HB4, HB8, and HB13, with *P* values of 1.15E-07, 8.8.E-09, 3.56E-07, and 3.93E-07, respectively. C, Accessions with contrasting haplotypes for the V/A trait were identified for HB1, HB4, and HB13. HB1 contains four sets of contrasting haplotypes, of which the CCAAT haplotype (HB1a; 58 accessions), the TCAAT haplotype (HB1b; 18 accessions), and the TCTCC haplotype (HB1c; 97 accessions) have 27%, 34%, and 22% higher V/A values, respectively, in comparison with the 80 accessions with the TATCC haplotype. For HB4, the AC (HB4a) and AT (HB4b) haplotypes are present in 82 and 95 accessions, respectively, and have V/A values 27% higher than the 92 accessions containing the TT haplotype. For HB13, the TA (HB13a) and TG (HB13b) haplotypes are present in 68 and 104 accessions, respectively, and have V/A values 24% and 26% higher than the AG haplotype (96 accessions). All haplotypes of HB1, HB4, and HB13 were significantly different at $\alpha = 0.05$ (for individual values, see Supplemental Table S6).

combination of MLMM (Table II) and conditional analysis (Fig. 2) suggests that SNP213665 tags the major causal polymorphism(s) for the V/A trait on chromosome 5.

To delimit the physical region of V/A association, LD analysis was performed between SNPs located within a 40-kb region centered on SNP213665. LD decay was defined as reaching nominal levels of $r^2 \leq 0.1$ and determined to be 19.1 kb, extending 3.9 kb downstream and 15.2 kb upstream of the peak SNP (Supplemental Fig. S2). This 19.1-kb LD region encompasses five genes: *ZEP*, two hypothetical proteins (At5g67020 and At5g67040), a lipase class 3 family protein (At5g67050), and the transcription factor *HECTATE1* (*HEC1*; At5g67060; Fig. 1B).

To obtain an overview of additional polymorphisms across this interval, the Arabidopsis 1001 Genome

Table I. *Pathway-level analysis identified significant marker-trait associations for 29 unique SNPs in the vicinity of five loci (including ZEP) associated with specific seed carotenoid traits at a significant 5% FDR threshold*

For the pathway-level analysis, 10,922 SNPs located within ±100 kb of the 42 candidate genes listed in Supplemental Table S2 were used. Raw GWAS *P* values were utilized and corrected to control for a 5% FDR using the Benjamini and Hochberg (1995) procedure. Asterisks denote genome-wide significant SNPs at the 5% FDR threshold, with the peak SNP, SNP213665, being shown in boldface (Fig. 1). Carotenoid traits are as follows: B/T-car = β-carotene/total carotenoids; B/L = β-carotene/lutein; B/(VNAZ) = β-carotene/(violaxanthin + neoxanthin + antheraxanthin + zeaxanthin); (ZVN)/A = (zeaxanthin + violaxanthin + neoxanthin)/antheraxanthin; (ZV)/A = (zeaxanthin + violaxanthin)/antheraxanthin; (BZVN)/A = (β-carotene + zeaxanthin + violaxanthin + neoxanthin)/antheraxanthin; (BZV)/(NA) = (β-carotene + zeaxanthin + violaxanthin)/(neoxanthin + antheraxanthin); and A/T-car = antheraxanthin/total carotenoids.

Gene	Association SNP	SNP Chromosome	SNP Position (bp)	Trait at 5% FDR
GGPS6	SNP30058	1	18,276,036	V/A
DXS3	SNP167260	5	3,572,584	V/A
βOHase2	SNP204021	5	21,356,578	B/(VNAZ)
	SNP204094	5	21,381,979	B/T-car; B/L; B/(VNAZ)
CruP	SNP72081	2	13,849,499	B/L
ZEP	SNP213607	5	26,727,756	V/A
	SNP213610	5	26,729,590	V/A
	SNP213611	5	26,729,945	V/A
	SNP213613	5	26,730,417	(ZVN)/A
	SNP21618	5	26,734,121	V/A; (ZVN)/A; (ZV)/A; (BZVN)/A; A/T-car
	SNP21619	5	26,734,331	(ZVN)/A; (ZV)/A; (BZVN)/A; A/T-car
	SNP213621	5	26,734,903	V/A
	SNP213623	5	26,737,771	V/A; (ZVN)/A; (ZV)/A; (BZVN)/A
	SNP213628	5	26,740,902	V/A
	SNP213636	5	26,742,617	V/A; (ZVN)/A; (ZV)/A; (BZVN)/A
	SNP213645	5	26,748,060	V/A; (ZVN)/A; (ZV)/A; (BZVN)/A
	SNP213651*	5	26,750,150	V/A; (ZVN)/A; (ZV)/A; (BZVN)/A; (BZV)/(NA)
	SNP213652	5	26,750,511	V/A
	SNP213653*	5	26,750,630	V/A
	SNP213654	5	26,750,649	V/A
	SNP213663	5	26,752,544	V/A
	SNP213665*	**5**	**26,753,237**	**V/A; (ZVN)/A; (ZV)/A; (BZVN)/A; (BZV)/(NA); A/T-car**
	SNP213666	5	26,753,820	V/A
	SNP213674	5	26,757,031	V/A
	SNP213677	5	26,758,064	V/A
	SNP213683	5	26,761,492	V/A; (ZVN)/A; (ZV)/A; (BZVN)/A; (BZV)/(NA); A/T-car
	SNP213689	5	26,766,420	V/A
	SNP213697	5	26,769,272	V/A
	SNP213713	5	26,776,446	V/A

Project database (http://signal.salk.edu/atg1001/3.0/gebrowser.php; Weigel and Mott, 2009) was assessed for 180 accessions of the 315-member diversity panel for which genome sequence was available at the time of analysis. This identified 29 new polymorphisms in the 19.1-kb LD interval in addition to the 42 scored on the 250K SNP array. Five of the new polymorphisms, including a 3-bp deletion, resulted in amino acid changes in the *ZEP* coding sequence: His-8 to Tyr, Ala-20 to Thr, Ser-27 to Pro, Val-28 to Gly, and Ile-47 to Leu (Supplemental Table S5). Three other potentially important polymorphisms were 3- and 6-bp deletions in the ZEP 5′ untranslated region and a large insertion 1,033 bp upstream of the *ZEP* transcription start site with three states (871-bp, 1-kb, and 1.6-kb insertions). To increase mapping resolution and power for statistical testing, an additional 89 accessions in the panel were scored for these new polymorphisms by sequencing or PCR. This resulted in 269 fully genotyped accessions for subsequent haplotype analysis.

The potential contributions of the 29 new polymorphisms to the V/A trait were determined by performing significance testing using a PC+K model and SNP213665 as the significance testing control (Supplemental Table S5). Of the 29 new polymorphisms, only three SNPs located in the 5′ upstream region of *ZEP* were significantly associated with seed V/A ($P = 6.7E-06$, $P = 1.7E-05$, and $P = 1.4E-05$; Supplemental Table S5). All four insertions/deletions (InDels; including the large multistate insertion 1,033 bp upstream) and the five nonsynonymous SNPs in the *ZEP* coding region lacked significant association with the trait. Moreover, while the addition of these new polymorphisms increased mapping resolution, none was more significantly associated than the peak GWAS SNP, SNP213665, which had $P = 1.09E-09$ in the 269 accessions analyzed, compared with $P = 4.64E-10$ in the full 315-member diversity panel (Supplemental Table S5).

To better delineate the molecular basis for V/A natural variation, a haploblock analysis of the 19.1-kb LD region was performed (Supplemental Table S5). A total

Table II. *MLMM analysis of the Arabidopsis seed V/A trait at genome-wide, chromosome-wide, and ±100-kb region levels encompassing the peak GWAS SNP, SNP213665*

For all levels, the optimal model has the lowest extended Bayesian information criterion (BIC) value (in boldface), which includes the peak SNP, SNP213665. NA, Not applicable.

MLMM Analysis	SNP Added to Model	Chromosome	Position	*P*	Extended BIC	Optimal Model
Genome	None	NA	NA	NA	−1,954.82	No
	SNP213665	5	26,753,237	7.12E-13	**−1,286.96**	**Yes**
Chromosome	None	NA	NA	NA	−1,267.62	No
	SNP213665	5	26,753,237	7.12E-13	**−1,289.75**	**Yes**
±100 kb	None	NA	NA	NA	−1,267.62	No
	SNP213665	5	26,753,237	7.12E-13	**−1,299.77**	**Yes**

of 14 haploblocks were identified, of which four were statistically significant (HB1, HB4, HB8, and HB13; Fig. 1B), with all except HB8 having statistically significant contrasting haplotypes (Fig. 1C; Supplemental Tables S4 and S5). It is noteworthy that the most significant haploblock, HB4 ($P = 8.85E-09$), is delimited by the peak association SNP, SNP213665, which itself is located at the 3′ intergenic region of *ZEP* (Fig. 1, A and B). These results are consistent with *ZEP* playing a major role in determining the V/A in Arabidopsis seed.

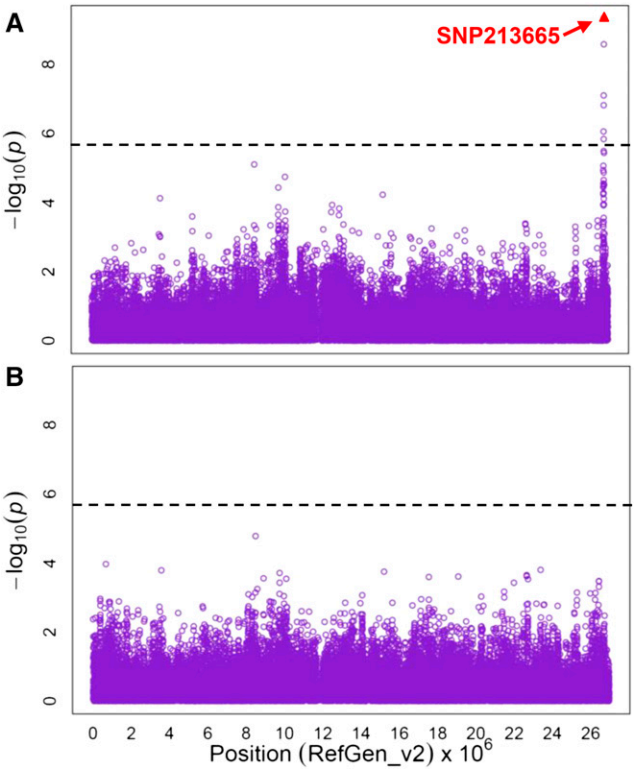

Figure 2. Manhattan plots of covariate analysis for seed V/A GWAS associations on chromosome 5. The *y* axis shows the negative log of *P* values for SNPs (purple circles), and the dashed black lines show the 5% FDR. The red triangle represents the peak SNP, SNP213665, at 26,753,237 bp on chromosome 5. A, Manhattan plot of seed V/A associations using a unified mixed linear model. B, Manhattan plot of seed V/A associations on chromosome 5 using a unified mixed linear model with SNP213665 included as a covariate. When SNP213665 was included as a covariate, none of the signals on chromosome 5 were statistically significant at 5% FDR.

QTL Mapping Supports ZEP Being the Major Locus Affecting the V/A Trait

To independently assess the role of *ZEP* in natural variation of the V/A trait in Arabidopsis seed, a QTL analysis was performed with Col-0 and Bla-1, two parental accessions with contrasting haplotypes for the most significant haploblock and haplotype, HB4a (Col-0 is superior and Bla-1 is inferior for the HB4a haplotype; Fig. 1C; Supplemental Table S6). A 142-line Col-0 × Bla-1 Arabidopsis recombinant inbred line population was scored for seed carotenoids, and QTL analysis using composite interval mapping for V/A was performed. Three V/A QTLs were identified on chromosomes 1, 4, and 5 (QTL1, QTL4, and QTL5, respectively), with log of the odds (LOD) scores of 6.72, 5.09, and 19.13, accounting for 19.7%, 15.3%, and 46.5% of variation, respectively (Fig. 3). QTL1 spans 1.53 Mb that contains 474 genes, none of which are known to impact carotenoids. QTL4 contains 588 genes in a 2.09-Mb region and also lacks known carotenoid-related genes (Fig. 3; Supplemental Table S2). QTL5 has both the highest LOD score and an explanation for phenotypic variation. QTL5 encompasses a 1.97-Mb interval that contains 619 genes, two of which are known to be associated with carotenoids: At5g62790 (*DXR*) and At5g67030, the *ZEP* gene (Fig. 3; Supplemental Table S2). While *DXR* and *ZEP* are both in the QTL confidence interval, they are separated by 1.53 Mb, with the markers flanking *DXR* having a maximum LOD score of 9, while those flanking *ZEP* are at the peak of the QTL, with a LOD score of 19. Together, these data are consistent with *ZEP* being the most significant contributor to the V/A trait in this population and the causal gene underlying QTL5 (Fig. 3), providing independent verification of the role of *ZEP* in natural variation for V/A levels in seeds.

ZEP Is Differentially Expressed in a Haplotype-Specific Fashion during Late Seed Development

Because the five nonsynonymous polymorphisms that changed amino acids in the ZEP coding frame

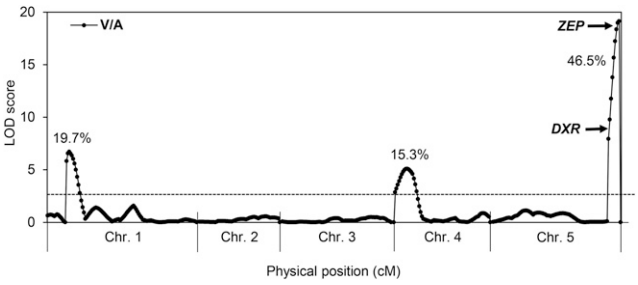

Figure 3. QTL analysis of the V/A trait in a Col-0 × Bla-1 recombinant inbred population. V/A QTL analysis was based on seed carotenoid data from 142 Col-0 × Bla-1 recombinant inbred lines. The dashed line indicates an $\alpha = 0.05$ significance threshold for QTL detection derived from running a permutation procedure (LOD > 2.5). The percentages indicate phenotypic variance explained by each QTL. *DXR*, *DXR* locus; *ZEP*, *ZEP* locus.

lacked significant associations for the V/A trait, we assessed whether gene expression differences might contribute to the large allelic effects of *ZEP* on seed carotenoid content. *ZEP* transcript levels in developing seed were analyzed by in silico and in vivo approaches. In silico gene expression in the Col-0 background showed that, relative to the housekeeping gene *PROTEIN PHOSPHATASE 2A SUBUNIT A3* (*PP2AA3*; Zuber et al., 2010), *ZEP* transcript levels increased strongly during the later stages of seed development (Fig. 4A; values obtained from AtGenExpress [http://jsp.weigelworld.org/expviz/expviz.jsp]).

To directly assess whether *ZEP* expression in members of the association panel had a pattern similar to Col-0, and to discern whether the timing or levels of expression were correlated with superior (exhibiting high phenotypic seed V/A content) and inferior (exhibiting low phenotypic seed V/A content) haplotypes at the *ZEP* locus, ZEP mRNA levels were quantified at three stages of seed development in accessions with contrasting haplotypes for the most significant V/A trait haplotype, HB4a (Fig. 4B). The 18-, 21-, and 42-DAP developmental stages tested correspond to mature seeds entering the desiccation period (ATGE_9 stage), desiccating seeds (ATGE_10 stage), and dry mature seeds (which is not represented in the ATGE database), respectively. Seed RNA for three accessions (Kl-5, Mt-0, and Nw-2) containing the HB4a AC haplotype, which confers high phenotypic V/A levels and is termed the superior haplotype, and four accessions (Gel-1, Mh-0, PHW-13, and Pr-0) containing the HB4a TT haplotype, which confers low phenotypic levels of V/A and is termed the inferior haplotype, were analyzed. At 18 DAP, *ZEP* transcript levels were relatively low and not significantly different between the two haplotype groups (Fig. 4B). However, at 21 and 42 DAP, *ZEP* expression in the superior haplotype lines was significantly higher than in the inferior haplotype lines, resulting in increased conversion of antheraxanthin to violaxanthin in the superior haplotype lines and, hence, a higher V/A (Figs. 1C and 4B).

ZEP activity is required for ABA synthesis due to its role in generating violaxanthin and neoxanthin, which are substrates for various NCEDs in ABA synthesis (Koornneef et al., 2002; Nambara and Marion-Poll, 2003, 2005; Lefebvre et al., 2006; Frey et al., 2012). To address whether the differences in *ZEP* gene expression between V/A inferior and superior haplotypes impacted violaxanthin and neoxanthin levels, and hence possibly indirectly ABA levels, violaxanthin and neoxanthin levels were assessed in Col-0 and the inferior and superior haplotype lines (Supplemental Table S7). The violaxanthin and neoxanthin contents of the inferior and superior haplotype lines were not significantly different between each other or Col-0, indicating that there is no shortage of ABA precursors and, hence, that ABA differences are unlikely to contribute indirectly to the carotenoid natural variation observed in mature dry seed.

Null mutant alleles for *ZEP* have been isolated previously and studied for the role of the enzyme in

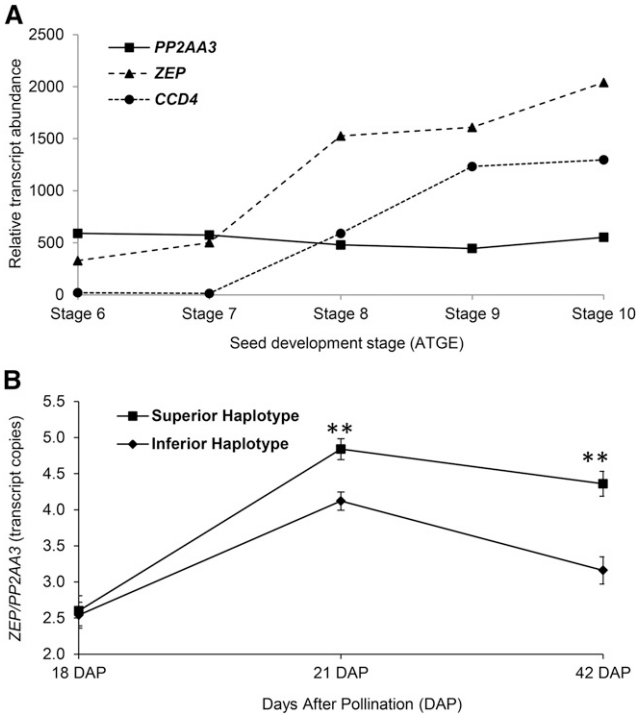

Figure 4. *ZEP* gene expression during seed development. A, Gene expression values for developing seed of the Col-0 accession were obtained from the AtGenExpress database (http://jsp.weigelworld.org/expviz/expviz.jsp). Seed development stages are as follows: 6, mid to late torpedo embryos; 7, late torpedo to early walking stick embryos; 8, walking stick to early curled cotyledon embryos; 9, curled cotyledon to early green cotyledon embryos; 10, green cotyledon embryos. *PP2AA3* was the control gene. B, Comparison of *ZEP* mRNA levels for four inferior (low phenotypic levels of the V/A trait) haplotype lines and three superior (high phenotypic levels of the V/A trait) haplotype lines in developing seeds at three stages of development: 18 d after pollination (DAP), 21 DAP, and 42 DAP (fully dry seed). Average values are shown with SE for $n = 3$. Asterisks indicate statistically significant differences (**, $\alpha = 0.05$) between inferior and superior haplotypes.

carotenogenesis and ABA synthesis in leaves or tubers (Galpaz et al., 2008; Park et al., 2008; Vallabhaneni and Wurtzel, 2009; Zhou et al., 2011; Lätari et al., 2015; Schwarz et al., 2015); however, their impact on carotenoid synthesis and accumulation in seed has not been assessed extensively (Frey et al., 2006; Gonzalez-Jorge et al., 2013; Owens et al., 2014; Schwarz et al., 2015; Suwarno et al., 2015). To determine the maximal possible effect of *ZEP* on carotenoid levels during seed development and maturation, a null transfer DNA (T-DNA) insertional mutant of *ZEP* was identified and characterized (Supplemental Fig. S3). Given that the *ZEP* locus has been known historically as *aba1* due to its initial identification based on the impact of severe alleles on ABA synthesis, we refer to this new allele as *aba1-7*. As expected in dry seed, loss of *ZEP* activity in *aba1-7* disrupts the epoxidation of zeaxanthin, increasing its levels 43-fold relative to the wild type, and reducing antheraxanthin, violaxanthin, and neoxanthin levels below detection. β-Carotene and lutein also were increased by 3.2- and 2.2-fold in *aba1-7* dry seed relative to Col-0, while total carotenoids were elevated 5.7-fold, predominantly due to the massive increase in the zeaxanthin content of seed (Fig. 5). Indeed, zeaxanthin accounts for 63% of total carotenoids in the *aba1-7* mutant, making it, rather than lutein, the most abundant seed carotenoid by a factor of 2. These observations were corroborated with a second allele, *aba1-6* (Niyogi et al., 1998), which also showed similarly increased carotenoid levels in dry seed (Supplemental Table S8).

To assess the impact of *ZEP* activity on the levels and turnover of carotenoids during seed development and desiccation, carotenoids were quantified in Col-0 and *aba1-7* lines at three different developmental time points: 15 DAP, fully mature green seed; 21 DAP, desiccating seed; and 42 DAP, fully mature dry seed (Fig. 6; Supplemental Fig. S4). From 15 to 42 DAP, both genotypes showed decreasing levels of all carotenoids, although the degree, timing, and severity of this decrease was strongly impacted by the *aba1-7* mutation (Fig. 6; Supplemental Fig. S4). At all stages of development, Col-0 contained lower levels of lutein, zeaxanthin, β-carotene, and total carotenoids than *aba1-7* (Fig. 6). It is also noteworthy that, relative to Col-0, the β-carotene level in *aba1-7* seed is 2.1-fold higher at 15 DAP but levels decrease precipitously and similarly in both lines, such that although β-carotene levels in *aba1-7* are 4.8-fold of that in Col-0, the final β-carotene level in *aba1-7* is only 11.2% of that at 15 DAP.

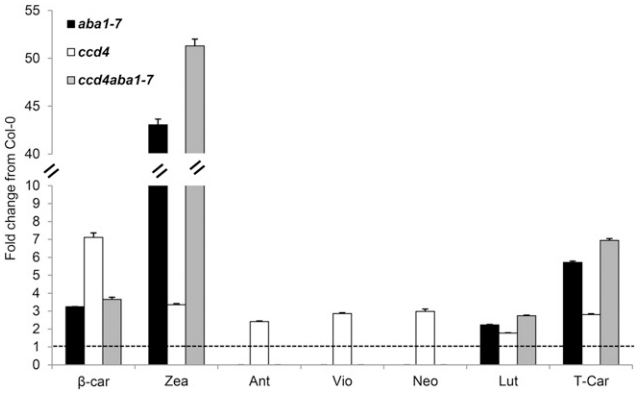

Figure 5. Fold increase of seed carotenoids in homozygous *aba1-7*, *ccd4*, and *ccd4aba1-7* mutants relative to wild-type Col-0. The fold change for carotenoids in each mutant is shown in relation to that in Col-0 seed. The dotted line indicates compound levels in Col-0. Average Col-0 levels in ng mg^{-1} are as follows: Neo (neoxanthin) = 0.77; Vio (violaxanthin) = 2.94; Ant (antheraxanthin) = 2.38; Lut (lutein) = 31.47; Zea (zeaxanthin) = 3.72; β-car (β-carotene) = 1.86; and T-Car (total carotenoids) = 43.61. For all genotypes, all carotenoid compounds were significantly different from Col-0 at $\alpha = 0.05$.

The higher levels of carotenoids other than zeaxanthin in *aba1-7* seed (Fig. 6), especially lutein and β-carotene, whose synthesis are *ZEP* independent (Supplemental Fig. S1), suggest that activities other than *ZEP* also might be pleiotropically affected in *aba1-7* and contribute to these aspects of the null mutant phenotype. That *aba1-7* exhibits the classic ABA-deficient wilting phenotype of other strong *aba1* alleles (Koornneef et al., 1982; Niyogi et al., 1998) suggests that the ABA deficiency in *aba1-7* might impact other pathway steps. We assessed this possibility by comparing in silico seed-specific gene expression for the 50 genes from the MEP pathway, carotenoid synthesis and degradation pathways, and ABA synthesis and signaling pathways in dry seed of Col-0 and the ABA-insensitive mutant *abi5-7*, as it has strong effects on ABA-dependent processes in developing seed (Supplemental Table S9; values from AtGenExpress [http://jsp.weigelworld.org/expviz/expviz.jsp]; Finkelstein and Lynch, 2000; Bensmihen et al., 2002; Brocard et al., 2002; Nakabayashi et al., 2005). We found previously that dry seed mRNA levels are an excellent proxy for mRNA levels during seed development (Angelovici et al., 2013; Gonzalez-Jorge et al., 2013). Consistent with this, the expression correlations (r^2) of the 50 target genes between dry Col-0 seed and seed developmental stages 8 and 10 (Supplemental Table S10) were 0.84 and 0.98, respectively, while that between dry Col-0 seed and dry *abi5-7* seed was 0.75. The majority of the 50 target genes were present at similar levels in dry seed of Col-0 and *abi5-7* with a few notable exceptions. Relative to Col-0, in the *abi5-7* mutant, *DXS2* mRNA (a MEP pathway gene) was increased 473% from a very low level, while *CCD4* mRNA was decreased 86% from a very high level; both are changes that could increase total carotenoid levels (Supplemental Table S9; Gonzalez-Jorge et al., 2013). *βOHase1* and *βOHase2* mRNA also were decreased 35% in *abi5-7* and likewise may contribute to the increased β-carotene levels (Davison et al., 2002; Fiore et al., 2006). Based on these data, it seems likely that, in addition to the large increase in zeaxanthin, resulting from the complete loss of *ZEP* activity in the *aba1-7* null allele (Fig. 5), the resulting ABA deficiency also contributes indirectly to the observed increased levels of other, non-*ZEP*-dependent carotenoids.

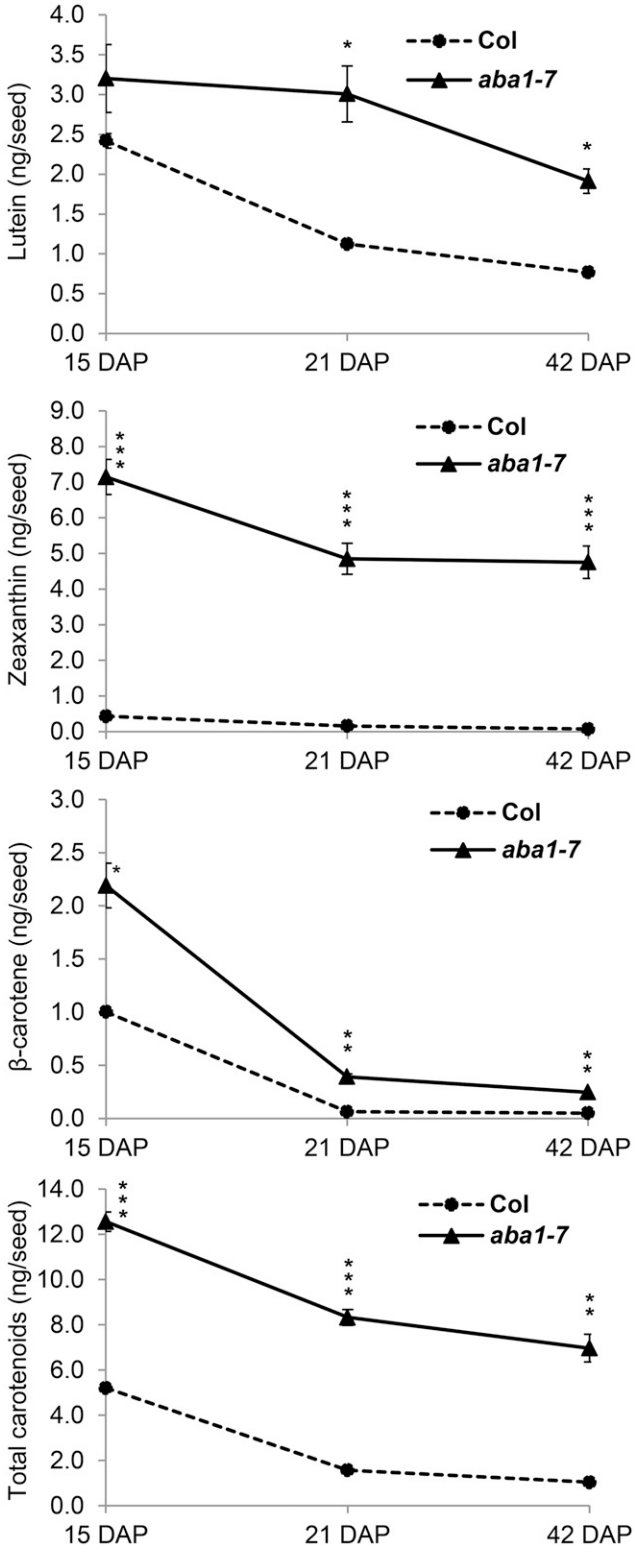

Figure 6. Carotenoid levels during seed development and desiccation in wild-type Col-0 and *aba1-7* mutant lines. Flowers were tagged at anthesis, seeds were harvested at specific DAP, and carotenoids were quantified by HPLC. Average carotenoid levels (ng seed^{-1}) are shown with SE for three biological replicates of 150 seeds for each time point and genotype. Asterisks indicate statistically significant differences at $\alpha = 0.05$ (*, $P < 0.01$; **, $P < 0.001$; and ***, $P < 0.0001$) relative to Col-0. Error bars smaller than the symbols are not shown.

Functional Analysis of Major Regulators of Carotenoid Accumulation in Seed

Our carotenoid pathway-level analysis identified eight loci in addition to *ZEP* that potentially contribute to carotenoid homeostasis in Arabidopsis seed: *DXS3*, *GGPS6*, *βOHase2*, *CruP*, *DXR*, *PSY*, *ZDS*, and *CCD4* (Table I; Supplemental Table S3). Null mutants of Arabidopsis *DXR*, *DXS3*, *ZDS*, and *PSY* are lethal (Estévez et al., 2001; Lindgren et al., 2003; Carretero-Paulet et al., 2006; Dong et al., 2007); *βOHase2* null mutants cause a modest reduction in seed xanthophylls (Tian et al., 2003; Kim et al., 2009); while *CCD4* null mutants have an extremely strong positive impact on all seed carotenoids (Gonzalez-Jorge et al., 2013). A T-DNA knockout of the remaining candidate *CruP*, a lycopene cyclase paralog, resulted in an increase in β-carotene epoxide levels in Arabidopsis leaves but did not significantly affect the levels of other leaf carotenoids (Bradbury et al., 2012). To determine if *CruP* activity affects carotenoid levels in Arabidopsis seed, a *cruP* null mutant was isolated and its seed carotenoids analyzed. *cruP* mutant seed had neoxanthin, antheraxanthin, and zeaxanthin levels indistinguishable from the wild type, while total carotenoids, lutein, violaxanthin, and β-carotene were elevated significantly by 12%, 14%, 18%, and 20%, respectively (Fig. 7), confirming the predicted involvement of this locus in influencing seed carotenoid levels (Table I; Supplemental Tables S3 and S4).

As highlighted previously, of the nine loci identified as contributors to natural variation in carotenoid levels in Arabidopsis seed (Table I; Supplemental Tables S3 and S4), null mutants of *ZEP* and *CCD4* have by far the greatest positive impact on traits. Therefore, we tested the hypothesis that combining null alleles for these two major controllers of seed carotenoid composition might have additive or synergistic impacts on seed carotenoid composition or content by generating and analyzing a *ccd4aba1-7* double mutant line (Fig. 5). Our results for *ccd4* are consistent with a prior report (Gonzalez-Jorge et al., 2013), where, relative to Col-0, total seed carotenoids were elevated by 2.8-fold, β-carotene by 7.1-fold, and lutein, antheraxanthin, violaxanthin, zeaxanthin, and neoxanthin by 1.7-, 2.4-, 2.8-, 3.3-, and 2.9-fold, respectively (Fig. 5). Introduction of *aba1-7* into the *ccd4* background had differing effects on individual seed carotenoids. The impact was synergistic for zeaxanthin, additive for total carotenoids and lutein, but only partially additive for β-carotene (Fig. 5).

DISCUSSION

Carotenoids serve many critical functions in seeds, including in ABA synthesis, light harvesting, and limiting damage by free radicals and lipid peroxidation, all

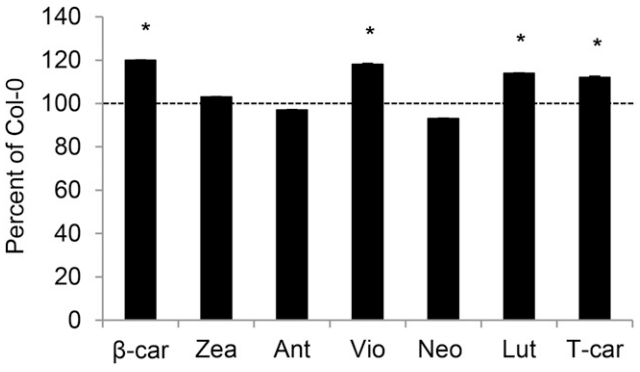

Figure 7. Dry seed carotenoid levels in homozygous *crup* mutants relative to wild-type Col-0. Carotenoid compounds are as follows: β-car, β-carotene; Zea, zeaxanthin; Ant, antheraxanthin; Vio, violaxanthin; Neo, neoxanthin; Lut, lutein; and T-Car, total carotenoids. Average values are shown with SE for $n = 6$. Error bars smaller than the symbols are not shown. Asterisks denote carotenoid compounds that were statistically significantly different at $\alpha = 0.05$. The dotted line indicates compound levels in Col-0.

of which ultimately impact seed aging, viability, and germination (Howitt and Pogson, 2006; Smolikova et al., 2011). Despite their importance to both seed biology and human health, our understanding of the regulatory mechanisms controlling carotenoid biosynthesis, sequestration, and degradation in seed remains limited compared with that in leaves and fruit (DellaPenna and Pogson, 2006; Cazzonelli and Pogson, 2010; Nisar et al., 2015). Our assessment of natural variation for seed carotenoids across a 315-member Arabidopsis diversity panel showed substantial variation for all traits, with violaxanthin and antheraxanthin having particularly high phenotypic variance and broad sense heritabilities (13-fold and 0.77 and 14-fold and 0.61, respectively; Supplemental Table S1). The strongest genome-wide significant association was for the V/A trait and identified *ZEP*, the enzyme responsible for the successive conversion of zeaxanthin to antheraxanthin and then violaxanthin, as the causal locus (Fig. 1; Supplemental Fig. S1). MLMM analysis (Table II), conditional GWAS (Fig. 2), and significance testing of polymorphisms (Supplemental Table S5) pinpointed SNP213665 (the peak SNP) as the SNP tagging natural variation for the V/A trait. Analysis of a QTL population with contrasting *ZEP* haplotypes independently confirmed the large impact of this locus on the V/A trait (Fig. 3).

Because nonsynonymous polymorphisms altering the ZEP protein sequence were not associated with the V/A trait, we assessed whether differences in gene expression contributed to its natural variation. *ZEP* expression is developmentally regulated during Arabidopsis seed development, with mRNA levels becoming elevated late in seed maturation, peaking at the start of desiccation (21 DAP), and then decreasing by the mature dry seed stage (42 DAP; Fig. 4). A comparison of *ZEP* RNA levels during this time frame in accessions with superior (exhibiting high phenotypic seed V/A content) and inferior (exhibiting low phenotypic V/A level) *ZEP* haplotypes (Fig. 4B) showed that the magnitude of *ZEP* expression is clearly correlated with the seed V/A trait. That levels of the carotenoid precursors of ABA, violaxanthin and neoxanthin (Supplemental Table S7), are not statistically different in inferior and superior *ZEP* haplotypes argues against a role for ABA in V/A natural variation in Arabidopsis seed, which is consistent with studies concluding that the regulatory control of ABA accumulation occurs downstream of epoxy-carotenoid synthesis (Frey et al., 2006). Our expression data concur with previous studies showing a positive correlation between *ZEP* expression and zeaxanthin content in fruit and leaves of transgenic and mutant lines of various plant species (Audran et al., 1998, 2001; Galpaz et al., 2008; Vallabhaneni and Wurtzel, 2009; Zhou et al., 2011; Couso et al., 2012). Indeed, altered gene expression appears to be an emerging mechanistic theme underlying natural variation in seed carotenoid traits, as, in addition to *ZEP*, expression changes also were responsible for β-carotene and β-xanthophyll associations with Arabidopsis *CCD4* and for β-carotene association with maize *LYCOPENE EPSILON-CYCLASE* and one of six B*ETA-CAROTENE HYDROXYLASES*, *CrtRB1* (Harjes et al., 2008; Yan et al., 2010; Gonzalez-Jorge et al., 2013).

To assess the potential contributions of other loci known from prior studies to be carotenoid associated, we performed a pathway-level analysis with 42 a priori candidate genes involved in carotenoid biosynthesis, degradation, and the plastidic MEP pathway. In addition to *ZEP*, eight other loci were identified to contribute to natural variation of seed carotenoids, especially β-xanthophyll traits: *DXS3*, *DXR*, *GGPS6*, *βOHase2*, *PSY*, *ZDS*, *CCD4*, and *CruP* (Table I; Supplemental Tables S3 and S4). With the exception of *CCD4*, these genes had not been associated previously with the natural variation of seed carotenoids in Arabidopsis, although most have been reported previously from mutant and transgenic studies to impact leaf and, in some cases, also seed carotenoid content. For example, overexpression of *DXR* and *DXS3* in Arabidopsis and *PSY* in Arabidopsis and *Brassica* spp. is positively correlated with increased levels of isoprenoids such as carotenoids and ABA in leaves and seed, respectively (Shewmaker et al., 1999; Estévez et al., 2001; Lindgren et al., 2003; Botella-Pavía et al., 2004; Carretero-Paulet et al., 2006), while null mutants of Arabidopsis *CCD4* and *βOHase2* impact seed carotenoid content positively and negatively, respectively (Tian et al., 2003; Kim et al., 2009; Gonzalez-Jorge et al., 2013). The *LYCOPENE CYCLASE PARALOG*, *CruP*, had not been studied previously in Arabidopsis seed, but in leaves, knockout and overexpression of *CruP* resulted in increased and decreased, respectively, levels of the leaf minor carotenoid, β-carotene-5,6-epoxide, without impacting leaf lutein, β-carotene, violaxanthin, and neoxanthin levels (Bradbury et al., 2012). It was hypothesized that *CruP*

prevents reactive oxygen species accumulation in leaves by reducing the levels of β-carotene-5,6-epoxide (Bradbury et al., 2012); however, in seed of the *CruP* knockout, β-carotene-5,6-epoxide was not detectable and, instead, β-carotene, violaxanthin, and lutein increased 20%, 18%, and 14%, respectively, relative to the wild type (Fig. 7), suggesting that *CruP* has additional functions in seeds.

Of the nine loci identified in this study as contributing to natural variation in Arabidopsis seed carotenoid levels, variation at the *ZEP* locus had the greatest impact (Figs. 1, A and C, 3, and 4). To assess the maximal potential impact of reducing *ZEP* activity on seed carotenoid traits, a null allele designated *aba1-7* was isolated and carotenoids were analyzed in dry seed (Fig. 5) and developing seed (Fig. 6; Supplemental Fig. S4). As expected, loss of *ZEP* activity led to an absence of the *ZEP*-dependent β-branch carotenoids, antheraxanthin, violaxanthin, and neoxanthin, and an astonishing 43-fold increase in zeaxanthin relative to the wild type, making it the most abundant carotenoid in *aba1-7* seed. However, β-carotene and lutein, which are not dependent on *ZEP* for their synthesis, also were increased unexpectedly in *aba1-7* seed (Fig. 5). There are at least two possible mechanisms to explain this result: a direct impact of the loss of *ZEP* activity on β-carotene and lutein and an indirect, pleiotropic effect due to ABA deficiency in *aba1-7*. In the direct scenario, portions of the β-carotene and lutein pools are normally epoxidated by *ZEP* in the wild type, which accelerates their degradation by carotenoid cleavage enzymes (Auldridge et al., 2006; Gonzalez-Jorge et al., 2013), and the loss of *ZEP* activity in *aba1-7* further increases the β-carotene and lutein pools. This possibility seems unlikely, as β-carotene-5,6-epoxide and lutein epoxide are not detectable in Col-0 and only trace amounts of lutein epoxide are present in *ccd1* and *ccd4* single and double mutant seed (Gonzalez-Jorge et al., 2013). As *CCD4* and *CCD1* are the major carotenoid-degrading enzymes in Arabidopsis (Auldridge et al., 2006; Gonzalez-Jorge et al., 2013), these genotypes would be expected to stabilize β-carotene-5,6-epoxide and lutein epoxide levels, as they do for the structurally similar epoxyxanthophylls violaxanthin and neoxanthin (Fig. 5).

An indirect, pleiotropic mechanism seems a more probable explanation for the increased β-carotene and lutein levels in *aba1-7* seed. Previous studies of ABA-insensitive mutants have shown how alterations to ABA-dependent processes impact seed maturation, dormancy, and embryo growth (Cheng et al., 2002; Frey et al., 2004). To investigate whether the ABA deficiency in *aba1-7* affects the regulation of carotenoid biosynthetic and degradation genes, the expression of 50 genes from four pathways (MEP, carotenoid biosynthesis and degradation, and ABA synthesis) was assessed in the ABA-insensitive mutant *abi5-7*, a major ABI locus impacting developing seed (Supplemental Table S9; Finkelstein and Lynch, 2000; Bensmihen et al., 2002; Brocard et al., 2002). It is particularly noteworthy that, relative to Col-0, the *CCD4* mRNA level was decreased by 86% in *abi5-7*. *CCD4* is a major negative regulator of seed carotenoid content in Arabidopsis (Gonzalez-Jorge et al., 2013) that is strongly induced during late seed maturation (Fig. 4A), when ABA levels peak. The reduction in *CCD4* expression in *abi5-7* is similar to that in Landsberg *erecta*, which is a poorly expressed, superior *CCD4* haplotype (Gonzalez-Jorge et al., 2013) that also exhibits 3- and 4-fold increases in dry seed β-carotene and lutein, relative to Col-0. Thus, it seems most likely that the increased β-carotene and lutein in *aba1-7* seed is a pleiotropic phenotype resulting from decreased *CCD4* activity as a result of ABA deficiency. This conclusion also is supported by carotenoid profiles in developing *aba1-7* seed (Fig. 6), which are consistent with reduced carotenoid turnover due to reduced *CCD4* expression/activity. In Col-0, the epoxy-xanthophylls, antheraxanthin, violaxanthin, and neoxanthin, total 1.35 ng seed^{-1} at 15 DAP and collectively decrease 90% by 42 DAP (Fig. 6). In *aba1-7*, the zeaxanthin that accumulates in their stead is 5.3-fold higher at 15 DAP (7.1 ng seed^{-1}) and decreases only 33% by 42 DAP. Lutein stability also is enhanced and decreases only 40% between 15 and 42 DAP versus 68% in the wild type.

This GWAS and a recent parallel GWAS of maize grain carotenoids (Owens et al., 2014) provide the most comprehensive genetic analyses to date in dicot and monocot seeds and allow, to our knowledge, the first direct comparison of the control of carotenoid natural variation in the two plant lineages. Seed development differs significantly in the two lineages, with dicots absorbing the endosperm during seed development and forming cotyledons that develop into photosynthetic organs upon germination, while in monocots the endosperm persists and accounts for the bulk of the grain that is subsequently degraded and mobilized during germination to support embryo growth. The Arabidopsis and maize GWAS identified a total of nine loci (Supplemental Tables S3 and S4) and six loci (Owens et al., 2014), respectively, that impact seed carotenoid content, of which four are in common: *ZEP*, *PSY*, *ZDS*, and *DXS3*. This suggests that, despite their differing ontologies, the underlying mechanisms and control points driving the natural variation of seed carotenoids in monocots and dicots are similar. This shared genetic architecture between these species has broad implications, as it suggests that comparative mapping of natural variation for essential dietary nutrients between monocots (maize, rice [*Oryza sativa*], and wheat, etc.) and dicots (Arabidopsis, soybean [*Glycine max*], and tomato, etc.) not only is possible but is likely to be a viable tool for defining and accelerating molecular breeding targets for these critical seed traits and their translation into breeding programs.

MATERIALS AND METHODS

Plant Materials

All plant material was grown under a 22°C/18°C 12-h light/dark photoperiod (100–120 μE) regimen. A total of 315 Arabidopsis (*Arabidopsis thaliana*)

accessions were collected from three individual biological outgrowths (replicates) of the 360 core collection (Platt et al., 2010). A single outgrowth of 142 lines from the Col-0 × Bla-1 recombinant inbred line population was utilized for QTL analysis (Simon et al., 2008). Several mutant lines were used in this study: *ZEP* gene (At5g67030), *aba1-6* (CS3772), *aba1-7* (CS25407, SALK_059469.55.50.x), *CCD1* gene (At3g63520, SAIL_390_C01, *ccd1-1*), *CCD4* gene (At4g19170, SALK_097984, *ccd4-1*), and *ccd1ccd4* mutant (Gonzalez-Jorge et al., 2013). The *aba1-7* homozygous line was identified by PCR genotyping with the following primers: *aba1-7*_KO_Forward, 5′-GATGTTGGTGGTGGAAAAATG-3′; *aba1-7*_KO_Reverse, 5′-ACGTTCAAGAGCATCGTCATC-3′; and T-DNA left border primer LBb1.3, 5′-ATTTTGCCGATTTCGGAAC-3′. All lines were obtained from the Arabidopsis Biological Resource Center at Ohio State University.

DNA Isolation and Genotyping

Extraction of genomic DNA for PCR genotyping was performed as described by Gilliland et al. (2006). InDel genotyping primers were as follows: *ZEP*_InDel_Sequencing_Forward, 5′-CCGTTAATCCGATCACCAGT-3′; *ZEP*_InDel_Sequencing_Reverse, 5′-GCTGATCACGTCACGTTCAA-3′; *ZEP*_InDel_Forward, 5′-TGCATGAGGAGGGTAAAAGG-3′; and *ZEP*_InDel_Forward, 5′-ACATCAACGAATGCGTGAAA-3′.

Seed Carotenoid Extraction and Quantification by HPLC

Except where indicated, all seed samples were harvested and allowed to dry for a period of at least 6 weeks prior to HPLC analysis. Sample preparation, extraction, and HPLC analysis were conducted as described by Gonzalez-Jorge et al. (2013).

Processing of Phenotypic Data

Phenotypic data from three biological outgrowths were processed using the procedures described by Lipka et al. (2013). Outliers were removed from phenotypic data using Studentized deleted residuals (Kutner et al., 2004) with a model including replicate and accession as random effects. Following outlier removal, the Box-Cox procedure (Box and Cox, 1964) was used to correct for the nonnormality of error terms and unequal variances. A mixed linear model fitted across the three replicates was utilized to calculate a best linear unbiased predictor for each line, which were then used as phenotypes. The aforementioned procedures were conducted in SAS version 9.2 (SAS Institute).

GWAS

GWAS of seed carotenoid best linear unbiased predictors was performed using 170,679 SNPs (out of the 250K Affymetrix SNP genotyping data set [call method 75]; Horton et al., 2012) with 5% or greater minor allele frequency for a total of 315 accessions. The unified mixed linear model (Yu et al., 2006) was used to test for associations between individual SNPs and traits of interest. This model includes PC (Price et al., 2006) as a fixed effect to account for population structure and a K matrix (Loiselle et al., 1995) accounting for familial relatedness. Utilization of the BIC (Schwarz, 1978) determined the optimal number of PCs to include in a mixed linear model by backward elimination (Supplemental Table S11). The PC and K matrix were calculated from 214,052 SNPs with a minor allele frequency greater than 10% among 1,307 accessions in the 250K SNP data set. The Benjamini and Hochberg (1995) procedure was used to control the FDR at 5%. The amount of phenotypic variation explained by the model was assessed using a likelihood ratio-based r^2 statistic (Sun et al., 2010). All analyses were conducted using the Genome Association and Prediction Integrated Tool software package (Lipka et al., 2012).

LD, Haploblock, and Haplotype Analysis

LD analysis was performed as described by Gonzalez-Jorge et al. (2013). Haploblocks for the 19.1-kb region surrounding *ZEP* were created in Haploview version 4.2 using the confidence interval method (Barrett et al., 2005). The association between each haploblock and trait of interest was assessed using a unified mixed linear model with corresponding PC and K (Yu et al., 2006). Models were fitted in SAS version 9.2 (SAS Institute) to allow the analysis of haplotype as a class variable. The LSMEANS statement in PROC MIXED was used to compare trait levels between haplotypes, and the Tukey-Kramer procedure was used to adjust for the multiple testing problem (Tukey, 1953; Kramer, 1956) at $\alpha = 0.05$.

MLMM Analysis

To resolve complex association signals surrounding a major-effect locus on chromosome 5, an MLMM analysis was conducted (Segura et al., 2012). In this analysis, the optimal model was determined at the genome level, where all 170,679 SNPs were included; at the chromosomal level, all chromosome 5 SNPs (totaling 42,538) were included, and at the local level, 287 SNPs within ±100 kb of the peak SNP on chromosome 5 were included. The extended BIC (Chen and Chen, 2008) was used to determine the optimal model.

Pathway-Level Analysis

Prior knowledge of carotenoid biosynthesis, degradation, and turnover was utilized to identify a subset of 42 genes whose enzymes are known to impact carotenoid levels (DellaPenna and Pogson, 2006). The association results from 10,921 markers located within ±100 kb of these 42 genes were then considered for a pathway-level analysis. For each trait analyzed, the Benjamini and Hochberg (1995) procedure was implemented to control the FDR at a maximum of 10%.

QTL Analysis

A 142-line Col-0 × Bla-1 recombinant inbred line population was used for QTL analysis using composite interval mapping performed with the PlabQTL package (Utz and Melchinger, 1996). The appropriate threshold LOD score for significance was determined by 1,000 permutation tests of the original data (Churchill and Doerge, 1994). A value of 2.5 LOD was calculated as the 5% significance threshold value for declaring a QTL.

Real-Time Quantitative Reverse Transcription-PCR

Total RNA extraction, DNase treatment, complementary DNA synthesis, and real-time quantitative reverse transcription-PCR were performed as described by Gonzalez-Jorge et al. (2013). Real-time quantitative reverse transcription-PCR primers were as follows: *PP2AA3*-Forward, 5′-ATCGCTTCTCGCTCCAGTAATG-3′; *PP2AA3*-Reverse, 5′-GACTATCGGAATGAGAGATTGC-3′; *ZEP*-Forward, 5′-GATGCAGCCAAATATGGGTCAAGG-3′; and *ZEP*-Reverse, 5′-GCCATTGCATGGATAATAGCGACTC-3′.

Carotenoid Degradation during Seed Development and Desiccation

Arabidopsis flowers of Col-0 and *aba1-7* plants were tagged at anthesis, and developing seeds at 15, 21, and 42 DAP were harvested. A total of 450 seeds were collected per developmental time point, specifically 150 seeds per line per time point in triplicate. These samples were then quantified for their carotenoid profiles by HPLC.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Pathways and associated genes leading to carotenoid synthesis and turnover in Arabidopsis.

Supplemental Figure S2. Genome-wide association and LD decay plot of the *ZEP* region.

Supplemental Figure S3. Identification of a null mutation for *ZEP*, *aba1-7*, and its location relative to the ethyl methanesulfonate mutant allele, *aba1-6*.

Supplemental Figure S4. Carotenoid levels during seed development and desiccation in wild type Col-0 and *aba1-7* mutant lines.

Supplemental Table S1. Mean and range of raw phenotypic values in ng mg^{-1} seed carotenoids across three biological outgrowths of the 315-member Arabidopsis diversity panel.

Supplemental Table S2. List of 42 a priori genes from three pathways utilized for the pathway-level analysis of carotenoid GWAS.

Supplemental Table S3. Pathway-level analysis of carotenoid GWAS at the 5% FDR threshold.

Supplemental Table S4. Pathway-level analysis of carotenoid GWAS at the 10% FDR threshold.

Supplemental Table S5. List of all SNPs and InDels identified in the 19.1-kb LD V/A association region surrounding the candidate gene, *ZEP*.

Supplemental Table S6. Contrasting haplotype analysis of the 19.1-kb *ZEP* LD region.

Supplemental Table S7. Carotenoid levels of neoxanthin and violaxanthin for the Col-0 inferior and superior *ZEP* haplotype lines.

Supplemental Table S8. Seed carotenoids in homozygous lines of wild-type Col-0 and an allelic series of *aba1* mutants.

Supplemental Table S9. Gene expression values for dry seed of the Col-0 accession and the ABA-insensitive mutant, *abi5-7*.

Supplemental Table S10. Gene expression values for late seed development stages of the Col-0 accession.

Supplemental Table S11. Model selection of carotenoid GWAS traits.

Received April 14, 2016; accepted May 3, 2016; published May 6, 2016.

LITERATURE CITED

Abbo S, Bonfil DJ, Kerem Z, Berkovitch Z, Reifen R (2010) Towards enhancing lutein concentration in chickpea, cultivar and management effects. Plant Breed **129:** 407–411

Abbo S, Molina C, Jungmann R, Grusak MA, Berkovitch Z, Reifen R, Kahl G, Winter P, Reifen R (2005) Quantitative trait loci governing carotenoid concentration and weight in seeds of chickpea (Cicer arietinum L.). Theor Appl Genet **111:** 185–195

Alboresi A, Dall'Osto L, Aprile A, Carillo P, Roncaglia E, Cattivelli L, Bassi R (2011) Reactive oxygen species and transcript analysis upon excess light treatment in wild-type Arabidopsis thaliana vs a photosensitive mutant lacking zeaxanthin and lutein. BMC Plant Biol **11:** 62–83

Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, Ghisla S, Bouwmeester H, Beyer P, Al-Babili S (2012) The path from β-carotene to carlactone, a strigolactone-like plant hormone. Science **335:** 1348–1351

Angelovici R, Lipka AE, Deason N, Gonzalez-Jorge S, Lin H, Cepela J, Buell R, Gore MA, DellaPenna D (2013) Genome-wide analysis of branched-chain amino acid levels in *Arabidopsis* seeds. Plant Cell **25:** 4827–4843

Arango J, Wüst F, Beyer P, Welsch R (2010) Characterization of phytoene synthases from cassava and their involvement in abiotic stress-mediated responses. Planta **232:** 1251–1262

Audran C, Borel C, Frey A, Sotta B, Meyer C, Simonneau T, Marion-Poll A (1998) Expression studies of the zeaxanthin epoxidase gene in *Nicotiana plumbaginifolia*. Plant Physiol **118:** 1021–1028

Audran C, Liotenberg S, Gonneau M, North H, Frey A, Tap-Waksman K, Vartanian N, Marion-Poll A (2001) Localisation and expression of zeaxanthin epoxidase mRNA in Arabidopsis in response to drought stress and during seed development. Funct Plant Biol **28:** 1161–1173

Auldridge ME, Block A, Vogel JT, Dabney-Smith C, Mila I, Bouzayen M, Magallanes-Lundback M, DellaPenna D, McCarty DR, Klee HJ (2006) Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. Plant J **45:** 982–993

Avendaño-Vázquez AO, Cordoba E, Llamas E, San Román C, Nisar N, De la Torre S, Ramos-Vega M, Gutiérrez-Nava MD, Cazzonelli CI, Pogson BJ, et al (2014) An uncharacterized apocarotenoid-derived signal generated in ζ-carotene desaturase mutants regulates leaf development and the expression of chloroplast and nuclear genes in *Arabidopsis*. Plant Cell **26:** 2524–2537

Bailey S, Grossman A (2008) Photoprotection in cyanobacteria: regulation of light harvesting. Photochem Photobiol **84:** 1410–1420

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics **21:** 263–265

Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E, Li Y, Bergelson J, Borevitz JO, Nordborg M, Vitek O, Salt DE (2010) A coastal cline in sodium accumulation in Arabidopsis thaliana is driven by natural variation of the sodium transporter AtHKT1;1. PLoS Genet **6:** e1001193

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B **57:** 289–300

Bensmihen S, Rippa S, Lambert G, Jublot D, Pautot V, Granier F, Giraudat J, Parcy F (2002) The homologous ABI5 and EEL transcription factors function antagonistically to fine-tune gene expression during late embryogenesis. Plant Cell **14:** 1391–1403

Blanco A, Colasuonno P, Gadaleta A, Mangini G, Schiavulli A, Simeone R, Digesù AM, De Vita P, Mastrangelo AM, Cattivelli L (2011) Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat. J Cereal Sci **54:** 255–264

Botella-Pavía P, Besumbes O, Phillips MA, Carretero-Paulet L, Boronat A, Rodríguez-Concepción M (2004) Regulation of carotenoid biosynthesis in plants: evidence for a key role of hydroxymethylbutenyl diphosphate reductase in controlling the supply of plastidial isoprenoid precursors. Plant J **40:** 188–199

Box GEP, Cox DR (1964) An analysis of transformations. J R Stat Soc Ser B Stat Methodol **26:** 211–252

Bradbury LM, Shumskaya M, Tzfadia O, Wu SB, Kennelly EJ, Wurtzel ET (2012) Lycopene cyclase paralog CruP protects against reactive oxygen species in oxygenic photosynthetic organisms. Proc Natl Acad Sci USA **109:** E1888–E1897

Brocard IM, Lynch TJ, Finkelstein RR (2002) Regulation and role of the Arabidopsis abscisic acid-insensitive 5 gene in abscisic acid, sugar, and stress response. Plant Physiol **129:** 1533–1543

Busch M, Seuter A, Hain R (2002) Functional analysis of the early steps of carotenoid biosynthesis in tobacco. Plant Physiol **128:** 439–453

Campbell R, Ducreux LJ, Morris WL, Morris JA, Suttle JC, Ramsay G, Bryan GJ, Hedley PE, Taylor MA (2010) The metabolic and developmental roles of carotenoid cleavage dioxygenase4 from potato. Plant Physiol **154:** 656–664

Carol P, Kuntz M (2001) A plastid terminal oxidase comes to light: implications for carotenoid biosynthesis and chlororespiration. Trends Plant Sci **6:** 31–36

Carretero-Paulet L, Cairó A, Botella-Pavía P, Besumbes O, Campos N, Boronat A, Rodríguez-Concepción M (2006) Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. Plant Mol Biol **62:** 683–695

Cazzonelli CI, Pogson BJ (2010) Source to sink: regulation of carotenoid biosynthesis in plants. Trends Plant Sci **15:** 266–274

Chandler K, Lipka AE, Owens BF, Li H, Buckler ES, Rocheford T, Gore MA (2013) Genetic analysis of visually scored orange kernel color in maize. Crop Sci **53:** 189–200

Chen J, Chen Z (2008) Extended Bayesian information criteria for model selection with large model spaces. Biometrika **95:** 759–771

Cheng WH, Endo A, Zhou L, Penney J, Chen HC, Arroyo A, Leon P, Nambara E, Asami T, Seo M, et al (2002) A unique short-chain dehydrogenase/reductase in *Arabidopsis* glucose signaling and abscisic acid biosynthesis and functions. Plant Cell **14:** 2723–2743

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971

Cook JP, McMullen MD, Holland JB, Tian F, Bradbury P, Ross-Ibarra J, Buckler ES, Flint-Garcia SA (2012) Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. Plant Physiol **158:** 824–834

Couso I, Cordero BF, Vargas MA, Rodríguez H (2012) Efficient heterologous transformation of Chlamydomonas reinhardtii npq2 mutant with the zeaxanthin epoxidase gene isolated and characterized from Chlorella zofingiensis. Mar Drugs **10:** 1955–1976

Cuttriss AJ, Cazzonelli CI, Wurtzel ET, Pogson BJ (2011) Carotenoids. Adv Bot Res **58:** 1–36

Dall'Osto L, Cazzaniga S, North H, Marion-Poll A, Bassi R (2007) The *Arabidopsis aba4-1* mutant reveals a specific function for neoxanthin in protection against photooxidative stress. Plant Cell **19:** 1048–1064

Davison PA, Hunter CN, Horton P (2002) Overexpression of beta-carotene hydroxylase enhances stress tolerance in Arabidopsis. Nature **418:** 203–206

DellaPenna D, Pogson BJ (2006) Vitamin synthesis in plants: tocopherols and carotenoids. Annu Rev Plant Biol **57:** 711–738

Dong H, Deng Y, Mu J, Lu Q, Wang Y, Xu Y, Chu C, Chong K, Lu C, Zuo J (2007) The Arabidopsis Spontaneous Cell Death1 gene, encoding a zeta-carotene desaturase essential for carotenoid biosynthesis, is involved in chloroplast development, photoprotection and retrograde signalling. Cell Res **17:** 458–470

Estévez JM, Cantero A, Reindl A, Reichler S, León P (2001) 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. J Biol Chem **276:** 22901–22909

Fang J, Chai C, Qian Q, Li C, Tang J, Sun L, Huang Z, Guo X, Sun C, Liu M, et al (2008) Mutations of genes in synthesis of the carotenoid precursors of ABA lead to pre-harvest sprouting and photo-oxidation in rice. Plant J **54:** 177–189

Fernandez MGS, Hamblin MT, Li L, Rooney WL, Tuinstra MR, Kresovich S (2008) Quantitative trait loci analysis of endosperm color and carotenoid content in sorghum grain. Crop Sci **48:** 1732–1743

Finkelstein RR, Lynch TJ (2000) The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. Plant Cell **12:** 599–609

Fiore A, Dall'Osto L, Fraser PD, Bassi R, Giuliano G (2006) Elucidation of the beta-carotene hydroxylation pathway in Arabidopsis thaliana. FEBS Lett **580:** 4718–4722

Fraser PD, Kiano JW, Truesdale MR, Schuch W, Bramley PM (1999) Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid synthesis in ripening fruit. Plant Mol Biol **40:** 687–698

Fray RG, Wallace A, Fraser PD, Valero D, Hedden P, Bramley PM, Grierson D (1995) Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. Plant J **8:** 693–701

Frey A, Boutin JP, Sotta B, Mercier R, Marion-Poll A (2006) Regulation of carotenoid and ABA accumulation during the development and germination of Nicotiana plumbaginifolia seeds. Planta **224:** 622–632

Frey A, Effroy D, Lefebvre V, Seo M, Perreau F, Berger A, Sechet J, To A, North HM, Marion-Poll A (2012) Epoxycarotenoid cleavage by NCED5 fine-tunes ABA accumulation and affects seed dormancy and drought tolerance with other NCED family members. Plant J **70:** 501–512

Frey A, Godin B, Bonnet M, Sotta B, Marion-Poll A (2004) Maternal synthesis of abscisic acid controls seed development and yield in Nicotiana plumbaginifolia. Planta **218:** 958–964

Galpaz N, Wang Q, Menda N, Zamir D, Hirschberg J (2008) Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. Plant J **53:** 717–730

Gilliland LU, Magallanes-Lundback M, Hemming C, Supplee A, Koornneef M, Bentsink L, DellaPenna D (2006) Genetic basis for natural variation in seed vitamin E levels in Arabidopsis thaliana. Proc Natl Acad Sci USA **103:** 18834–18841

Giuliano G, Bartley GE, Scolnik PA (1993) Regulation of carotenoid biosynthesis during tomato development. Plant Cell **5:** 379–387

Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pagès V, Dun EA, Pillot JP, Letisse F, Matusova R, Danoun S, Portais JC, et al (2008) Strigolactone inhibition of shoot branching. Nature **455:** 189–194

Gonzalez-Jorge S, Ha SH, Magallanes-Lundback M, Gilliland LU, Zhou A, Lipka AE, Nguyen YN, Angelovici R, Lin H, Cepela J, et al (2013) Carotenoid cleavage dioxygenase4 is a negative regulator of β-carotene content in *Arabidopsis* seeds. Plant Cell **25:** 4812–4826

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

Hobbs RP, Bernstein PS (2014) Nutrient supplementation for age-related macular degeneration, cataract, and dry eye. J Ophthalmic Vis Res **9:** 487–493

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjálmsson BJ, et al (2012) Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. Nat Genet **44:** 212–216

Howitt CA, Pogson BJ (2006) Carotenoid accumulation and function in seeds and non-green tissues. Plant Cell Environ **29:** 435–445

Jahns P, Holzwarth AR (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. Biochim Biophys Acta **1817:** 182–193

Jourdan M, Gagné S, Dubois-Laurent C, Maghraoui M, Huet S, Suel A, Hamama L, Briard M, Peltier D, Geoffriau E (2015) Carotenoid content and root color of cultivated carrot: a candidate-gene association study using an original broad unstructured population. PLoS ONE **10:** e0116674

Kandianis CB, Stevens R, Liu W, Palacios N, Montgomery K, Pixley K, White WS, Rocheford T (2013) Genetic architecture controlling variation in grain carotenoid composition and concentrations in two maize populations. Theor Appl Genet **126:** 2879–2895

Kim J, Smith JJ, Tian L, DellaPenna D (2009) The evolution and function of carotenoid hydroxylases in Arabidopsis. Plant Cell Physiol **50:** 463–479

Koornneef M, Bentsink L, Hilhorst H (2002) Seed dormancy and germination. Curr Opin Plant Biol **5:** 33–36

Koornneef M, Jorna ML, Brinkhorst-van der Swan DL, Karssen CM (1982) The isolation of abscisic acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberellin sensitive lines of Arabidopsis thaliana (L.) Heynh. Theor Appl Genet **61:** 385–393

Kramer CY (1956) Extension of multiple range tests to group means with unequal numbers of replications. Biometrics **12:** 307–310

Krinsky NI, Johnson EJ (2005) Carotenoid actions and their relation to health and disease. Mol Aspects Med **26:** 459–516

Kutner MH, Nachtsheim CJ, Neter J (2004) Applied Linear Regression Models. McGraw-Hill Education, Boston

Landrum JT, Bone RA (2004) Dietary lutein & zeaxanthin: reducing the risk for macular degeneration. Agro Food Ind Hi-Tech **15:** 22–25

Lätari K, Wüst F, Hübner M, Schaub P, Beisel KG, Matsubara S, Beyer P, Welsch R (2015) Tissue-specific apocarotenoid glycosylation contributes to carotenoid homeostasis in Arabidopsis leaves. Plant Physiol **168:** 1550–1562

Lefebvre V, North H, Frey A, Sotta B, Seo M, Okamoto M, Nambara E, Marion-Poll A (2006) Functional analysis of Arabidopsis NCED6 and NCED9 genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. Plant J **45:** 309–319

Lindgren LO, Stålberg KG, Höglund AS (2003) Seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll, and abscisic acid. Plant Physiol **132:** 779–785

Lipka AE, Gore MA, Magallanes-Lundback M, Mesberg A, Lin H, Tiede T, Chen C, Buell CR, Buckler ES, Rocheford T, et al (2013) Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. G3 (Bethesda) **3:** 1287–1299

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: Genome Association and Prediction Integrated Tool. Bioinformatics **28:** 2397–2399

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic-structure of a tropical understory shrub, Psychotria officinalis (Rubiaceae). Am J Bot **82:** 1420–1425

Marin E, Nussaume L, Quesada A, Gonneau M, Sotta B, Hugueney P, Frey A, Marion-Poll A (1996) Molecular identification of zeaxanthin epoxidase of Nicotiana plumbaginifolia, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of Arabidopsis thaliana. EMBO J **15:** 2331–2342

McCarty DR (1995) Genetic control and integration of maturation and germination pathways in seed development. Annu Rev Plant Physiol Plant Mol Biol **46:** 71–93

Nakabayashi K, Okamoto M, Koshiba T, Kamiya Y, Nambara E (2005) Genome-wide profiling of stored mRNA in Arabidopsis thaliana seed germination: epigenetic and genetic regulation of transcription in seed. Plant J **41:** 697–709

Nambara E, Marion-Poll A (2003) ABA action and interactions in seeds. Trends Plant Sci **8:** 213–217

Nambara E, Marion-Poll A (2005) Abscisic acid biosynthesis and catabolism. Annu Rev Plant Biol **56:** 165–185

Nisar N, Li L, Lu S, Khin NC, Pogson BJ (2015) Carotenoid metabolism in plants. Mol Plant **8:** 68–82

Niyogi KK (1999) Photoprotection revisited: genetic and molecular approaches. Annu Rev Plant Physiol Plant Mol Biol **50:** 333–359

Niyogi KK, Grossman AR, Björkman O (1998) *Arabidopsis* mutants define a central role for the xanthophyll cycle in the regulation of photosynthetic energy conversion. Plant Cell **10:** 1121–1134

Owens BF, Lipka AE, Magallanes-Lundback M, Tiede T, Diepenbrock CH, Kandianis CB, Kim E, Cepela J, Mateos-Hernandez M, Buell CR,

et al (2014) A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels. Genetics **198:** 1699–1716

Paran I, van der Knaap E (2007) Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. J Exp Bot **58:** 3841–3852

Park HY, Seok HY, Park BK, Kim SH, Goh CH, Lee BH, Lee CH, Moon YH (2008) Overexpression of Arabidopsis ZEP enhances tolerance to osmotic stress. Biochem Biophys Res Commun **375:** 80–85

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, et al (2010) The scale of population structure in Arabidopsis thaliana. PLoS Genet **6:** e1000843

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

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet **38:** 904–909

Qin G, Gu H, Ma L, Peng Y, Deng XW, Chen Z, Qu LJ (2007) Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in Arabidopsis by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. Cell Res **17:** 471–482

Rock CD, Zeevaart JA (1991) The aba mutant of Arabidopsis thaliana is impaired in epoxy-carotenoid biosynthesis. Proc Natl Acad Sci USA **88:** 7496–7499

Ruiz-Sola MA, Rodríguez-Concepción M (2012) Carotenoid biosynthesis in Arabidopsis: a colorful pathway. The Arabidopsis Book **10:** e0158, doi: 10.1199/tab.0158

Schwarz G (1978) Estimating the dimension of a model. Ann Stat **6:** 461–464

Schwarz N, Armbruster U, Iven T, Brückle L, Melzer M, Feussner I, Jahns P (2015) Tissue-specific accumulation and regulation of zeaxanthin epoxidase in Arabidopsis reflect the multiple functions of the enzyme in plastids. Plant Cell Physiol **56:** 346–357

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet **44:** 825–830

Shewmaker CK, Sheehy JA, Daley M, Colburn S, Ke DY (1999) Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. Plant J **20:** 401–412X

Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal FX, Durand-Tardif M, Pelletier G, Camilleri C (2008) Quantitative trait loci mapping in five new large recombinant inbred line populations of Arabidopsis thaliana genotyped with consensus single-nucleotide polymorphism markers. Genetics **178:** 2253–2264

Smolikova GN, Laman NA, Boriskevich OV (2011) Role of chlorophylls and carotenoids in seed tolerance to abiotic stressors. Russ J Plant Physiol **58:** 965–973

Sonah H, O'Donoughue L, Cober E, Rajcan I, Belzile F (2015) Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotechnol J **13:** 211–221

Sun G, Zhu C, Kramer MH, Yang SS, Song W, Piepho HP, Yu J (2010) Variation explained in mixed-model association mapping. Heredity (Edinb) **105:** 333–340

Suwarno WB, Pixley KV, Palacios-Rojas N, Kaeppler SM, Babu R (2015) Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. Theor Appl Genet **128:** 851–864

Tian L, Magallanes-Lundback M, Musetti V, DellaPenna D (2003) Functional analysis of beta- and epsilon-ring carotenoid hydroxylases in *Arabidopsis*. Plant Cell **15:** 1320–1332

Toledo-Ortiz G, Huq E, Rodríguez-Concepción M (2010) Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. Proc Natl Acad Sci USA **107:** 11626–11631

Tukey JW (1953) The problem of multiple comparisons. *In* The Collected Works of John W. Tukey, **Vol 8**, 1994. Chapman & Hall, New York

Utz HF, Melchinger AE (1996) PLABQTL: a program for composite interval mapping of QTL. Journal of Quantitative Trait Loci **2:** 1–5

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

Weigel D, Mott R (2009) The 1001 Genomes Project for Arabidopsis thaliana. Genome Biol **10:** 107–112

Welsch R, Arango J, Bär C, Salazar B, Al-Babili S, Beltrán J, Chavarriaga P, Ceballos H, Tohme J, Beyer P (2010) Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. Plant Cell **22:** 3348–3356

Wen W, Li D, Li X, Gao Y, Li W, Li H, Liu J, Liu H, Chen W, Luo J, et al (2014) Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. Nat Commun **5:** 3438–3448

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

Yan J, Kandianis CB, Harjes CE, Bai L, Kim EH, Yang X, Skinner DJ, Fu Z, Mitchell S, Li Q, et al (2010) Rare genetic variation at Zea mays crtRB1 increases beta-carotene in maize grain. Nat Genet **42:** 322–327

Young AJ (1993) Factors that affect the carotenoid composition of higher plants and algae. *In* Carotenoids in Photosynthesis, AJ Young and G Britton, eds. Springer, Dordrecht, The Netherlands, pp 160–205

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet **38:** 203–208

Zhou X, McQuinn R, Fei Z, Wolters AM, van Eck J, Brown C, Giovannoni JJ, Li L (2011) Regulatory control of high levels of carotenoid accumulation in potato tubers. Plant Cell Environ **34:** 1020–1030

Zhou Y, Han Y, Li Z, Fu Y, Fu Z, Xu S, Li J, Yan J, Yang X (2012) ZmcrtRB3 encodes a carotenoid hydroxylase that affects the accumulation of α-carotene in maize kernel. J Integr Plant Biol **54:** 260–269

Zuber H, Davidian JC, Wirtz M, Hell R, Belghazi M, Thompson R, Gallardo K (2010) Sultr4;1 mutant seeds of Arabidopsis have an enhanced sulphate content and modified proteome suggesting metabolic adaptations to altered sulphate compartmentalization. BMC Plant Biol **10:** 78–91