Title
Genome-Wide Association Study and Pathway-Level Analysis of Kernel Color in Maize.

Permalink
https://escholarship.org/uc/item/6d01c1z7

Journal
G3 (Bethesda, Md.), 9(6)

ISSN
2160-1836

Authors
Owens, Brenda F
Mathew, Deepu
Diepenbrock, Christine H
et al.

Publication Date
2019-06-05

DOI
10.1534/g3.119.400040

Peer reviewed
Genome-Wide Association Study and Pathway-Level Analysis of Kernel Color in Maize

Brenda F. Owens,*,1,2 Deepu Mathew,*,3 Christine H. Diepenbrock,†,1,4 Tyler Tiede,*,5 Di Wu,† Maria Mateos-Hernandez,*,6 Michael A. Gore,† and Torbert Rocheford*  
*Department of Agronomy, Purdue University, West Lafayette, Indiana 47907, and †Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853  
ORCID ID: 0000-0001-6896-8024 (M.A.G.)

ABSTRACT  Rapid development and adoption of biofortified, provitamin A-dense orange maize (Zea mays L.) varieties could be facilitated by a greater understanding of the natural variation underlying kernel color, including as it relates to carotenoid biosynthesis and retention in maize grain. Greater abundance of carotenoids in maize kernels is generally accompanied by deeper orange color, useful for distinguishing provitamin A-dense varieties to consumers. While kernel color can be scored and selected with high-throughput, low-cost phenotypic methods within breeding selection programs, it remains to be well established as to what would be the logical genetic loci to target for selection for kernel color. We conducted a genome-wide association study of maize kernel color, as determined by colorimetry, in 1,651 yellow and orange inbreds from the Ames maize inbred panel. Associations were found with y1, encoding the first committed step in carotenoid biosynthesis, and with dxs2, which encodes the enzyme responsible for the first committed step in the biosynthesis of the isoprenoid precursors of carotenoids. These genes logically could contribute to overall carotenoid abundance and thus kernel color. The lcyE and zep1 genes, which can affect carotenoid composition, were also found to be associated with colorimeter values. A pathway-level analysis, focused on genes with a priori evidence of involvement in carotenoid biosynthesis and retention, revealed associations for dxs3 and dmes1, involved in isoprenoid biosynthesis; ps1 and vp5, within the core carotenoid pathway; and vp14, involved in cleavage of carotenoids. Collectively, these identified genes appear relevant to the accumulation of kernel color.

KEYWORDS  colorimeter  carotenoid  isoprenoid  genome-wide association study  biofortification

Malnutrition, or hidden hunger, remains a serious issue, even as increased agricultural productivity has helped to provide more energy and calories on a global scale (Welch and Graham 1999). As much as half of the world’s population may be deficient in one or more micronutrients, with 125–130 million pre-school children and 7 million pregnant women suffering from vitamin A deficiency (VAD) (Stevens et al. 2015). Biofortification, the improvement of crop nutritional quality through breeding and/or agronomics, has been proposed as a sustainable tool to help with addressing micronutrient malnutrition (Bouis and Welch 2010), and has been found to be cost-effective (Meenakshi et al. 2012; Bouis and Hunt 1999; Qaim et al. 2007). Improvement of provitamin A carotenoid levels is generally a promising target, given that naturally occurring yellow and orange-pigmented accessions have been identified for many commonly white-pigmented, starchy staple foods such as maize, cassava, banana, and sweet potato (Amorim et al. 2009; Carvalho et al. 2016; Takahata et al. 1993). For biofortification to be effective, micronutrient densities must reach levels that impact human health, and the varieties and final food products must be acceptable to growers and consumers. Through decades of technical and broader contextual work, the international breeding organizations of CIMMYT, IITA and HarvestPlus, and partners have achieved the successful development of provitamin A-dense maize varieties, nearing target nutrient levels, which also have local and regional adaptation and relevance (Pixley et al. 2013, Menkir et al. 2017). Specifically, there has been a need to develop maize with distinctly orange kernel color for enhanced product recognition and enhanced consumer acceptance, including in certain sub-Saharan African nations where white maize is preferred but outreach and educational initiatives have successfully linked enhanced nutritional properties to the novel orange color (Meenakshi et al. 2012; Muzhingi et al. 2008; reviewed in Simpungwe et al. 2017). For the consistent and facilitated development of biofortified, provitamin A-dense maize varieties that meet target nutrient levels and also have strongly orange endosperm, it
Carotenoids, including the provitamin A compounds α-carotene, β-carotene, and β-criptoxanthin, are members of a large group of isoprenoid compounds synthesized in plants. Deoxy-xylulose 5-phosphate (DOXP) is formed by deoxy-xylulose 5-phosphate synthase (DXS) in the first step of the non-mevalonate (or mevalenyl-throtol 4-phosphate, hereafter MEP) pathway for isoprenoid biosynthesis in plastids. Seven more reactions are needed for the formation of the immediate carotenoid precursor, geranylgeranyl pyrophosphate (GGPP) from isopentenyl pyrophosphate (IPP) (Figure 1) (Hirschberg 2001; Rodriguez-Concepcion and Boronat 2002; Hunter 2007; Rodriguez-Concepcion et al. 2013; Vranová et al. 2013). The first committed step in carotenoid biosynthesis involves the formation of phytoene from two molecules of GGPP by phytoene synthase (PSY) (Buckner et al. 1996). Four more steps result in the biosynthesis of lycopene, after which there resides a key branch point in the pathway. For the biosynthesis of α-carotene branch carotenoids, lycopene can be cyclized by lycopene β-cyclase (LCYB) at one end and by lycopene ε-cyclase (LCYE) at the other end to form α-carotene; from there, hydroxylation of the β-ring produces zeaxanthin, and subsequent hydroxylation of the ε-ring produces lutein. Alternatively, for the biosynthesis of β-carotene branch carotenoids, lycopene can be cyclized by LCYB at both ends to form β-carotene; from there, hydroxylation of one β-ring produces β-criptoxanthin, and subsequent hydroxylation of the other β-ring produces zeaxanthin. Zeaxanthin can be further epoxidated to antheraxanthin and violaxanthin (Figure 2) (Hirschberg 2001; DellaPenna and Pogson 2006). A number of apocarotenoid metabolites are additionally formed from the oxidative cleavage of carotenoids by carotenoid cleavage dioxygenases (CCDs) and 9-cis-epoxycarotenoid dioxygenases (NCEDs), including strigolactones, abscisic acid (ABA), and various aromatic volatile compounds (Tan et al. 1997; Schwartz et al. 1997; Schwartz et al. 2001; Matusova et al. 2005; Sun et al. 2008; Vogel et al. 2008; Messing et al. 2010; Vallabhaneni et al. 2010; reviewed by Auldridge et al. 2006).

Many carotenoid compounds have yellow-to-red coloration dependent on functional groups and the length of their conjugated double bond systems (Khoo et al. 2011). Lutein and zeaxanthin, the two most abundant carotenoid compounds in maize grain (Owens et al. 2014), have been reported as light yellow and yellow-orange, respectively (Weber 1987; Meléndez-Martínez et al. 2007). Within the maize kernel, carotenoids predominantly accumulate in the vitreous portion of the endosperm (Weber 1987), though ABA which is derived from carotenoids plays a key role in the embryo in seed dormancy (McCarty 1995; Kermode 2005). The genes described in the MEP pathway and carotenoid biosynthetic pathway are logical a priori candidates for the genetic control of kernel color, given that their gene action could feasibly impact the hue and/or intensity of maize endosperm coloration.

Carotenoid composition, or relative abundance of individual carotenoid compounds, is typically quantified using high-performance liquid chromatography (HPLC). However, HPLC is cost- and labor-intensive and may not be amenable to the high-throughput measurements called for in certain stages of a breeding program (Diepenbrock and Gore 2015). For example, measurement methodologies that are still quantitative but less resource-intensive may have particular utility in the initial stages of breeding, in which large numbers of progeny are typically evaluated (Jaramillo et al. 2018, Ikeogu et al. 2017, Lozano-Alejo et al. 2007). However, it is important to understand the genetic loci underlying kernel color traits, particularly if colorimetry is to be used as a pre-screening tool in early-stage selections, so as not to select against favorable alleles at loci controlling provitamin A levels (or other compositional traits of importance to human health and nutrition).

Gradation in orange kernel color was previously visually scored on an ordinal scale, on bulks of kernels sampled from maize ears of 10 recombinant inbred line (RIL) families of the U.S. maize nested association mapping (NAM) population (McMullen et al. 2009). This study identified QTL for kernel color, which also mapped to regions containing carotenoid biosynthetic pathway genes (Chandler et al. 2013). Breeding for carotenoid levels based on visual selection for deep orange kernel color (and allele mining from exotic accessions) has also been carried out (Burt et al. 2011). Three QTL studies in other cereals identified intervals that were associated with colorimeter measurements of wheat endosperm, wheat flour, and sorghum endosperm, and that were in the vicinity of genes with putative involvement in carotenoid accumulation (Fernandez et al. 2008, Blanco et al. 2011, Zhao et al. 2013). These findings, combined with the rapid, inexpensive, and quantitative nature of colorimetric measurements, suggest that colorimetry may be a feasible method for quantification of maize kernel color in breeding programs, including for genetic analyses.

A colorimeter is an instrument that converts reflectance measurements into values that correspond to human perception of color. The CIELAB (L’,a’,b’) system is based on color-opponent theory, or color being perceived by the following pairs of opposites (Hunter and Harold 1987). The L axis represents a light to dark scale where positive values are lighter and negative values are darker. The “a” axis represents a greenness to redness scale where positive values are more red and negative values are more green. The “b” axis represents a yellowness to blueness scale where positive values are more yellow and negative values are more blue. Chroma is calculated from “a” and “b” values (Berger-Schunn 1994). Chroma represents the saturation or vividness of color, and hue represents the basic perceived color (whether the color would be called green or orange, for example) (Darrigues et al. 2008). Thus, hue and Chroma convert the a’ and b’ values to scores that represent a place in the color space to which humans have assigned a color name. Colorimeter values offer certain advantages over visual scoring given that they are quantitative, providing a more continuous
Figure 1 The plastidic methylerythritol 4-phosphate (MEP) biosynthetic pathway in plants. Compound names and abbreviations are as listed in the figure. Enzyme names and abbreviations: DOXP synthase (DXS), DOXP reductoisomerase (DXR), MEP cytidylyltransferase (MCT), CDP-ME synthase (DMES), DP-ME kinase (CMK), MECP synthase (MDS, MECS), 4-hydroxy-3-methylbut-2-enyl-diphosphate [HMBBP] synthase (HDS), HMBBP reductase (HDR), isopentenyl pyrophosphate isomerase (IPPI), geranylgeranyl pyrophosphate synthase (GGPS).
Figure 2 The carotenoid biosynthetic pathway in plants. Enzyme names and abbreviations: phytoene synthase (PSY), phytoene desaturase (PDS), ζ-carotene isomerase (ZDS), ζ-carotene desaturase (ZDS), carotenoid isomerase (Crt-ISO), lycopene β-cyclase (LCYB), lycopene ε-cyclase (LCYE), β-carotene hydroxylase (CRTRB), P450 carotenoid β-ring hydroxylase (CYP97A), P450 carotenoid epsilon-ring hydroxylase (CYP97C), zeaxanthin epoxidase (ZEP).
scale of measurement, objective, allowing values to be compared across breeding populations over time, and representative of multiple components of kernel color.

Colorimetric methods were used in this study to genetically dissect the kernel color of 1,651 inbred lines from the Ames maize inbred panel (Romay et al. 2013). This study was conducted to 1) investigate the regions of the maize genome influencing kernel color using a genome-wide association study (GWAS), and 2) determine whether pathway-level analysis reveals additional associations with carotenoid-related genes.

**MATERIALS AND METHODS**

**Experimental Design and Phenotypic Data**

We grew a 2,448 experimental inbred line subset of a population consisting of 2,815 maize inbred lines maintained by the National Plant Germplasm System (Romay et al. 2013), hereafter referred to as the Ames maize inbred panel. Seed was provided by the North Central Regional Plant Introduction Station (NCRPIS) in Ames, IA, and grown as a single replicate at the Purdue Agromony Center for Research and Education (ACRE) in West Lafayette, IN, in 2012 and 2013. The inbred lines were grouped into six sets based on maturity (i.e., flowering time) to facilitate pollination, harvesting, and phenotyping efforts. Each set was arranged in a 20 × 24 incomplete block design. Each block within each set was augmented with an experiment-wide check line (B73) plot in a random position, and six other check lines of varying maturities based on flowering time (P39, Mo17, B97, NC358, Mo18W, CML247 in 2012 and PH40, Mo17, PHG35, PHG39, CML247, DK311H6 in 2013) were included twice per block in random positions. An experimental unit consisted of a one-row plot, 3.81 m in length containing approximately 15 plants. Plots had a spacing between rows of 0.762 m. Efforts were made to hand-pollinate up to six plants per plot. Self-pollinated ears were hand-harvested and dried for 72 h with forced hot air. After drying, ears were stored away from light in burlap sacks at ambient winter temperatures in West Lafayette, IN, for up to four months until measurements could be taken.

Inbred lines that were sweet corn or popcorn, or with white, red or blue endosperm color were removed from the data set because the kernels have characteristics that interfere with comparison of color measurements. Red and blue lines have pericarp color due to anthocyanins that are unrelated to carotenoid content, and white lines have very little carotenoid content. Popcorn and sweet corn have different kernel shapes than dent corn that may alter reflectance. This removal process resulted in 1,769 yellow and orange inbreds from the Ames panel that were analyzed by colorimetry.

To quantify kernel color, a Konica Minolta CR-400 Chroma Meter was used. This instrument is also called a colorimeter by the manufacturer and is described to perform colorimetry (https://sensing.konicaminolta.asia/product/chroma-meter-cr-400/). We will use the term colorimeter and colorimetry henceforth. The color values L*, a*, b*, and hue (h) were measured. Chroma (C*) values were not provided by the colorimeter, thus this value was calculated according to the formula Chroma = (a* ² + b* ²) ¹/₂ (McLaren 1976). These measurements and values correspond to the CIELAB L*a*b* system and the L*C*h system mathematically derived from it. Colorimeter settings used the standard illuminant D65 and an observer angle of 2° during the measurements. Three well-filled maize ears per plot were measured, with five random positions on each ear used for colorimeter recordings. The colorimeter was calibrated relative to a white reference before beginning measurements, and again every 15 min while measurements were conducted. Measurement of an ear required approximately 30 sec.

**Phenotypic Data Analysis**

To identify and remove significant outliers, a mixed linear model was fitted for each kernel colorimeter trait in ASReml-R version 3.0 (Gilmour et al. 2009). The full model fitted to the data was the following:

\[ Y_{ijklmn} = \mu + \text{check}_i + \text{genotype}_j + \text{year}_k + \text{genotype} \times \text{year}_j + \text{set(year)}_{kl} + \text{row(year)}_{lm} + \text{block(set \times year)}_{klm} + e_{ijklmn} \]  

where \( Y_{ijklmn} \) is an individual phenotypic observation; \( \mu \) represents the grand mean; \( \text{check}_i \); \( \text{genotype}_j \); \( \text{year}_k \); \( \text{genotype} \times \text{year}_j \); \( \text{set(year)}_{kl} \); \( \text{row(year)}_{lm} \); and \( \text{block(set \times year)}_{klm} \) are fixed effects; and \( e_{ijklmn} \) is the residual (or random error term) for individual phenotypic observation \( n \). The residuals were assumed to be independent and identically distributed, normal random variables with mean zero and variance \( \sigma_e^2 \), that is, ~ \( N(0, \sigma_e^2) \). The Kenward-Roger approximation was applied to calculate degrees of freedom (Kenward and Roger 1997). With the exception of the grand mean and check term, all other terms were fitted as random effects according to ~ \( N(0, \sigma_i^2) \). Studentized deleted residuals (Neter et al. 1996) were then calculated, and observations determined to be significant outliers based on the Bonferroni correction (corresponding to \( \alpha = 0.05 \)) were removed. Plot-level averages were then calculated for each colorimeter trait.

For each given trait, the calculated 2012 and 2013 plot-level averages were used as the response variable in an iterative mixed linear model fitting procedure using the full model (Equation 1) in ASReml-R version 3.0 (Gilmour et al. 2009). The final, best-fit model for each trait was obtained by removing all random terms from the model that were not significant at \( \alpha = 0.05 \) in a likelihood ratio test (Littell et al. 2006). This final model was used to generate a best linear unbiased predictor (BLUP) for each genotype (Table S1).

Variance component estimates from the full model (Equation 1) were used for the estimation of heritability on a line-mean basis (Hung et al. 2012; Holland et al. 2003). Standard errors for these heritability estimates were calculated using the delta method (Holland et al. 2003). The Pearson’s correlation coefficient (r) between the BLUP values for each pair of colorimeter traits was calculated to assess the degree of their association (at \( \alpha = 0.05 \)), using the ‘cor’ function in R version 3.5.1 (R Core Team 2018).

Prior to conducting the GWAS, the Box-Cox power transformation (Box and Cox 1964) was used on the BLUP values for each trait to correct for unequal variance and non-normality of the residual error term (Table S2). The Box-Cox procedure was performed using the MASS package version 7.3-50 in R. Lambda values ranging from -2 to +2 were evaluated in increments of 0.5 to determine the optimal convenient lambda for each trait, which was then used for the transformation. A lambda value of ‘2’ (square transformation) was obtained for hue and \( L^* \), whereas a lambda value of ‘1’ (no transformation) was obtained for \( a^* \), \( b^* \), and \( C^* \).

**Genome-wide association study**

A GWAS was conducted for each of the five traits using the single-nucleotide polymorphism (SNP) data set developed using the genotyping-by-sequencing (GBS) platform for the Ames panel (Romay et al. 2013). The GBS marker data set used in this study consisted...
of partially imputed SNP genotypic data with B73 AGPv4 coordinates (ZeaGBSv27_publicSamples_imputedV5_AGPV4-161010.h5, available on CyVerse at http://datacommons.cyverse.org/browse/plant/home/shared/panzea/genotypes/GBS/v27). Additional quality filters were imposed to retain SNPs with a call rate greater than 70%, minor allele frequency (MAF) greater than 2%, and inbreeding coefficient greater than 80%, resulting in a final dataset of 268,006 high-quality SNPs. In addition, inbred lines with a call rate lower than 40% were excluded, given that missing genotype scores were still present in the SNP data set after partial imputation.

For each kernel colorimeter trait, the GWAS was conducted using a mixed linear model that included the population parameters previously determined (Zhang et al. 2010), using hypothesis testing to examine this data set for associations between the genotype scores of each of the 268,006 SNPs and BLUP values from the 1,651 experimental inbred lines having both genotypic and phenotypic data, including after the above-described quality control steps. The R package GAPIT, version 2017.08.18 (Lipka et al. 2012), was used to conduct this GWAS. To control for population structure and unequal relatedness, the mixed linear models that were fit in GWAS included principal components (PCs) (Price et al. 2006) and a kinship matrix based on VanRaden’s method 1 (VanRaden 2008) that was calculated using the full set of 268,006 partially imputed SNPs. Before performing the GWAS, the missing genotypes remaining for all SNP markers were imputed with a conservative, middle value, corresponding to a heterozygous state at that SNP. The Bayesian information criterion (Schwarz 1978) was used to determine the optimal number of PCs to include as covariates in the mixed linear model for each trait. The extent of phenotypic variation accounted for by the model (or coefficient of determination) was estimated with a likelihood-ratio-based $R^2$ statistic ($R^2_{LD}$) (Sun et al. 2010). The Benjamini–Hochberg procedure (Benjamini and Hochberg 1995) was used to control the false discovery rate (FDR) at 5% in the presence of multiple comparisons (hypothesis tests).

Pathway-level analysis

A set of 58 genes related to the biosynthesis and retention of carotenoids in maize was determined based on homology with known genes in Arabidopsis thaliana, and was previously used for a pathway-level analysis of carotenoid HPLC measurements in a small ($n = 201$) maize association panel (Owens et al. 2014). These same 58 genes, with the addition of $\zeta$-carotene isomerase ($\zeta$-iso) and homogentisate solanesyl transferase ($\text{sst}$), are referred to as pathway genes or a priori candidate genes in this study. Pathway-level analysis was used to reduce the number of association tests conducted, thus using a priori knowledge of the pathway to reduce the magnitude of the correction used to control the FDR at 5% (Califano et al. 2012; Owens et al. 2014). The set of 2,339 SNPs within ±50 kb of the coding regions of the 60 a priori candidate genes was used in pathway-level analysis. The interval of ±50 kb was a conservative estimate based on a previous finding in the Ames maize inbred panel of rapid decay of mean linkage disequilibrium in genic regions, reaching an average $r^2 = 0.2$ within 1 kb, with large variance due to population structure, among other factors (Romay et al. 2013).

Data availability

Phenotypes are provided in Tables S1 and S2 in the form of untransformed and transformed BLUPs. The GBS sequencing data are available at NCBI SRA (study accession number SRP021921). The SNP marker data are available on CyVerse as previously specified, and accession names are listed in Tables S1 and S2. Supplemental material available at FigShare: https://doi.org/10.25387/g3.7638590.

RESULTS

All of the colorimeter traits were highly heritable, with line-mean heritabilities ranging from 0.75 to 0.89 (Table 1). Hue values were positively correlated with $L^*$ ($r = 0.75$) and negatively correlated with $a^*$ ($r = -0.94$). Chroma and $b^*$ values were strongly positively correlated ($r = 0.99$) (Table 2). This correlation is likely due to $b^*$ values contributing most to Chroma (intensity of color), given the larger magnitude of $b^*$ relative to $a^*$ and the equal weighting of these two traits in the calculation of Chroma, whereas $a^*$ values corresponded more to hue (perceived color) in this data set.

A total of 27 unique SNPs were identified in GWAS for the five kernel colorimeter traits at an FDR-adjusted P-value of 5% (Table S3). Manhattan plots for each trait are presented in Figure S1. Associations were detected for two genes involved in the provision of substrate for carotenoid biosynthesis. A single SNP was detected within (i.e., in the coding region of) a gene encoding 1-deoxy-D-xylulose 5-phosphate synthase (dxs2), the first and committed step in the MEP pathway, with significant associations for $a^*$ and hue (Table 3). Two SNPs significantly associated with $a^*$ were detected within a gene encoding phytoene synthase ($y1$), the first and committed step in the biosynthesis of carotenoids.

Two genes in the core carotenoid pathway were also identified. Two significant SNP associations were detected for hue within the gene encoding lycopene $\varepsilon$-cyclase ($lcyE$), which affects the partitioning of

### Table 1

| Trait | Lines | Mean | Range | Std. Dev. | Estimate | Standard Error |
|-------|-------|------|-------|-----------|-----------|----------------|
| hue   | 1768  | 82.49| 61.78 – 93.08 | 4.52 | 0.89 | 0.01 |
| $L^*$ | 1768  | 67.19| 54.01 – 78.22  | 4.09 | 0.75 | 0.01 |
| $a^*$ | 1768  | 7.36 | -1.78 – 20.66   | 3.18 | 0.87 | 0.01 |
| $b^*$ | 1768  | 55.62| 36.54 – 74.36   | 6.22 | 0.80 | 0.01 |
| Chroma| 1769  | 56.08| 36.27 – 74.29   | 5.86 | 0.79 | 0.01 |

| Trait | BLUPs | Heritabilities |
|-------|-------|---------------|
| | | Estimate | Standard Error |
| hue | $L^*$ | $a^*$ | $b^*$ | Chroma |
| $r = 1.00$ | 0.75 | -0.94 | 0.54 | 0.44 |
| $r = 1.00$ | 1.00 | -0.68 | 0.58 | 0.52 |
| $r = 1.00$ | 0.54 | -0.28 | -0.18 |
| $r = 0.99$ | 1.00 | 0.99 |

All correlations were significant at $\alpha = 0.01$. 

### Table 2

Pearson’s correlation coefficients of untransformed BLUP values for five kernel colorimeter traits evaluated in the Ames maize inbred panel

| Trait | hue | $L^*$ | $a^*$ | $b^*$ | Chroma |
|-------|-----|-------|-------|-------|--------|
| $r$   | 1.00| 0.75  | -0.94 | 0.54  | 0.44   |
| $r$   | 1.00| 1.00  | -0.68 | 0.58  | 0.52   |
| $r$   | 0.54| 0.99  | -0.18 |
| $r$   | 0.99| 1.00  |

All correlations were significant at $\alpha = 0.01$. 

---

1. **Table 1** Means, ranges, and standard deviations (Std. Dev.) of untransformed BLUP values for five kernel colorimeter traits evaluated in the Ames maize inbred panel and estimated heritability on a line-mean basis across two years
A significant SNP associated with α was located near the gene encoding zeaxanthin epoxidase (zept1), approximately 25 kb downstream of the gene. Zeaxanthin epoxidase converts zeaxanthin to antheraxanthin and subsequently violaxanthin, all within the β-branch of the pathway.

Twenty-one SNPs having significant associations with one or more traits did not have an *a priori* candidate gene within the ±50 kb search space. These search spaces were subsequently examined, in case they contained other genes having plausible biological involvement with carotenoid color. Briefer analyses were subsequently examined, in case they yielded additional associations for colorimeter traits with all four of the carotenoid genes identified in GWAS: two SNPs in the coding region of *dxs2*, four SNPs in the coding region of *y1*, nine SNPs in the coding region of *lcyE*, and three SNPs proximal to *zept1* (Table 4, Table S4).

Additional associations were identified through pathway analysis in regions proximal to a number of genes not identified in GWAS. An association was found for *Chroma* in the vicinity of another gene that encodes DXS (*dxs3*, chr. 9). Two SNPs were significant for hue in the vicinity of 4-diphosphocytidyl-2C-methyl-D-erythritol synthase (*dmes1*, chr. 3), another gene in the MEP pathway. Within the core carotenoid pathway, two additional genes were identified for α*: 9-cis-lycopene β-cyclase (**lcyB**, *ps1*, *vp7*, chr. 5) and phytoene desaturase (*vp5*, chr. 1). Finally, a gene related to carotenoid cleavage, encoding 9-cis-epoxycarotenoid dioxygenase (NCED) (*vp14*, chr. 1), was identified for L*.

**DISCUSSION**

A colorimeter was used to quantify kernel color in a large, diverse maize inbred panel. Color visual scoring has shown effectiveness in biparental crosses, where only a few classes of kernel color are segregating (Chandler et al. 2013), but is not suitable or tractable for large diversity panels with continuous gradients of kernel color. The most significant association in this study was detected for a SNP in the coding region of *dxs2*—one of three genes in the maize genome encoding DXS, the first enzyme in the MEP pathway (Cordoba et al. 2011). Significant associations were also detected in the coding region of phytoene synthase

---

**Table 3** Carotenoid-related genes identified through genome-wide association study of five kernel colorimeter traits in the Ames maize inbred panel, and the most significant SNP for each trait-by-gene combination

| Gene ID    | Gene Trait SNP ID | Chr Position of SNP | P-value | FDR-adjusted P-value | MAF | R²LR | R²LR SNP |
|------------|-------------------|---------------------|---------|----------------------|-----|------|---------|
| Zm00001d003512 zept1 a’ | S2_44473748          | 2                   | 46,329,893 | 7.30E-07             | 2.44E-02 | 0.141 | 0.414 | 0.423 |
| Zm00001d003634 y1 a’ | S6_82020346          | 6                   | 85,064,521 | 2.65E-07             | 1.58E-02 | 0.061 | 0.414 | 0.423 |
| Zm00001d019060 dxs2 a’ | S7_14078791          | 7                   | 14,495,640 | 3.25E-10             | 8.72E-05 | 0.050 | 0.414 | 0.428 |
| Zm00001d019060 dxs2 hue | S7_14078791          | 7                   | 14,495,640 | 4.74E-08             | 6.35E-03 | 0.050 | 0.493 | 0.502 |
| Zm00001d011210 lcyE hue | S8_138888278          | 8                   | 143,026,247 | 2.52E-07             | 2.25E-02 | 0.416 | 0.493 | 0.501 |

**Table 4** Most significant SNP for each trait-by-gene combination within 50 kb of carotenoid-related genes identified through pathway-level association analyses of five kernel colorimeter traits in the Ames maize inbred panel

| Gene ID    | Gene Trait SNP ID | Chr Position of SNP | P-value | FDR-adjusted P-value | MAF | R²LR | R²LR SNP |
|------------|-------------------|---------------------|---------|----------------------|-----|------|---------|
| Zm00001d027936 vp5 a’ | S1_17625344          | 1                   | 17,930,201 | 6.16E-05             | 1.02E-02 | 0.185 | 0.414 | 0.420 |
| Zm00001d033222 vp14 L’ | S1_250915992         | 1                   | 255,044,348 | 1.59E-05             | 3.68E-02 | 0.161 | 0.336 | 0.344 |
| Zm00001d003513 zept1 a’ | S2_44443991          | 2                   | 46,299,837 | 9.89E-06             | 3.81E-03 | 0.401 | 0.414 | 0.421 |
| Zm00001d003513 zept1 hue | S2_44443991          | 2                   | 46,299,837 | 1.47E-04             | 2.83E-02 | 0.401 | 0.493 | 0.498 |
| Zm00001d042584 dmes1 hue | S3_170107623         | 3                   | 172,731,046 | 3.47E-04             | 4.72E-02 | 0.180 | 0.493 | 0.497 |
| Zm00001d015651 ps1 a’ | S5_100735811          | 5                   | 103,264,914 | 2.53E-04             | 3.45E-02 | 0.096 | 0.414 | 0.419 |
| Zm00001d036345 y1 b’ | S6_82018091           | 6                   | 85,062,266 | 2.85E-05             | 3.30E-02 | 0.062 | 0.288 | 0.295 |
| Zm00001d036345 y1 Chroma | S6_82018091          | 6                   | 85,062,266 | 1.38E-05             | 2.96E-02 | 0.062 | 0.253 | 0.262 |
| Zm00001d036345 y1 a’ | S6_82019628           | 6                   | 85,063,803 | 1.83E-06             | 1.06E-03 | 0.040 | 0.414 | 0.422 |
| Zm00001d036345 y1 hue | S6_82020346           | 6                   | 85,064,521 | 3.12E-04             | 4.56E-02 | 0.061 | 0.493 | 0.497 |
| Zm00001d019060 dxs2 a’ | S7_14078791           | 7                   | 14,495,640 | 3.25E-10             | 7.53E-07 | 0.050 | 0.414 | 0.428 |
| Zm00001d019060 dxs2 hue | S7_14078791           | 7                   | 14,495,640 | 4.74E-08             | 1.10E-04 | 0.050 | 0.493 | 0.502 |
| Zm00001d011210 lcyE a’ | S8_138882509          | 8                   | 143,020,478 | 1.87E-05             | 3.93E-03 | 0.397 | 0.414 | 0.420 |
| Zm00001d011210 lcyE hue | S8_138882509          | 8                   | 143,020,478 | 8.60E-06             | 2.49E-03 | 0.397 | 0.493 | 0.499 |
| Zm00001d011210 lcyE b’ | S8_138887654          | 8                   | 143,024,723 | 6.91E-06             | 1.60E-02 | 0.366 | 0.288 | 0.297 |
| Zm00001d045538 dxs3 Chroma | S9_20472920          | 9                   | 20,252,034 | 1.65E-05             | 4.75E-02 | 0.179 | 0.253 | 0.261 |

**Gene ID**: Gene designation and position of SNPs from B73 RefGen_v4 (www.maizegdb.org); **Gene**: Annotated gene containing SNP, except zept1, which is 25 kb upstream of the SNP; **Position of SNP**: Genomic position (bp) of the SNP from B73 Refgen_v4; **FDR-adjusted P-value**: False discovery rate adjusted P-value; **MAF**: Minor-allele frequency; **R²LR**: R² likelihood ratio value of model without SNP; **R²LR SNP**: R² likelihood ratio value of model with SNP.
(y1), a gene that controls the first committed step in carotenoid biosynthesis (Buckner et al. 1996; Cunningham and Gantt 1998; DellaPenna and Pogson 2006).

Although joint linkage analysis of visual color score data detected a QTL in the vicinity of y1 (Chandler et al. 2013), neither y1 nor dss2 were strong hits in a genome-wide association study of HPLC carotenoid data in 201 inbreds with yellow to orange kernel color from the Goodman-Buckler diversity panel (Owens et al. 2014). In the present study of kernel color in a large association panel of 1,651 inbreds, significant associations were detected in the coding regions of both of these genes. PSY has been considered to be the key enzyme limiting carotenoid accumulation in maize endosperm (Zhu et al. 2008). The identification of dss2 and y1 in this study indicates that genetic variation at these loci is associated with kernel color, likely due to the role of these genes in substrate provision for the biosynthesis of pigmented carotenoids. These genes merit further examination given that dss2 and y1 respectively encode the first and committed steps in the MEP pathway and core carotenoid pathway, and showed the most significant statistical associations in this study. In particular, investigation of the main effects and any interaction effects of these two genes in maize, as well as their expression dynamics through kernel development and upon the expression or knockdown of one or both genes, may provide further insight into the extent to which their association with kernel color (and potentially carotenoids) is separate vs. coordinated.

Associations in the regions of lcyE and zep1—genes affecting flux within and through the core carotenoid pathway—were identified both in this study of kernel color and in the prior study of carotenoid HPLC values in the Goodman-Buckler panel (Owens et al. 2014). Notably, signals in the vicinity of three of the genes identified in our GWAS—lcyE, zep1, and y1—were also detected in a previous joint-linkage analysis of visual scores for gradation in orange kernel color in 10 families of the U.S. maize NAM population (Chandler et al. 2013). Carotenoid compounds in the α- vs. β-branches have different spectral properties that influence color, due to differing numbers of double bonds in their structures. Specifically, the β-branch compounds (β-carotene, β-cryptoxanthin, and zeaxanthin) have 11 conjugated double bonds and correspondingly have lower a' values and higher b' values than α-carotene and lutein, which have 10 conjugated double bonds (Meléndez-Martinez et al. 2007; Khoo et al. 2011). Thus, a shift in the relative concentrations of these compounds has the potential to affect kernel color.

For lcyE, encoding a protein that acts at the key pathway branch point, associations were indeed seen in the Goodman-Buckler panel for two ratio traits (β-branch to α-carotenoid carotenoids, and β-branch to α-branch xanthophylls) as well as lutein, zeaxanthin, total α-xanthophylls, and total β-xanthophylls. An allele of lcyE with reduced expression was found to result in the formation of fewer ε-rings and a reduction in α-branch compounds relative to β-branch compounds (Harjes et al. 2008). Similarly to lcyE, associations with zep1—encoding a protein that acts within the β-pathway branch—were seen in the Goodman-Buckler panel for the ratio trait of β-branch to α-branch xanthophylls, as well as zeaxanthin and total β-xanthophylls.

Taken together, the identification of dss2 and y1 (genes involved in overall substrate provision) in the present study suggests that kernel color can be utilized to select for greater carotenoid abundance in general. However, the simultaneous identification of lcyE and zep1 (genes involved in carotenoid composition) suggests that the relative abundance of individual carotenoid compounds is likely to also be affected when selecting on kernel color. Therefore, the levels of individual carotenoid compounds will need to be monitored when colorimetry is applied as an early selection tool for lines having favorable orange color, to ensure that the favorable genetic variants needed for the maintenance or improvement of provitamin A levels are also retained. For example, the concentrations of the more abundant provitamin A carotenoids in maize grain, β-carotene and β-cryptoxanthin, might be increased simultaneously with orange kernel color if substrate were to be modulated via lcyE to flow preferentially through the β-branch of the pathway. Alternatively or in addition, favorable alleles of the gene encoding β-carotene hydrolase (crtRB1), which converts β-carotene to β-cryptoxanthin to zeaxanthin, could be selected that favor accumulation and retention of these provitamin A compounds while also producing sufficient zeaxanthin to obtain the vivid orange color.

While there are many cytochrome P450s in the maize genome, the high level of homology between the product of GRMZM2G063663 and CYP14, which acts within the α-branch of the carotenoid pathway, suggests that this gene is a candidate for further examination. Regarding isopentenyl transferase (IPT), its activity has been found in maize to affect the distribution of aleurone vs. starchy endosperm layers (Geisler-Lee and Gallie 2005). Certain aleurone-deficient mutants have been found to be deficient in carotenoids, and it has been suggested that there may be some functional connection between aleurone differentiation and carotenoid biosynthesis (reviewed in Gontarek and Bercat 2017). The finding of signals proximal to ipt10 in this study for kernel color suggests a potential genetic target for the further investigation of that hypothesis. Finally, the product of enolase—phosphoenolpyruvate (PEP)—has many potential metabolic routes. Nevertheless, the action of enolase resides only two steps prior to that of DXS (which takes pyruvate as one of its substrates), and PEP is an important precursor for isoprenoid biosynthesis. An engineering strategy in E. coli that increased PEP concentrations was found to elevate levels of lycopene, the carotenoid compound that sits at the pathway branch point (Zhang et al. 2013). While enolase1 may have underlaid associations with kernel color in this GWAS, it may not be a viable breeding target given the relatively higher likelihood of complex and/or unfavorable pleiotropic effects within central metabolism. There are other physical properties of the kernel—such as pericarp thickness or kernel flintiness or relative density (Lozano-Alejo et al. 2007)—which may affect perceived color and merit further examination.

The pathway-level analysis conducted in this study revealed a number of additional genes significantly associated with kernel color. Notably, an association with dss3 suggests that this gene, in addition to dss2, may play a role in the accumulation of carotenoids in the maize kernel. An association was found with dnes2, which encodes 4-diphosphocytidyl-2C-methyl-D-erythritol synthase, the third step in the MEP pathway. The gene encoding this enzyme in A. thaliana, present in a single copy and termed MCT, has been found along with certain other MEP pathway genes to have very low seed expression levels in certain developmental stages, in a manner that may be limiting to carotenoid biosynthesis (Meier et al. 2011). In this study, the associations with MEP pathway genes are an indication that the genetic control of the provision of IPP, a precursor for biosynthesis of carotenoids and other isoprenoids, is relevant to kernel color.

Three genes underlying classical viviparous maize mutants were identified in this study: vp5, encoding PDS (Hable et al. 1998); vp7, encoding LCYB (Singh et al. 2003); and vp14, encoding NCED (Tan et al. 1997). These three genes were previously recognized as Class Two viviparous mutants, which in addition to viviparous (precocious germination) exhibit altered endosperm and seedling color due to effects on carotenoid and chlorophyll biosynthesis (Robertson 1955). These three mutants have also been found to be deficient in ABA (McCarty 1995;
Schwartz et al. 1997). The action of PDS and LCYB takes place prior to and coincident with the pathway branch point, respectively. The two corresponding mutants are also deficient in carotenoids (McCarty 1995), which would tend to affect kernel color if the pigmented carotenoids are among those depleted. NCED acts within the β-pathway branch, cleaving 9-cis-xanthophylls to xanthoxin (Tan et al. 1997), which is then converted to ABA. The vp14 mutant was found to have reduced levels of zeaxanthin compared to wild type, though levels of the immediate substrates of NCED were unaffected (Tan et al. 1997). Given the finding of an effect on zeaxanthin levels, and the general action of NCED in the portion of the pathway corresponding to pigmented β-branched carotenoids and their derivatives, the association of genetic variation at vp14 with kernel color is not entirely surprising.

Another cleavage enzyme, CCD—encoded by one or more copies of ccd1 within the White Cap (Wc) locus in maize (Tan et al. 2017)—was not detected as being associated with natural variation in this study. The Wc locus was created in some maize accessions by a macrotransposon insertion, with subsequent tandem duplications resulting in the amplification of ccd1 copy number in a subset of those accessions, and has been found to impact endosperm color through the degradation of carotenoids by CCD. Notably, the Wc locus was likely identified in the previously conducted analysis of visual scores for gradation in orange kernel color in 10 U.S. maize NAM families. While the ccd1 progenitor locus (Ccd1r) was not contained in the QTL support interval identified on chromosome 9 (149.54 to 151.48 Mb, AGP v2), the macrotransposon insertion that created Wc was subsequently characterized in Tan et al. (2017), and appears to have been included in the interval. This QTL putatively corresponding to Wc was only significant in two of the 10 NAM families analyzed (Chandler et al. 2013), suggesting the possibility of rare variation at the Wc locus which may have precluded its identification in the present study. Additionally, given the tandem duplications inherent to Wc in some accessions, potentially informative paralogous SNP markers in this region may have been excluded in the SNP filtering process in the present study. Alternatively, the localization of variation relating to CCD may have been dispersed at the genetic level among a varying number of ccd1 copies within Wc (in addition to the Ccd1r progenitor locus itself). This dispersion could present particular difficulties for the detection of genetic signal in the presence of low SNP coverage and/or rare variation. Finally, given that only lines with yellow to orange endosperm were analyzed in this study, it could be that the variation in ccd1 copy number was too constrained (with yellow-endosperm lines being on the lower end of the dynamic range in copy number; Tan et al. 2017) for a genetic association with loci encoding CCD to be present and/or identified in this panel.

Notably, dxs2 and lcyE having been detected in association with hue at a genome-wide level, along with other genes in the pathway-level analysis, suggests that the allelic state at each of these loci has been associated with natural variation in perceived kernel color. Hue angle is measured counterclockwise from the +a° axis (at 0°), which corresponds to pure red, with the +b° axis (at 90°) corresponding to pure yellow. The hue angles observed in this study ranged from 61.78 to 93.08° (Table 1). Given this observed range, selecting for an allele that tends to decrease the hue angle could be expected to shift the average perceived kernel color in the direction of pure red (at 0°), which would also correspond in this case to a perception of more orangeness.

Further studies are needed to determine whether natural variation at the loci identified in these analyses corresponds to differences in transcription levels, post-translational regulation, and/or enzyme activity. A GWAS using kernel color phenotypes and HPLC-based carotenoid values for the same set of materials may enable the identification of alleles that are favorable for kernel color as well as carotenoid composition and concentration. An increasing knowledge of the genetic mechanisms affecting kernel color, and the potential relationships between color values and carotenoid values, will be useful in coordinating breeding efforts to improve both sets of phenotypes. Establishing optimal ranges for each colorimeter trait for use in a selection index could provide a useful and inexpensive breeding tool, particularly to screen for kernel color and total carotenoid levels in the early stages of breeding. Some of the evaluation, selection, and elimination could potentially be done while the ears are still on the plants, or in a harvest pile at the end of a nursery row. This would save labor and reduce handling of non-selected ears. Selection of favorable alleles of the loci detected in this study, particularly y1 and dxs2, in conjunction with the previously established alleles of lcyE and ctrlR1, provide a logical and promising strategy for the rapid development of provitamin A-dense maize lines that also produce a recognizable and desirable orange kernel color.

**Acknowledgments**

This research was supported by the National Science Foundation (NSF) IOS-0922493 (TRR) and IOS-1546657 (MAG), Harvest Plus (TRR, MAG), Purdue Patterson Chair funds (TRR), Cornell University startup funds (MAG), and by a USDA National Needs Fellowship (CHD). We thank Jerry Chandler and Chris Hoagland for assistance with fieldwork and seed handling.

**Literature Cited**

Amorim, E. P., A. D. Vilarinhos, K. O. Cohen, V. B. O. Amorim, J. A. dos Santos-Serejo et al., 2009 Genetic diversity of carotenoid-rich bananas evaluated by Diversity Arrays Technology (DAFT). Genet. Mol. Biol. 32: 96–103. https://doi.org/10.1590/S1415-47572009003000024

Anderson, C. M., E. K. Cannon, J. L. Portwood, J. M. Gardiner, L. C. Harper et al., 2016 MaizeGDB update: new tools, data and interface for the maize model organism database. Nucleic Acids Res. 44: D1195–D1201. https://doi.org/10.1093/nar/gkv1007

Auldridge, M. E., D. R. McCarty, and H. J. Klee, 2006 Plant carotenoid cleavage oxygenases and their apocarotenoid products. Curr. Opin. Plant Biol. 9: 315–321. https://doi.org/10.1016/j.pbi.2006.03.005

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

Berger-Schunn, A., 1994 Practical Color Measurement: A Primer for the Beginner, A Reminder for the Expert, Wiley, New York.

Blanco, A., P. Colasunno, A. Gadaleta, G. Mangini, A. Schiavulli et al., 2011 Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat. J. Cereal Sci. 54: 255–264. https://doi.org/10.1016/j.jcs.2011.07.002

Bouis, H., and J. Hunt, 1999 Linking food and nutrition security: past lessons and future opportunities. Asian devel. Rev. 17: 168–213.

Bouis, H. E., and R. M. Welch, 2010 Biofortification—a sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. Crop Sci. 50: S20–S32. https://doi.org/10.2135/cropsci2009.09.0531

Box, G. E. P., and D. R. Cox, 1964 An analysis of transformations. J. Roy. Stat. Soc. B. Met. 26: 211–252.

Buckner, B., P. S. Miguel, D. Janick-Buckner, and J. L. Bennetzen, 1996 The y1 gene of maize codes for phytoene synthase. Genetics 143: 479–488.

Burt, A. J., C. R. Maingey, M. P. Smid, B. J. Shelp, and E. A. Lee, 2011 Allele mining of exotic maize germplasm to enhance macular carotenoids. Crop Sci. 51: 991–1004. https://doi.org/10.2135/cropsci2010.06.0335

Califano, A., A. J. Butte, S. Friend, T. Ideker, and E. Schadt, 2012 Leveraging models of cell regulation and GWAS data in integrative network-based association studies. Nat. Genet. 44: 841–847. https://doi.org/10.1038/ng.2355

Carvalho, L. J. C. B., M. A. V. Agustini, J. V. Anderson, E. A. Vieira, C. R. B. de Souza et al., 2016 Natural variation in expression of genes...
Owens, B. F., A. E. Lipka, M. Magallanes-Lundback, T. Tiede, C. H. Diepenbrock 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. https://doi.org/10.1534 genetics.114.169979

Pixley, K., N. Palacios Rojas, R. Babu, R. Mutale, R. Surles et al., 2013 Biofortification of maize with provitamin A carotenoids, pp. 271–292 in Carotenoids in Human Health, Nutrition and Health, edited by Tanumihardjo, S. A. Springer Science and Business Media, New York. https://doi.org/10.1007/978-1-62703-203-2_17

Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick et al., 2006 Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38: 904–909. https://doi.org/10.1038/ng1847

Quim, M., A. Stein, and J. V. Meenan, 2007 Economics of biofortification. Agric. Econ. 37: 119–133. https://doi.org/10.1111/j.1574-0862.2007.00239.x

R Core Team, 2018 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robertson, D. S., 1955 The genetics of vivipary in maize. Genetics 40: 745–760.

Rodríguez-Concepción, M., and A. Boronat, 2002 Elucidation of the methylenetetrol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiol. 130: 1079–1089. https://doi.org/10.1104/pp.007138

Rodríguez-Concepción, M., N. Campos, A. Ferrer, and A. Boronat, 2013 Biosynthesis of isoprenoid precursors in Arabidopsis, pp. 439–456 in Isoprenoid Synthesis in Plants and Microorganisms, edited by Bach, T. J., and M. Rohmer. Springer, New York.

Romay, M. C., M. J. Millard, J. C. Glaubitz, J. A. Peiffer, K. L. Swarts et al., 2013 Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14: R55. https://doi.org/10.1186/gb-2013-14-6-r55

Schwartz, S. H., B. C. Tan, D. A. Gage, J. A. Zeevaart, and D. R. McCarty, 1997 Specific oxidative cleavage of carotenoids by VP14 of maize. Science 276: 1872–1874. https://doi.org/10.1126/science.276.5320.1872

Schwartz, S. H., X. Qin, and J. A. D. Zeevaart, 2001 Characterization of a novel carotenoid cleavage dioxygenase from plants. J. Biol. Chem. 276: 25208–25211. https://doi.org/10.1074/jbc.M102146200

Schwarz, G., 1978 Estimating the dimension of a model. Ann. Stat. 6: 461–464. https://doi.org/10.1214/aos/1176344136

Simpangwe, E., T. Dhlilwayo, M. Palenberg, V. Taleon, E. Birol et al., 2017 Orange maize in Zambia: crop development and delivery experience. Afr. J. Food Agric. Nutr. Dev. 17: 11973–11999. https://doi.org/10.18697/afrfund.78.8HarvestPlus08

Singh, M., P. E. Lewis, K. Hardeman, L. Bai, J. K. C. Rose et al., 2003 Activator mutagenesis of the pink scutellum/liviparous7 locus of maize. Plant Cell 15: 874–884. https://doi.org/10.1105/tpc.1012049

Stevens, G. A., J. E. Bennett, Q. Hennoçq, Y. Lu, L. M. De-Regil et al., 2015 Trends and mortality effects of vitamin A deficiency in children in 138 low-income and middle-income countries between 1991 and 2013: a pooled analysis of population-based surveys. Lancet Glob. Health 3: e528–e536. https://doi.org/10.1016/s2214-109x(15)00039-x

Sun, G., C. Zhu, M. H. Kramer, S. S. Yang, W. Song et al., 2010 Variation explained in mixed-model association mapping. Heredity 105: 333–340. https://doi.org/10.1038/hdy.2010.11

Sun, Z., J. Hans, M. H. Walter, R. Matusova, J. Beckwilder et al., 2008 Cloning and characterisation of a maize carotenoid cleavage dioxygenase (ZmCCD1) and its involvement in the biosynthesis of apocarotenoids with various roles in mutualistic and parasitic interactions. Planta 228: 789–801. https://doi.org/10.1007/s00425-008-0781-6

Takahata, Y., T. Noda, and T. Nagata, 1993 HPLC determination of β-carotene content of sweet potato cultivars and its relationship with color values. Japan. J. Breed. 43: 421–427. https://doi.org/10.1270/jsbbs1951.43.421

Tan, B. C., J. C. Guan, S. Ding, S. Wu, J. W. Saunders et al., 2017 Structure and origin of the White Cap locus and its role in evolution of grain color in maize. Genetics 206: 135–150. https://doi.org/10.1534 genetics.116.198911

Tan, B. C., S. H. Schwartz, J. A. D. Zeevaart, and D. R. McCarty, 1997 Genetic control of abscisic acid biosynthesis in maize. Proc. Natl. Acad. Sci. USA 94: 12235–12240. https://doi.org/10.1073/pnas.94.22.12235

VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci. 91: 4414–4423. https://doi.org/10.3168/jds.2007-0980

Valabhaneni, R., L. M. Bradbury, and E. T. Wurtzel, 2010 The carotenoid dioxygenase gene family in maize, sorghum, and rice. Arch. Biochem. Biophys. 504: 104–111. https://doi.org/10.1016/j.jb.2010.07.019

Vogel, J. T., B. C. Tan, D. R. McCarty, and H. J. Klee, 2008 The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. J. Biol. Chem. 283: 11364–11373. https://doi.org/10.1074/jbc.M71016200

Vranova, E., D. Coman, and W. Gruissem, 2013 Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annu. Rev. Plant Biol. 64: 665–700. https://doi.org/10.1146/annurev-arplant-050312-120116

Weber, E. J., 1987 Carotenoids and tococols of corn grain determined by HPLC. J. Am. Oil Chem. Soc. 64: 1129–1134. https://doi.org/10.1007/BF02612988

Welch, R. M., and R. D. Graham, 1999 A new paradigm for world agriculture: meeting human needs: Productive, sustainable, nutritious. Field Crops Res. 60: 1–10. https://doi.org/10.1016/S0378-4290(98)00129-4

Yan, J., C. B. Kandianis, C. E. Harjes, L. Bai, E. Kim et al., 2010 Rare genetic variation at Zea mays crtBBI increases β-carotene in maize grain. Nat. Genet. 42: 322–327. https://doi.org/10.1038/ng.551

Zhang, C., X. Chen, R. Zou, K. Zhou, G. Stephanopoulos et al., 2013 Combining genotype improvement and statistical media optimization for isoprenoid production in E. coli. PLoS One 8: e75164. https://doi.org/10.1371/journal.pone.0075164

Zhang, Z., E. Ersoz, C. Q. Lai, R. J. Todhunter, H. K. Tiwari et al., 2010 Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. 42: 355–360. https://doi.org/10.1038/NG.546

Zhao, Y., H. Sun, Y. Wang, Y. Pu, F. Kong et al., 2013 QTL mapping for the color, carotenoids and polyphenol oxidase activity of flour in recombinant inbred lines of wheat. Aust. J. Crop Sci. 7: 328–337.

Zhu, C., S. Naqvi, J. Breitenbach, G. Sandmann, P. Christou et al., 2008 Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. Proc. Natl. Acad. Sci. USA 105: 18232–18237. https://doi.org/10.1073/pnas.0809737105

Communicating editor: R. Wisser