Published in final edited form as:
Science. 2008 January 18; 319(5861): 330–333. doi:10.1126/science.1150255.

Natural Genetic Variation in *Lycopene Epsilon Cyclase* Tapped for Maize Biofortification

Carlos E. Harjes[1,*], **Torbert R. Rocheford**[2,†], **Ling Bai**[3], **Thomas P. Brutnell**[3], **Catherine Bermudez Kandianis**[2], **Stephen G. Sowinski**[4], **Ann E. Stapleton**[5], **Ratnakar Vallabhaneni**[6,7], **Mark Williams**[4], **Eleanore T. Wurtzel**[6,7], **Jianbing Yan**[8], and **Edward S. Buckler**[1,9,10,†]

[1]Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853, USA.

[2]Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA.

[3]Boyce Thompson Institute, Ithaca, NY 14853, USA.

[4]DuPont Crop Genetics Research, Wilmington, DE 19880, USA.

[5]Department of Biology and Marine Biology, University of North Carolina, Wilmington, NC 28403, USA.

[6]Department of Biological Sciences, Lehman College, City University of New York (CUNY), Bronx, NY 10468, USA.

[7]The Graduate School and University Center, City University of New York (CUNY), New York, NY 10016, USA.

[8]International Maize and Wheat Improvement Center (CIMMYT), Apartado Postal 6-64, 06600 Mexico, DF, Mexico.

[9]U.S. Department of Agriculture, Agricultural Research Service, Plant, Soil and Nutrition Research Unit, Ithaca, NY 14853, USA.

[10]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY 14853, USA.

Abstract

Dietary vitamin A deficiency causes eye disease in 40 million children each year and places 140 to 250 million at risk for health disorders. Many children in sub-Saharan Africa subsist on maize-based diets. Maize displays considerable natural variation for carotenoid composition, including vitamin A precursors α-carotene, β-carotene, and β-cryptoxanthin. Through association analysis, linkage mapping, expression analysis, and mutagenesis, we show that variation at the *lycopene epsilon cyclase* (*lcyE*) locus alters flux down α-carotene versus β-carotene branches of the carotenoid pathway. Four natural *lcyE* polymorphisms explained 58% of the variation in these two branches and a threefold difference in provitamin A compounds. Selection of favorable *lcyE* alleles with inexpensive molecular markers will now enable developing-country breeders to more effectively produce maize grain with higher provitamin A levels.

†To whom correspondence should be addressed. esb33@cornell.edu (E.S.B); trochefo@uiuc.edu (T.R.R.).
*Present address: Monsanto Company, Leesburg, GA 30903, USA.

Supporting Online Material www.sciencemag.org/cgi/content/full/319/5861/330/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S5
References

Maize is the dominant subsistence crop in much of sub-Saharan Africa and the Americas, where between 17 and 30% of children under age of 5 are vitamin A–deficient. This results in xerophthalmia (progressive blindness), increased infant morbidity and mortality, and depressed immunological responses (1). Vitamin A deficiency starts with inadequate provitamin A or vitamin A content or bioavailability in foods and is exacerbated by disease-induced malabsorption.

Diet diversification, food fortification, and supplementation (2–4) have all been used to combat dietary micronutrient deficiencies. Ideally, all children would have access to a varied diet rich in fruits and vegetables, but diet diversification is often limited by crop seasonality, expense, and low bioavailability of green leafy plant carotenoids (5,6). Poor infrastructure in developing countries has limited widespread use of direct vitamin supplementation. Perhaps the most feasible approach to eradicating death and disease caused by dietary deficiencies is biofortification, a process by which staple crops are purposefully bred for higher nutritional density (7,8). Although biofortified foods can potentially be an inexpensive, locally adaptable, and long-term solution to diet deficiencies, cultural preferences may limit their acceptance. This may be particularly true for those crops where transgenics are the only alternative to boost provitamin A content, given limited acceptance of genetically modified organisms in developing countries.

Carotenoids are derived from the isoprenoid biosynthetic pathway and are precursors of the plant hormone abscisic acid and of other apocarotenoids (9). The first committed step of this pathway [as recently revised (10)] is formation of phytoene from geranylgeranyl diphosphate by phytoene synthase (*y1/psy1*) (Fig. 1) (11). Recent studies in maize suggest that the *psy1* locus has been the target of a selective sweep following selection for endosperm-accumulating carotenoids and shift from white to yellow kernels (12). The first branch point of this pathway (Fig. 1) occurs at cyclization of lycopene where action of lycopene beta cyclase (LCYB) at both ends of linear lycopene produces a molecule with two β rings. Alternatively, the coaction of LCYB and lycopene epsilon cyclase (LCYE) generates a β,ε-carotene that is a precursor to lutein (13). Relative activities of LCYB and LCYE are hypothesized to regulate the proportion of carotenes directed to each branch of this pathway (13–15). Indeed, transgenic manipulations of LCYE expression in *Arabidopsis*, potato, and *Brassica* increase the pool of β ring–containing carotenes and xanthophylls (13,16–18).

Maize exhibits considerable natural variation for kernel carotenoids, with some lines accumulating as much as 66 μg/g. The predominant carotenoids in maize kernels, in decreasing order of concentration, are lutein, zeaxanthin, β-carotene, β-cryptoxanthin, and α-carotene. β-Carotene contains two provitamin A structures (two nonhydroxylated β-ionone rings) and β-cryptoxanthin and α-carotene one each (single nonhydroxylated β-ionone ring). Among lines included in our diverse maize panel, β-carotene levels reached 13.6 μg/g. However, most yellow maize grown and consumed throughout the world has only 0.5 to 1.5 μg/g β-carotene. Comparisons between β-carotene and total carotenoids with grain color (scaled according to shade of yellow) revealed poor correlations with low R^2 values (Fig. 2), which indicated that marker-assisted selection (MAS) may prove much more efficient than selection based on color alone.

To dissect the phenotypic diversity, we used an association-mapping approach that exploits the genetic diversity of maize to provide resolution within 2000 base pairs (bp) (19–21). In the context of plant breeding, this has the added advantage of identifying the most favorable allele within a diverse genetic background, which provides the necessary genotypic information to facilitate the design of efficient maize introgression and selection schemes throughout the world. We complemented the association mapping with linkage mapping to evaluate the effects

in a genetically less complex background and with a mutagenesis program to isolate novel allelic variation within an elite near-isogenic background.

To evaluate functional diversity (Fig. 1), eight candidate genes representing select members of gene families encoding biosynthetic enzymes of the carotenoid pathway were sampled across a diverse panel of 288 maize lines, of which 204 were yellow. Subsets of yellow lines were grown in four different years and surveyed for whole-kernel carotenoids by high-performance liquid chromatography (HPLC). The yellow lines averaged 23 μg/g for total carotenoids (range 5.5 to 66.0 μg/g) and 1.7 μg/g for β-carotene (range 0.06 to 13.6 μg/g).

For association analysis, we used a mixed-model approach that controlled for complex population and pedigree relationships (22). Among our current sampling of candidate genes, *lycopene epsilon cyclase* (*lcyE*) (14) had the largest effect on partitioning the two branches of carotenoids and, consequently, on β-carotene and β-cryptoxanthin content. In maize, the single-copy *lcyE* gene consists of 10 exons spanning 3640 bp (Fig. 3). After initial association and screening for polymorphisms in key haplotypes, four regions were selected and scored across the entire panel. On the basis of the position of LCYE in the biochemical pathway, we predicted that the ratio of the sum of kernel carotenoids from each pathway branch would form the strongest association. Indeed, this was confirmed (Table 1), with the strength of the association confirming that *lcyE* plays a key role in controlling this ratio. Correspondingly, levels of predominant provitamin A compounds bcarotene and β-cryptoxanthin were also highly associated with *lcyE*.

Subsequent haplotype analysis revealed several probable causative polymorphisms for the ratio of α- and β-carotene branches for the 2003 field season (table S1). A large promoter indel and an amino acid substitution in exon 1 explain most of the variation (R^2 =36%; n = 135; P = 1.27 × 10^{-12}) with a 5.2-fold effect. A second indel in the 3′ UTR also has a significant 3.3-fold effect and contributes to variation not explained by the promoter polymorphism (type III SS; P =1.9 × 10^{-4}). The fourth significant polymorphism at position 2238 in intron 4 was associated with a 2.5-fold effect (type III SS; P = 0.0003). The overall, four-term model explains 58% of the variation (P = 9.2 × 10^{-17}). These significant polymorphisms exhibit some linkage disequilibrium (LD), and only nine haplotypic classes exist in our sample, which limits full differentiation of the effects of each polymorphism. Overall, there is a ninefold difference between two of the more differentiated haplotype classes, and sixfold between two more common haplotypes (table S2). There was a threefold increase in the proportion of β-carotene and β-cryptoxanthin between the common haplotypes. Verification of these results was provided by significant associations in subsequent field seasons (Table 1).

Expression analysis indicated that *lcyE* is preferentially expressed in the endosperm relative to the embryo (fig. S1). Expression profiling of kernels at 15 and 20 days after pollination (DAP) indicated expression levels correlated well with the ratio of carotenoids from each pathway branch, explaining 70 to 76% of the variance. Lines with transposon insertions near the start site had much lower expression levels [in 15 DAP and 20 DAP lower bya factor of 3.7 and 13, respectively (fig. S2)]. The 3′ indel may also have expression effects, but our statistical tests lacked the power to confirm this hypothesis. A quantitative trait locus (QTL) experiment that examined segregation of B73-Mo17 alleles in leaves found significant variation in the cis-regulation of *lcyE* expression, along with several other regions that also contribute to expression level control of *lcyE* (fig. S3).

In a previous study, three major QTL were identified for accumulation of carotenoids in maize (23). Two of these QTL colocalized with *y1* and *zeta carotene desaturase* (*zds*); the third QTL mapped to a region without a candidate gene. We mapped *lcyE* to chromosome 8 bin 5, near marker bnlg1599, and it colocalized with this previously undetermined QTL. This QTL showed

significant effects for modification of the ratio of α to β branch carotenoids [logarithm of the odds ratio for linkage (or lod) score of 34.05; R^2 54.4%] and explained 31.7% of the variation for lutein (lod 16.5). The magnitude of effects was not as large as in association or mutagenesis analysis. However, this biparental QTL population only segregated for the amino substitution (at codon 216) and a modest promoter polymorphism and does not segregate for the 3′ polymorphism. Notably, this QTL was not significant for total carotenoids, which further supports the conclusion that variation within *lcyE* gene underlies this QTL for carotenoid composition.

To confirm association and QTL results, mutagenesis induced by ethane methyl sulfonate (EMS) was conducted to isolate additional alleles of *lcyE*. Two M_2 ears of inbred Q×47 segregated for a distinct change in endosperm color from yellow to orange, with orange recessive to yellow (these color changes were apparent in the inbred isogenic background, but not in diverse breeding materials). HPLC analysis of orange and yellow kernels confirmed a shift in the zeaxanthin:lutein ratio in the direction of zeaxanthin. This orange endosperm mutation was backcrossed into the standard genetic inbred line B73, and *lcyE* was tested as a candidate gene, which revealed that the Q×47 *lcyE* haplotype cosegregates with orange endosperm and ratio of α-carotene versus β-carotene branch carotenoids (fig. S4).

The most favorable haplotype for higher βcarotene branch carotenoids included both the large promoter insertion and 3′ 8-bp insertion. In the diverse panel we tested, this haplotype occurs in 5% of temperate inbreds and 16% of tropical inbreds. MAS at this locus should be effective for several reasons: (i) The most favorable haplotype is found with at least modest frequency in different germ plasm sources and thus breeders can select donors from their relatively more adapted sources. (ii) The favorable haplotype has a large effect. (iii) Visual selection is ineffective for differentiating carotenoid composition and selecting provitamin A compounds. (iv) In comparison with HPLC analysis of carotenoids, polymerase chain reaction (PCR) scoring of the *lcyE* locusismuch less expensive (costing perhaps 1/1000th that of HPLC) and more accessible to developing countries with greatest need for provitamin A.

An approach that empowers local breeder involvement through inexpensive visual selection for darker yellow to orange kernels to enhance flux into carotenoid pathway, and also incorporates MAS for *lcyE*, should result in increased levels of provitamin A compounds. To expedite creation of improved germ plasm globally, we provide information on PCR-based markers (fig. S5). Donor inbreds and improved breeding lines derived at the International Maize and Wheat Improvement Center (CIMMYT) from synthetics of diverse panel inbreds with higher β-carotene are available by contacting T.R.R. This will facilitate selection worldwide of the most favorable *lcyE* alleles, which we have begun in our program. We are screening tropical breeding germ plasm collections in collaboration with CIMMYT.

To date, MAS for natural variation has been limited by resolution and scope (germ plasm diversity). Alleles have generally been characterized in the limited genetic background and resolution of biparental QTL studies, leaving in question their relevance to broader germ plasm (24), particularly for germ plasm outside of the temperate United States. As a result, the primary use for MAS is backcross breeding of transgenic traits. In contrast, the association mapping approach used here allows for rapid generation of selectable markers based on performance of diverse germ plasm. This provides markers more relevant in a broad genetic background, and that enables breeders to search for favorable alleles in their locally adapted germ plasm sources.

In ongoing studies, we are attempting to identify alleles for other genes in the pathway that increase total carotenoids and that slow the conversion of β-carotene to β-cryptoxanthin and zeaxanthin, to exploit more fully the natural genetic variation potential in provitamin A

biofortification of maize. These results will then be further incorporated in breeding efforts to create a healthier maize crop for the world's poorest people.

Although the genetic results and strategy presented here are encouraging, they need to be placed in context as part of an overall biofortification effort encompassing breeding infrastructure, seed distribution, societal acceptance, dietary habits, and nutritional impact. Information now available on some of these issues is encouraging. Results from an animal model for human vitamin A metabolism indicated vitamin A activity of provitamin A in orange maize was greater than assumed by a factor of about four (25). A successful intervention to introduce βcarotene–rich, orange sweet potato in Mozambique, where only white sweet potato was previously cultivated, suggests that orange-colored staple foods can be acceptable, and their regular consumption results in improved vitamin A status (26). Related follow-up acceptance studies of yellow and orange maize in Mozambique and Zimbabwe are in progress with initial results encouraging (27). The dietary habits of many Africans, in which maize is consumed for all three meals a day, indicates that maize is a good target for biofortification (28). The recent positive nutritional and acceptance results will need to be coordinated with comprehensive breeding and seed distribution efforts to realize the potential of provitamin A–biofortified maize, as, for example, is coordinated by the HarvestPlus Global Challenge Program.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References and Notes

1. Underwood BA. J. Nutr 2004;134:231S. [PubMed: 14704325]
2. Mora JO. J. Nutr 2003;133:2990S. [PubMed: 12949398]
3. West CE. Nutr. Rev 2000;58:341. [PubMed: 11140905]
4. Dary O, Mora JO. J. Nutr 2002;132:2927S. [PubMed: 12221271]
5. West CE, Eilander A, van Lieshout M. J. Nutr 2002;132:2920S. [PubMed: 12221270]
6. van Lieshout M, de Pee S. Am. J. Clin. Nutr 2005;81:943. [PubMed: 15817879]
7. Graham RD, Welch RM, Bouis HE. Adv. Agron 2001;70:77.
8. Fraser PD, Bramley PM. Prog. Lipid Res 2004;43:228. [PubMed: 15003396]
9. DellaPenna D, Pogson BJ. Annu. Rev. Plant Biol 2006;57:711. [PubMed: 16669779]
10. Li F, Murillo C, Wurtzel ET. Plant Physiol 2007;144:1181. [PubMed: 17434985]
11. Buckner B, Miguel PS, Janick-Buckner D, Bennetzen JL. Genetics 1996;143:479. [PubMed: 8722797]
12. Palaisa K, Morgante M, Tingey S, Rafalski A. Proc. Natl. Acad. Sci. U.S.A 2004;101:9885. [PubMed: 15161968]
13. Pogson B, McDonald KA, Truong M, Britton G, DellaPenna D. Plant Cell 1996;8:1627. [PubMed: 8837513]
14. Cunningham FX Jr. et al. Plant Cell 1996;8:1613. [PubMed: 8837512]
15. Pecker I, Gabbay R, Cunningham FX Jr. Hirschberg J. Plant Mol. Biol 1996;30:807. [PubMed: 8624411]
16. Yu B, Lydiate D, Young L, Schäfer U, Hannoufa A. Transgenic Res. 2007 10.1007/s11248.
17. Diretto G, et al. BMC Plant Biol 2006;6:13. [PubMed: 16800876]
18. Pogson BJ, Rissler HM. Philos. Trans. R. Soc. London B Biol. Sci 2000;355:1395. [PubMed: 11127994]
19. Flint-Garcia SA, et al. Plant J 2005;44:1054. [PubMed: 16359397]
20. Remington DL, et al. Proc. Natl. Acad. Sci. U.S.A 2001;98:11479. [PubMed: 11562485]
21. Materials and methods are available as supporting material on *Science* Online.
22. Yu J, et al. Nat. Genet 2006;38:203. [PubMed: 16380716]

23. Wong JC, Lambert RJ, Wurtzel ET, Rocheford TR. Theor. Appl. Genet 2004;108:349. [PubMed: 14523521]
24. Asíns MJ. Plant Breed. Rev 2002;121:281.
25. Howe JA, Tanumihardjo SA. J. Nutr 2006;136:2562. [PubMed: 16988127]
26. Low JW, et al. J. Nutr 2007;137:1320. [PubMed: 17449599]
27. Stevens RA, Winter-Nelson A. Food Policy. in press.
28. Li S, Tayie FAK, Young MF, Rocheford T, White WS, Agric J. Food Chem 2007;55:10744.
29. Matthews, PD.; Wurtzel, ET. Food Colorants: Chemical and Functional Properties. Socaciu, C., editor. CRC Press; Boca Raton, FL: 2007. p. 347-398.
30. We thank S. Islam, C. Paul, W. Liu, and W. White for running HPLC on samples; H. Yates for support of molecular genetics; and N. Stevens for technical editing of the manuscript. This work was supported by NSF DBI-0321467 (to E.S.B.), U.S. Agency for International Development (to T.R.R.), HarvestPlus (to T.R.R.), NIH (S06-GM08225 to E.T.W.), Professional Staff Congress—CUNY research award (to E.T.W.), New York State (to E.T.W.), NSF DBI-0604923 (to T.R.R), TRIAD Foundation (L.B. and T.P.B.), U.S. Department of Agriculture (USDA) Cooperative State Research, Education, and Extension Service (CSREES), National Research Initiative grant 2003-00745 (to A.E.S.), and USDA—Agricultural Research Service (to E.S.B.). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

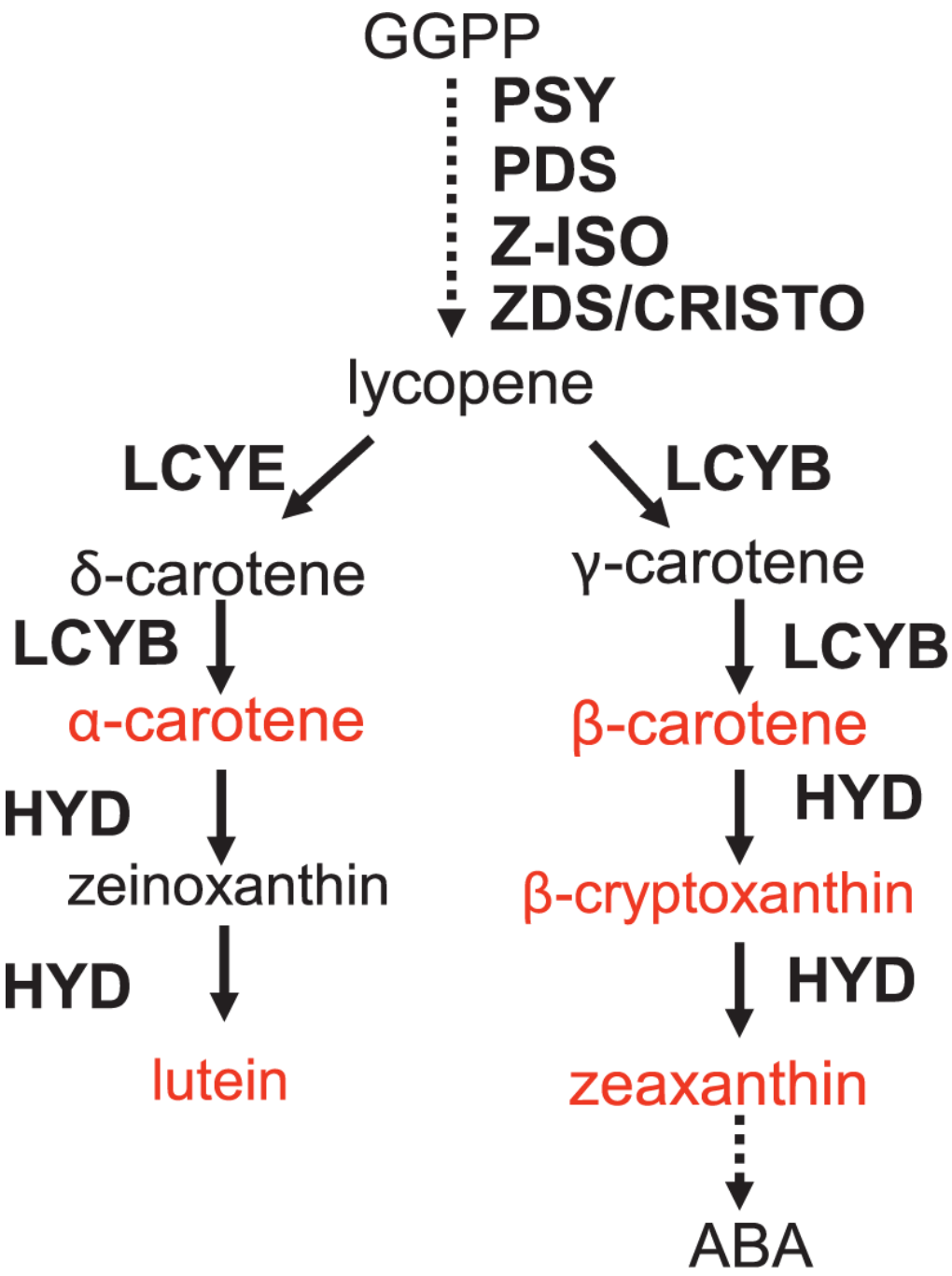

Fig. 1.

Simplified carotenoid biosynthetic pathway in plants (29). Enzymatic reactions are represented by arrows, dashed lines represent multiple enzymatic steps. Substrates in red were evaluated in this study. Compounds: GGPP, geranylgeranyl diphosphate; ABA, abscisic acid. Enzymes: PSY, phytoene synthase; PDS, phytoene desaturase; Z-ISO, 15-*cis* zetacarotene isomerase; ZDS, zetacarotene desaturase; CRTISO, carotene isomerase; HYD, carotene hydroxylase enzymes, which include ε- and β-ring hydroxylases.

Total carotenoids 37.36 µg/g β-*carotene* 4.19 µg/g

Total carotenoids 8.48 µg/g β-*carotene* 5.93 µg/g

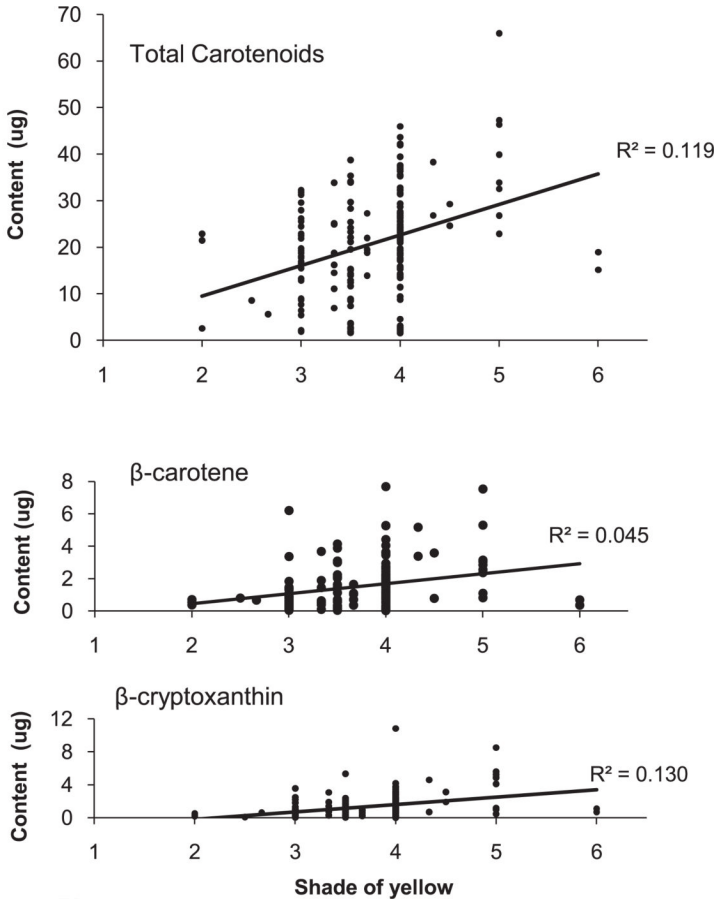

Fig. 2.

Grain color and carotenoid content. The graphs depict the low correlation between visual grain color and total carotenoids, β-carotene, and β-cryptoxanthin in diverse inbreds. In these kernels, the shade of yellow ranges from white (score of 1) to dark orange (score of 6). White kernels were excluded from the analysis. The difficulty in visual selection for β-carotene content is further exemplified by the images on the left, where the yellow maize below has higher β-carotene than the orange variety above. These correlations are across the diverse panel of 228 maize inbreds; correlations for grain color and total carotenoids are higher when scored across segregating populations and narrow ranges of germ plasm, but correlations for β-carotene and β-cryptoxanthin remain low.

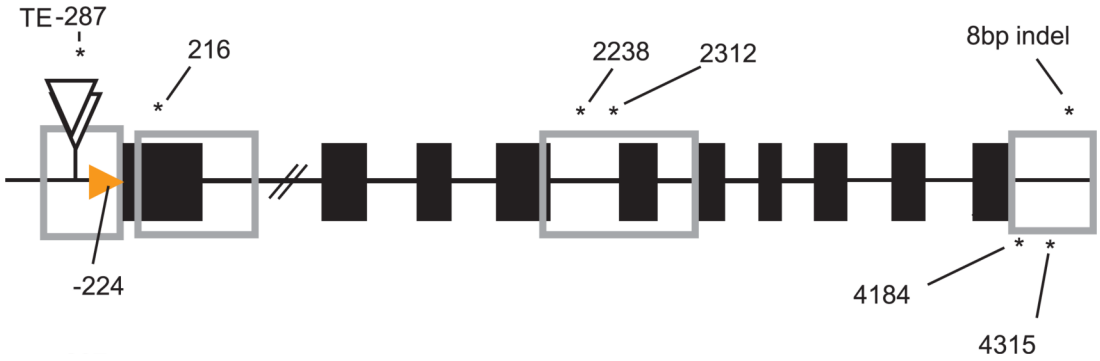

Fig. 3.

Schematic diagram of the maize *lcyE*. Putative transcription start sites are depicted with orange arrow, translated exons as black squares, and the sampled regions as gray boxes. Polymorphisms that significantly associated with changes in flux between the lutein and zeaxanthin branches of the pathway are labeled with asterisks. The 5′ transposable element insertion(s) are represented by the white triangles. Positions relative to the sequence alignment are indicated numerically above the polymorphisms.

Table 1

lcyE associations across seasons. Association results for significant polymorphisms identified in the four regions sampled along the *lcyE* gene. Each polymorphism is labeled numerically by its position on the alignment relative to the exon 1 start codon. Followed by the favorable allele (bold)/unfavorable allele at the site. An initial scan for association using both β-carotene and the ratio of the two pathway branches was conducted using the mixed model incorporating population structure and kinship. Subsequently a simpler general linear model (GLM) was used to evaluate data sets from additional years, including population structure (Q), given the oligogenic behavior of the trait the change in flux estimates for 2003 do not include Q. Avg., average; n.c., nonconvergence; n.s., not significant.

	***lcyE* association (*P*), mixed model of**						
	β-Carotene/all	**Ratio of branches**		***lcyE* association as a ratio across environments (GLM) (*P*)**			**Fold change in flux**
Environment (year)	(2003)	(2003)	(2002)	(2003)	(2004)	(2005)	(2003)
Avg. observation no.	(157)	(154)	(44)	(156)	(154)	(156)	
Polymorphic site							
5′ TE **1**+**4**/2/3	5.42×10^{-4}	3.96×10^{-11}	0.024	8.05×10^{-11}	0.008	8.61×10^{-9}	6.5
216 **G** / T	n.c.	1.35×10^{-10}	0.059	1.24×10^{-10}	0.003	2.93×10^{-10}	2.8
2238 **G** / T	1.22×10^{-4}	1.69×10^{-9}	0.008	2.12×10^{-10}	0.023	1.08×10^{-9}	2.7
2312 **A** / T	1.70×10^{-3}	n.c.	n.s.	6.84×10^{-4}	0.026	0.005	2.9
4184 **G** / A	3.06×10^{-4}	n.c.	8.87×10^{-4}	2.23×10^{-10}	0.019	1.13×10^{-8}	2.6
4315 **C** / G	1.84×10^{-4}	7.01×10^{-10}	0.012	3.07×10^{-9}	5.75×10^{-4}	6.79×10^{-7}	2.6
3′Indel **8**/0	4.80×10^{-3}	2.75×10^{-9}	n.s.	1.46×10^{-8}	8.97×10^{-4}	4.13×10^{-6}	3.5