JIPB Journal of Integrative Plant Biology

Journal of Integrative Plant Biology 2012, **54** (4): 260–269

Research Article

ZmcrtRB3 Encodes a Carotenoid Hydroxylase that Affects the Accumulation of α-carotene in Maize Kernel

Yi Zhou[1], Yingjia Han[1], Zhigang Li[1], Yang Fu[1], Zhiyuan Fu[1,2], Shutu Xu[1], Jiansheng Li[1], Jianbing Yan[1,3] and Xiaohong Yang[1*]

[1]National Maize Improvement Center, Key Laboratory of Crop Genomics and Genetic Improvement, China Agricultural University, 100193 Beijing, China

[2]Zhengzhou Sub-center of Chinese National Maize Improvement Center, Henan Agricultural University, Zhengzhou 450002, China

[3]National Key Laboratory of Crop Improvement, Huazhong Agricultural University, Wuhan 430070, China

[*]Corresponding author

Tel: +86 10 6273 2400; Fax: +86 10 6273 3808; E-mail: redyx2005@126.com

Articles can be viewed online without a subscription.

Available online on 20 February 2012 at www.jipb.net and www.wileyonlinelibrary.com/journal/jipb

doi: 10.1111/j.1744-7909.2012.01106.x

Abstract

α-carotene is one of the important components of pro-vitamin A, which is able to be converted into vitamin A in the human body. One maize (*Zea mays* L.) ortholog of carotenoid hydroxylases in *Arabidopsis thaliana*, *ZmcrtRB3*, was cloned and its role in carotenoid hydrolyzations was addressed. *ZmcrtRB3* was mapped in a quantitative trait locus (QTL) cluster for carotenoid-related traits on chromosome 2 (bin 2.03) in a recombinant inbred line (RIL) population derived from By804 and B73. Candidate-gene association analysis identified 18 polymorphic sites in *ZmcrtRB3* significantly associated with one or more carotenoid-related traits in 126 diverse yellow maize inbred lines. These results indicate that the enzyme ZmcrtRB3 plays a role in hydrolyzing both α- and β-carotenes, while polymorphisms in *ZmcrtRB3* contributed more variation in α-carotene than that in β-carotene. Two single nucleotide polymorphisms (SNPs), SNP1343 in 5′untranslated region and SNP2172 in the second intron, consistently had effects on α-carotene content and composition with explained phenotypic variations ranging from 8.7% to 34.8%. There was 1.7- to 3.7-fold change between the inferior and superior haplotype for α-carotene content and composition. Thus, SNP1343 and SNP2172 are potential polymorphic sites to develop functional markers for applying marker-assisted selection in the improvement of pro-vitamin A carotenoids in maize kernels.

Keywords: α-carotene; association analysis; maize; pro-vitamin A; *ZmcrtRB3*.

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**(4), 260–269.

Introduction

Maize (*Zea mays* L.) is one of the most widely grown crops in the world, and plays a critical role as the sources of human food, animal feed and bioenergy. Maize kernels contain a wide range of genetic variability for pro-vitamin A carotenoids (α-carotene, β-carotene and β-cryptoxanthin), having vitamin A activity in the human body. Epidemiological studies demonstrated that food rich in these and other carotenoids are extremely beneficial for human and animal health due to their antioxidant properties. Furthermore, inadequate intake of vitamin A may cause many serious diseases, such as cardiovascular disease, cancer and cataracts (Fraser and Bramley 2004). According to a World Health Organization (WHO) survey during the period between 1995 and 2005, vitamin A deficiency (VAD) continues to be a prevalent problem in more than 100 developing countries affecting an estimated 200 million preschool children and 30 million pregnant women, among which around 5.2 million of preschool children and 9.8 million pregnant women had the risk of night blindness. It is a great challenge to conquer the health

problems caused by VAD in developing countries. Contrarily to food fortification and pills supplementation, biofortification of staple crops such as maize with enhanced levels of pro-vitamin A carotenoids seems to be an economical and sustainable way to combat VAD. Recently, some elite maize inbreds having a wide range of pro-vitamin A carotenoids, especially β-carotene (about 200-fold change between maize kernels with the highest and lowest β-carotene content), have been identified in a diverse germplasm (Harjes et al. 2008). This makes it available to perform biofortification by mining favorable alleles associated with pro-vitamin A carotenoids.

In plants, the carotenoid biosynthesis pathway begins with the conversion of geranylgeranyl pyrophosphate (GGPP) into lycopene by a series of enzymes including phytoene synthase (PSY), phytoene dehydrogenase (PDS), ξ-carotene dehydrogenase (ZDS) and carotenoid isomerase (CRTISO) (Buckner et al. 1996; Hable et al. 1998; Park et al. 2002; Matthews et al. 2003). Lycopene represents a branch point (Zhu et al. 2003) and the pathway diverges towards two alternate routes to produce (i) β-carotene by lycopene β-cyclase (LCYB) (Hirschberg 2001), from which β-cryptoxanthin and zeaxanthin are further synthesized by β-ring hydroxylase (crtRB) (DellaPenna et al. 2006), and (ii) α-carotene by lycopene ε-cyclase (LCYE), or in combination with LCYB, is further converted into lutein by crtRB and ε-ring hydroxylase (Tian et al. 2004). Among the known enzymes in the carotenoid biosynthesis pathway, crtRB catalyzes the key steps of converting pro-vitamin A carotenoids to non-pro-vitamin A carotenoids, and thus, significantly influences the accumulation of pro-vitamin A in maize kernels.

The carotenoid biosynthesis pathway is also well characterized in maize. Several key enzymes in the pathway have been extensively studied in maize such as *PSY1* on chromosome 6 (6.01 bin), that encodes the first rate-limiting enzyme in the pathway. *PSY1* was cloned by transposon tagging, and is positively correlated with carotenoid content in maize endosperm (Buckner et al. 1996). Two constitutively expressing single copy genes, *PDS* and *ZDS*, were mapped on chromosome 1 (1.02 bin) and 7 (7.02 bin), respectively (Hable et al. 1998; Matthews et al. 2003). *LCYB* was cloned by transposon tagging and mapped on chromosome 5 (5.04 bin) (Singh et al. 2003). Recently, vitamin A pathway-driven association studies have identified the significant associations of natural polymorphisms in *LCYE*, *crtRB1*, *PSY1*, with the levels of carotenoid derivatives in maize kernels, and polymerase chain reaction (PCR)-based functional markers have been developed for molecular breeding of pro-vitamin A carotenoids particularly for β-carotene (Harjes et al. 2008; Yan et al. 2010; ZY Fu et al., unpubl. data). *PSY1* controls the flux of substrates into the carotenoid pathway (JY Fu et al. unpubl. data), *LCYE* directs the flow of substrates towards α-carotene versus β-carotene branches of the carotenoid pathway (Harjes et al. 2008), and *crtRB1* effects on the levels of β-carotene content and conversion (Yan et al. 2010). Thus, it will be perfect to design molecular breeding for pro-vitamin A carotenoids in maize kernels when natural variations contributing to α-carotene, β-cryptoxanthin content and conversion are further identified.

The objectives of this study were to clone a gene, *ZmcrtRB3*, encoding carotenoid hydroxylase; and to characterize the contribution and putative function of *ZmcrtRB3* with carotenoid content and composition in maize kernels.

Results

Gene cloning and gene structure of *ZmcrtRB3*

Two non-heme di-iron enzymes in *Arabidopsis thaliana*, BCH1 and BCH2, exhibit the activities in the hydroxylation of α-carotene, β-carotene and β-cryptoxanthin (Kim et al. 2009). Tblastn searches with BCH1 and BCH2 protein sequences (The Arabidopsis Information Resource (TAIR), AT4G25700.1 and AT5G52570.1, respectively) against the maize high throughput genomic sequence database in National Center for Biotechnology Information (NCBI) identified over 10 hits of maize homologous genes of carotenoid hydroxylase distributing on the chromosomes 1, 2, 4, 9 and 10 in the maize genome (Vallabhaneni et al. 2009; Yan et al. 2010). The maize complementary DNA (cDNA) sequence, AY844956.1, had the highest identity with BCH1 and BCH2 in *A. thaliana* (BCH1, E value = 4.00E-110; BCH2, E value = 3.00E-110), and was named as *ZmcrtRB3* and sequentially selected as the candidate gene.

The maize bacterial artificial chromosome (BAC) clone, AC196442.3, from chromosome 2 was identified to contain sequences (27 937–31 850 bp) with high similarity to *ZmcrtRB3* by blast searches with AY844956.1 in Maize Assembled Genomic Island database (MAGI). Function domain analysis of this genomic fragment in InterProScan confirmed that *ZmcrtRB3* is highly consistent with the function of BCH1 and BCH2 in *A. thaliana* and belongs to the fatty acid hydroxylase gene superfamily (FA_hydroxylase superfamily). The gene structure of *ZmcrtRB3* was determined using two maize cDNA sequences (Genbank, AY844956.1 and FL024732.1), and confirmed by cloning the cDNA sequence from maize inbred B73 embryo at 20 days after pollination (DAP) (Genbank, JQ420132). The *ZmcrtRB3* gene in B73 has six exons and the full length cDNA sequence is of 1 392 bp, encoding 309 amino acids. The length of 5′untranslated region (UTR) and 3′UTR are 107 bp and 358 bp, respectively. The coding sequence of *ZmcrtRB3* in B73 has 100% of identity with *ZmBCH1* (Genbank, GQ131287; Li et al. 2010). In addition, *ZmcrtRB3* was located within contig73 on chromosome bin 2.03 in maize genome by using blast searches with *ZmcrtRB3* genomic sequence in Maize Sequence database.

ZmcrtRB3 maps to a QTL cluster for carotenoid derivatives

A 60 bp InDel was detected from the genomic sequence alignment between By804 and B73 in the 3′UTR of *ZmcrtRB3*, and thus a marker, M1, was developed in the flanking region of the 60 bp polymorphic site. The marker was mapped between simple sequence repeat (SSR) markers umc1422 and umc1776 on chromosome bin 2.03 in By804/B73 recombinant inbred line (RIL) population (Chander et al. 2008). The genetic distance of umc1422 and umc1776 from M1 was 2.8 cM and 13.9 cM, respectively. Following the addition of M1 marker in linkage map, a novel quantitative trait locus (QTL) for β-carotene composition (ratio of β-carotene to total carotenoids; βC/TC), *qβC2*, was detected on chromosome 2 (Figure S1). The QTL, *qβC2,* explained 8.5% of phenotypic variation for β-carotene composition with logarithm of the odds (LOD) value of 5.9. The By804 allele at this locus had an additive effect of 0.007 for increased βC/TC. Another QTL (LOD = 4.1) for α/β-carotene ratio (αC/βC) was also mapped within the same region (Figure S1), and accounted for 4.7% of the αC/βC variation. The B73 allele contributed to the increase of αC/βC by 0.19. *ZmcrtRB3* was mapped within the QTL cluster (Figure S1), suggesting the function of *ZmcrtRB3* might be related to carotenoid hydroxylase.

Nucleotide diversity and LD extension in *ZmcrtRB3*

Among the 126 yellow maize inbred lines, 194 nucleotide polymorphisms (41 InDels and 153 single nucleotide polymorphisms (SNPs), **Table 1**) were detected in a 3 563 bp genomic region harboring *ZmcrtRB3*. On average, one in every 18 base pairs was polymorphic. Additionally, the distribution of the nucleotide diversity was not even across *ZmcrtRB3*. The 5′UTR region had the most abundant nucleotide diversity, followed by 3′UTR, introns and exons (**Table 1**). Among the polymorphisms in the coding region, five SNPs leading to synonymous changes and two InDels causing four and one amino acid insertion/deletion occurred in the first exon.

Linkage disequilibrium (LD) of *ZmcrtRB3* decayed rapidly within 400 bp (r^2= 0.16), but extended to over 3 500 bp in 126 yellow maize inbred lines (Figure S2). It indicated that the resolution of association mapping was likely not to reach up to the gene level. Furthermore, two potential LD blocks were identified in 5′UTR and 3′UTR of *ZmcrtRB3* (Figure S3). The LD block in 5′UTR (S207-S1195, 77 polymorphic sites) covers nearly one third of the measured regions, while that in 3′UTR (S3147-S3168, 13 polymorphic sites) fell within a small region spanning 20 bp and the levels of LD between these two LD blocks were high (Figure S3).

Table 1. Distributions of polymorphisms in *ZmcrtRB3* across 126 maize inbred lines

Polymorphisms	Total	SNP	Indels	Synonymous	Non-synonymous
Promoter and 5′UTR	103	88	15		
Exon1	13	11	2	5	6
Intron1	18	10	8		
Exon2	1	1	0	0	1
Intron2	10	7	3		
Exon3	2	2	0	0	2
Intron3	9	5	4		
Exon4	1	1	0	0	1
Intron4	6	4	2		
Exon5	3	3	0	0	3
Intron5	1	0	1		
Exon6	0	0	0	0	0
3′UTR	27	21	6		

SNP, single nucleotide polymorphism; UTR, untranslated region.

Natural variations of *ZmcrtRB3* contribute to α-carotene

Of the 194 polymorphic sites in *ZmcrtRB3*, 18 (16 SNPs and two InDels) showed significant associations ($P \leq 0.01$) with one or more traits (Table S1). The most significant polymorphic site was SNP2172, a SNP located in the second intron containing a G-to-C transition (**Figure 1**, **Table 2**, Table S1). SNP2172 was in weak LD with other significant polymorphic sites detected in this study (**Figure 1C**), indicating SNP2172 was a unique site associated with α-carotene content and composition. SNP2172 showed a highly significant effect on α-carotene content (αC, $P = 1.1 \times 10^{-9}$) and ratio of α-carotene to total carotenoids (αC/TC, $P = 3.7 \times 10^{-8}$), whereas lower magnitude of effect was observed on ratio of α-carotene to lutein (αC/lut, $P = 5.8 \times 10^{-3}$). The explained phenotypic variations ranged from 9.0% (αC/lut) to 29.9% (αC), and the maximum fold change range from 1.3 (αC/lut) to 2.6 (αC/TC) (**Table 2**). SNP1343 (C/T), localized in the 5′UTR, was also significantly associated with αC, αC/TC and αC/lut ($P = 1.4 \times 10^{-3}$, $P = 3.4 \times 10^{-3}$, $P = 1.2 \times 10^{-3}$, respectively). The levels of LD between SNP1343 and 11 SNPs as well as SNP2172 were low, whereas SNP1343 showed a moderate level of LD with the polymorphic sites in the coding region of *ZmcrtRB3* (**Figure 1C**). SNP1343 explained 8.7% to 15.0% of the phenotypic variations, and there was 1.4 to 1.8-fold difference between the favorable (T) and the unfavorable allele (C) (**Table 2**). The effects of these two SNPs suggest that *ZmcrtRB3* contributes mainly to α-carotene compared with β-carotene or β-cryptoxanthin in maize kernels.

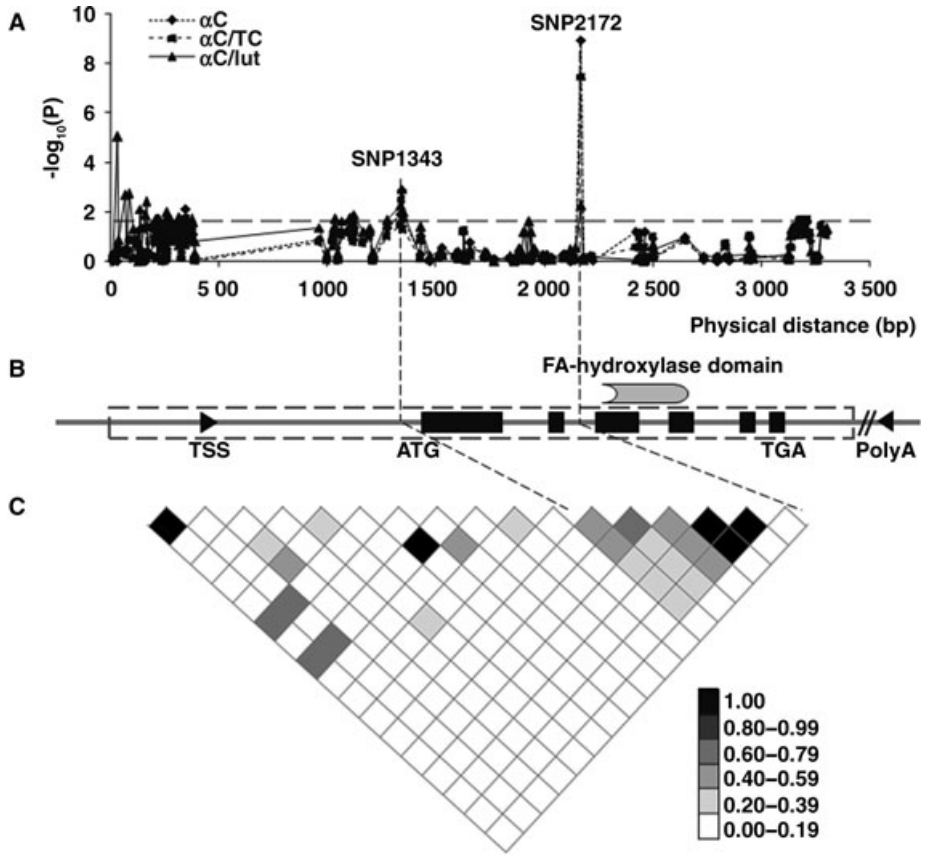

Figure 1. Association mapping results and LD pattern of *ZmcrtRB3*.

(A) Associations between polymorphic sites in *ZmcrtRB3* and αC, αC/TC, αC/lut. αC, α-carotene; αC/TC, the ratio of α-carotene to total carotenoids; αC/lut, the ratio of α-carotene to lutein.
(B) Gene structure of *ZmcrtRB3*. The black rectangles represent position of exons in *ZmcrtRB3*.
(C) LD pattern of the polymorphic sites in *ZmcrtRB3* significantly associated with 16 carotenoid related traits.
The dotted lines across the whole figure indicate the positions of two significant single nucleotide polymorphisms in the gene and LD matrix.

Table 2. Summary of associations between polymorphic sites in *ZmcrtRB3* with α-carotenoid related traits in maize kernel

Polymorphic sites[a]	Alleles[b]	Frequencies	Traits[c]	*P*-values[d]	R^2 (%)[e]	Fold change[f]
SNP1343	C/**T**	98/43	αC	1.4×10^{-3}	8.7	1.7
			αC/TC	3.4×10^{-3}	8.8	1.8
			αC/lut	1.2×10^{-3}	15.0	1.4
SNP2172	**G**/C	16/95	αC	1.1×10^{-9}	29.9	2.3
			αC/TC	3.7×10^{-8}	34.8	2.6
			αC/lut	5.8×10^{-3}	9.0	1.3

[a]Only polymorphic sites significantly associated with αC, αC/ALL, αC/lut are shown.
[b]The favorable alleles associated with higher αC are in bold and underlined.
[c]αC, α-carotene; αC/TC, the ratio of α-carotene to total carotenoids; αC/lut, the ratio of α-carotene to lutein.
[d]*P* values for association analysis are calculated using a mixed model incorporating population structure and kinship in TASSEL 2.1.0.
[e]R^2 from ANOVA shows the percentage of phenotypic variation explained.
[f]Fold change was calculated between the favorable allele and the unfavorable allele.

Table 3. Haplotype effects of SNP1343 and SNP2172 in *ZmcrtRB3*

Haplotypes (SNP1343/ SNP2172)[a]	*n*	Means ± *SD*[b] αC (μg/g)	αC/TC	αC/lut
C/C	87	0.123 ± 0.007	0.011 ± 0.001	0.030 ± 0.001
T/C	7	0.129 ± 0.009	0.010 ± 0.002	0.038 ± 0.008
C/**G**	11	0.253 ± 0.023	0.024 ± 0.004	0.037 ± 0.002
T/**G**	5	0.345 ± 0.020	0.036 ± 0.009	0.049 ± 0.007
R^2 (%)[c]		34.7	40.9	21.4
P value[d]		7.9×10^{-10}	4.2×10^{-12}	1.1×10^{-5}
Fold change[e]		2.8	3.7	1.7

[a]The favorable alleles for αC are in bold and underlined.

[b]αC, α-carotene; αC/TC, the ratio of α-carotene to total carotenoids. αC/lut, the ratio of α-carotene to lutein.

[c]R^2 from ANOVA shows the percentage of phenotypic variation explained.

[d]*P*-values for association analysis are calculated using ANOVA.

[e]Fold change was calculated between the most favorable haplotype and the least favorable haplotype.

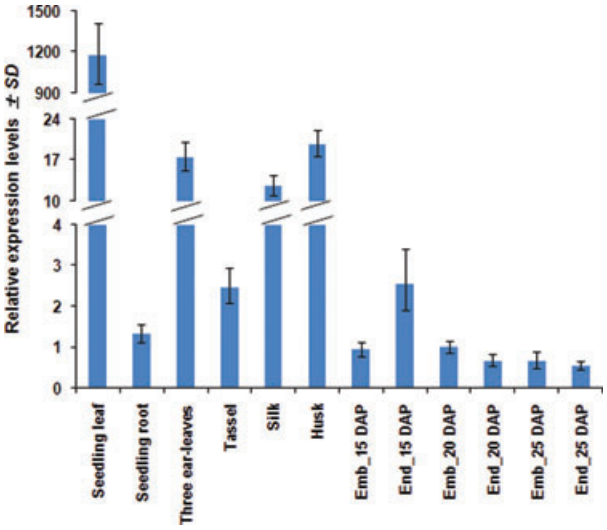

Figure 2. Expression pattern of *ZmcrtRB3* in various tissues in B73.

DAP, days after pollination; Emb, embryo; End, endosperm.

Haplotype analysis revealed that four haplotypic classes exist in the present analyzed association mapping panel as SNP1343 and SNP2172 were in weak LD, which allows the full differentiation of the effects of each polymorphism. Of the four haplotype classes, the haplotypes with both favorable alleles of two sites (T/G) were the most favorable haplotypes for αC, αC/TC and αC/lut, which with both unfavorable alleles (C/C) were the least favorable haplotypes for αC and αC/lut, whereas two haplotype classes (C/C, T/C) had similar effects on αC/TC (**Table 3**). Overall, the most favorable haplotypes exhibited a 2.8-, 3.7- and 1.7-fold increase over the least favorable haplotypes for αC, αC/TC and αC/lut, respectively. In addition, the haplotype from two polymorphic sites in *ZmcrtRB3* accounted for 34.7% of the phenotypic variation for αC, 40.9% for αC/TC, and 21.4% for αC/lut.

Sequentially, a molecular marker based on cleaved amplified polymorphic sequences (CAPS) was developed for SNP2172 for the molecular breeding of pro-vitamin A carotenoids in maize kernels (Figure S4). The searches for restriction endonuclease sites found that no restriction sites were available for SNP1343 while they did exist for SNP2172. For SNP2172, the PCR products were digested into 48 bp and 293 bp fragments by MvaI when the alleles were C, whereas they were not cleaved when the alleles were G. Thus, the polymorphisms reflecting on agarose gels can be applied in selecting the favorable alleles of *ZmcrtRB3* in high pro-vitamin A molecular breeding.

Expression profile of *ZmcrtRB3*

Real-time quantitative reverse transcription (qRT)-PCR of 12 different tissues in B73 was performed to address the expression profile of *ZmcrtRB3*. The expression levels of *ZmcrtRB3* were extremely high in seedling leaf, whereas relatively low expression was detected in root, embryo and endosperm (**Figure 2**). Further investigation of *ZmcrtRB3* expression carried out in the embryo and endosperm at 15, 20 and 25 DAP of five additional inbreds showed that the expression of *ZmcrtRB3* was genotype dependent (Figure S5). On average, the expression level was the highest in embryo and endosperm at 15 DAP and 20 DAP, respectively. To address whether the increase of α-carotene content and composition in maize inbreds was due to SNP1343 in the 5′UTR region of *ZmcrtRB3*, the transcript level of *ZmcrtRB3* in embryo and endosperm at 20 DAP was quantified, but no correlations were observed between *ZmcrtRB3* expression level and α-carotene composition (αC/TC and αC/lut). However, a very weak association was detected only for α-carotene content (**Figure 3**). These results imply that different alleles of *ZmcrtRB3* identified in the present study might not be associated with the transcription. However, we do not rule out the effect of this SNP on *ZmcrtRB3* at post-transcriptional levels as different genotypes having SNP1343 did not show variation in the coding region.

Discussion

On the basis of retinal structure, α-carotene, β-carotene and β-cryptoxanthin are considered as pro-vitamin A carotenoids due to their vitamin A activity. To address VAD by biofortification,

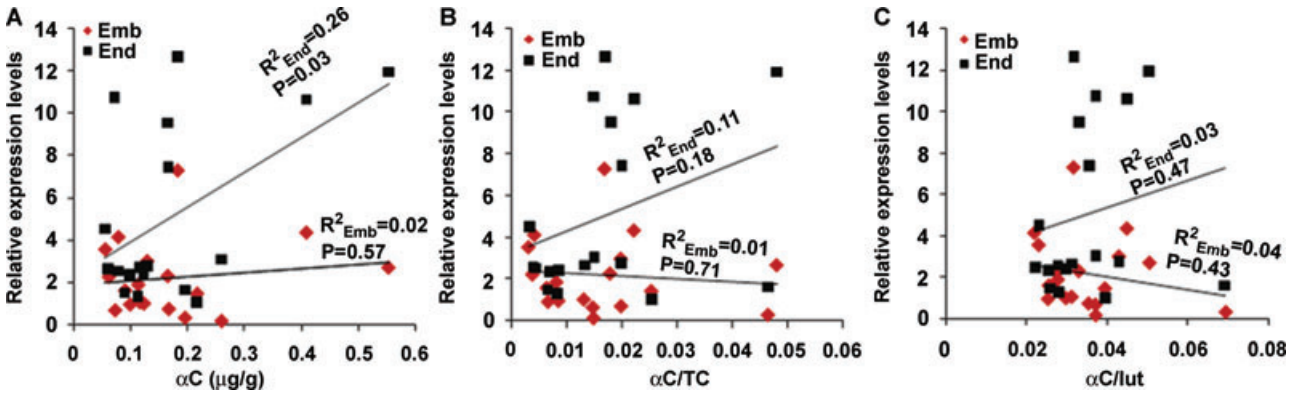

Figure 3. Correlation between *ZmcrtRB3* expression levels in embryo and endosperm at 20 DAP and α-carotenoid related traits.

(A) Correlation between *ZmcrtRB3* expression levels and αC.
(B) Correlation between *ZmcrtRB3* expression levels and αC/TC.
(C) Correlation between *ZmcrtRB3* expression levels and αC/lut.
αC, α-carotene; αC/TC, the ratio of α-carotene to total carotenoids; αC/lut, the ratio of α-carotene to lutein.

it is necessary to understand the accumulation mechanisms of pro-vitamin A carotenoids including biosynthesis and degradation in maize kernels. Prior studies pointed out that four carotenoid hydroxylases in two classes, the cytochrome P450 enzymes (CYP97A3 and CYP97C1) and non-heme di-iron enzymes (BCH1 and BCH2), catalyze the formation of non pro-vitamin A carotenoids (lutein and zeaxanthin) from pro-vitamin A carotenoids (Sun et al. 1996; Tian and DellaPenna 2001; Tian et al. 2003). Further investigation in *A. thaliana* showed that carotenoid hydroxylases (BCH1 and BCH2) co-regulated the metabolism of α-carotene and β-carotene (Kim et al. 2009). However, carotenoid hydroxylases are primarily responsible for catalyzing hydroxylation of β-carotene metabolism in maize (Vallabhaneni et al. 2009; Yan et al. 2010; Li et al. 2010). Here, we found that a gene, *ZmcrtRB3*, encoding a carotenoid hydroxylase, was able to catalyze α-carotene in maize kernels. QTL mapping and candidate-gene association mapping in multiple environments also indicated that the enzyme ZmcrtRB3 has a role in hydrolyzing β-carotene, whereas natural variations in *ZmcrtRB3* exhibited more association with the level of α-carotene than β-carotene. It seems reasonable as crtRB1 have extremely high activity in hydrolyzing β-carotene to form β-cryptoxanthin in maize kernel (Yan et al. 2010). Additionally, the high expression of *ZmcrtRB3* in seedling leaves suggests that *ZmcrtRB3* may play a more important role in carotenoid hydrolyzation in leaves than that in maize kernel as carotenoids play vital roles in photosynthesis, especially in photoprotection and light collection (Demmig-Adams et al. 1996). These results will broaden our knowledge on the biosynthesis mechanism of pro-vitamin A carotenoids and will help to understand how ZmcrtRB3 enhancing the hydroxylation activity in hydrolyzing α-carotene. Further analysis of the correlation between *ZmcrtRB3* expression level and α-carotene-related traits did not support the biological function of *ZmcrtRB3* in the carotenoid metabolism pathway. This may be due to limited samples in limited environments for analyzing the expression pattern of *ZmcrtRB3*. Thus, continuing function observations, such as gene silencing and enzyme activity, are necessary to unravel the biological mechanism of *ZmcrtRB3* in hydrolyzing α-carotene.

Two SNPs, SNP1343 and SNP2172, in the non-coding region of *ZmcrtRB3* were significantly associated with α-carotene content and composition. As SNP1343 was located in 5′UTR, it might regulate the hydroxylation of α-carotene by increasing the expression level of *ZmcrtRB3* to improve the activity of its enzyme. However, this hypothesis was not supported by the transcript profile of *ZmcrtRB3* in endosperm. Thus, the role of SNP1434 in the regulation of the α-carotene level in maize kernel is still unclear. SNP2172 located in the second intron seems to be a real functional site in *ZmcrtRB3* as numerous studies have validated the role of natural variations in introns to human disease (Sasabe et al. 2007; Saxena et al. 2007; Belbin et al. 2008). Additionally, it is possible that these two SNPs are in strong LD with upstream regulatory elements, which were not detected in this study. Actually, the effects of genes on targeted traits by the regulation of long-distance upstream regulatory elements were common in maize such as long distance cis-acting regulatory element >41 kb upstream of *Teosinte branched1* (*tb1*) acting in altering *tb1* transcription to affect phenotypic variations (Clark et al. 2006; Studer et al. 2011; Zhou et al. 2011); and another cis-acting regulatory element, *Vgt1*, positioned 70kb upstream of an *Ap2*-like transcription factor is a key component for flowering time in maize (Salvi et al. 2007).

From a purely breeding point of view, SNP1343 and SNP2172 are still potential polymorphic sites to develop functional markers that could be used in marker assisted breeding for pro-vitamin A in maize. The functional sites of *PSY1*, *LCYE* and *crtRB1* were well characterized in maize, and the corresponding PCR-based friendly markers were available (Harjes et al. 2008; Yan et al. 2010; ZY Fu et al. unpubl. data). The combination of favorable alleles of *LCYE* and *crtRB1* has been exhibited to efficiently improve the level of pro-vitamin A carotenoids by increasing β-carotene content (Yan et al. 2010), and the integration of another favorable allele of *PSY1* further demonstrated the improvement of all carotenoids derivatives by increasing more substrates to carotenoid biosynthetic pathway in maize kernels (ZY Fu et al. unpubl. data). Therefore, pyramiding of favorable alleles of carotenoids biosynthetic pathway genes in maize including *ZmcrtRB3* will definitely enhance the level of pro-vitamin A carotenoids, particularly α-carotene in maize kernels as *ZmcrtRB3* performed functions in hydrolyzing α-carotene. Yet, the pro-vitamin A carotenoids are in the dynamic carotenoid metabolism pathway and interact in a complex way. A key enzyme might regulate the natural variations of over two carotenoids. For example, *ZmcrtRB3* had a functional signal of hydrolyzing β-carotene besides α-carotene. To obtain the highest levels of pro-vitamin A carotenoids by combining the favorable alleles of *PSY1*, *LCYE*, *crtRB1* and *ZmcrtRB3*, further investigation should be conducted to assess their combinatorial effects on carotenoid content and composition especially pro-vitamin A carotenoids in maize kernels.

Materials and Methods

Plant materials and phenotyping

A RIL population derived from a single cross between By804 and B73 was used to map *ZmcrtRB3*. The details of experimental condition and carotenoids measurement from 245 RILs have been described in our previous study (Chander et al. 2008). A maize association panel consisting of 155 inbred lines (Yang et al. 2010) was used to detect the associations between the nucleotide polymorphisms of *ZmcrtRB3* and carotenoid content as well as compositions in maize kernel. The panel was planted in a randomized complete block design with two replications at the farm of China Agricultural University during summer 2006 and 2007 (Beijing, BJ, E 116°11′, N40°81′), and winter 2007 (Hainan, HN, E 108°56′, N 18°19′) whereas one replication was planted at the farm of the Sichuan Agricultural University in summer 2009 (Ya'an, YA, E 102°59′, N 29°58′). Each genotype was grown in a single row of 3 m length, spaced 0.67 m apart with a planting density of 45 000 plants/ha. More than six ears in each line were self-pollinated and equal amounts of kernels from selfed ears in individual lines from each replication were bulked for chemical analysis. Among 155 inbred lines, only 126 lines having yellow kernel color were phenotyped and genotyped. The best linear unbiased predictors (BLUPs) for individual traits in each line were calculated using SAS 8.02 (SAS Institute 1999) in both RIL and association mapping populations. BLUPs for each line across environments were used for the overall analysis.

Mining candidate genes

Based on the knowledge of carotenoid biosynthesis pathway (EMP, Enzymes and Metabolic Pathways database, http://www.plantcyc.org), the protein sequences of carotenoid hydroxylase homologous genes *BCH1* and *BCH2* in *A. thaliana* were retrieved from TAIR (http://www.arabidopsis.org). The homologous sequences of BCH1 and BCH2 in maize were obtained via tblastn in NCBI (http://www.ncbi.nlm.nih.gov). The related BAC sequences in maize were identified in MAGI database (http://magi.plantgenomics.iastate.edu) and *in silico* mapped in maize sequence database (http://www.maizesequence.org). The best matched BAC sequences located the targeted genomic regions were used for sequence analysis.

RNA preparation and reverse transcription

Total RNA was extracted from different plant tissues in various lines (Table S2) using a RNA extraction kit (BioTeke, Beijing, China). The first-strand cDNA was synthesized in a volume of 20 μL containing 2 μg RNA, 40 U RNasin (TaKaRa, Kyoto, Japan), 0.5 μg oligo (dT) primers and 200 U M-MLV reverse transcriptase (Promega, Madison, WI, USA).

Gene structure and cloning of full-length cDNA

To define the gene structure of *ZmcrtRB3*, the InterProScan Sequence Search (http://www.ebi.ac.uk/Tools/pfa/iprscan) and the Gene Finding program (http://mendel.cs.rhul.ac.uk) were used. The gene structure was further confirmed by the sequences of full-length cDNA.

The first-strand cDNA from embryo at 20 DAP in B73 was used to obtain the full-length cDNA of *ZmcrtRB3* using Rapid Amplification of cDNA 3′Ends (3′RACE) (Invitrogen Carlsbad, CA, USA). To ensure the specificity of RACE primers, five orthologs of carotenoid hydroxylase in maize, *ZmcrtRB2* (1.01bin), *ZmcrtRB3* (2.03bin), *ZmcrtRB4* (4.09bin), *ZmcrtRB5* (9.07bin) and *ZmcrtRB1* (10.05bin) (Yan et al. 2010), were used for sequence alignments. The alignments of cDNA sequences showed the differences between *ZmcrtRB3* and other orthologs, and two specific forward primers (*ZmcrtRB3* 1F, *ZmcrtRB3* 2F, Table S3) for 3′RACE were developed using Primer 5.0. The 3′RACE products were amplified and

sequenced after purification with agarose gel DNA purification kit (TianGen Biotech, Beijing, China).

Genotyping and sequence analysis

The genomic sequence of *ZmcrtRB3* was used to design primers by Primer 5.0. Four primers (Table S3) covering the whole gene were used to sequence *ZmcrtRB3* in 126 maize lines. The sequences were assembled using ContigExpress in Vector NTI Advance 10 (Invitrogen), aligned using MUSCLE (Edgar 2004), and manually corrected by BioEdit (Hall 1999).

Linkage mapping and QTL mapping

The polymorphisms between two parents (By804 and B73) of the RIL population were identified using the alignment of *ZmcrtRB3* sequences. For mapping of *ZmcrtRB3*, primers (M1F/R, Table S3) were designed in the flanking region of the polymorphic sites, and used to genotype 245 lines in the By804/B73 RIL population. MAPMAKER/EXP 3.0 (Lincoln et al. 1992) was used to integrate *ZmcrtRB3* locus into the previously constructed linkage map of By804/B73 RIL population (Chander et al. 2008). The Kosambi mapping function was used for converting recombination values to map distances.

The BLUPs of 16 carotenoid related traits for each line across two environments in the By804/B73 RIL population (Chander et al. 2008) were used for QTL mapping. These traits included α-carotene, β-carotene, β-cryptoxanthin, lutein, zeaxanthin content and composition, total carotenoid content, pro-vitamin A content (sum of α-carotene, β-carotene and β-cryptoxanthin), the α/β-carotene ratio, the α-carotene/lutein ratio, the β-carotene/β-cryptoxanthin ratio, and the β-cryptoxanthin/zeaxanthin ratio. The composite interval mapping model (Zeng et al. 1994) implemented in QTL Cartographer V2.5 (Wang et al. 2005) was used to map QTL for carotenoid-related traits in the RIL population using the linkage map integrated with the M1 marker (*ZmcrtRB3*). Model 6 was used to detect QTL with a window size of 10 cM, five control markers and the forward regression method. For each trait, a genome-wide threshold for declaring putative QTL was estimated by 1 000 random permutations (Doerge and Churchill 1996) at a significance level of 0.05.

Linkage disequilibrium and candidate-gene association analysis

The polymorphic sites (SNPs and InDels) of *ZmcrtRB3* with minor allele frequency (MAR) $\geq$0.05 in 126 maize lines were extracted using TASSEL 2.1.0 (Bradbury et al. 2007). The levels of LD between two sites were calculated in TASSEL 2.1.0. The LD plot of *ZmcrtRB3* was painted by the averaged r^2 within a sliding window with a window size of 400 bp and a step size of 200 bp. The associations between the extracted SNPs and InDels with MAF $\geq$0.05 and 16 carotenoid-related traits were carried out using a mixed model (Yu et al. 2006) incorporating population structure and kinship (Yang et al. 2010) in TASSEL 2.1.0. The phenotypic variations explained by individual SNPs and the effects of haplotypes were estimated by one-way of analysis of variance (ANOVA) in EXCEL.

Real-time qRT-PCR

Twelve different tissues (Table S2) from maize inbred line B73, and embryo and endosperm at 15, 20 and 25 DAP in five selected inbred lines, namely Zheng58, Chang7–2, SC55, CI7 and DE.EX were collected and used for the expression of *ZmcrtRB3*. Furthermore, the embryo and endosperm at 20 DAP in 12 additional inbred lines (Table S2) were also collected to investigate the associations between the expression levels of *ZmcrtRB3* and carotenoid related traits in maize kernel. *ZmcrtRB3* specific primers (Exp 2F/1R, Table S3) were used to perform real-time qRT-PCR using the Ex Taq premix kit (TaKaRa Shuzo, Japan) in ABI 7500 Real-Time FAST PCR System (Applied Biosystems Foster City, CA, USA). Three replicates of each sample were obtained and the $2^{-\Delta\Delta C}{}_T$ method (Livak and Schmittgen 2001) was used to calculate the relative expression levels of *ZmcrtRB3* with actin as the endogenous control. The endosperm at 20 DAP of B73 was used as a reference sample. Correlation analysis was performed using PROC CORR in SAS 8.02 (SAS Institute 1999).

Accession numbers

GenBank sequences of *ZmcrtRB3* in this study are deposited under the accession numbers JQ420132, JQ288306-JQ288417.

Acknowledgements

We thank Dr Subhash Chander for useful comments on this manuscript. This research was supported by the National Hi-Tech Research and Development Program of China (863 Program, 2012AA10A307), the National Natural Science Foundation of China (30821140352, 31123009) and the HarvestPlus Challenge Program in China.

Received 16 Dec. 2011 Accepted 7 Feb. 2012

References

Belbin O, Dunn JL, Chappell S, Ritchie AE, Ling Y, Morgana L, Pritchard A, Warden DR, Lendon CL, Lehmann DJ, Mann DM, Smith AD, Kalsheker N, Morgan K (2008) A SNP in the ACT

gene associated with astrocytosis and rapid cognitive decline in AD. *Neurobiol. Aging* **29,** 1167–1176.

Bradbury PJ, Zhang ZW, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23,** 2633–2635.

Buckner B, Miguel PS, Janick-Buckner J, Bennetzen JL (1996) The y1 gene of maize codes for phytoene synthase. *Genetics* **143,** 479–488.

Chander S, Guo YQ, Yang XH, Zhang J, Lu XQ, Yan JB, Song TM, Rocheford TR, Li JS (2008) Using molecular markers to identify two major loci controlling carotenoid contents in maize grain. *Theor. Appl. Genet.* **116,** 223–233.

Clark RM, Wagler TN, Quijada P, Doebley J (2006) A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet.* **38,** 594–597.

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

Demmig-Adams B, Gilmore AM, Adams WW (1996) Carotenoids 3: *In vivo* function of carotenoids in higher plants. *FASEB. J.* **10,** 403–412.

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142,** 285–294.

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* **32,** 1792–1797.

Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. *Progr. Lipid. Res.* **43,** 228–265.

Hable WE, Oishi KK, Schumaker KS (1998) *Viviparous-5* encodes phytoene desaturase, an enzyme essential for abscisic acid (ABA) accumulation and seed development in maize. *Mol. Gen. Genet.* **257,** 167–176.

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41,** 95–98.

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

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. *Curr. Opin. Plant Biol.* **4,** 210–218.

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

Li QR, Farre G, Naqvi S, Breitenbach J, Sanahuja G, Bai C, Sandmann G, Capell T, Christou P, Zhu CF (2010) Cloning and functional characterization of the maize carotenoid isomerase and β-carotene hydroxylase genes and their regulation during endosperm maturation. *Transgenic Res.* **19,** 1053–1068.

Lincoln S, Daly M, Lander E (1992) Mapping genetic mapping with MAPMAKER/EXP3.0. Whitehead Institute Technical Report, Cambridge, MA, USA.

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}$CT method. *Methods* **25,** 402–408.

Matthews PD, Luo RB, Wurtzel ET (2003) Maize phytoene desaturase and ζ-carotene desaturase catalyse a poly-Z desaturation pathway: Implications for genetic engineering of carotenoid content among cereal crops. *J. Exp. Bot.* **54,** 2215–2230.

Park H, Kreunen SS, Cuttriss AJ, DellaPenna D, Barry JP (2002) Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis. *Plant Cell* **14,** 321–332.

Salvi S, Sponza G, Morgante M, Tomes D, Niu XM, Fengler KA, Meeley R, Ananiev EV, Svitashev S, Bruggemann E, Li BL, Hainey CF, Radovic S, Zaina G, Rafalski JA, Tingey SV, Miao GH, Phillips RL, Tuberosa R (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. USA* **104,** 11376–1138.

SAS Institute (1999) SAS Software, Cary, NC.

Sasabe T, Furukawa A, Matsusita S, Higuchi S, Ishiura S (2007) Association analysis of the dopamine receptor D2 (DRD2) SNP rs1076560 in alcoholic patients. *Neurosc. Lett.* **412,** 139–142.

Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Boström K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Råstam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjögren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316,** 1333–1336.

Singh M, Lewis PE, Hardeman K, Ling B, Jocelyn KC, Mazourek M, Chomet P, Brutnell TP (2003) Activator mutagenesis of the pink scutellum1/viviparous7 locus of maize. *Plant Cell* **15,** 874–884.

Sun Z, Gantt E, Cunningham FXJr (1996) Cloning and functional analysis of the β-carotene hydroxylase of *Arabidopsis thaliana*. *J. Biol. Chem.* **271,** 24349–24352.

Tian L, DellaPenna D (2001) Characterization of a second carotenoid β-hydroxylase gene from *Arabidopsis* and its relationship to the *LUT1* locus. *Plant Mol. Biol.* **47,** 379–388.

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

Tian L, Musetti V, Joonyul K, Magallanes-Lundback M, DellaPenna D (2004) The Arabidopsis *LUT1* locus encodes a member of the cytochrome P450 family that is required for carotenoid epsilon-ring hydroxylation activity. *Proc. Natl. Acad. Sci. USA* **101,** 402–407.

Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat. Genet.* **43,** 1160–1163.

Vallabhaneni R, Gallagher CE, Licciardello N, Cuttriss AJ, Quinlan RF, Wurtzel ET (2009) Metabolite sorting of a germplasm collection reveals the *hydroxylase 3* locus as a new target for maize provitamin A biofortification. *Plant Physiol.* **151,** 1635–1645.

Wang S, Basten CJ, Zeng ZB (2005) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Yan JB, Kandianis CB, Harjes CE, Bai L, Kim EH, Yang XH, Skinner DJ, Fu ZY, Mitchell S, Li Q, Fernandez MG, Zaharieva M, Babu R, Fu Y, Palacios N, Li JS, DellaPenna D, Brutnell T, Buckler ES, Warburton ML, Rocheford T (2010) Rare genetic variation at *Zea mays crtRB1* increases β-carotene in maize grain. *Nat. Genet.* **42,** 322–327.

Yang XH, Yan JB, Shah T, Warburton ML, Li Q, Li L, Gao YF, Chai YC, Fu ZY, Zhou Y, Xu ST, Bai GH, Meng YJ, Zheng YP, Li JS (2010) Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. *Theor. Appl. Genet.* **121,** 417–431.

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38,** 203–208.

Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* **136,** 1457–1468.

Zhou LL, Zhang JY, Yan JB, Song RT (2011) Two transposable element insertions are causative mutations for the major domestication gene *teosinte branched 1* in modern maize. *Cell Res.* **21,** 1267–1270.

Zhu CF, Yamamura S, Nishihara M, Koiwa H, Sandmann G (2003) cDNAs for the synthesis of cyclic carotenoids in petals of *Gentiana lutea* and their regulation during flower development. *Biochim. Biophys. Acta* **1625,** 305–308.

(Co-Editor: Jose Luis Araus)

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Summary of significant associations between polymorphisms of *ZmcrtRB3* and carotenoid related traits in maize kernel.

Table S2. List of materials used for expression analysis.

Table S3. List of primers used in this study.

Figure S1. *ZmcrtRB3* was located within a quantitative trait locus (QTL) cluster controlling carotenoid related traits in the By804/B73 recombinant inbred line (RIL) population.

Figure S2. Linkage disequilibrium (LD) plot of the average r^2 against the physical distance for *ZmcrtRB3* in 126 yellow maize lines.

Figure S3. Gene structure and linkage disequilibrium (LD) matrix of *ZmcrtRB3*.

Figure S4. Polymerase chain reaction (PCR) assays for SNP2172 in *ZmcrtRB3*.

Figure S5. Expression pattern of *ZmcrtRB3* in embryo and endosperm at 15, 20 and 25 d after pollination (DAP) in six maize inbred lines.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.