GENETIC DISSECTION OF CAROTENOID CONCENTRATION AND COMPOSITIONAL TRAITS IN MAIZE GRAIN

BY

CATHERINE KANDIANIS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Doctoral Committee:

Professor Torbert R. Rocheford, Chair
Professor Martin O. Bohn
Professor G. Richard Johnson
Professor Michael R. Murphy

ABSTRACT

Carotenoid compounds are derivatives of a well described secondary metabolic pathway in plants, participating in a diverse array of physiological functions, and are nutritionally valued vitamin precursors in the human diet. With the goal of enhancing the quantity (concentration) and quality (composition) of carotenoids in consumable plant tissues such as grain, breeding approaches have sampled from the extensive phenotypic variation that exists for these traits among maize inbreds. The predominant carotenoids in maize grain include lutein and zeaxanthin, which are collectively called di-hydroxy xanthophyll carotenoids, as well as β-carotene, α-carotene and β-cryptoxanthin, which are proVitamin A carotenoids. While phenotypic sampling and recombination across diverse maize germplasm has been successful, response to phenotypic selection has not been large enough to satisfy target nutritional levels. Greater phenotypic gain may be more predictably achieved if: 1) the genetic network regulating biological functions which contribute to carotenoid accumulation was better understood; and 2) the relative effect of the genetic loci controlling this quantitative trait was known, especially in varying genetic backgrounds. To this end, an investigation of the genetic basis for variation in carotenoid concentration and composition was performed in the studies presented here.

Two QTL analyses revealed gene networks likely involved in carotenoid biosynthesis, conversion, and degradation as the primary drivers of variation in carotenoid concentration and composition. Investigation of one QTL hotspot on maize chromosome 9, observed to account for significant phenotypic variation in almost all carotenoid intermediates, revealed an allelic series associated with a large reduction in the major carotenoid intermediates of maize grain. This QTL, encoding *carotenoid cleavage dioxygenase* 1 (*ccd1*), showed a substrate preference for all measured carotenoids except β-carotene. We further evaluated allelic variation affecting carotenoid composition focusing on two genes within the carotenoid biosynthesis pathway at *lycopene epsilon cyclase* (*lcyε*) and *beta-carotene hydroxylase* (*crtRB1*). Variation at each of these genes was found to significantly affect carotenoid ratios or intermediates hypothesized to change through predicted substrate-enzyme interactions. Several other traits including total carotenoid concentration, were unexpectedly affected. Using allele-specific marker assisted selection at *lcyε* and *crtRB1* in synthetic populations developed for high total carotenoid content lead to a 3.95-11.33-fold improvement in β-carotene and 1.45-3.22-fold improvement in

proVitamin A for the favorable genotypes. The combined information from these studies highlights new genetic targets for further improvement of carotenoid concentration and composition, and provides guidelines for the selection and recombination of desirable genetic variation in breeding germplasm.

For Trish and Michael, my "familee"

ACKNOWLEDGEMENTS

This research could not have been possible without the guidance and encouragement of my advisor, Dr. Torbert Rocheford, who provided me with an opportunity follow my interests in basic biochemistry and molecular biology while helping me to gain an appreciation for applied genetics through breeding. I greatly appreciate the help of Dr. Dick Johnson, who spent many occasions helping me to fill the gaps in my understanding of breeding and genetics. As instructors of my quantitative genetics and mathematical modeling classes, and as members of my committee, I also appreciate the help of Dr. Martin Bohn and Dr. Michael Murphy in equipping me with tools that were necessary for the completion of this dissertation.

I could not have accomplished the field work, laboratory experiments and data analysis without an entire troop of past and present Rocheford and Bohn lab members, especially the carotenoid project members including Patricia Bermudez, Robyn Stevens, Kristin Chandler, Pete Ortegel. Dr. Debra Skinner and Dr. Kevin Montgomery. Soon to be Dr. Kevin Armstrong was a comrade in the daily struggle of statistics, experimental design, and Turner Hall life! Colleagues from the Buckler Lab at Cornell University, the DellaPenna Lab at Michigan State University and the Maize Breeding Group supervised by Dr. Kevin Pixley at CIMMYT facilitated many aspects of this project, and made the collaboration extremely enjoyable. I would also like to thank the College of ACES for sponsoring my research through the Jonathan Baldwin Turner Fellowship, as well as the Harvest Plus program for research funds and workshops.

To my family, I cannot adequately represent the joy and encouragement that you have brought to my life. When complete, this dissertation will have the mark of several Bermudez and Kandianis family members, through help in field transplanting and collection, greenhouse pollinations, discussion of database construction, midnight mathematical modeling forays, or fantastic HPLC troubleshooting. The mere fact that you proudly display orange corn in your homes always makes me smile widely.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

Chapter 1

LIST OF FIGURES

Appendix I

CHAPTER 1

Genetic Architecture Controlling Maize Grain Carotenoid Accumulation in Two QTL Mapping Populations

1.1 Abstract

The genetic basis for variation in maize grain carotenoid concentrations was investigated in two $F_{2:3}$ populations, DEexp x CI7 and A619 x SC55, which represent high total carotenoid and high β-carotene inbred lines. Significantly higher carotenoid concentrations were found in grain samples from both populations that were harvested from a subtropical environment, El Batan, Mexico, in comparison to a temperate environment, Urbana, Illinois. Genotype by environment interactions were significant for most carotenoid traits. Using data from the subtropical environment, a total of over 50 quantitative trait loci (QTL) were identified for single carotenoid traits and ratios in each population, using additive, dominance and additive by additive genetic models. A multivariate approach for these correlated traits was taken, using carotenoid trait principal components (PCs) that jointly accounted for 97% or more of trait variation. Component loadings for carotenoid PCs could be interpreted in the context of known pathway relationships. Importantly, QTL for univariate and multivariate traits were found to cluster in close proximity to map locations of known methyl-erythritol pathway (MEP), isoprenoid pathway and carotenoid pathway biosynthesis genes. Several of these genes, including *lycopene epsilon cyclase*, *carotenoid cleavage dioxygenase1* and *beta-carotene hydroxylase*, were mapped in the segregating populations. These three genes showed pleiotropic effects on α-branch carotenoids, total carotenoid profile and β-branch carotenoids, respectively. By comparing the known map positions of genes involved in the MEP, isoprenoid and carotenoid pathways with the genetic location of detected QTL, we speculated on the function of these functionally unidentified trait effects. Our results confirmed that several QTL are involved in the modification of carotenoid profiles, and suggest potential genetic targets that could be used for improvement of total carotenoid and β-carotene in future breeding populations.

1.2 Introduction

Selection of yellow grain varieties is noted as a hallmark of maize domestication [1]. The yellow pigmentation, attributed to an accumulation of carotenoids in the endosperm has resulted

from a gain of function mutation in the primary biosynthesis reaction at the *y1* or *psy1* locus, which encodes the first, rate-limiting enzyme in the carotenoid pathway, phytoene synthase [2] (Figure 1.1). Although this change was phenotypically dramatic, it did not have any perceivable effect on overall plant physiology [3]. White and yellow maize production in the US was approximately equivalent until discovery of the advantage of increased Vitamin A nutrition in yellow maize grain in the 1930s [4], which led to the selection of pigmented grain as the feed of choice for animals [5, 6].

Plant based carotenoids are now widely recognized for their antioxidant and nutritional qualities, which include provitamin A (proVA) activity [7]. Upon symmetrical cleavage, carotenoids with proVA activity can produce one or two retinyl groups which are the structural base for Vitamin A molecules (Figure 1.2). ProVA carotenoids can be found in many plant based foods and include β-carotene (2 retinyl groups), β-cryptoxanthin (1 retinyl group) and α-carotene (1 retinyl group). Vitamin A, also known as retinol, is involved in immune function, protection of vision systems, and cellular differentiation [8, 9]. The conversion of proVA carotenoids to Vitamin A is facilitated by cleavage reactions specific to animal metabolism [10, 11]. Therefore, pre-formed Vitamin A can only be obtained by consumption of animal food products or synthetic supplements.

Vitamin A yield from consumed proVA is dependent upon the starting concentrations of proVA in plant foods, bioavailabilty of the carotenoid from the food matrix, and bioconversion of proVA through degradation [12]. The reduction in Vitamin A attributed to these processes would appear to require the consumption of substantial amounts of carotenoids to satisfy human dietary Vitamin A needs. However, daily Vitamin A requirements can be satisfied by plant based carotenoids through several servings of fruits, vegetables or cereals [13, 14]. There is no upper limit to the amount of plant based carotenoids that can be safely ingested, in contrast to the toxic levels caused by excessive consumption of Vitamin A [15]. Additional health benefits result from consumption of lutein and zeaxanthin carotenoids, which have been shown to aid in the prevention of macular degeneration [16, 17]. Therefore, from a nutritional standpoint both the quantity and quality of carotenoid profiles are important when plant based foods are being used to satisfy dietary requirements.

Maize grain carotenoid concentrations are among the highest produced in cereals [18]. Maize exhibits considerable diversity in the composition of the major carotenoids, lutein and

zeaxanthin, in addition to those of the proVA carotenoids [19]. Compositional and concentration differences in seed carotenoids are observed to be quantitatively inherited [20]. Carotenoid production in maize appears to be largely regulated by phytoene synthase at the *y1* locus [2, 21]. Evidence from model plant species suggests that upstream isoprenoid biosynthesis may play a substantial role in regulating substrate allocation to the carotenoid pathway [22, 23]. Many endosperm color mutants accumulating various intermediates of the carotenoid pathway have been identified and some have been biochemically characterized. These mutants have been summarized by Robertson [24] and include viviparous2 (*vp2), viviparous5 (vp5), viviparous9 (vp9), pink scutellum (ps), white3 (w3), albescent (al), yellow9 (y9)* and *yellow8* (*y8)*. Quantitative trait locus (QTL) analysis has been employed in various studies using populations that segregate both in seed color [25] and carotenoid composition [26-28]. Results indicate that both total carotenoid and profile differences arise from allelic variation at several QTL, and that some of these loci exhibit pleiotropic effects on multiple carotenoids. Evidence that pleiotropic QTL effects on carotenoids can originate from within the biosynthetic pathway was shown for carotenoid biosynthesis genes *lycopene epsilon cyclase* (*lcyε*) [19] and *beta carotene hydroxylase* (*crtRB1*), [29] where association mapping found multiple ratio and branch traits to be affected by a single locus. This suggests that further investigation of the genetic basis for carotenoid trait QTL should exploit the information known for the maize carotenoid pathway.

Carotenoid production occurs within plastid organelles, in close proximity to the site of glycolysis. As depicted in Figure 1.1, three-carbon molecules from glycolysis are committed to the methyl-erythritol phosphate (MEP) pathway through deoxyxylulose synthase (DXS) and deoxyxylulose reductase (DXR) [30]. The MEP pathway synthesizes five-carbon isoprenoids in the form of isopentyl diphosphate (IPP), which isomerizes through IPPI to substrate DMAPP; both isomers form a common pool shared by secondary metabolite pathways synthesizing plant sterols, terpenes, gibberellins, and carotenoids [23]. Condensed IPP units that form geranylgeranylpyrophosphate (GGPP) through a synthase reaction (GGPPS) are committed to the carotenoid pathway upon their condensation by phytoene synthase (PSY). This is the first reaction leading to the synthesis of carotenoids and it is highly regulated both in terms of gene expression and enzymatic activity [21, 31].

Within the carotenoid pathway, PSY produces a 40-carbon unit that undergoes desaturation and isomerization to form lycopene. The resultant conjugated double bond system

contributes to the light absorptive properties, which is characteristic of colored carotenoids found in yellow maize grain [32]. Lycopene undergoes cyclization through lycopene beta cyclase (LCYβ). Subsequent cyclization by LCYβ leads to β-carotene and its hydroxylated derivatives β-cryptoxanthin and zeaxanthin. Lycopene epsilon cyclase (LCYε) produces α-carotene, zeinoxanthin and lutein. Products downstream of zeaxanthin can undergo reversible epoxidation reactions through zeaxanthin epoxidase (ZEP) in the photosynthetic xanthophyll cycle and can continue through a series of cleavage reactions to form the plant hormone abscisic acid. Carotenoid intermediates are reported to also exit the pathway through a degradation mechanism through carotenoid cleavage dioxygenases (CCDs), as shown in Figure 1.1.

Maize homologs encoding most pathway steps have been well characterized in recent years. These include: phytoene synthase 1 [21, 33] (PSY1, locus *y1,* bin 6.02), phytoene desaturase [34, 35] (PDS, locus *vp5*, bin 1.02), ζ-carotene desaturase [36] (ZDS, locus *vp9,* bin 7.02), lycopene β-cyclase [37] (LCYβ, locus *ps1/vp7*, bin 5.04), carotenoid Z-isomerase [38] (Z-ISO, locus, *y9/y12*, bin 10.02), carotenoid cleavage dioxygenase 1 (CCD1, locus *wc,* bin 9.07), lycopene ε-cyclase [39] (LCYε, locus *lcyε*, bin 8.05) and β-carotene hydroxylase 1 [29, 40] (CRTRB1, locus *crtRB1/hyd3*, bin 10.06). Several other genes encoding maize enzymes have been cloned and mapped using homology to model crop species including ZEP, CRTISO, DXS, DXR, HDR, HDS, GGPPS, and IPPI as listed in Figure 1.1 [31]. Among these studies, it has been observed that pathway regulation is tightly controlled by PSY and DXS, and DXR [23, 41, 42], but that compositional differences can be affected by each reaction step.

Considering the known carotenoid pathway framework and maize genetic studies conducted thus far, it is evident that naturally occurring allelic variation exists for at least some of these pathway steps, which has been shown to affect phenotypic variation. It is also highly probable that some QTL for carotenoid traits without known functional roles likely are genes within or upstream of the carotenoid biosynthesis pathway. If so, pathway QTL could affect a suite of traits, making these loci prime targets for modification of carotenoid quantity or composition. Under these assumptions and hypotheses we carried out a study to identify QTL responsible for variation in carotenoids using genetic populations that were derived from high total carotenoid and high β-carotene parent inbreds, and included three molecular markers to map *lcyε*, *ccd1* and *crtRB1* loci. We evaluated the effects of subtropical and temperate environments on carotenoid trait profiles, variance components and QTL effects, and employed a

multivariate approach to QTL mapping of carotenoid traits as a possible method to identify major regulatory control points for carotenoid accumulation.

1.3 Materials and Methods

1.3.1. Genetic Materials

Two mapping populations were developed from maize inbred lines selected for superior and complementary carotenoid composition and total carotenoid profiles. This information is based on multi-year surveys of the Goodman-Buckler Diversity Panel [43], otherwise known as the association panel (see Appendix A).

The first population consists of $F_{2:3}$ progeny derived from the cross of DEexp x CI7. DEexp is a full season maturity line developed by Dr. Jim Hawk at the University of Delaware by selfing from a southeastern U.S. region Pioneer Hi-bred F_1 hybrid. CI7 is a late maturing line, developed by the USDA-ARS and derived from a backcross of (L317 x 33-16) L317. Both lines are among the Goodman-Buckler Diversity lines highest in β-carotene concentrations, and also have high total carotenoid concentrations. Specific combining effects for β-carotene have been observed in the DEexp x CI7 hybrid, including high-parent heterosis for β-carotene in some years (T.R. Rocheford, personal communication).

The second population consists of $F_{2:3}$ progeny derived from the cross of A619 x SC55. A619 is an early maturing line developed in Minnesota, derived from a backcross of (A171 x OH43) OH43, and demonstrates a high total carotenoid profile relative to most lines in the Goodman Diversity Panels in 2003 and 2005 (top 25% and 20% respectively). SC55 is a late maturing line developed in South Carolina with proportionally high β-carotene. Favorable specific combining ability for high β-carotene concentration was observed in hybrid seed of the A619 x SC55 cross as described in Stevens [44] as well as reports by Egesel [45].

The above F_1 crosses were self pollinated to create the F_2 generation. The F_2 seed from a single ear for each cross was planted, and plants were pollinated to create ears of $F_{2:3}$ derived seed. The final population for DEexp x CI7 constituted of 103 $F_{2:3}$ families and A619 x SC55 constituted 227 $F_{2:3}$ families. The population size for the DEexp x CI7 population was only 103 families due to field flooding problems during a winter nursery in Hawaii.

1.3.2. Field Evaluation

In 2005, the 103 $F_{2:3}$ lines of DEexp x CI7 and 227 $F_{2:3}$ lines of A619 x SC55 were grown in two replicates in each of two environments, the University of Illinois Urbana-Champaign (Env. 1) and the International Maize and Wheat Improvement Center (CIMMYT), El Batan, Mexico (Env. 2). Average monthly temperature and precipitation recorded at both locations in 2005 is listed in Table 1.1. The populations were replicated in both environments and the experimental design was an alpha (0,1) incomplete block design. The families were planted in single row plots of 5m length, with 76 cm spacing between rows. Each plot was thinned to a density of approximately 15 plants per 5m, or 43,000 plants ha^{-1}. Seven to nine plants were sib-pollinated within each row. After storage at room temperature for approximately four months, seed was bulked within each row at shelling and an aliquot of approximately 10g of seed from each row was stored at -80^{o}C until vitamin extraction could be performed on a plot mean basis.

1.3.3. DNA Extraction and Genotyping

DNA extraction was performed as described by Stevens [44]. A DNA sample was prepared from a bulk of seed from each $F_{2:3}$ family, and was used for genotyping. A parental survey using 748 publicly available microsatellite markers from MaizeGDB was performed. Pioneer Hi-Bred performed in-kind genotyping on both DEexp x CI7 and A619 x SC55 populations using polymorphic markers identified in the parental survey in addition to several proprietary markers, designated by the prefix "pio". Several SSR markers genotyped at the University of Illinois were added to the genetic map. Allele-specific markers *lcyε*-MZA, *crtRB1*-InDel4, and *ccd1*-5p were used to map three genes in the carotenoid biosynthesis pathway: lycopene epsilon cyclase (*lcyε)* beta-carotene hydroxylase (*crtRB1),* and carotenoid cleavage dioxygenase 1 (*ccd1*), respectively. Marker *lcy*ε-MZA was designed to distinguish a three-nucleotide difference in exon 1 of LCYε, using MZA clone sequence donated in-kind by Pioneer. The targeted polymorphism differed between A619 and SC55 and therefore was used to genotype $F_{2:3}$ progeny of only this cross. Primers for the *lcy*ε-MZA assay included: *LYCe-MZA-P1-L(SC55 specific primer)*, ATT TTT CTG GTA TTT ATT CAG C; *LYCe-MZA-P2(A619 specific primer)*, AAG GCT ACT ACC TCC ATG AAA; *LYCe-MZA-All-R1*, AAT GAG AAT AGT ATG AGA TCG. Marker *crtRB1*-InDel4, was designed to detect a 12 bp indel in exon1 of the *crtRB1* gene segregating in both populations. Primers for the *crtRB1*-InDel4 assay included: *crtRB1-D4-F2*, ACC GTC ACG

TGC TTC GTG CC; *crtRB1-D4-R1*, CTT CCG CGC CTC CTT CTC. Marker CCD1-pro was designed to distinguish a promoter polymorphism in Zm*ccd1;* this marker was only used in the A619 x SC55 population. Primers in the *ccd1*-5p assay used to detect the allelic series of CCD1-pro included: *ccd1-WC-L1*, CCG TGC TCG GAC AGA ATA GT; *ccd1-B73-rev-L1*, CTC ACA CGT GTC AAC GCC; *ccd1-ALL-R1*, GTC GTT TCG GTG GCT GTC. Further assay information can be found in the Appendices B and C.

1.3.4. Phenotypic Data Collection

Carotenoid extraction from whole maize kernels was performed according to the method of Granado [46]. Gradient separation and quantification was performed as described by Stevens [44] at Iowa State University (W. White). Briefly, this method uses an intermediate scale (5 mL) organic extraction (hexane: methylene chloride) and a heated saponification step. Carotenoid separation was performed on a C-30 YMC Waters Carotenoid column, with gradient separation (methanol/ methyl-tert-butyl ether) and photodiode array detection. Measured carotenoid traits included lutein, zeaxanthin, zeinoxanthin, β-cryptoxanthin, α-carotene, trans and cis isomers of β-carotene, phytoene and trans-phytofluene. The summation of all carotenoids with absorption spectra in the visible range (lutein, zeaxanthin, zeinoxanthin, β-cryptoxanthin, α-carotene, total β-carotene) is termed total colored carotenoid in subsequent analyses. The ratio of colorless: colored carotenoids is defined by the proportion of phytoene plus phytofluene to total colored carotenoid. Proportion of alpha branch carotenoids (lutein, zeinoxanthin, α-carotene) to beta branch carotenoids (zeaxanthin, β-cryptoxanthin, β-carotene) is defined by α/β branch. All other ratios are comprised of absolute concentrations.

1.3.5. Genetic Map Construction

Genetic maps for both populations were constructed using JoinMap® Version 3 [47]. Genotype data was screened for segregation distortion; markers with genotype segregation significantly different from the predicted 1:2:1 Mendelian ratio were removed. Maps were created from individual linkage groups using a LOD threshold of 0.001 and recombination frequency of 0.49 as described in Stevens [44] using Haldane's map function. Genetic maps for individual chromosomes were re-calculated with the addition of allele-specific markers. The final linkage map for DEexp x CI7 consisted of 109 markers, spaced with an average interval

length of 15 cM over a total of 1486 cM. The final linkage map for A619 x SC55 consisted of 117 markers, spaced with an average interval length of 16.2 cM over a total of 1728.6 cM. Genetic maps are shown in Figures 1.3 and 1.4.

1.3.6. Phenotypic Data Analyses

All analyses were carried out using statistical procedures in SAS version 9.2 [48] Trait means and ranges by environment and for combined environments were obtained through Proc MEANS. Product-moment correlations of raw trait data were performed using Proc CORR. Analysis of variance for trait data was modeled according to $y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \delta_{k(ij)} + \gamma_l + (\alpha\gamma)_{il} + e_{ijkl}$ where y is the carotenoid phenotype (μg g^{-1}) of an individual in the i^{th} environment, j^{th} replicate, k^{th} block and l^{th} genotype, μ is the population mean, α is the effect of the i^{th} environment, β is the effect of the j^{th} replication in the i^{th} environment, δ is the effect of the k^{th} block in j^{th} replication of the i^{th} environment, γ is the effect of the l^{th} genotype (or family), $(\alpha\gamma)$ is the interaction effect of the i^{th} environment and the l^{th} genotype, and ε is the experimental error. The analyses were performed with Proc MIXED and all model effects were considered random. Variance component estimates for genetic variance (σ_g^2), genotype by environment interaction variance (σ_{ge}^2), and error variance (σ^2) were performed using Proc VARCOMP. Heritability estimates were calculated on an entry mean basis as described by Hallauer and Miranda [49], p.90 by:

$$h^2 = \frac{\hat{\sigma}_g^2}{\frac{\hat{\sigma}^2}{re} + \frac{\hat{\sigma}_{ge}^2}{e} + \sigma_g^2}$$

Best linear unbiased predictors (BLUPs) of $F_{2:3}$ families for all traits were calculated in each population by environment for further use in principal component and linkage analyses. Trait data was modeled according to: $y_{ijkl} = \mu + \beta_j + \gamma_l + e_{jl}$, where β is the effect of the j^{th} replication, γ is the effect of l^{th} genotype (or family), and ε is the experimental error. Block effects were not included as they were not found to significantly account for phenotypic variation in previous models. BLUPs for the Illinois location of A619 x SC55 were not generated and information from this population/ location combination will not be presented, as identical traits were not found to be correlated between Illinois replicates nor between Illinois and Mexico

replicates. This suggests there was a potential plot frameshift error in field planting, harvest, or HPLC analysis.

1.3.7. Principal Component Analysis

Principal component analysis (PCA) was used to reduce a set of highly correlated traits into a smaller set of uncorrelated variables. Using the variance-covariance matrix Σ derived from vector X consisting of p trait BLUPs where: $X' = X'_1 + X'_2 + ... + X'_p$, an orthogonal set of linear combinations (PCs) to account for the maximum variance in Σ was determined. The percentage of variance explained by a PC is equal to the size of the associated eigenvalue, $\hat{\lambda}_i$, as compared to the total variation represented across all PCs. The linear combination is defined by the associated eigenvector $\hat{a}_i$. Elements within the eigenvector $\hat{a}_i$ are component loadings and were used to estimate PC scores for quantitative trait locus mapping, such that: $y_i = \hat{a}_i X'_i$ where y_i represents the PC score, $\hat{a}_i$ the associated eigenvector and X'_i the trait BLUP vector for the ith family.

Trait BLUPs for DEexp x CI7 (Mexico environment) and A619 x SC55 (Mexico environment) were used to generate PCs. Two PCA were run on each population using Proc PRINCOMP in SASv9.2: one consisted of the matrix of trait BLUPs for colored carotenoids (lutein, zeaxanthin, zeinoxanthin, β-cryptoxanthin, α-carotene, β-carotene) in which resultant principal components were labeled by PC-vis (visible); the other PCA consisted of colored carotenoids plus phytoene and phytofluene, in which the resultant principal components were labeled by PC-path (pathway). An evaluation of two trait sets was conducted to determine if there were differences between the generated PCs, as PC-vis consisted strictly of traits typically used for phenotypic selection, whereas PC-path included colorless carotenoids that could have a large effect on trait loadings but typically are not involved in phenotypic selection.

1.3.8. Linkage Analysis by Composite Interval Mapping

Quantitative trait locus analysis was performed using composite interval mapping (CIM) with PLABQTL software [50]. This mapping methodology is based on the Haley-Knott regression method for marker intervals [51], and is supplemented by cofactor selection[52, 53]. Univariate (single

trait BLUP values) and multivariate (PC scores) traits were mapped according to the following model: $y_j = a + B_i x_{ij} + \sum_{k \neq i} g_k z_{kj}$ where y is the trait value of family j (μg g^{-1}), a is the intercept, B_i is the genetic effect of the QTL located within the interval between marker i and marker i+1, x_i is a coded variable indicating the genotype of the marker interval (where maximally the value of parent 1 is -1, and parent 2 is 1), $\sum_{k \neq i} g_k z_{kj}$ is the summation of selected cofactor effects where g_k is the partial regression coefficient of the trait value on marker cofactor k and z_{kj} is a coded variable for the genotype at cofactor k.

A LOD threshold equivalent to an experiment wise Type I error rate of α=0.25 was used as the initial model selection criterion [54] for both populations; to obtain the threshold, 1000 permutations of the trait data were run using a model of specified gene action, cofactor selection [55], and a 2 cM scan. This corresponded to a threshold of LOD 3.5 for DEexp x CI7 analyses and LOD 2.9 for A619 x SC55 analyses. Regression models accounting for digenic additive by additive interactions or dominance effects were tested using the marker set selected by stepwise regression. Model fit criterion including Akaike's Information Criterion (AIC) and adj R^2 were used to select models accounting for greatest phenotypic variation with fewest estimated parameters. For most traits in both populations, an additive model produced an adequate fit to the data, unless otherwise indicated by an additive by additive interaction in the results or dominance effects listed in the caption.

Adjusted R^2 values from the final multiple regression model express the proportion of phenotypic variation adjusted by the number of parameters estimated in the multiple regression model. Additive effects are noted for each QTL, and reflect the partial regression coefficient of the marker genotype on the phenotype. These effects are expressed as a deviation from the second inbred parents, which in this study are CI7 and SC55.

1.4. Results

1.4.1. Comparison of Descriptive Statistics for Traits Across Locations

Means for parents DEexp and CI7 were found to be most variable for lutein, β-carotene, and total colored carotenoid concentrations as well as the ratios for colorless: colored carotenoids, β-carotene: β-cryptoxanthin and α:β branch (Table 1.2). The majority of total

colored carotenoid composition was derived from lutein and β-carotene in both parental lines. A comparison of environments indicated that carotenoid levels in DEexp and CI7 were higher in the Mexico environment than in the Illinois environment for all colored carotenoids. Effects of magnitude and direction were seen in β-carotene concentration for DEexp; β-carotene in DEexp was observed to be 0.61 μg g^{-1} lower than CI7 in Illinois, but it was 2.88 μg g^{-1} higher than CI7 in Mexico. A similar effect of magnitude between environments was observed for carotenoid concentrations within the $F_{2:3}$ populations, such that all analyzed carotenoid traits except α-carotene were found to be significantly different between the $F_{2:3}$ populations grown in Mexico and Illinois environments (α=0.05). All colored carotenoid traits were higher in the Mexico environment. Colorless carotenoids (phytoene and phytofluene) were higher in the Illinois environment for both parent inbreds and the population. Trait concentrations and ratios were normally distributed, and transgressive segregation for all traits was observed in both locations.

Genetic variance components for all traits in DEexp x CI7 population were significant (α=0.05), and large in comparison to genotype by environment and error variance components (Table 1.2). Heritabilities on an entry mean basis for carotenoid concentrations (h^2) ranged from 0.59 for β-carotene to 0.82 for zeinoxanthin. Trait ratios had heritabilities within the same range, and total colored carotenoids were found to have moderate heritability (h^2=0.60). Genotype by environment interaction variances were found to be significant for lutein, β-carotene, phytoene, phytofluene and total colored carotenoids, in addition to the ratios of colorless:colored carotenoids and α/β branch carotenoids.

A619 and SC55 parent means for the Illinois location differed in all carotenoid concentrations and ratios with the exception of zeinoxanthin (Table 1.3). A619 had higher concentrations than SC55 in all carotenoids except β-carotene. Comparing the $F_{2:3}$ population traits across environments, higher carotenoid concentrations were generally observed in the Mexico location. Trait distributions for nearly all carotenoids were skewed toward the direction of the SC55 parent profile. Transgressive segregation was observed for all carotenoids at the higher end of trait distributions.

1.4.2. Trait Correlations

Correlations of carotenoid trait concentrations from the Mexico location for the DEexp x CI7 $F_{2:3}$ families were positive and significant (α=0.05) among nearly all colored carotenoids (Table 1.4). A notable exception was the negative correlation between β-carotene and lutein (*r* = -0.32). α-branch carotenoids (lutein, zeinoxanthin, α-carotene) were highly correlated. Within the β-branch, there was a significant correlation between zeaxanthin and β-cryptoxanthin (r = 0.70) but not zeaxanthin and β-carotene. A low but significant negative correlation was observed between colorless carotenoids (phytoene and phytofluene) and several colored carotenoids.

A619 x SC55 $F_{2:3}$ family carotenoid concentration correlations were positive and significant among nearly all colored and colorless carotenoids. Correlations were found to be insignificant for α-carotene and β-carotene with other carotenoids (Table 1.5).

1.4.3. Principal Component Analysis

Variation within a six trait set of colored carotenoids (lutein, zeaxanthin, zeinoxanthin, β-cryptoxanthin, α-carotene and β-carotene) or an eight trait set of six colored and two colorless (phytoene and phytofluene) carotenoids was summarized using principal components designated as PC-vis and PC-path, respectively. For DEexp x CI7 $F_{2:3}$ progeny, two PCs were found to account for 98% of the total trait variance in the colored trait set and three PCs accounted for 97% of the total variance in the eight trait set (Table 1.6). PC1-vis accounted for 95% of the variance (λ=8.56) in which lutein was the only substantial loading; PC2-vis accounted for 4% of the variance (λ=0.36) and had substantial positive loadings for zeaxanthin (0.18) and β-carotene (0.98). For the path traits, PC1-path accounted for the majority of the variance (90%, λ=8.59) almost entirely through variation in lutein (0.99). PC2-path explained 5% of the trait variance (λ=0.51) through substantial positive loadings in colorless carotenoids (phytoene, 0.91; phytofluene, 0.33) and β-carotene (0.22), whereas PC3-path explained nearly as much variation (4%, λ=0.36) through a positive loadings for lutein (0.23) and β-carotene (0.95), and a negative loading for phytoene (-0.18).

The variance explained across PCs identified in trait matrices for A619 x SC55 $F_{2:3}$ progeny was distributed over three and four PCs for colored and pathway traits, respectively.

PC1-vis accounted for 79% of the variance (λ=5.20) through substantial loadings for lutein (0.91) and zeaxanthin (0.40). PC2-vis and PC3-vis had much smaller eigenvectors and explained variation through loadings in lutein, zeaxanthin and β-carotene (Table 1.7). Of the four PC traits identified for the pathway traits, PC1-path and PC2-path each accounted for 49% (λ=5.92) and 39% (λ=4.73) of the trait variance, respectively, through substantial loadings in lutein, zeaxanthin and phytoene. PC3-path and PC4-path accounted for much less variation (7% and 3% respectively) and had substantial loadings for lutein, zeaxanthin and β-carotene. Eigenvectors for the identified A619 x SC55 PCs demonstrated both synergistic and antagonistic effects through positive and negative loadings.

As both the eigenvalues and eigenvectors appeared to be redundant between the visible and pathway PCs, a correlation of PC trait scores was performed. For both populations, highly positive and substantial correlations were typically found between one PC-vis trait with one PC-path trait, indicating that similar variation was represented by the identified PCs in both trait sets. As pathway PCs were found to account for nearly all the variation explained by the visible PCs in both populations (Table 1.8), only pathway PCs were used for further analysis.

Loadings for path PCs were roughly interpreted in the context of the biochemical pathway. In DEexp x CI7, lutein accounted for most of the pathway variation, whereas smaller trait correlations appeared to be involved in apportioning carotenoid branch precursors to β-carotene (PC2-path) or could be involved in the competition between phytoene precursors and downstream carotenoids (PC3-path). A619 x SC55 PCs were interpreted as a description of synergistic fluxes (PC1-path) or competitive allocation (PC2-path) of phytoene substrate to the largest carotenoid pools (lutein and zeaxanthin). PC3-path described a competition between lutein/ β-carotene versus zeaxanthin, whereas PC4 described a competition between α and β branch carotenoids.

1.4.4. Quantitative Trait Locus Mapping in DEexp x CI7 Population

For the DEexp x CI7 population, marker-genotype association with trait BLUPs from Mexico identified a range of 2-8 QTL that explain significant variation in carotenoid concentrations and ratios (Table 1.9). Multiple regression models accounted for 20.3-55.3% (R^2_{adj}) of phenotypic variation in individual carotenoid concentrations, and 31.8-58.6% (R^2_{adj}) of variation in ratios. Variation for carotenoid concentrations and ratios from the Illinois

environment were modeled at most by a single, large QTL explaining 15.4-38.9% of the variation in concentrations and 20.3-40.4% of the variation for ratios (Table 1.10). For several traits including α-carotene, β-carotene, phytoene and phytofluene, no QTL were found to explain the trait variation among the $F_{2:3}$ progeny grown in Illinois. QTL detected in the Illinois location (Table 1.10) were often found to overlap with those detected in the Mexico location, but not in the case of every trait (Table 1.9). Additive effects for QTL shared between environments for the same trait were found to be of similar magnitude.

Although DEexp and CI7 parent inbred profiles did not vary greatly (Table 1.2), a reasonably large number of QTL were detected for several traits in the derived $F_{2:3}$ progeny for the Mexico environment (Table 1.9). Variation in colorless carotenoids, including both raw concentrations and proportion to colored carotenoids, was attributed to 4-7 main effect QTL which predominantly were located on chromosomes 3 and 4. A large proportion of the favorable additive effects at these QTL were contributed by DEexp. Eight QTL were detected for β-carotene, a number of which individually accounted for more than 10% of the variation for this trait; most of the positive additive effects for these QTL were also derived from DEexp. Variation in lutein, the carotenoid differing most between inbreds DEexp and CI7, was explained by only two large QTL were detected at chromosome 3 (position 70, positive CI7 effect of 1.19 $\mu g\ g^{-1}$) and chromosome 8 (position 66, positive DEexp effect of 2.81 $\mu g\ g^{-1}$). Variation in α/β branch ratio was affected by five main effect QTL, the two largest positive effects originating from the DEexp alleles at chromosome 4 (position 98, R^2_{part}=19.9%) and chromosome 8 (position 64, R^2_{part}=42.1%). A substantial effect on this trait (ratio effect=0.26, R^2_{part}=9.1%) was attributed to an epistatic interaction between chromosome 4 and an effect from CI7 at chromosome 5 (position 100, R^2_{part}=0.5%).. Dominance effects were not found to significantly account for variation in any of the analyzed traits in the DEexp x CI7 population.

Detected QTL in final regression models were found to overlap among carotenoid traits known to be related through branch associations (main or β–branch) or through extent of hydroxylation (carotenes or xanthophylls). For the main pathway branch, phytoene and phytofluene shared many common QTL at chromosome 3 (position 74; position 142), chromosome 4 (position 52), and chromosome 10 (position 54). β-branch carotenoids shared common QTL at chromosome 2 (position 66), chromosome 8 (position 80), and chromosome 10 (positions 56-58). The chromosome 10 QTL interestingly increased the β-carotene: β-

cryptoxanthin ratio through the DEexp allele (ratio effect=0.96, R^2_{part}=26.3%). Conversely, increased zeaxanthin was derived from the CI7 allele by the same QTL on chromosome 10, indicating that this QTL likely affected the transition of substrate through this branch. α–and β-carotenes were both affected by a QTL on chromosome 2 (position 64-66) and another on chromosome 5 (position 82-84) with both positive effects derived from the DEexp alleles. The most pervasive pleiotropic effects on the entire pathway were detected at chromosome 2 (position 32-40) and chromosome 3 (position 68-72), which both increased total colored carotenoids in addition to many individual carotenoids.

QTL mapping for PC1-path, the multivariate trait accounting for 90% of the variance among the carotenoid trait matrix, revealed two major QTL at chromosomes 3 (position 70) and 8 (position 66) to account for 37.8% of the total phenotypic variation (Table 1.11). These effects were nearly identical to those found for lutein alone. PC2-path and PC3-path traits (accounting for a total of 7% of the variation in the carotenoid trait matrix) were found to be significantly affected by 11 (R^2_{adj}=52.1%) and 9 (R^2_{adj}=68.0%) non-overlapping QTL, with much smaller additive effects. Most of these effects were contributed by DEexp, and coincided with univariate trait QTL. Regression models for the PC traits noticeably improved the amount of variation explained in both environments, indicating that covariation among carotenoid traits was more consistent across environments than absolute concentrations.

1.4.5. Quantitative Trait Locus Mapping in A619 x SC55 Population

Trait BLUPs from the Mexico location were used to identify significant marker genotype-phenotype associations for the A619 x SC55 population. A range of 3-6 QTL was found to explain 21.6-56.2% of the variation for individual carotenoid concentrations, and 37.2-47.4% for carotenoid ratios (Table 1.12). A model for total colored carotenoids explained 53.9% of the phenotypic variation through four main effects and one epistatic interaction. Fewer additive by additive interactions were detected in the A619 x SC55 population than in DEexp x CI7. Variation in two traits, β-carotene and β-carotene: β-cryptoxanthin ratio, was found to be better explained by models including dominance effects (Table 1.13 rather than only additive effects (Table 1.12).

As parent inbred profiles between A619 and SC55 differed in almost all carotenoids, significant QTL were examined for pleiotropic effects. A major effect located on chromosome 9

(position 162-166) was found to significantly affect phytoene (1.33 μg g^{-1}, R^2_{part}=19.6%), phytofluene (0.25 μg g^{-1}, R^2_{part}=19.6%), zeinoxanthin (0.08 μg g^{-1}, R^2_{part}=3.3%), lutein (1.74 μg g^{-1}, R^2_{part}=18.5%), β-cryptoxanthin (0.21 μg g^{-1}, R^2_{part}=17.6%) and zeaxanthin (0.46 μg g^{-1}, R^2_{part}=10.4%). Reduction in each of these carotenoid traits was associated with the SC55 allele, which agreed with the low carotenoid profile of inbred SC55.

Pleiotropic effects were observed within pathway branches. A major effect on chromosome 7 (position 46) was found to explain 23.8% of the variation in phytoene and 32.5 % of the variation in phytofluene, with the A619 allele associated with higher levels. A QTL at chromosome 5 (position 70-76) augmented total colored carotenoids (1.00 μg g^{-1}, R^2_{part}=8.1%) and β-carotene (0.38 μg g^{-1}, R^2_{part}=17.5%) through the SC55 allele, and also decreased colorless:colored carotenoids through the same allele (0.12 μg g^{-1}, R^2_{part}=7.3%). A similar antagonistic relationship was observed in the β-branch by a QTL on chromosome 10 (position 88-94), which was observed to increase β-carotene (0.37 μg g^{-1}, R^2_{part}=14.0%) and β-carotene: β-cryptoxanthin (ratio effect = 4.57, R^2_{part}=40%), while decreasing β-cryptoxanthin (0.21 μg g^{-1}, R^2_{part}=17.6%) and zeaxanthin (0.46 μg g^{-1}, R^2_{part}=10.4%) through the SC55 allele. α-branch carotenoids were significantly affected by a QTL located on chromosome 8 (position 88-92), which accounted for significant variation in α-carotene, lutein and α/β branch ratio.

Use of allele specific, or functional, markers for specific carotenoid biosynthesis genes permitted the comparison of gene map locations with detected QTL effects. *Lycopene epsilon cyclase (lcyε)* was mapped to chromosome 8, 92 cM, coinciding with trait effects for several α-branch traits including α-carotene (position 88), lutein (position 88) and α/β branch ratio (position 92). *Beta-carotene hydroxylase* (*crtRB1*) mapped to chromosome 10, 93 cM, close to QTL effects for β-cryptoxanthin (position 88), zeaxanthin (position 88), β-carotene: β-cryptoxanthin (position 88) and β-carotene (position 94). *Carotenoid cleavage dioxygenase 1* (*ccd1*) mapped to chromosome 9, 145 cM, between two clusters of QTL at positions 136 (affecting lutein and zeaxanthin) and positions 162-166 (affecting phytoene, phytofluene, β-cryptoxanthin, zeaxanthin, zeinoxanthin, lutein and total carotenoids).

Genetic effects accounting for variation in four multivariate traits were mapped in A619 x SC55. QTL associated with variation in all PC-path traits consisted of models with 4-5 main effect QTL and a few epistatic interactions (Table 1.14). Models for multivariate traits were

found to account for more phenotypic variation (range 31.1-60.4%) than the univariate carotenoid traits (21.6-56.2%). Variation in PC1-path, which consisted of significant and positive loadings for phytoene, lutein and zeaxanthin, was explained by five QTL with positive effects predominantly originating from the A619 genome. The QTL at chromosome 9 (position 166) with large, pleiotropic effects was observed to affect this PC trait. PC2-path, which consisted of significant positive loadings for lutein and zeaxanthin, and a negative loading for phytoene was explained by five QTL, including one observed to have pleiotropic effects on β-branch carotenoids at chromosome 5 (position 76). PC3-path, which describes the antagonistic variation between zeaxanthin with lutein and β-carotene, was affected by four QTL, one which mapped in close proximity to the *crtRB1* locus on chromosome 10, and another which affected β-branch traits on chromosome 2 (position 128). Competition between α- and β-branches, described by PC4-path, was affected by five main effect QTL and two digenic interactions. Positive effects at all of these QTL were small, and were contributed by SC55 (Table 1.12).

1.4.6. Comparison of Detected QTL to Known Biosynthesis Genes

A total of 63 main effect QTL for carotenoid univariate traits were found to span 29 marker intervals for the DEexp x CI7 $F_{2:3}$ population, in contrast to 22 unique main effects for the multivariate principal component traits. For the A619 x SC55 population, a total of 50 main effect QTL across 29 marker intervals were found to explain phenotypic variation in univariate traits, as compared to 19 main effect QTL for multivariate traits across 18 marker intervals. The map locations for univariate and PC traits often overlapped, but QTL that were unique to PC traits were detected in both populations (Figures 1.5 and 1.6). Many of the detected QTL clusters were found within proximity of mapped carotenoid biosynthesis genes, listed in Table 1.15. Although genetic maps for both populations did not provide sufficient resolution to definitively conclude that the known genes and detected QTL mapped to the same location, inferences for these gene-phenotype relationships were strengthened by comparing the affected traits with the known biochemical functions of previously identified genes (Figure 1.1).

Overall substrate increases to the pathway in population DEexp x CI7 were described by PC2-path, which mapped in close proximity to genes involved in carbon substrate allocation to the MEP pathway (HDR, bin 1.09; DXR, bin 3.04) and isoprenoid substrate allocation to the carotenoid pathway (PSY1, bin 6.01). PC3-path described the balance between colorless

carotenoids and colored pathway intermediates. This trait was found to map near genes involved with carotenoid conversion within the main branch (ZDS, bin 7.02) in addition to conversion of main branch carotenoids to α and β branch carotenoids (LCYβ, bin 5.04). PSY2, an isoform of phytoene synthase in bin 8.07, was also found to map near PC3-path QTL. The largest trait variation, which was observed for lutein, was represented by the PC1-path trait, and coincided with the map locations for DXR (bin 3.04) and LCYe (bin 8.05),, which are both involved in committing carbon substrate to the carotenoid pathway, and to the α-branch, respectively..

Multivariate traits for A619 x SC55 also described overall substrate increases to the pathway through PC1-path and allocation between colorless substrate and colored intermediates through PC2-path. These traits mapped close to genes that are involved in the allocation of substrate precursors to the carotenoid pathway (HDS, bin 5.03; GGPPS1, bin 7.04) as well as one that removes carotenoids from the pathway through degradation (CCD1/*wc1*, bin 9.07). Balance between the accumulation of β-carotene and zeaxanthin explained by PC3-path had associated QTL mapping close to genes involved in the conversion of zeaxanthin including ZEP1(bin 2.04) and crtRB1(bin 10.06) as well as colorless carotenoid conversion by CrtISO1 (bin.4.08). PC4-path described competition between the α– and β– pathway branches; although QTL for this trait were found to map within the same chromosome bins containing genes allocating substrate to the entire pathway (GGPPS1, bin 2.08; DXS1, bin 6.05).

1.5 Discussion

1.5.1 Inferences to Metabolic Functions for Detected QTL

The genetic architecture for high β-carotene and total carotenoids as well as various individual and compositional traits was investigated in two $F_{2:3}$ populations, DEexp x CI7 and A619 x SC55. Through map location comparisons of the detected QTL with those reported for genes of the MEP, isoprenoid and carotenoid pathways, we assessed that a large proportion of the effects detected in this study are likely derived from metabolic QTL. Confidence in the association of known biosynthesis genes with detected QTL was likely to be higher if the trait affected by a particular QTL was one known to be a product or substrate of the associated biosynthesis gene. It must be emphasized, however, that these are putative associations, but could be further tested through association analysis, TILLING or other genetic mutant approaches. As such, functional associations for the detected suite of QTL suggest that

carotenoid concentrations are likely influenced by 1) the allocation of carbon substrate to the carotenoid pathway, and 2) the effect of two known degradation mechanisms at CCD1 and ZEP1. Variation in β-carotene concentration, on the other hand, is largely attributed to 1) allocation of carotenoid substrate from the main branch supplied by phytoene through the bifurcation point at lycopene, 2) competition between α– and β– branches and 3) conversion within the β–branch.

The two investigated populations differed in some of the identified genetic variation controlling total colored carotenoid concentrations. DEexp x CI7 highlighted the effects of possible genetic variation in metabolic steps providing substrate precursors to the carotenoid pathway, including those within the MEP pathway associated with HDR and DXR [22, 23], as well as reactions that alter allocation within the carotenoid pathway including LCYe. PSY1, the first committed step of the carotenoid pathway [21], was found to map near a QTL accounting for variation in phytoene, which provides the precursor substrate of the major colored carotenoids in grain. A large interaction between chromosomes 2 and 3 affecting the total carotenoid trait may have originated from precursor synthesis at DXR and removal of zeaxanthin through ZEP [31]. These single trait QTL were confirmed in multitrait analyses, in which PC traits detected genetic variation describing control of overall flux to the pathway (as with DXR and HDR), or allocation of substrate to competing pathway branches (as with LCYe and ZEP). Phenotypic variation for total colored carotenoids within the A619 x SC55 population, on the other hand, was dominated by a large effect proximal to the CCD1 locus. Considering the origin of the deleterious effect came from the SC55 allele, this may explain why skewed trait distributions were observed for nearly all carotenoids in the $F_{2:3}$ progeny of this cross. Interestingly, multitrait analyses for this population highlighted the potential importance of variation in precursor pathways as compared to minimal detection of QTL associated with biosynthesis pathways upstream of the carotenoid pathway in single trait analyses.

The allocation of carotenoid substrate between branches or its inter-conversion within a branch was found to be heavily controlled by loci mapping to the location of known biosynthesis genes. Both populations shared some of these trait effects including those which mapped in proximity to *crtRB1*. As the functional alleles at these genes are known to segregate between parent inbred pairs (see Goodman Buckler Diversity panel genotypes in Appendix A), it is not surprising to detect this effect in both populations. Variation in β-carotene was regulated by

more QTL than for most other traits. This trait also appeared to be affected by genes associated with precursor biosynthesis, suggesting that both β-carotene and total carotenoids could be simultaneously increased by selecting favorable genetic variation in the MEP and isoprenoid pathways.

We have broadly assumed that many of the detected QTL map within reasonable proximity to MEP, isoprenoid or carotenoid biosynthesis genes through the comparison of physical map or chromosome bin locations. It is plausible for genes other than those in the biosynthesis pathway to cause this variation; however, as the map locations and the expected trait effects closely overlap between the detected QTL and the known genes, there is substantial evidence to suggest that these are metabolic QTL. This has in part been confirmed by association analyses for *lcyε* [19] and *crtRB1* [29] as well as CCD1 (Kandianis, Chapter 2). Future testing of biosynthesis genes hypothesized to underlie the QTL detected in this (and other) studies may likely come to the conclusion that few non-metabolic QTL appear to affect carotenoid trait variation. Some examples of non-metabolic effects modifying carotenoid concentration and composition have been noted in other plant species. Published reports of mutations in DNA binding domain proteins in cauliflower [56] and plastid development in tomato [57] demonstrate that regulatory and developmental factors can also control carotenoid biochemistry in developing curd and fruit, respecitvely. With available maize genomic sequence, it should be possible to search for maize homologs of these genes for further hypothesis testing.

1.5.2 Comparison of Study Results to Other Carotenoid Trait Mapping Studies

Comparison with other published and unpublished mapping studies confirmed that many of the largest detected QTL were found in the vicinity of carotenoid biosynthesis genes. In a QTL analysis of the backcross progeny of (IHO x B73) B73 by Islam [25], a large effect was mapped to chromosome 6, coinciding with the map location of a marker distinguishing an allelic series at phytoene synthase [58] (*y1SSR*). The detection of this effect was not surprising, as the cross was made between a yellow (B73) and white (IHO) endosperm line. However, the significant and large epistatic effects of this locus on other QTL in chromosome bin 8.05 (presumably LCYε) and bin 9.07 (presumably CCD1) were quite prominent and surprising, demonstrating a case where statistical and biological epistasis arise from the same example.

Chander [28] used a set of recombinant inbred lines derived from BY804 x B73 to demonstrate that kernel carotenoids were also highly influenced by the allelic state of *y1SSR* at chromosome 6, as well as QTL in chromosome bin 5.03 (presumably LCYβ) and bin 10.06 (presumably CRTRB1) which largely affected compositional differences. Yan and Kandianis [29] later mapped a functional marker for *crtRB1* to the peak of the trait effect in bin 10.06 for this population. In an $F_{2:3}$ derived cross of W64a x A632, Wong [26] performed a QTL experiment that highlighted QTL to affect compositional differences at chromosome bin 7.02 (reported as ZDS) and bin 8.05, which was later found to be attributed to LCYe by Harjes [19].

Similar to the studies conducted by Wong [26] and Chander [28], gene-specific markers were used in the analyses presented in this chapter to determine if trait effects jointly mapped to the genetic location of specific biosynthesis genes. Functional markers such as those employed in this study are designed to detect sequence variation leading to a phenotypic change [59], and thus are in complete (or near complete) linkage disequilibrium with the alleles associated with trait changes. As the chance for marker-QTL recombination is almost negligible, these markers provide a preliminary test for the functional assignment of a QTL. In this study, allele-specific functional markers were used for *ccd1, crtRB1* and *lcyε* in population A619 x SC55, and for *crtRB1* in population DEexp x CI7. Trait effects typically mapped within six centimorgans of each functional marker, and in some cases mapped to the same position. Several QTL did not map directly onto the marker positions, even though it was highly probable that the biosynthesis gene was the QTL. It is possible that the calculation of incorrect marker order or recombination distances when calculating the genetic map could have resulted from missing functional marker genotype data. In addition, missing or imprecise phenotype data from a small population size could imprecisely position detected QTL effects in the interval mapping analysis. Indeed, markers with more complete map data such as *lcyε* and *crtRB1* typically had smaller marker-QTL recombination distances than those with more missing data such as *ccd1*. With the availability of the B73 maize genome sequence (www.maizesequence.org), it will be possible to design functional markers for proposed biosynthesis gene targets, and to use a similar mapping approach for these genes. As previously mentioned, it is important to keep in mind that this test does not provide strong evidence of causality, and must be used with caution when using small population sizes or sparse marker coverage.

1.5.3 Utility of Multivariate Analysis for Carotenoid Traits

Most published QTL analyses approach discovery of genetic architecture for a given trait using single trait linkage analyses. As the intent for many of these investigations is to address the genetic architecture for a larger biological question, it is quite common for multiple, related trait measurements to be made. Jiang and Zeng [60] reported that information describing the size and system of effects from pleiotropic QTL could be lost through the estimation of single trait effects. They presented an alternative to simultaneously test correlated traits through an extension of composite interval mapping with maximum likelihood [53, 61]. Multitrait mapping approaches using multivariate techniques [62] or transformation of correlated traits to canonical variables which are then used univariately [63] have also been successfully employed. Use of trait correlations in QTL analyses has been reported to result in increased power of QTL resolution, increased precision of parameter estimates, and the simplification of simultaneous hypothesis testing for multiple traits [64]. While multivariate analyses for QTL mapping are largely favorable, it has also been noted that the loss of QTL detection can also occur if the pleiotropic effects are small [63]. Thus, there is good reason to use multitrait and single trait analyses in a complementary manner.

Principal components have been used to generate linear combinations for traits of interest that can undergo univariate analysis by composite interval mapping. In most cases, the simplification of a large system of traits can help in the identification of single QTL controlling variation within the trait system. A small number of these multitrait QTL potentially could be employed more efficiently in marker assisted selection than a multitude of single trait QTL. Upadayayula [65] effectively used multivariate analysis to determine major QTL governing the variation for correlated inflorescence architecture traits. In a study on the genetic basis of tomato fruit shape, Brewer [66] genetically mapped principal components describing the relationship among physical fruit attributes to determine the greatest sources of variation among populations. Principal components for sunflower grain oil content and composition were used by Ebrahimi [67] to identify constitutive versus adaptive QTL across water regimes. In all cases, the dimensionality of a larger body of correlated traits was condensed into variables that represented the largest proportion of trait variation through the fewest dimensions.

In this study, a series of principal components (PCs) were identified using the variance-covariance matrix for eight traits. All of the traits measured in this study were related by origin

from the same metabolic pathway and were therefore correlated. Competition within the pathway and rate limitations at certain reaction steps provided sufficient variation to warrant use of a multivariate approach. Three to four PCs were found to explain 97-98% of the total variance, with a single PC explaining as much as 90% of the relevant variance. Composite interval analysis of multivariate traits detected 2-8 QTL for each principal component. Interestingly, a large number of QTL were found to explain small PC traits constituting 4-5% of the trait variation in the multivariate analysis. As several of the detected QTL mapped close to biosynthesis genes that could plausibly affect the carotenoid variation represented by the PC loadings, these QTL did not appear to be artifactual. It is possible that removal of the variation attributed to the first and largest PCs reveal background variation that is attributed to many small genetic effects. Multitrait and more than one single trait QTL were frequently found to cluster at the same map location which was suggestive of a pleiotropic QTL. Alternatively, closely linked QTL was perceived to be less likely, as a single candidate gene derived from the MEP, isoprenoid or carotenoid pathway was typically located within a QTL cluster. Several loci with significant effects were uniquely detected for multivariate traits in the A619 x SC55 population, suggesting that single trait analyses lacked the power to detect these loci.

1.5.4 Environmental Effects on Grain Carotenoids and Implications on Physiology

A comparison of trait means and distributions for both populations across two environments revealed significantly higher carotenoid concentrations for samples harvested from the Mexico environment than the Illinois environment. Menkir [68] and Chander [69] reported significant variance components for genotype by environment interaction for carotenoid concentrations in adapted germplasm. Menkir [70] noted that environmental effects were more frequently found to occur in germplasm that was not adapted to the test environments. Results from the study presented here, as well as those from hybrid performance studies across subtropical and temperate environments presented in Chapter 3, suggest that carotenoid concentrations in maize grain grown in subtropical (El Batan, Mexico) or tropical (Puerto Vallarta, Mexico) environments are higher than in temperate environments (Urbana, Illinois) for the same tested genotypes. Although there are no conclusive reasons for this difference, it is evident that differences in day/night temperature, precipitation, UV index, and soil quality, in

addition to cultural practices, can lead to differences of magnitude and direction for carotenoid profiles of various genotypes.

The sensitivity of carotenoid production to environmental stimuli has been reported in various basic physiological studies of *psy1*, the first committed step of the carotenoid pathway, and a major driver of total carotenoid concentrations. In leaves, PSY1 protein has been noted to be mildly induced by heat stress, qualified as an increase from 20°C to 37°C [71]. More strikingly, carotenoid accumulation, or carotenogenesis, occurs in the dark through PSY1 activity, and can be dramatically increased if heat stress is imposed during darkness. This implies that carotenogenesis rates could substantially increase if day/ night temperatures were similar and high. Comparing the average day/night temperatures observed during the 2005 Mexico season and the 2005 Illinois season, it appears that the population grown at the Mexico location likely experienced less heat stress than that grown in the Illinois environment, suggesting that heat-induced carotenogenesis does not necessarily contribute to higher grain carotenoids from the Mexico environment (Table 1.1). The timing of carotenogenesis in developing maize endosperm is well-programmed in inbreds, and has been found to be correlated to PSY1 transcript profiles [72]. This implies that environmental effects occurring during the period of grain fill may modify carotenoid accumulation. As the temperatures are lower at the Mexico location, there is a longer grain filling period, suggesting that there is a longer window for carotenogenesis to be disturbed by environmental stimuli.

There are several interesting observations related to effects of environment on grain carotenoid accumulation that warrant further study. Firstly, ear size for samples grown in Mexico has been noted to be smaller than for the same genotypes grown in Illinois (Kandianis, unpublished observation). If this proves to be the case, it could be possible that the locations differ on a concentration basis, but are more similar on a total content basis. Physiological parameters typically measured for yield and development studies including kernel weights, number of kernels per row, number of kernel rows per ear and ear length should routinely be employed when comparing QTL studies across environments. This information could be used to generate covariate adjusted trait scores. If there is a physiological basis for the elevation of carotenoids in maize grain harvested from tropical or subtropical environments, further investigation of QTL x environment interactions could benefit breeding programs in these regions that are developing high-carotenoid lines for increased nutrient density.

1.5.5 Recommendations of Genetic Targets for Future Carotenoid Breeding Studies

The QTL results presented here are consistent with the concept that modification of grain carotenoid profiles primarily originates from variation in the carotenoid biosynthesis pathway, or metabolism that directly creates and allocates substrate to the carotenoid pathway. As more thoroughly described in Chapter 2, association mapping of sequence variation in candidate genes from the carotenoid biosynthesis pathway has revealed that carotenoid composition is significantly altered by genes regulating α- and β-branch allocation through *lcyε* and β-branch conversion through *crtRB1*. From the associations of mapped QTL effects with map locations of known biosynthesis genes, there is reason to suspect that increased total carotenoids can be achieved through altered metabolism, and that biochemical reactions upstream of the carotenoid pathway must be targeted in order to attain this increase. Characterizing allelic variation representative of MEP and isoprenoid pathways through loci such as DXR should be highly informative and useful in association analysis tests and marker assisted selection studies for total carotenoid grain profiles. Some of the genes in these pathways were detected by PCs representing significant and positive loadings of the major colored carotenoids and phytoene; however, in the case where phytoene and colored carotenoids negatively covary, it may be necessary to include measurement of colorless carotenoids (phytoene and phytofluene) for further genetic analyses.

Degradation components appear to affect total carotenoid accumulation, through select carotenoids on a specific pathway branch (ZEP) or through non-specific reduction of nearly all carotenoids (CCD1). Compositional differences likely can be achieved through modification of intra-pathway steps including cyclization reactions (LCYε, LCYβ) and hydroxylation reactions (CRTRB1). Given that most of the phenotypic variation for grain carotenoids in maize exists in lutein, zeaxanthin, β-cryptoxanthin and β-carotene, the steps identified here primarily affect these intermediates and will likely account for most of the phenotypic variation upon selection of the desired alleles. Marker assisted selection strategies for β-carotene and proVitamin A concentrations are already employing variation at *lcyε* and *crtRB1* to attain desired carotenoid profiles. Given the results of this study, there is merit to further investigate the ZEP locus and *lcyβ* for compositional modification.

1.6 Tables

Table 1.1**:** Average temperatures (^{o}C) and rainfall (cm) of Urbana, IL and El Batan, Mexico environments in 2005

		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
High Temperature (^{o}C)	Urbana	0	3	9	17	23	28	29	28	26	18	10	3
	El Batan	21	22	24	25	26	24	23	23	22	22	22	21
Low Temperature (^{o}C)	Urbana	-9	-6	-1	4	11	16	18	17	12	6	1	-6
	El Batan	7	7	10	11	12	13	13	13	12	11	8	7
Precipitation (cm)	Urbana	5	5	8	9	12	11	12	11	8	7	9	7
	El Batan	0	0	1	2	4	10	12	12	10	4	1	0

Note: Data obtained from CIMMYT (www.cimmyt.org) and The Weather Channel online (www.weather.com)

Table 1.2: Means, variance components and heritabilities of carotenoid traits in DEexp x CI7 $F_{2:3}$ population grown in El Batan, Mexico and Urbana, Illinois in 2005

Trait Means (stdev)		**Lutein**	**Zeaxanthin**	**Zeinoxanthin**	**β-cryptoxanthin**	**α-carotene**	**β-carotene**
Parents							
Env 1	DEexp	14.39	2.16	0.81	0.30	0.51	8.10
	CI7	5.54	2.18	0.59	0.31	0.37	5.22
Env 2	DEexp	5.92	1.50	0.42	0.23	0.28	3.55
	CI7	3.43	1.33	0.34	0.22	0.28	4.16
F_2 Population							
Env 1	Mexico	12.26 (4.09)	2.52 (0.64)	0.71 (0.21)	0.30 (0.07)	0.38 (0.12)	5.41 (1.02)
Env 2	Illinois	8.25 (2.66)	1.73 (0.58)	0.69 (0.21)	0.28 (0.11)	0.39 (0.11)	4.13 (0.49)
	Combined	10.28 (4.00)	2.13 (0.73)	0.70 (0.21)	0.29 (0.09)	0.39 (0.11)	4.77 (1.03)
Trait Ranges (F_2)							
Env 1	Mexico	2.85 - 26.25	1.15 - 6.37	0.31 - 1.74	0.15 - 0.58	0.15 - 0.76	3.14 - 9.28
Env 2	Illinois	1.95 - 17.39	0.86 - 6.51	0.34 - 1.57	0.14 - 1.21	0.16 - 0.80	3.08 - 5.62
Estimated Variance Components (F_2)							
	σ^2_g	5.701**	0.156**	0.024**	0.003**	0.005**	0.179**
	σ^2_{ge}	1.533**	-0.014	0.002	-0.001	0.001	0.052**
	σ^2	4.774	0.233	0.016	0.006	0.006	0.401
Heritability (F2)							
	$h^2_{\text{entry mean}}$	0.74	0.75	0.82	0.79	0.71	0.59

Parent profiles are from one replicate from each environment. Environment designations are: Env 1, El Batan, Mexico; Env 2, Urbana, Illinois. Means and ranges are listed in μg g^{-1} DW. Variance components significant at α=0.01 are indicated (**).

Table 1.2:, continued

Trait Means (stdev)		Phytoene	Phytofluene	Total Colored	Colorless:Colored	β-car: β-crypto	α:β Branch
Parents							
Env 1	DEexp	1.05	0.51	26.28	0.13	27.66	1.49
	CI7	1.88	0.69	14.21	0.18	16.74	0.84
Env 2	DEexp	1.87	0.57	11.89	0.21	15.53	1.26
	CI7	2.21	0.70	9.75	0.30	19.03	0.71
F_2 Population							
Env 1	Mexico	1.34 (0.95)	0.60 (0.35)	21.60 (4.35)	0.10 (0.07)	18.27 (4.48)	1.66 (0.61)
Env 2	Illinois	2.13 (1.09)	0.65 (0.24)	15.49 (2.99)	0.19 (0.12)	15.85 (4.59)	1.54 (0.51)
	Combined	1.73 (1.09)	0.62 (0.30)	18.58 (4.83)	0.14 (0.11)	17.07 (4.69)	1.60 (0.57)
Trait Ranges (F_2)							
Env 1	Mexico	0.17 - 4.54	0.10 - 1.85	10.28 - 36.35	0.01 - 0.43	8.68 - 35.78	0.48 - 3.67
Env 2	Illinois	0.46 - 5.05	0.17 - 1.44	7.12 - 26.19	0.04 - 0.67	2.92 - 31.14	0.41 - 3.02
Estimated Variance Components (F_2)							
	σ^2_g	0.490**	0.043**	5.883**	0.004**	6.091**	0.187**
	σ^2_{ge}	0.362**	0.026**	2.277**	0.004**	-3.754	0.018**
	σ^2	0.191	0.019	5.909	0.002	13.851	0.117
Heritability (F_2)							
	$h^2_{\text{entry mean}}$	0.68	0.71	0.69	0.60	0.79	0.83

Environment designations are: Env 1, El Batan, Mexico; Env 2, Urbana, Illinois. Parent profiles are from one replicate from each environment. Means and ranges are listed in μg g^{-1} DW. Variance components significant at α=0.01 are indicated (**).

Table 1.3: Means of carotenoid traits for A619 x SC55 $F_{2:3}$ population grown in El Batan, Mexico and Urbana, Illinois in 2005

Trait Means (stdev)		**Lutein**	**Zeaxanthin**	**Zeinoxanthin**	**β-cryptoxanthin**	**α-carotene**	**β-carotene**
Parents							
Env 2	A619	6.90	5.78	0.94	1.12	3.54	0.21
	SC55	0.38	0.29	0.90	0.28	0.38	3.09
F_2 Population							
Env 1	Mexico	4.07 (2.58)	1.37 (1.40)	0.92 (0.49)	0.72 (0.34)	0.34 (0.12)	4.94 (1.17)
Env 2	Illinois	2.84 (2.60)	0.92 (0.90)	1.09 (0.67)	0.48 (0.29)	0.33 (0.09)	3.52 (0.52)
	Combined	3.45 (2.67)	1.14 (1.19)	1.01 (0.60)	0.60 (0.34)	0.34 (0.10)	4.21 (1.10)
Trait Ranges (F_2)							
Env 1	Mexico	0.58 - 14.31	0.16 - 8.06	0.25 - 3.78	0.17 - 2.08	0.16 - 0.82	2.86 - 9.48
Env 2	Illinois	0.37 - 10.92	0.22 - 5.24	0.21 - 3.71	0.14 - 1.79	0.17 - 0.62	2.53 - 5.05

Trait Means (stdev)		**Phytoene**	**Phytofluene**	**Total Colored**	**Colorless:Colored**	**β-car: β-crypto**	**α:β Branch**
Parents							
Env 2	A619	4.29	1.57	18.48	0.32	0.18	1.60
	SC55	0.39	0.16	5.33	0.10	11.12	0.46
F_2 Population							
Env 1	Mexico	4.74 (2.87)	1.34 (0.61)	12.37 (4.39)	0.53 (0.34)	8.60 (5.04)	0.74 (0.30)
Env 2	Illinois	2.79 (1.70)	0.97 (0.41)	9.20 (4.18)	0.45 (0.27)	9.84 (5.48)	0.80 (0.44)
	Combined	3.71 (2.48)	1.15 (0.54)	10.77 (4.56)	0.66 (0.48)	7.74 (4.97)	0.77 (0.38)
Trait Ranges (F_2)							
Env 1	Mexico	0.36 - 14.57	0.14 - 3.37	5.54 - 32.39	0.05 - 1.63	2.00 - 32.17	0.21 - 1.88
Env 2	Illinois	0.55 - 9.80	0.30 - 2.27	4.06 - 21.69	0.08 - 1.57	2.11 - 27.02	0.23 - 2.73

Environment designations are: Env 1, El Batan, Mexico; Env 2, Urbana, Illinois. Parent profiles are from one replicate from each environment; only parent profiles from Env 2 were obtained. Variance components and heritabilities are not shown due to issues of genotype assignment as listed in methods.

Table 1.4: Carotenoid trait correlations for DEexp x CI7 $F_{2:3}$ population grown in El Batan, Mexico in 2005

	lut	**zea**	**zein**	**bcry**	**acar**	**bcar**	**total**	**phyene**
zea	0.15 (ns)							
zein	0.79	0.41						
bcry	-0.18 ns	0.70	0.03 ns					
acar	0.48	0.19	0.48	0.19				
bcar	-0.32	0.15 ns	-0.12 ns	0.46	0.32			
total	0.95	0.35	0.85	0.05 ns	0.61	-0.04 ns		
phyene	-0.23	-0.28	-0.28	-0.13 ns	-0.16 ns	0.17 ns	-0.24	
phyflu	-0.19	-0.24	-0.24	-0.10 ns	-0.15 ns	0.14 ns	-0.19	0.97

All correlation coefficients are significant at α=0.05 unless otherwise indicated (ns: non-significant).
Traits are coded for lutein (lut), zeaxanthin (zea), zeinoxanthin (zein), β-cryptoxanthin (bcry), α-carotene (acar), β-carotene (bcar), total colored carotenoid (total), phytoene (phyene) and phytofluene (phyflu).

Table 1.5: Carotenoid trait correlations for A619 x SC55 $F_{2:3}$ population grown in El Batan, Mexico in 2005

	lut	**zea**	**zein**	**bcry**	**acar**	**bcar**	**total**	**phyene**
zea	0.67							
zein	0.36	0.31						
bcry	0.45	0.70	0.33					
acar	0.13 ns	-0.10 ns	0.43	0.15				
bcar	0.14	-0.10 ns	0.03 ns	-0.11 ns	0.51			
total	0.93	0.79	0.47	0.59	0.26	0.30		
phyene	0.45	0.70	0.33	0.99	0.15	-0.11 ns	0.59	
phyfl	0.52	0.45	0.21	0.44	0.33	0.61	0.67	0.44

All correlation coefficients are significant at α=0.05 unless otherwise indicated (ns: non-significant).
Traits are coded for lutein (lut), zeaxanthin (zea), zeinoxanthin (zein), β-cryptoxanthin (bcry), α-carotene (acar), β-carotene (bcar), total colored carotenoid (total), phytoene (phyene) and phytofluene (phyflu).

Table 1.6: Principal component parameters for DEexp x CI7 $F_{2:3}$ population describing carotenoid trait variation

	Colored Carotenoids		All Carotenoids		
	PC1-vis	PC2-vis	PC1-path	PC2-path	PC3-path
Eigenvalues (λ)	8.56	0.36	8.59	0.51	0.36
Total Variance (%)	0.95	0.04	0.90	0.05	0.04
Eigenvectors (a)					
Lutein	1.00	0.06	0.99	0.08	0.05
Zeaxanthin	0.02	0.18	0.02	-0.12	0.23
Zeinoxanthin	0.04	0.04	0.04	-0.01	0.05
β-cryptoxanthin	0.00	0.03	0.00	-0.01	0.03
α-carotene	0.01	0.05	0.01	0.00	0.05
β-carotene	-0.07	0.98	-0.07	0.22	0.95
Phytoene			-0.06	0.91	-0.18
Phytofluene			-0.02	0.33	-0.07

Carotenoid concentrations included in principal components derived for colored carotenoids (vis) and all carotenoids (path) are listed. Substantial component loadings > 0.15 are highlighted in light gray; those < -0.15 are highlighted in dark gray.

Table 1.7: Principal component parameters for A619 x SC55 $F_{2:3}$ population describing carotenoid trait variation

	Colored Carotenoids			All Carotenoids			
	PC1-vis	PC2-vis	PC3-vis	PC1-path	PC2-path	PC3-path	PC4-path
Eigenvalues (λ)	5.20	0.80	0.43	5.92	4.73	0.80	0.42
Total Variance (%)	0.79	0.12	0.06	0.49	0.39	0.07	0.03
Eigenvectors (a)							
Lutein	0.91	0.32	-0.26	0.57	0.71	-0.33	-0.25
Zeaxanthin	0.40	-0.75	0.49	0.26	0.31	0.75	0.49
Zeinoxanthin	0.06	-0.02	0.02	0.03	0.06	0.02	0.02
β-cryptoxanthin	0.07	-0.14	0.08	0.05	0.04	0.15	0.08
α-carotene	0.00	0.04	0.03	0.01	0.00	-0.04	0.03
β-carotene	0.04	0.56	0.83	0.05	0.00	-0.56	0.83
Phytoene				0.76	-0.63	0.00	-0.03
Phytofluene				0.15	-0.09	0.04	-0.05

Carotenoid concentrations included in principal components derived for colored carotenoids (vis) and all carotenoids (path) are listed. Substantial component loadings > 0.15 are highlighted in light gray; those < -0.15 are highlighted in dark gray.

Table 1.8: Correlation of principal component scores from visible carotenoid traits versus all pathway traits within DEexp x CI7 and A619 x SC55 $F_{2:3}$ mapping populations

Population: DEexp x CI7 (n=102)

	PC1-path	PC2-path	PC3-path
PC1-vis	1.00**	0.01	0.00
PC2-vis	0.00	0.23*	0.97**

Population: A619 x SC55 (n=227)

	PC1-path	PC2-path	PC3-path	PC4-path
PC1-vis	0.82**	0.88**	-0.05	-0.02
PC2-vis	0.04	0.04	-1.00**	-0.31**
PC3-vis	0.05	-0.09	0.31**	0.99**

Scores were calculated using Mexico BLUPs.
Significance of Pearson correlation coefficients given for α=0.01 (**), α=0.05 (*). Highly correlated PCs are highlighted in gray.

Table 1.9: QTL detected by composite interval mapping for carotenoid composition traits in DEexp x CI7 population, Mexico environment, 2005

					Phytoene			**Phytofluene**			**Colorless: Colored**		
Chr.	Chr. Bin	Contig	Interval	Pos.	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)
3	3.02/04	111/127	umc1814 - umc1025	40	5.43	0.23	8.6						
	3.04/05	120/125	umc1683 - umc1102	74	11.66	-0.40	20.9	7.68	-0.12	12.6	9.99	-0.02	13.2
	3.08	146	umc1273 - pio6	142	5.06	0.20	9.6	4.62	0.06	3.7	5.49	0.01	4.5
4	4.00		pio7 - pio8	0	4.18	-0.22	11.6						
	4.02/03	156/158	phi295450 - adh2*	48							4.52	-0.02	12.9
	4.03/04	158/164	adh2 - umc2061	52	5.91	-0.23	9.8	4.93	-0.14	17.7			
5	5.04/05	235/239	bnlg1208 - phi333597*	116							3.91	0.03	14.2
6	6.00/02	260/271	umc1018 - umc1083	46	9.11	-0.30	16.7				5.75	-0.01	7.0
8	8.01/02	326/329	umc1483 - umc1034	22							4.92	0.01	5.2
9	9.07/08	391	bnlg1375 - umc1505	112	5.18	0.23	9.3						
10	10.06	415	CrtRB1 - bnlg1028	54	10.75	-0.41	23.0	9.17	-0.13	13.6	8.77	-0.02	12.4
			*Interaction									-0.02	6.5
Model: R^2adj							**48.4**			**27.8**			**41.6**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

Table 1.9:, continued

					β-carotene			**β-cryptoxanthin**			**Zeaxanthin**		
Chr.	Chr. Bin	Contig	Interval	Pos.	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)
2	2.02/04	70/75	umc1756 - umc1026	42				7.95	-0.02	22.2			
	2.06	91	pio4 - umc2194	66	4.8	-0.18	8.7				5.13	-0.23	16.2
3	3.04/05	120/125	umc1683 - umc1102	76	4.2	-0.20	9.7						
4	4.04/05	164/172	umc2061 - umc1895	66	3.67	0.15	7.0						
5	5.03	210/219	umc2060 - umc1692	82	6.89	-0.31	18.8						
	5.06/09	251/254	umc2198 - umc2209	178	3.92	-0.28	18.3						
7	7.02	298/301	umc2327 - phi034	72	8.16	-0.34	24.6						
	7.03	318	bnlg1070 - pio11	92				4.06	-0.02	10.6			
8	8.04/07	353/363	umc1343 - bnlg1828	80	6.82	0.35	18.4	10.43	0.03	25.6			
9	9.01	368/371	bnlg2122 - umc1588	14	3.68	-0.11	3.3						
10	10.06	415	CrtRB1 - bnlg1028	56							3.79	0.16	8.9
	10.06	415	bnlg1028 - phi323152	58				4.02	0.01	5.1			
Model: R^2adj							**55.6**			**37.6**			**20.3**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

Table 1.9:, continued

					α-carotene			**Zeinoxanthin**			**Lutein**		
Chr.	Chr. Bin	Contig	Interval	Pos.	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)
1	1.09	57	umc2028 - pio3	188				4.06	0.05	5.7			
2	2.02	70/75	umc1756 - umc1026*	30				4.47	-0.04	3.7			
	2.04	77	umc1465 - pio4	64	7.59	-0.02	8.5						
	2.06/08	101/103	umc2205 - umc1745	112				3.66	-0.05	8.6			
3	3.04	116/120	umc1025 - bnlg1019*	68				5.04	0.09	18.7			
	3.04	116/123	bnlg1019 - umc1683	70							6.24	1.35	12.7
5	5.03	210/219	umc2060 - umc1692	84	5.71	-0.04	15.5						
6	6.06/07	288/289	umc1296 - 1897	154							3.89	1.02	5.3
8	8.04/07	353/363	umc1343 - bnlg1828	66							5.23	-3.01	34.1
9	9.01	368/371	bnlg2122 - umc1588	14	6.25	-0.02	4.7						
			*Interaction						-0.06	4.3			
Model: R^2adj							**22.0**			**21.7**			**39.5**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

Table 1.9:, continued

					α: β Branch			**β-carotene: β-cryptoxanthin**			**Total Colored**		
Chr.	Chr. Bin	Contig	Interval	Pos.	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)
2	2.02	70/75	umc1756 - umc1026**	36				5.14	0.96	13.2	5.3	-1.33	10.2
3	3.04	116/120	bnlg1019 - umc1683**	70	11.33	0.19	14.8				5.06	1.04	7.8
4	4.08	188	bnlg2162 - umc2187*	98	6.12	-0.24	19.9						
5	5.03/04	219/235	umc1692 - bnlg1208*	100	3.73	0.04	0.5						
7	7.02	298/301	umc2327 - phi034	78	3.64	0.14	8.8						
8	8.03	340	pio13 - bnlg1863	44							3.97	-2.41	29.2
	8.04/07	353/363	umc1343 - bnlg1828	64	10.66	-0.45	42.1						
9	9.01/04	371	umc1588 - umc1107	16							4.61	-0.88	5.4
10	10.06	415	CrtRB1 - bnlg1028	56				9.4	-1.45	26.3			
			*Interaction			-0.26	9.1						
Model: R^2adj							**58.6**			**31.8**			**35.6**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

Table 1.10: QTL detected by composite interval mapping for carotenoid composition traits in DEexp x CI7 population, Illinois environment, 2005

Chromosome		**chr. 2**			**chr. 3**			**chr. 5**			**chr. 8**					
Bin		2.02-2.04			3.04			5.05			8.01			8.01-8.02		
Flanking Markers		umc1756 - umc1026			bnlg1019 - umc1683			phi333597 - umc1941			bnlg1863 -umc1343			umc1343 - bnlg1828		
Chromosome Position (cM)		positions 34-36			position 72			position 126			positions 52-54			position 66		
Genetic Parameters	Model R^2_{adj}	LOD	Add.	R^2	LOD	Add.	R^2	LOD	Add.	R^2	LOD	Add.	R^2	LOD	Add.	R^2
Lutein	38.9										10.63	-2.28	40.3			
Zeaxanthin	15.7	4.14	-0.16	18.0												
Zeinoxanthin	20.2										5.16	-0.10	22.2			
β-cryptoxanthin	15.4	4.03	-0.02	17.2												
α-carotene	*															
β-carotene	*															
Phytoene	*															
Phytofluene	*															
Total colored	30.4										8.09	-2.10	32.4			
Colorless:colored	20.3										5.28	0.07	22.6			
βcar:βcry	28.2	7.72	1.65	31.0												
α/β branch ratio	40.4										11.1	-0.46	41.6			
PC1-path	46.4				3.54	1.48	14.9							12.4	3.05	43.8
PC2-path	*															
PC3-path	20.9							5.58	-0.39	22.5						

Stepwise regression models not found to have any significant regression coefficients (QTL) are marked with *. Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination).

Table 1.11: QTL detected by composite interval mapping for carotenoid principal components in DEexp x CI7 population, Mexico environment, 2005

			PCA1-path			**PCA2-path**			**PCA3-path**		
Chr	Flanking Markers	Pos.	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)	LOD	Add	R^2 (%)
3	bnlg1019 - umc1683	70	6.19	1.20	10.2						
8	umc1343 - bnlg1828	66	4.05	-2.83	31.2						
1	phi109275 - umc2217	100				3.52	-0.22	8.3			
2	umc2246 - bnlg1017	0				3.74	-0.23	9.6			
3	umc1892 - umc1814	28				7.47	0.31	13.5			
	umc1683 - umc1102	74				3.99	-0.64	24.7			
	umc2265 - pio5	90				3.97	0.40	11.3			
	umc1273 - pio6	142				10.07	0.25	14.5			
4	adh2 - umc2061	50				4.92	-0.29	16.0			
6	umc1018 - umc1083	40				9.24	-0.28	18.1			
7	pio9 - umc2327	66				4.53	-0.13	3.3			
9	bnlg1375 - umc1505	112				4.34	0.28	12.8			
10	CrtRB1 - bnlg1028	54				15.62	-0.53	34.2			
2	umc2246 - bnlg1017	0							3.78	-0.18	11.2
	umc1465 - pio4	64							9.14	-0.30	29.5
4	umc1895 - umc2027	68							4.33	0.22	19.7
5	umc2060 - umc1692	82							12.54	-0.42	38.8
	umc2198 - umc2209	178							4.53	-0.25	22.0
7	umc2327 - phi034	74							8.60	-0.33	31.7
8	bnlg1828 - umc1663	94							3.56	0.25	15.1
9	umc1107 - umc1094	52							3.76	0.18	10.6
10	umc1576 - umc1367	22							4.10	-0.25	16.3
Model: R^2adj					**37.8**			**52.1**			**68.0**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination).

Table 1.12: QTL detected by composite interval mapping for carotenoid composition traits in A619 x SC55 population, Mexico environment, 2005

					Phytoene			**Phytofluene**			**Colorless: Colored**		
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)
1	1.11	64	umc2242-umc1979	178	3.30	-0.50	3.2	3.13	-0.16	7.9			
3	3.07	142	pio_5-umc1489	134				5.91	-0.16	8.3			
4	-		pio_7-pio_8	2							3.71	-0.08	8.0
5	5.03	212/217	umc2295-bnlg1892	76							6.09	-0.12	14.7
	5.06	251	pio_9-umc2013*	144				2.91	0.06	1.4			
7	7.02	297	umc1068-bnlg1094*	46	13.10	-1.45	23.8	22.07	-0.34	32.5	6.90	-0.12	13.8
	7.04	323/325	umc1944-umc1125	120	4.30	0.80	9.1				3.68	0.07	5.3
9	9.07	391	pio_13-umc1675	120							3.37	-1.33	3.5
	9.08	391	zct128-umc1505	166	6.59	-1.33	19.6	8.07	-0.25	19.6			
10	10.01	392/393	umc2053-phi059	8							3.33	0.10	8.4
			*Interaction						-0.11	2.3			
Model: R^2adj							**38.1**			**48.0**			**37.2**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

Table 1.12:, continued

					β-carotene[a]			**β-cryptoxanthin**			**Zeaxanthin**		
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)
2	2.04	77	umc1541-pio_4	128				5.63	-0.12	10.1			
	2.05	90	pio_4-umc1459*	130							6.74	-1.00	33.5
	2.07	91/105	bnlg1396-dupssr25	172	5.97	0.34	10.2						
4	4.02	156	pio_08-phi295450	32	6.26	0.39	13.5						
	4.03	158/176	adh2-umc1142	68	4.34	-0.27	6.6						
5	5.03	212/217	umc2035-umc2295	72	10.07	0.38	17.5						
6	6.05	285/287	umc1805-umc1859	90	5.03	0.30	10.9						
7	7.00	293/296	umc1241-umc1068	46				3.24	-0.08	6.7			
9	9.07	391	umc1675-umc2099*	136							4.32	-0.49	11.6
	9.08	391	zct128-umc1505	164				4.07	-0.11	12.4	3.25	-0.41	7.4
10	10.03	400	umc2017-pio_14	62				3.33	-0.06	2.5			
	10.05	414	umc1506-CrtR-B1	88				10.20	-0.21	17.6	4.20	-0.46	10.4
	10.06	417	CrtR-B1-umc1993	94	6.43	0.37	14.0						
			*Interaction									0.79	19.4
Model: R^2adj							**37.1**			**45.5**			**56.2**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

a: indicates that dominance model in Table 1.13 provided a better fit to the β-carotene trait data (AIC: -196.94, BIC:-122.25 , AICC: -190.28, R^2_{adj}: 42.9, k=23) than the additive model (AIC: -186.66, BIC: -141.21, AICC: -184.26, R^2_{adj}:37.1, k=14).

Table 1.12:, continued

					α-carotene			Zeinoxanthin			Lutein		
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)
1	1.03	-	phi339017-pio_1	72				5.75	0.20	14.6			
4	-	-	pio_7-pio_8	0	6.84	0.03	11.7						
5	5.02	208/212	umc1587-umc2035*	64	6.06	0.03	8.3	9.29	0.20	16.0			
6	6.05	283	pio_10-umc1114*	76				3.85	-0.10	4.5			
8	8.04	-	umc1343-LCYe	88	5.93	0.03	9.1						
	8.05	354	LCYe-umc1340	92							5.21	0.64	10.0
9	9.07	391	umc1675-umc2099**	136							3.77	-0.46	1.4
	9.07	391	CCD1-zct128**	162				3.57	-0.08	3.3	9.78	-1.74	18.5
			*Interaction						-0.13	3.5			
			**Interaction									0.65	2.2
Model: R^2adj							**21.6**			**27.0**			**53.6**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

Table 1.12:, continued

					α: β Branch			β-carotene: β-cryptoxanthin[a]			Total Colored		
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)
1	1.01	2/4	umc1177-umc1071*	0	2.93	-0.08	2.3						
2	2.05	77	umc1541-pio_4	128	3.57	0.07	6.7						
4	-	-	pio_7-pio_8	2							3.36	0.75	4.8
5	5.03	212/217	umc2035-umc2295	70							4.43	1.00	8.1
6	6.05	283	pio_10-umc1114*	74	5.05	-0.07	0.7						
8	8.05	354	LCYe-umc1340	92	4.66	0.09	12.9						
9	9.07	391	pio_13-umc1675**	120							6.74	-1.33	7.0
	9.07	391	CCD1-zct128**	162	10.28	-0.17	31.9				12.76	-2.32	23.9
10	10.05	414	umc1506-CrtR-B1	88				10.63	4.57	40.0			
			*Interaction			-0.05	0.4						
			**Interaction									1.32	5.7
Model: R^2adj							**47.4**			**39.4**			**53.9**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

a: indicates that dominance model in Table 1.13 provided a better fit to the β-carotene: β-cryptoxanthin trait data (AIC: 457.96, BIC:474.38, AICC: 458.27, R^2_{adj}: 42.7, k=5) than the additive model (AIC: 467.98, BIC: 481.11, AICC: 468.18, R^2_{adj}:39.4.1, k=4).

Table 1.13: Dominance models for β-carotene and β-carotene: β-cryptoxanthin ratio in A619 x SC55, Mexico environment, 2005

					β-carotene					**β-carotene: β-cryptoxanthin**				
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	Dom.	R^2 (%)	LOD	Add.	R^2 (%)	Dom.	R^2 (%)
2	2.07	91/105	bnlg1396-dupssr25	174	6.10	0.32	10.7	-0.54	0.2					
4	4.02	156	pio_08-phi295450	32	6.31	0.38	13.9	-0.38	0.1					
	4.03	158/176	adh2-umc1142	68	4.34	-0.33	10.6	0.00	0.0					
5	5.03	212/217	umc2035-umc2295	72	10.40	0.39	19.5	0.08	0.1					
6	6.02/04	271/276	umc1178 - umc1918	34	4.13	0.13	2.7	-3.93	11.0					
	6.05	285/287	umc1805-umc1859	90	5.15	0.25	8.1	0.11	0.1					
10	10.05	414	umc1506-crtRB1	88						10.63	4.73	42.90	-1.95	5.9
	10.06	417	crtR-B1-umc1993	94	6.59	0.34	13.1	0.11	0.1					
Model: R^2adj									**42.9**					**42.7**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination).

Table 1.14: QTL detected by composite interval mapping for carotenoid principal components in A619 x SC55 population, Mexico environment, 2005

					PCA1-path			PCA2-path		
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)
1	1.11	64	umc2242-umc1979	182	5.18	-0.55	4.8			
7	7.00	293-296	umc1241-umc1068	44	14.02	-1.30	24.3			
7	7.04	323/325	umc1944-umc1125	122	2.93	0.49	4.8			
9	9.07	391	umc1675-umc2099	138	3.81	-0.95	11.9			
9	9.08	391	zct128-umc1505	166	11.04	-1.63	23.7			
4	-	-	pio_7-pio_8	6				5.79	0.63	5.1
5	5.03	212/217	umc2295-bnlg1892	76				4.15	0.85	9.4
7	7.04	323	pio_11-umc1944	118				4.06	-0.63	5.8
8	8.02	329/345	umc1034-phi115	64				3.36	0.70	6.6
9	9.07	391	pio_13-umc1675	116				3.95	-1.43	20.4
Model: R^2adj							**60.4**			**31.1**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination).

Table 1.14:, continued

					PCA3-path			PCA4-path		
Chr.	Chr. Bin	Ctg	Interval	Pos.	LOD	Add.	R^2 (%)	LOD	Add.	R^2 (%)
2	2.02	70	umc1934-zca381	88	3.29	-0.29	6.6			
2	2.04	77	umc1541-pio_4*	128	8.94	-0.68	23.3			
4	4.08	188/192	umc2187-umc1559*	140	3.07	0.24	4.9			
10	10.06	417	CrtR-B1-umc1993	94	5.32	-0.37	10.3			
			Interaction*			-0.31	3.1			
2	2.08	108	dupssr25-umc1736*	184				4.51	0.2	4.2
4	4.02	156	pio_08-phi295450**	28				4.78	0.2	6.8
5	5.03	212/217	umc2035-umc2295*	72				8.02	0.3	16.2
6	6.02	271/276	umc1178-umc1918**	34				4.23	0.2	4.6
6	6.05	285/287	umc1805-umc1859	90				6.01	0.3	11
			Interaction*						0.3	5.9
			Interaction**						0.2	2.9
Model: R^2adj							**40.9**			**36.1**

Genetic location, significance and effect of selected QTL for each trait are listed. Indicated are LOD (Logarithm of Odds), Add (Additive effect) and R^2 (%) (coefficient of partial determination). Significant digenic interactions detected between main effect QTL are indicated at the bottom of the list of main effects; contributing main effects are marked by asterisk(s).

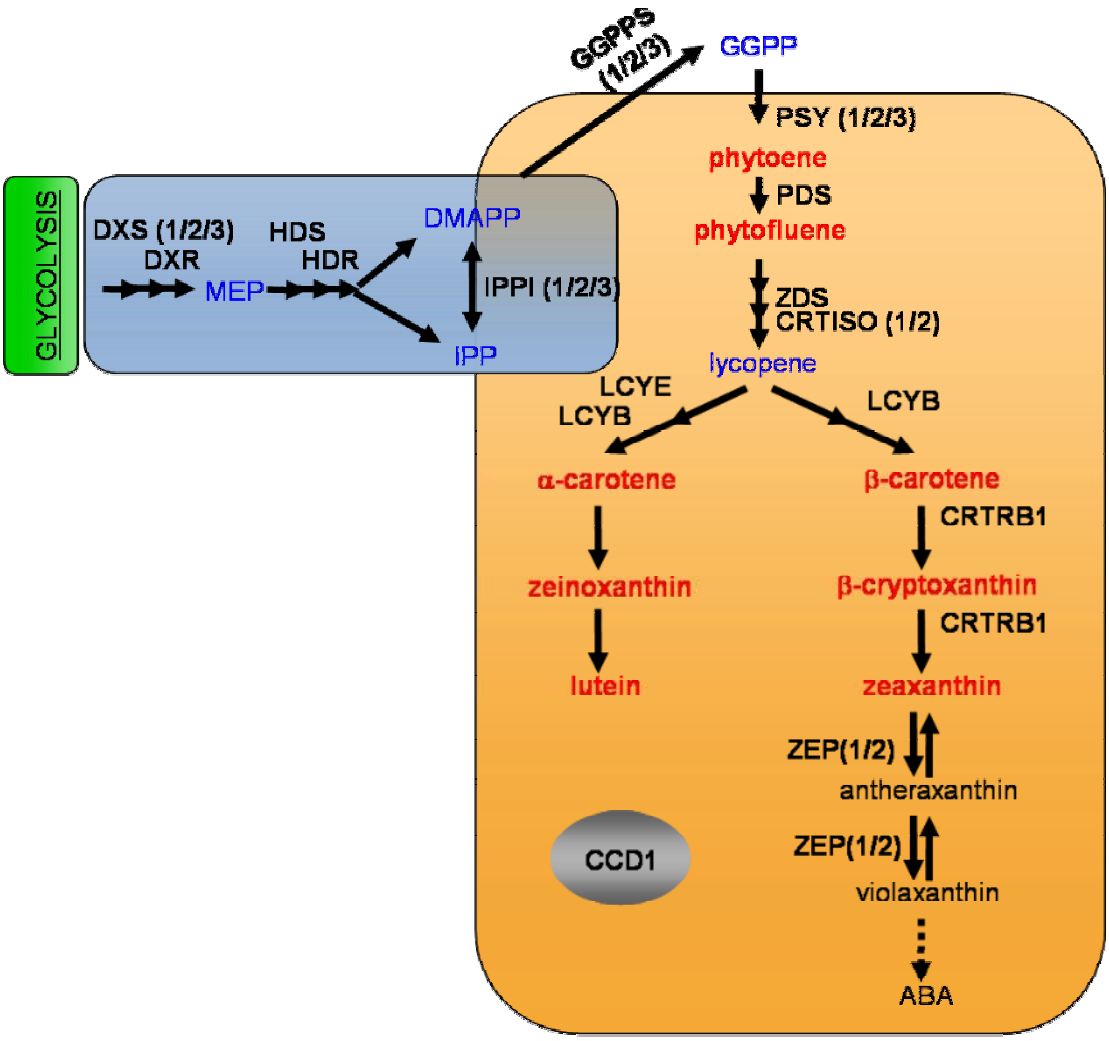
Table 1.15: Map location of carotenoid biosynthesis genes in *Zea mays* and clusters of carotenoid QTL identified in this study

Pathway Enzyme	Gene Location		Traits affected by QTL Mapped to Region	
	Chr. Bin	Contig	DEexp x CI7 $F_{2:3}$	A619 x SC55 $F_{2:3}$
PDS	1.02	ctg8		
HDR	1.09		PC2, zein	
ZEP1	2.04		zea, αcar	PC3, βcry, βcar
GGPPS1	2.08		zein	PC4
CrtISO2	2.09			
DXR	3.04	ctg117	PC1, phyene, phyfl, zein, βcar, lut	
DXS3	4.06			
CrtISO1	4.08			PC3
HDS	5.03		αcar, βcar	PC2, PC4, βcar, αcar, zein
LCYb	5.04	ctg225	PC3, αcar, βcar	
PSY1	6.01	ctg270	PC2, phyene	
IPPI3	6.05			βcar, zein
DXS1	6.05	ctg285		PC4, βcar, zein
ZDS	7.02	ctg297	PC3, βcar	phyene, phyfl
PSY3	7.03	ctg320	βcry	
IPPI1	7.03			
GGPPS2	7.04			PC1, PC2
GGPPS3	8.01			
IPPI2	8.03		total	
LCYe	8.05	ctg355	PC1, βcar, βcry, lut	lut, αcry
PSY2	8.07	ctg363	PC3, βcar, βcry	
DXS2	9.03	ctg376		
CCD1	9.07	ctg391	PC2, phyene	PC1, PC2, phyene, phyfl, βcry, zea, zein, lut, total
ZEP2	10.04			
crtRB1	10.06	ctg415	PC2, phyene, phyfl, βcar, zea	PC3, βcar, βcry, zea,

Pathway enzyme abbreviations are listed in Figure 1.1 caption. Traits are coded for lutein (lut), zeaxanthin (zea), zeinoxanthin (zein), β-cryptoxanthin (βcry), α-carotene (αcar), β-carotene (βcar), total colored carotenoid (total), phytoene (phyene) and phytofluene (phyflu).

1.7 Figures

Figure 1.1: Carotenoid biosynthesis pathway in *Zea mays*



Carotenoids are derived from products of glycolysis (green) and isoprenoid biosynthesis (blue). Substrate is committed to the carotenoid biosynthesis pathway (orange) by phytoene synthase (PSY). Pathway enzymes are shown in black, carotenoid intermediates measured in this study in red, and all other intermediates in blue. Enzymes defined as: DXS, 1-deoxy-D-xylulose-5-phosphate synthase; DXR, DXP reductoisomerase; HDS, 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase; IPPI, isopentyl pyrophosphate isomerase; GGPPS, geranyl geranyl pyrophosphate synthase; PSY, phytoene synthase; PDS, phytoene desaturase; ZDS, zeta carotene desaturase, CRTISO, carotenoid isomerase; LCYe, lycopene epsilon cyclase; LCYb, lycopene beta cyclase; CRTRB1, beta carotene hydroxylase; ZEP, zeaxanthin epoxidase; CCD1, carotenoid cleavage dioxygenase 1. Abbreviated intermediates are: MEP, methyl erythritol 4-phophate; DMAPP, dimethylallyl diphosphate; IPP, isopentyl diphosphate; GGPP, geranyl geranyl pyrophosphate.

Figure 1.2: Molecular structures of Vitamin A and proVitamin A carotenoids

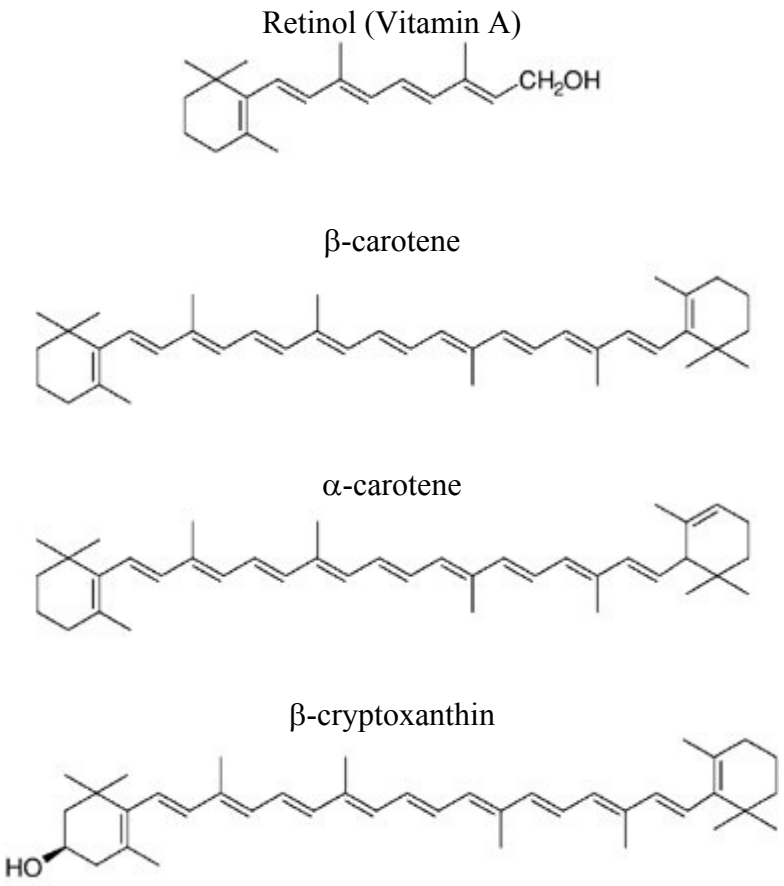


Figure 1.3: Genetic map for DEexp x CI7 $F_{2:3}$ population

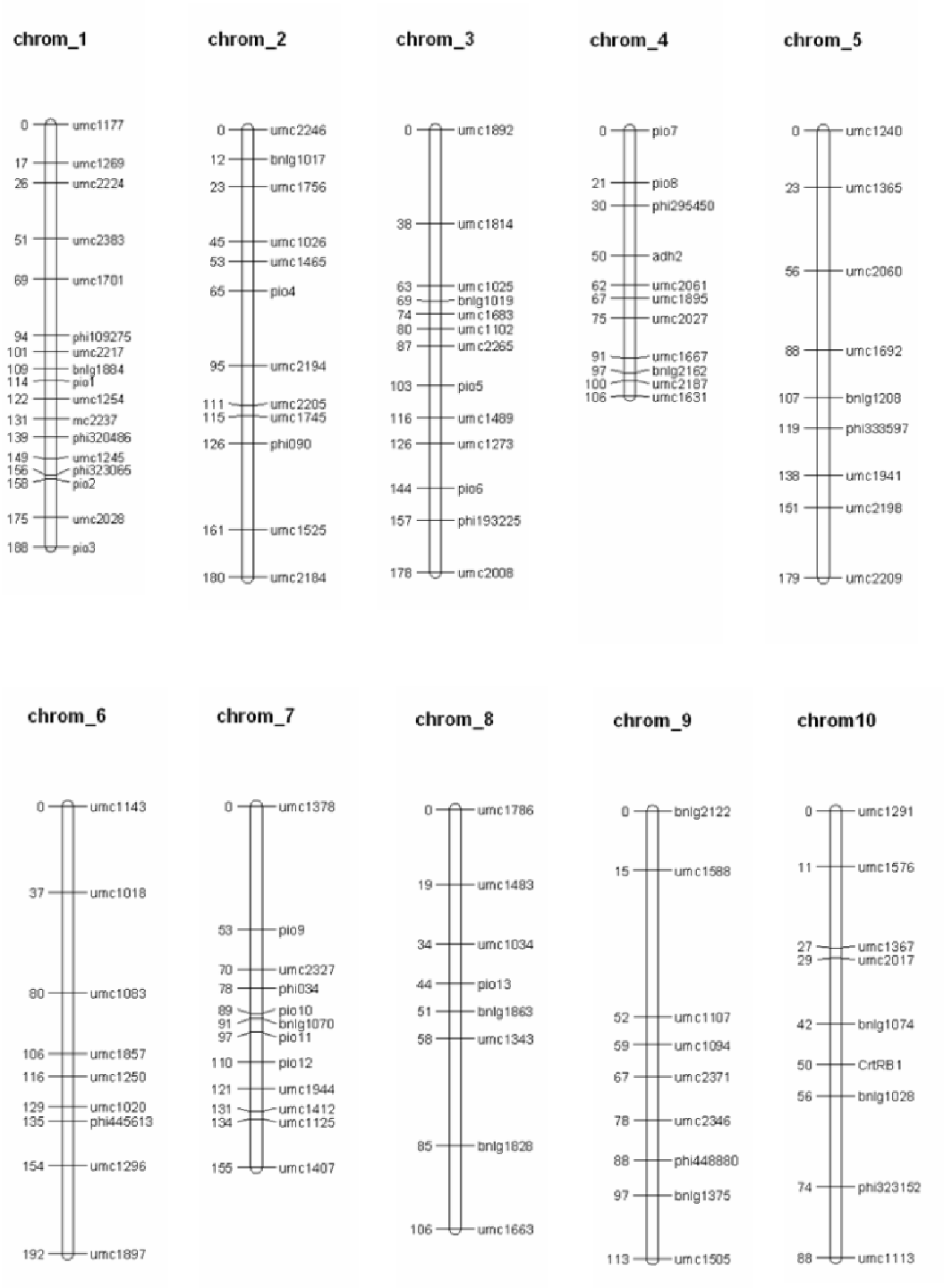


Figure 1.4: Genetic map for A619 x SC55 $F_{2:3}$ population

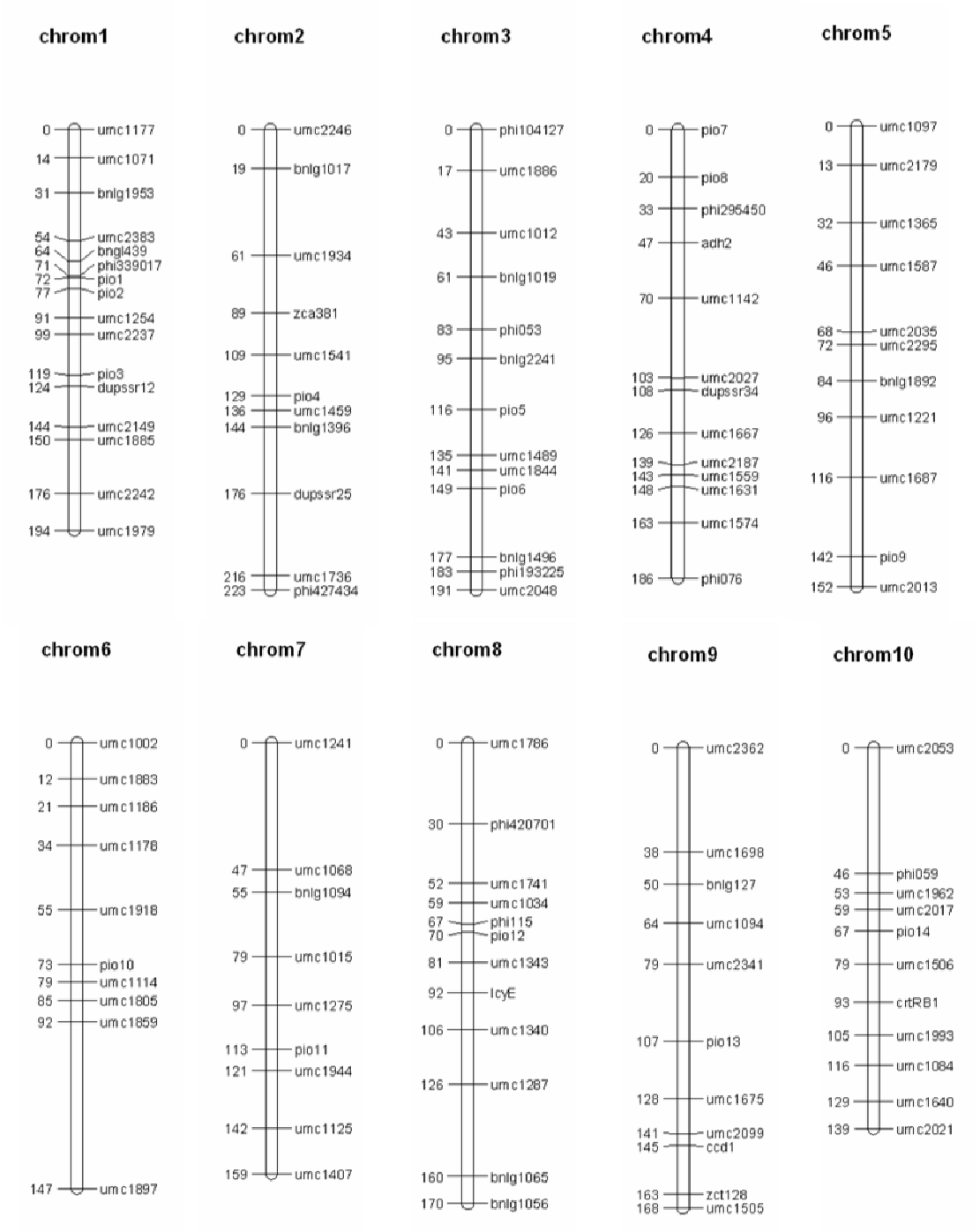
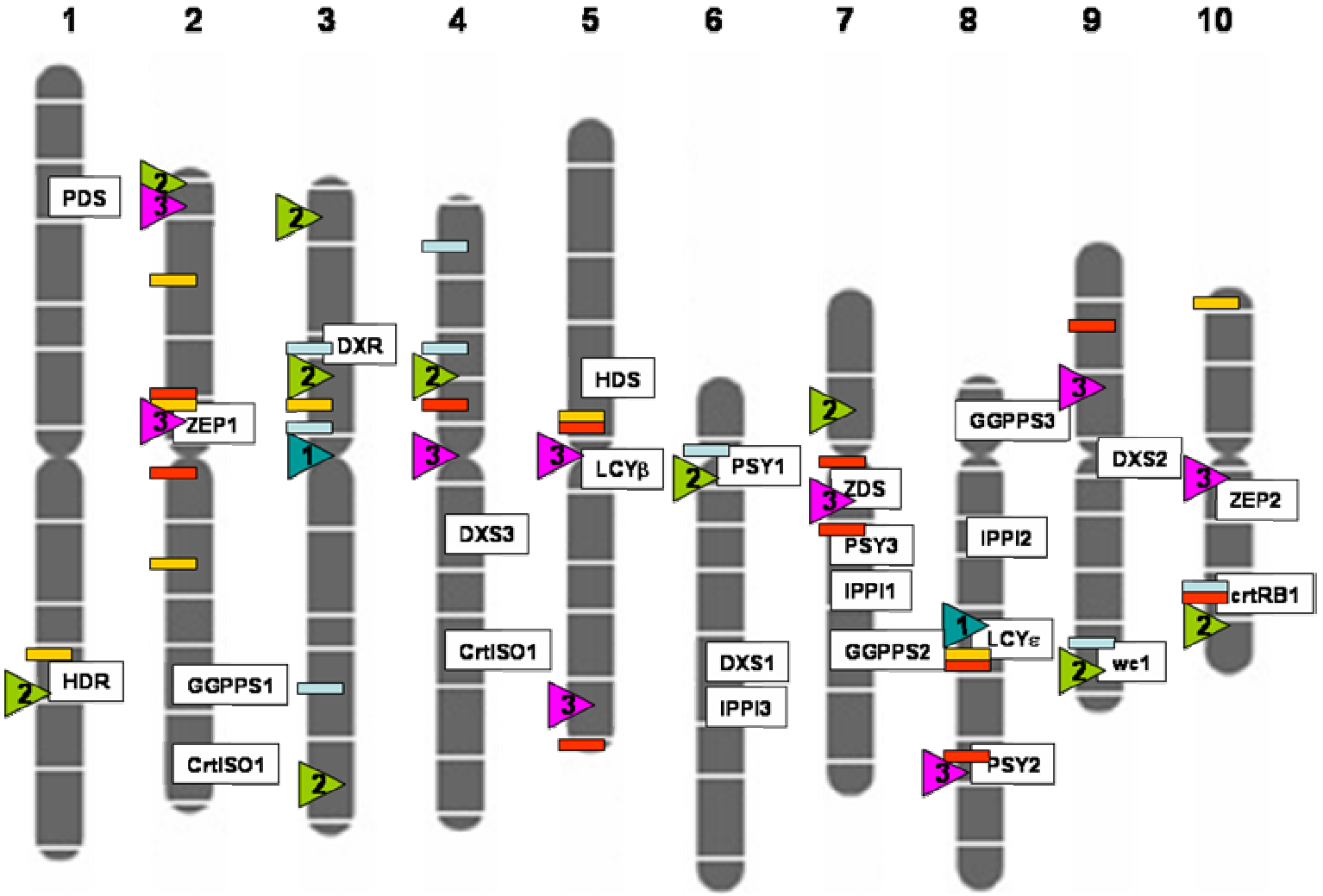
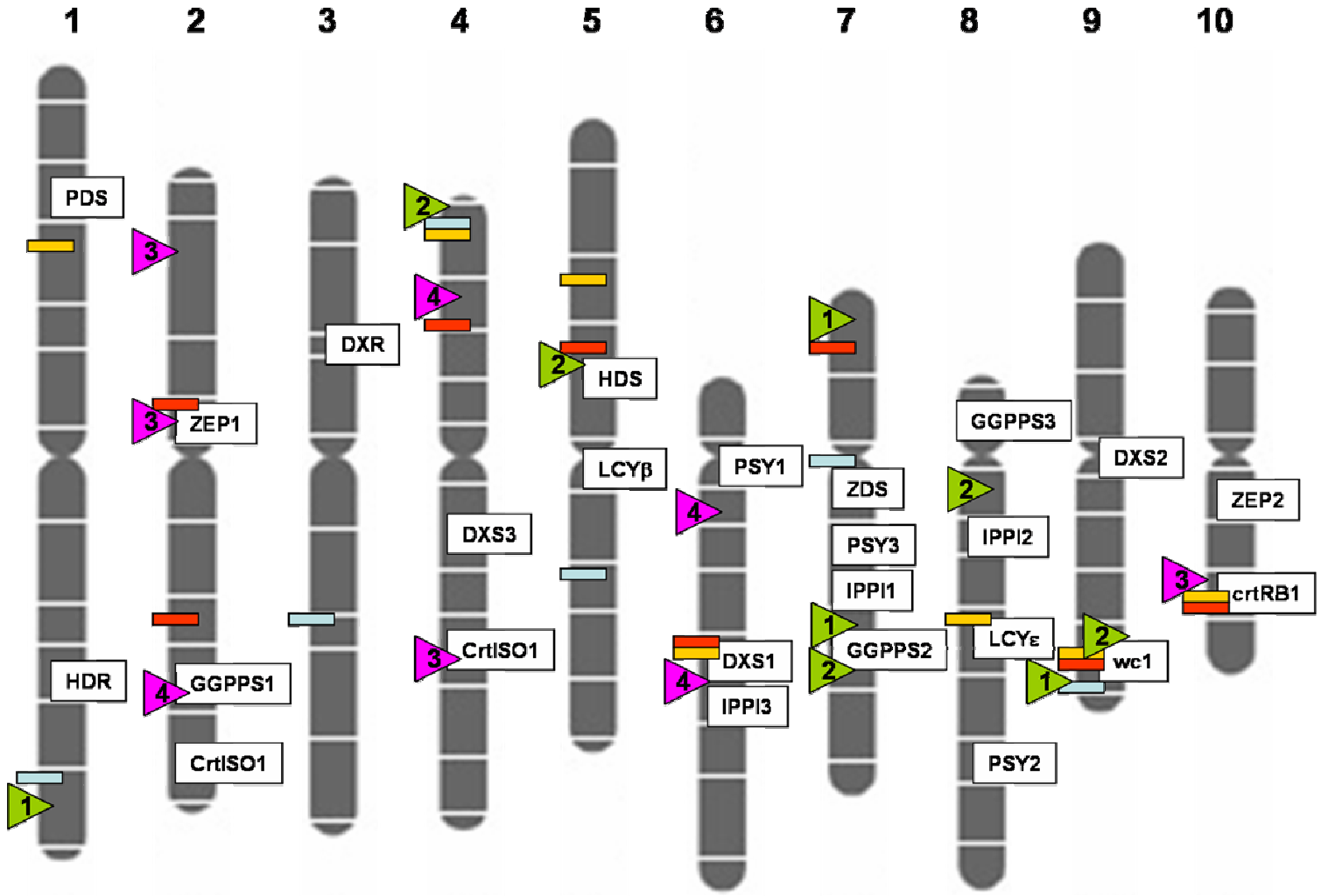


Figure 1.5: Chromosome bin locations of mapped QTL for carotenoid and PCA traits in DEexp x CI7 population



Boxes indicate location of significant univariate QTL for colorless carotenoids (blue), α-branch carotenoids (yellow) and β-branch carotenoids (orange). Triangles indicate location of significant multivariate QTL for PCs predominated by lutein (blue, 1); β-carotene and colorless carotenoids (green, 2), and phytoene versus β-branch carotenoids (magenta, 3). White boxes indicate known location of pathway biosynthesis genes (see methods for complete list).

Figure 1.6: Chromosome bin locations of mapped QTL for carotenoid and PCA traits in A619 x SC55 population



Boxes indicate location of significant univariate QTL for colorless carotenoids (blue), α-branch carotenoids (yellow) and β-branch carotenoids (orange). Triangles indicate location of significant multivariate QTL for PCs predominated by dihydroxy-xanthophylls and phytoene (green, 1 and 2), and β-carotene and dihydroxy-xanthophylls (magenta, 3 and 4). White boxes indicate known location of pathway biosynthesis genes (see methods for complete list)

CHAPTER 2

Identification and Characterization of *Zmccd1*, a Major QTL Underlying Carotenoid Reduction in Maize Endosperm

2.1 Abstract

Previous QTL analyses showed that a genetically mapped locus in maize chromosome bin 9.07 significantly affects carotenoid traits in multiple populations. Based on QTL effects and location, we proposed that genetically controlled carotenoid degradation processes could be responsible for these major alterations of carotenoid profiles. We tested Zm*ccd1*, a maize homolog of the carotenoid cleavage dioxygenase family, as the causal factor. A polymorphism in the promoter of Zm*ccd1* was associated with changes in total carotenoid concentration as well as lutein. Allele-specific markers were developed and used for linkage mapping analysis of carotenoid QTL in the A619 x SC55 $F_{2:3}$ population. Results indicate that the position of a significant and large effect for lutein and total carotenoid concentrations lies at the Zm*ccd1* map location, and the allele associated with a strong degradation effect is dominant. Use of the dominant mutant also revealed that Zm*ccd1* degradation processes likely predominate during the end of the grain filling period in dent corn.

2.2 Introduction

Malnutrition is a driving force for the development of micronutrient dense staple crop varieties. Enhancement of mineral and nutrient concentrations in crops has been accomplished by the recombination of favorable and rare genetics to develop superior varieties [73]. To address health needs in a reasonable timescale, the adoption of a product development approach focused on targeted genetic improvements is required. The approach of biofortification in maize is being taken by HarvestPlus, a program established by the Consultative Group on International Agriculture Research (CGIAR), which focuses on the improvement of provitamin A, iron and zinc concentrations in grain through conventional plant breeding and modern biotechnology [14]. Thus far, both strategies have made significant improvements in provitamin A concentration by exploiting the genetic and biochemical information known for the carotenoid biosynthesis pathway. Using a similar approach in the United States, breeding programs for maize grain

targeting increased xanthophyll concentrations have been created to satisfy dietary requirements that aid in the prevention of macular degeneration [17].

All plant tissues synthesize carotenoids, and compositional profiles and amounts vary according to the biological roles these compounds perform. In leaves, carotenoids are involved in photoprotection, where they serve as electron acceptors in the process of non-photochemical quenching [74]. Carotenoids are a substrate for the formation of abscisic acid, which plays an integral part in plant water and gas status through its regulation of stomatal aperture, as well as a role in seed germination. During maize root development, carotenoid degradation has been implicated as a defense mechanism against pathogens [75]. A gain of function mutation in a pathway enzyme expressed in seed endosperm resulted in the ability to select for yellow grain commonly grown today [76]. The biological function of carotenoids does not appear to have imparted a selective advantage on the viability of yellow versus white seed, as corn hybrids of both color classes perform well. However, human preference has selected for the presence of grain carotenoids, largely due to their association with animal health [4]. For the carotenoid pathway to satisfy a diversity of roles, tissue specific transcriptional regulation and redundancy in biochemical function through extensive gene families have been observed for several carotenoid pathway steps[31, 77].

Formation of carotenoid pathway precursors competes with the production of chlorophyll[22]. Substrate is committed to carotenoid biosynthesis once chemically modified by phytoene synthase, an enzyme encoded by the *y1*/*psy1* locus in maize [21]. The main branch leading away from phytoene (Figure 2.1) is irreversible and produces several colorless carotenoid precursors including phytofluene, zeta-carotene isomers, and neurosporene isomers (not listed in Figure 2.1). The production of lycopene marks the first pigmented carotenoid in the pathway, and serves as the main substrate to a bifurcated pathway that leads to the α- and β-carotenoid branches. Shunt of carbon substrate to either branch is controlled by the lycopene cyclases. Chemical modification by *lycopene beta cyclase* (LCYβ) is required for both branches, whereas modification by *lycopene epsilon cyclase* (LCYε) only affects the α-branch. Within each branch, carotenoids undergo a series of chemical reductions which transforms the molecules from highly fat soluble carotenes (α-carotene and β-carotene) to the slightly less fat soluble xanthophylls (zeinoxanthin, lutein, β-cryptoxanthin, and zeaxanthin). In maize grain, lutein and zeaxanthin are the predominant carotenoids. Compared to other cereals, maize has the

greatest phenotypic diversity for concentration of β-carotene, β-cryptoxanthin and α-carotene, which all contain provitamin A structures [19]. Characterization of these pathway steps at the genetic and biochemical levels have utilized various model species and tools including mutagenesis, recombinant expression, transcriptional expression and biochemical profiling [78, 79].

Basic genetic studies in model species have provided a wealth of knowledge that is being applied to improve carotenoid levels in crop species. The question exists if a known biochemical framework in model systems can be used to help explain the basis QTL in maize. Using the characterized biochemical frameworks from model plant species, Harjes [19] and colleagues investigated the control of substrate flux through the pathway bifurcation (see Figure 2.1) which was hypothesized to contribute to the regulation of the β-carotene containing branch. Surveys of a genetically diverse set of inbreds revealed three significant polymorphisms in the gene encoding lycopene epsilon cyclase (*lcyε*). Strong statistical associations led to the conclusion that genetic variation at this locus significantly altered the ratio of the branches, leading to increased proportion of β-carotene as part of total carotenoids. Supporting evidence from an eQTL experiment, carotenoid QTL analysis, and a mutagenesis study all pointed to *lcyε* as the causal factor for the modification of substrate flux. These results allowed the design of PCR-based markers targeted to the three polymorphisms which are currently being used in provitamin A breeding programs. Since report of *lcyε* as a causal QTL, discovery of *β-carotene hydroxylase* (*HYD3*/*crtRB1*), has led to the definition of a second major QTL affecting the conversion of β-carotene to downstream substrates and to the identification of a rare allele contributing to higher provitamin A [29, 40].

Manipulation of carotenoid synthesis and conversion is critical to making phenotypic gain for these traits, yet it is also critical to preserve the levels of synthesized carotenoid by reducing endogenous degradation processes. Grain carotenoid concentrations can be dramatically reduced in cereals during grain fill [80] and post-harvest storage (P.Beyer, M.Grusak, personal communications). Therefore, we explored the concept of genetically controlled degradation by maize orthologs of the carotenoid cleavage dioxygenase family (CCD) which have been biochemically characterized in Arabidopsis [81]. In this model species, members of the CCD family have been found to differ in substrate preferences; AtCCD1 is reported to reduce β-carotene and lutein seed levels. AtCCD4 has been observed to have an effect on all carotenoid substrates (D.DellaPenna, personal communication), the NCED subfamily has been implicated in

the formation of downstream substrate abscisic acid from zeaxanthin [82, 83]. Homologs of the CCD family have not been studied to a large extent in maize because of the complexity of the locus [84], the inadequacy of in vitro expression system [85], and until recently, the lack of maize genomic sequence. Molecular characterization of the *White Cap 1* (Wc1) maize mutant which is associated with a reduction in kernel color, facilitated cloning of the maize *ccd1* locus [84]. The maize and *Arabidopsis* CCD gene families appear to have similar functions in the removal of carotenoids. Therefore, a specific focus of this research is to determine whether AtCCD1 and Zm*ccd1* share substrate specificity for β-carotene. If so, this would make the Zm*ccd1* locus a primary target of selection in maize provitamin A breeding programs.

An integrative approach was taken to determine the in vivo function of Zm*ccd1* (herein noted as *ccd1*), evaluate its map location relative to the bin 9.07 location of the *wc1* locus and determine its interaction with loci responsible for the accumulation of carotenoid metabolites. The combined results of association mapping of *ccd1*, QTL mapping in the A619 x SC55 $F_{2:3}$ population segregating at loci *lcyε*, *crtRB1* and *ccd1*, and metabolic profiling of developing maize endosperm will be discussed. An evaluation of the *ccd1* locus in breeding programs for increased levels of provitamin A and total carotenoids will be presented.

2.3. Materials and Methods

2.3.1 Germplasm Development, Field Evaluation and Sample Collection

A diverse association panel of 281 inbred lines was grown in one-row plots in a randomized alpha(0,1) incomplete block design in Champaign-Urbana, Illinois during the summers of 2002, 2003, 2004, and 2005. The maize association panel was selected as a genotypically diverse germplasm set for which there was no selection of inbreds based on phenotype [43, 86]. Inbreds were selected from temperate, tropical and subtropical backgrounds to maximize allelic diversity within the panel and minimize linkage disequilibrium (LD) among lines.

An $F_{2:3}$ population comprising 227 families was derived from parental inbreds A619 and SC55, which were selected on the basis of their relatively high total carotenoid and high β-carotene levels, respectively. The population was grown during the summer of 2005 in two environments, the University of Illinois Urbana-Champaign and the International Maize and Wheat Improvement Center (CIMMYT), El Batan, Mexico. The experimental design at both

locations was a randomized alpha (0,1) incomplete block design with two replicates. The families were planted in single row plots of 5m rows, with 76 cm between rows. Each plot was thinned to a density of approximately 15 plants per 5m, or 43000 plants ha^{-1}. Seven to nine plants were sib-pollinated within each row, using a plant only once as a pollen source. Seed was bulked within each row after shelling. Seed was stored at room temperature for approximately four months, and then an aliquot of approximately 10g of seed from each row was stored at -80^{o}C until vitamin extraction could be performed.

Maize inbreds A619, SC55, CI7, DEexp, KUI3 and B77 were planted in blocks based on desired genotypic contrasts of the carotenoid biosynthetic genes *lcyε*, *crtRB1* and *ccd1*. The five contrasting blocks were: (1) A619, SC55 (2) KUI3, SC55 (3) CI7, DE3 (4) CI7, KUI3 (5) KUI3, B77. Within each block the inbreds were planted in two-row plots. The experiment was grown in a randomized split plot design with three replicates in Champaign-Urbana, Illinois during the summer of 2008. In each block, pollinations were made to obtain selfed ears of each inbred (P1 and P2), ears of the hybrid P1 x P2, and ears of the reciprocal cross P2 x P1. Developing ears were harvested on the basis of days after pollination (DAP) according to a 15 point time-course including 12, 14, 16, 18, 21, 24, 27, 30, 33, 35, 38, 40, 45 DAP, a fresh final harvest point (55-65 DAP) and a final harvest point with heat treatment of approximately 37° F in a dryer for five days. Whole ears from all time points were flash frozen in liquid nitrogen and stored at -80° C. Frozen kernels were removed from ears and divided into four portions to 1) stage the reproductive maturity of the kernels (R1 through R4), 2) obtain fresh and dry five-kernel weights (3 replicates), 3) measure kernel volume (7 replicates), and 4) reserve a bulk seed sample for carotenoid and mRNA transcript profiling. Only carotenoid profiles from the self-pollinated inbreds will be discussed.

2.3.2 Carotenoid Extraction and Quantification

Grain samples from the association panel and A619 x SC55 $F_{2:3}$ progeny were phenotyped for carotenoid concentrations by High Performance Liquid Chromatography described in Harjes [19]. Concentrations were represented on a dry weight basis (micrograms carotenoid per gram kernel dry weight).

Kernel samples from the 2008 planting were used for carotenoid profiling, embryos from 4-6 kernels were removed, and the remaining frozen endosperm was coarsely homogenized.

Further homogenization was performed on 30-50 mg of sample using a QIAGEN Tissuelyser II homogenizer and 4 mm steel ball bearings. Carotenoids were extracted in microcentrifuge tubes using 600 µl of 2:1 methanol:chloroform containing BHT (1 mg/mL) and tocol as an internal standard by further homogenization. After addition of 400 µl water and 200 µl chloroform, the samples were vortexed for 15 minutes and spun at 12 000 g for 10 minutes. The bottom fraction was collected, dried, and resuspended in 200 µl injection buffer (95:5 acetonitrile: ethyl acetate). HPLC analysis was carried out on 50 µl of the final extract.

Carotenoids were separated by HPLC on a C18 column (Spherisorb ODS2 5 micron, 150 × 2.1 mm, Column Engineering, Ontario) with a Shimadzu LC-20AD HPLC at variable flow rates with solvent A (acetonitrile: water [9:1 v/v]) and solvent B (ethyl acetate) and the following gradient: 0–20 min, 5% to 77% B, 1.0 mL/min; 20-20.2 min, 77 to 100% B, 1.0 mL/min; 20.2-22.2 min, 100% B, 1.5 mL/min, 22.2-22.4 min, 100-5% B, 1.5 mL/min; 22.4-25 min, 5% B, 1.0 mL/min. HPLC peak areas were integrated at 450 nm.

2.3.3 Genotypic Data Collection and Molecular Marker Design

DNA sample collection and PCR-based genotyping for the A619 x SC55 $F_{2:3}$ families are described in Stevens [44]. A total of 114 microsatellite markers were assayed, which was done in-kind by Pioneer. Of these markers, 101 are publicly available on MaizeGDB. The remaining 13 herein assigned a "pio" prefix are proprietary Pioneer markers.

Three markers specific to maize genes *lcyε* (*lcyε*-MZA), *crtRB1 (crtRB1*-InDel4) and *ccd1* (*ccd1*-5p) were used to genotype the population. Marker *lcyε*-MZA was designed to distinguish a three-nucleotide difference in exon 1 of *lcyε* between A619 and SC55 (*lcyε-MZA-P1-L(SC55)*, ATT TTT CTG GTA TTT ATT CAG C; *lcyε-MZA-P2(A619)*, AAG GCT ACT ACC TCC ATG AAA; *lcyε-MZA-All-R1*, AAT GAG AAT AGT ATG AGA TCG). This was accomplished using inbred-specific sequence kindly provided by Pioneer. Marker *crtRB1*-InDel4, detects a 12 bp indel segregating between the inbreds (*crtRB1-D4-F2*, ACC GTC ACG TGC TTC GTG CC; *crtRB1-D4-R1*, CTT CCG CGC CTC CTT CTC). Marker *ccd1*-5p was designed to distinguish three allelic states of the ccd1 promoter in the association panel which were b73 (identical to the B73 reference allele), Wc (allele from the dominant White Cap mutant) and wt (wild type). Inbreds A619 and SC55 differed in the alleles of the promoter polymorphism, allowing this marker to be used for both associationand QTL analyses. The *ccd1*

promoter marker was designed to have a conserved forward primer and allele-specific reverse primer to enable detection of heterozygotes, and a Wc and b73 chimera arising from intragenic recombination within a highly mutable locus [84]. Primers in the four-marker assay used to detect the allelic series of ccd1-5p are as follows: *ccd1-wt-L1*, ACT CAC TCG TAC TCA TCT ATC CAA; *ccd1-Wc-L1*, CCG TGC TCG GAC AGA ATA GT; *ccd1-b73-rev-L1*, CTC ACA CGT GTC AAC GCC; *ccd1-ALL-R1*, GTC GTT TCG GTG GCT GTC. Cycling conditions for all assays are listed in Appendix C.

Genotype information of the *psy1*-InDel388 polymorphism for all lines in the association panel was kindly provided by Y. Fu. Carotenoid QTL genotypes for members of the association panel evaluated with markers from this study (except *lcyε*-MZA) are listed in the Appendix A.

2.3.4 Genetic Map Construction and Composite Interval Mapping

Linkage maps were generated using JoinMap® Version 3 as described in Stevens [44]. The marker order is consistent with the physical map locations of the SSR/PCR primers found in the publicly available maize genome sequence. Comprised of the 117 markers described above, the total map length is 1727.8 cM, with an average of 16.1 cM between markers.

Both raw concentration data and ratios of raw concentrations (derived traits) were used for analyses. Derived traits expressed as ratios or sums of the direct estimates were calculated. The derived traits more accurately describe the genetic effects on product-substrate conversion, competition between pathway branches or metabolites, and changes in pools with similar chemical structures. Best linear unbiased predictors (BLUPs) of all direct and derived traits were generated for the two replicates of the Mexico location of the QTL mapping experiment. Phenotypic data for the Illinois location was not included in this analysis. QTL mapping was done by composite interval mapping (CIM) using stepwise regression for cofactor regression with PLABQTL software [50]. This software is based on the Haley-Knott regression method [87] and permits the evaluation of models varying in gene action. Regression models are evaluated on the basis of model fit criterion such as AIC and adj R^2. An additive model was used to fit the data A threshold corresponding to an experiment-wise Type I error rate of α=0.25 was used for QTL selection of each trait (approximately LOD 2.9). This error rate has been accepted as a suitable genome-wide threshold for exploratory QTL analyses [50].

Analyses of variance and correlations were performed using SAS version 9.2 [48]. Trait variation was modeled with single and all two and three factor interactions of known genes *lcyε, crtRB1* and *ccd1* with Proc GLM. A Bonferroni corrected experiment wise error rate (α=0.01) was used to test significance for model components. Contrasts and least square means were drawn to evaluate gene action models for *lcyε, crtRB1* and *ccd1* when they were found to be associated with a given trait effect. Pearson correlation coefficients for catergorical data were obtained using Proc CORR.

2.3.5 Association Analyses

Association analysis was conducted using a general linear model incorporating population structure as implemented in TASSEL [88] (www.maizegenetics.net). This approach accounts for the multiple levels of relatedness based on random genetic markers that are used to establish population structure. Using a general linear model, a statistical association between trait variation and Zm*ccd1* marker genotype was evaluated using the following linear model: $y_i = u + x_i + \mathrm{B}_1(x_{i1} - \overline{x_{.1}}) + \mathrm{B}_2(x_{i2} - \overline{x_{.2}}) + e_i$ where y_i is the carotenoid concentration for a given trait, μ is the pedigree/line mean, x_i is the effect of marker genotype of the ith allele, $\mathrm{B}_1(x_{i1} - \overline{x_{.1}})$ is the coefficient of linear regression of y_i on population structure covariate 1 (non-stiff stalk), $\mathrm{B}_2(x_{i2} - \overline{x_{.2}})$ is the coefficient of linear regression of y_{ij} on population structure covariate 2 (stiff stalk), and e_i is the random experimental error. Three allelic classes were tested in the marker genotype term: b73, wt (wild type) and Wc (white cap). A fourth class of b73-like + Wc was collapsed into the Wc class, as no significant difference was found between these two class means. Pedigrees that were considered to have "white" endosperm color were excluded from the analysis. Missing marker data also limited the number of individuals that were used in the single-gene analysis (year 2001, n=29; year 2002, n=40; year 2003, n=77; year 2005, n=98), reducing the power of analyses. Therefore, a single-year analysis was run only for the year 2005. In addition, a combined analysis including all four years was run, using the BLUPs as the trait means. An experiment-wise type I error rate α=0.05 cutoff was used for significance.

2.4 Results

2.4.1 Comparison of Trait Distributions for Germplasm

Two different types of maize populations were evaluated for the effect of *ccd1* on maize grain carotenoids. The association mapping panel contains a genetically diverse germplasm set that minimizes linkage disequilibrium and maximizes phenotypic diversity, whereas the A619 x SC55 $F_{2:3}$ mapping population has less phenotypic variability but enables linkage analysis through a narrow based family structure. Comparison of the carotenoid profiles between the A619 x SC55 $F_{2:3}$ mapping population and the maize association panel are shown in Figure 2.2. White lines in the association panel constituted the lower end of the trait distributions, in contrast to the absence of white lines in the mapping population. More trait variation was observed in the association panel in comparison to the mapping population. Exceptions to this trend were observed in β-carotene and the colorless carotenoids, phytoene and phytofluene. This was consistent with the greater genetic diversity of the association panel, but also illustrated that genetic composition of A619 and SC55 affecting β-carotene and the colorless carotenoids provided complementary levels of variation in the $F_{2:3}$ progeny.

The skewness of the trait distributions in both populations distinctly differed for most traits. The frequency distributions of trait classes for nearly all carotenoids were roughly normal in the association panel, in contrast the mapping population frequency distributions for lutein, zeinoxanthin, α-carotene, zeaxanthin, phytoene and phytofluene were all found to be skewed to the upper end of the distribution. The results indicated that reduced carotenoid concentrations were dominant in this population, which was highly suggestive of the effect of a single, large QTL.

2.4.2 Characterization of Association Panel for Zm*ccd1* Promoter Allelic Frequencies

Full length sequence for maize *ccd1* was obtained using genomic clone AC200748 and was found to contain 23 exons spanning 9949 bp (Figure 2.3, panel a). A promoter polymorphism consisting of a variable insertion was found approximately 500 bp upstream of the transcription start site (*ccd1*-5p). Previous study of Zm*ccd1* indicated that a wild type (wt) version of the locus is most commonly found in teosinte species [84]; as this allele appears to be the ancestral sequence variant, it is designated as the reference allele. Two other alleles, b73 (found in maize inbred B73) and Wc (in reference to the classical *white cap* locus) were also found in

diverse maize lines and were presumed to be orthologs of the wt allele, which contained no insertions. The b73 and Wc alleles are each marked by a distinctly different transposable element insertion immediately 5' to the start site (Figure 2.3 panel b). In the case of Wc, this insertion is accompanied by a tandem duplication of 12-24 copies of *ccd1* [84] (D. McCarty, personal communication).

The b73 allele was the most frequent *ccd1* promoter variant in the association panel with a frequency of 0.57, and the Wc class showed a frequency of 0.24 (Table 2.1). White germplasm had higher frequency of the Wc allele, suggesting a relationship between allelic states of *ccd1* and *psy1*, the locus primarily responsible for presence of seed endosperm carotenoids. This was supported by correlation analysis of functional sequence variation at *psy1* (InDel 388) with variation at *ccd1*-5p (Table 2.2), in which the deletion at *psy1*-InDel388 associated with a reduction in carotenoids frequently co-occurred with the Wc allele at *ccd1*. As expected, variation contributing to reduced *psy1* function was highly correlated to white endosperm color (r=0.76, p<0.0001). Unexpectedly, a high correlation between *psy1* and *ccd1*-5p alleles that promoted absence of pigment was observed (r=0.50, p<0.0001). In a two-way analysis of variance, *psy1* and *ccd1* explained approximately 69% of the variance in color (p>0.0001), as compared to either locus alone (61% and 32%, respectively). Presence of Wc, however, did not completely remove carotenoids from maize endosperm as more than a dozen yellow lines were found to have the Wc allele.

2.4.3 Association of Zm*ccd1* and Grain Xanthophyll Concentration

Statistical associations between the allelic classes of ccd1-5p with all absolute and derived traits were examined to determine if the polymorphism was linked to a reduction in carotenoid levels consistent with the proposed function of *ccd1* in *Arabidopsis* [89]. Analyses of statistical associations with 2005 phenotypic data indicated that allelic differences in the Zm*ccd*1 promoter explained significant variation in lutein (p=0.006) and were marginally significant for total colored carotenoid (p=0.07), which is a trait derived from the summation of the xanthophylls and carotenes (Table 2.3). No significant effect of marker class was found for β-carotene, or for any of the other absolute or derived traits. A combined year analysis indicated the marker polymorphism to account for considerable variation in lutein (p=0.0076) as well as derived traits for total colored carotenoid (p=0.0036), β-carotene/total (p=0.0115), α-

carotene/lutein (p=0.0025) and β-carotene/zeaxanthin (p=0.0035). Marginal associations were found for zeaxanthin (p=0.067) and lutein/total (p=0.069) at α=0.05. Using a comparison of the trait means for each allelic class, the Wc allele resulted in a 35% decrease in lutein for the combined analysis, and a 40% reduction in the year 2005 analysis (Table 2.4). Decrease of total colored carotenoid was largely due to the reduction in lutein as judged by the trait distributions (Figure 2.2). Effects seen for derived traits, which utilize total carotenoid concentration in the ratio, were likely attributed to fluctuations in the large proportion of lutein that is typically present in the total carotenoid amount.

2.4.4 Localization of *ccd1* to QTL Effects for Total Carotenoids

Significant QTL for 8 direct and 4 derived traits were found in the A619 x SC55 $F_{2:3}$ population and are listed in greater detail in Chapter 1. A range of 3-6 significant loci per trait explaining a large proportion of the variation for most traits (adj R^2=0.37-0.59) were found using additive models. Significant QTL were found on all chromosomes. QTL were often found to be pleiotropic, affecting more than one trait within the same branch or between parallel branches in a pattern consistent with enzyme specificity at a parallel step in the carotenoid biosynthesis pathway.

For both α-branch and β-branch carotenoids, a QTL was detected on chromosome 9 in an interval flanked by *ccd1* (Table 2.5). Contribution from SC55 exhibited a pronounced negative effect on lutein (reduction of 1.79 μg g^{-1}), as well as reductions in zeaxanthin (0.86 μg g^{-1}), β-cryptoxanthin (0.11 μg g^{-1}) and zeinoxanthin (0.07 μg g^{-1}). The QTL accounted for as much as 29% of the variation in individual xanthophyll compounds and contributed a 40-60% reduction in lutein and zeaxanthin.

β-branch carotenoids, consisting of β-carotene, β-cryptoxanthin and zeaxanthin, were affected by QTL on chromosome 10 within the umc1506 - *crtRB1* – umc1993 interval with intra-interval marker positions of 79 cM, 93 cM and 104 cM, respectively (Table 2.5). The QTL positions for each of the three traits were at 94 cM, 88 cM, and 88 cM, respectively. As previously mentioned, the *crtRB1* enzyme controls conversion of β-carotene to β-cryptoxanthin and β-cryptoxanthin to zeaxanthin. Additive effects estimates indicate that the SC55 allele at this locus is correlated with an increase in β-carotene (0.37 μg g^{-1}), and a decrease in downstream β-cryptoxanthin (0.20 μg g^{-1}) and zeaxanthin (0.50 μg g^{-1}) substrates. A similar

result was observed for the β-carotene/β-cryptoxanthin ratio (2.74 μg g^{-1}). A similar chemical conversion could be catalyzed by a hydroxylase enzyme such as *crtRB1* in the α-branch (consisting of α-carotene, zeinoxanthin and lutein). This QTL did not have a significant effect on the α-branch carotenoids individually, but results showed that *crtRB1* accounts for a significant amount of variation in the α-carotene/ zeinoxanthin ratio (12.2%, 0.08 μg g^{-1}).

Control of the branch bifurcation by *lcyε* was also evident in this population. Approximately 15% of the variation for the derived trait α/ β branch was explained by this locus (position of QTL effect = 92 cM; position of *lcyε* = 91). *lcyε* was also found to explain a large proportion of the variation for lutein (11.3%).

2.4.5 Allelic Interactions Within Known QTL

Additive and dominance effects models were applied to the *lcyε, crtRB1* and *ccd1* loci. Only traits significantly affected by these genes were examined. Linear and quadratic contrasts were calculated to assess the significance of additive and dominance gene action within loci. *lcyε* and *crtRB1* were found to act only in an additive manner on lutein and total carotenoids, where the linear component was found to be significant at α=0.05 (Table 2.6). *ccd1* exhibited significant dominance as well as additive effects as seen by the linear and quadratic components for lutein, zeaxanthin and total carotenoids. Inspection of least squares means by marker class indicated that the SC55 allele for *ccd1* is dominant to the A619 allele. This agrees with mutant studies of white cap (P.Stinard, personal communication), where the mutant condition (Wc) is typically found to be dominant.

2.4.6 Carotenoid Production During Kernel Development

To validate significant QTL effects, an experiment was initiated in which inbred lines with contrasting carotenoid *lcyε, crtRB1* and *ccd1* QTL haplotypes were evaluated for carotenoid production potential during kernel development, with *ccd1* being one of these loci. The carotenoid profile from 12 DAP to harvest (65 DAP) indicated that carotenoid production and accumulation were highly dependent upon the haplotype of the line (Figure 2.5). Inbreds in this experiment were selected to contrast known *ccd1* promoter allelic states, with b73 allele representing a weak, less active degradation enzyme, and Wc allele representing a strong,

overactive version of *ccd1*. Comparison of developmental profiles for lines containing *ccd1*-5p-b73 (CI7 and KUI3) with a line containing *ccd1*-5p-Wc (SC55) revealed a significant decrease in lutein for the Wc line, but not the b73-like lines (Figure 2.5). Profiles of zeaxanthin provided evidence of the Wc allelic effect as well. To our knowledge, this is first in vivo evidence of CCD1 substrate specificity in maize and provides a biochemical validation of the QTL mapping results.

Total carotenoid concentration continually increased over time of kernel development. A rapid rate of carotenoid accumulation occurred from 12 to 27 DAP followed by a slower increase after 27 DAP. Genotypic differences are apparent in total carotenoid profiles as well as those for individual carotenoids.

2.5 Discussion

This study provided very strong evidence that *Zmccd1,* encoding the maize carotenoid cleavage dioxygenase 1 protein, is contributing to large, pleiotropic effects mapped to bin 9.07 in several mapping experiments for carotenoid traits [25, 26, 28, 44]. Segregation for *ccd1* was found to account for significant phenotypic variation in total colored carotenoid concentration which is likely caused by a significant decrease in lutein, with smaller effects on zeaxanthin, zeinoxanthin, and β-cryptoxanthin. These results agree in part with in vitro tests for the substrate specificity of ZmCCD1, where β-cryptoxanthin and zeaxanthin were found to decrease in the presence of the enzyme [90]. An analysis of the lutein substrate was not possible with the published recombinant system; therefore, this is the first report of lutein as a substrate of ZmCCD1. Considering that biochemical homology between *Arabidopsis* and maize was assumed, the data does not entirely agree with AtCCD1 substrate specificity, as β-carotene did not associate with *Zmccd1* sequence variation. Of all cleavage dioxygenases in *Arabidopsis*, AtCCD1 is most similar to ZmCCD1 in amino acid identity (76.55% identical), making the difference in substrate specificity a somewhat unexpected finding. Comparison of least square means for the allelic variants of the *Zmccd1* promoter indicated that the Wc allele had a strong degradation effect. Dominance of the strong Wc allele as well as negative epistatic interactions between *ccd1* and *lcyε* revealed that a single copy of the Wc allele could substantially reduce total carotenoids. The profiling results demonstrated that a negative effect on total carotenoids by the *ccd1*-5p-Wc genotype is largely attributed to a removal of lutein. The effect of this

degradation is most severe around 30 DAP suggesting that there is a temporal separation between peak synthesis and degradation [31, 77] . This is in stark contrast to carotenogenesis during the maturation process of sorghum grain where degradation gradually out-competes synthesis, leaving little carotenoid accumulation at harvest [80]. Genetically controlled carotenoid synthesis and degradation mechanisms in maize may differ substantially from those in sorghum.

The enhanced removal of carotenoids associated with the *ccd1*-Wc allele is hypothesized to originate from the multi-copy gene repeat connected to the promoter polymorphism, rather than through the promoter polymorphism. Although we did not quantitatively evaluate the effect of copy number variation here, PCR amplification with the marker developed in this study was observed to yield more product with Wc-containing templates. This asymmetry in allele copy number often competitively interfered in the detection of heterozygotes (data not shown). The origin of the tandem repeat is likely due to the insertion of helitron transposable elements which have been shown to significantly alter genome microcollinearity in a number of eukaryotes including *Arabidopsis*, rice [91] and maize [92, 93].. Characteristic of these elements is their large size (8-20 kb) which is attributed to the capture of gene pieces [94] as the element moves through the genome and reinserts through a rolling circle insertion mechanism [91]. Although it is often stated that insertions have not been seen to duplicate host sequence, mutated rolling circle insertion mechanisms have been reported to insert several copies of the host sequence leaving a tandem repeat [95].

Most reports of helitron insertions would suggest that *ccd1* repeats in Wc may not necessarily be full length or even sequential; this is quite possible, as attempts to amplify full length *ccd1*-5p-b73 sequence for association mapping (primer sites indicated in Figure 2.2) in this study largely failed due to exact repeats of specific exons in Chromosomes 1 and 10 (data not shown). The Wc+b73 chimera found in the association mapping panel is evidence of multiple insertion events within the same location. While we posit that the *ccd1* effect in Wc could be due to heightened degradation activity leading to reduction of endosperm color, the possibility of gene silencing through duplicated gene copies also exists [94] . We surmise, however, that this process would reduce *ccd1* activity rather than increase it.

Historically, colorless or white maize has been preferred for human consumption whereas yellow maize has been primarily used for animal feed due to nutrition [96]. Sensory characteristics of the grain including texture, scent and flavor are a function of the protein/starch matrix as well

as the degradation of carotenoids which occurs during food preparation [97]. Though most white maize breeding has been performed by the private sector and therefore few papers are published on the subject [3] the genetics of maize color with particular reference to carotenoids is well known [3, 98]. Of the dozen or so loci that are involved in controlling maize endosperm color and have been used in breeding white maize, the white cap locus, or *ccd1*, is not listed among them. Results here have indicated that alleles contributing to reduced kernel color in *psy1* and *ccd1* are co-inherited more times than expected by random chance, and therefore have been selected in tandem. It is possible that the *ccd1* locus has not been a direct selection target, but instead has been a modifier locus which conditions for pure white endosperm and often accompanies selection.

Development of high total carotenoid populations can also benefit from a marker assisted selection strategy for *ccd1*. High total carotenoid breeding can be simplified through the visual selection of more yellow/orange kernel pigmentation since there is a high correlation between total carotenoid and kernel color within breeding populations. Previous experimentation has shown that the effect from bin 9.07 reduces kernel color [44,27]; based on results from this study, it is likely that *ccd1*is responsible for this effect. High carotenoid breeding programs in developing countries have crossed donor lines with enhanced orange pigment to predominantly white Mexican and African maize varieties in an effort to combine the alleles for higher total carotenoid with favorable agronomic traits (K. Pixley, personal communication). If the frequencies of the Wc- *ccd1*allele in the Mexican and African populations are similar to the white germplasm in the maize diversity association panel, the recurrent parents may harbor deleterious alleles at both *ccd1*and *psy1* which will dilute the high carotenoid effect upon recombination. Therefore, use of markers for *ccd1*alleles should be useful to enhance recovery of the desirable donor genetics while excluding heritable degradation effects.

The genetic architecture of the carotenoid trait system appears to be predominated by biochemical QTL that explain much of the trait variation and exhibit pleiotropic effects [26, 28, 44] (Chapter 1). Association mapping can be quite useful in cases where candidate gene prediction is desired for QTL with large trait effects, as has been observed for *lcyε, crtRB1* and most recently *ccd1*. Statistical associations between quantitatively inherited traits and causal loci have been performed for many biochemical, regulatory and developmental genes with traits including starch composition/quality [99], plant and inflorescence architecture [100], and flowering time [101, 102].

It is critical to note that each report selected candidate loci prior to conducting the study, meaning that some prior knowledge of the biology affecting the trait was required. In many instances, candidate genes cannot be postulated due to the complexity of the trait arising from multiple QTL and the involvement of many pathways. Genome-wide association studies are becoming available in maize, and should enable investigators to evaluate loci that would not otherwise have a biological basis for testing [103]. As there is precedent for the involvement of structural genes in carotenoid accumulation in addition to genes involved in the biosynthesis pathway [56, 104], this method will be extremely helpful in identifying new QTL. In addition, pairing of association analysis and linkage mapping has proved to yield complementary results for QTL validation here as well as various other studies [39, 105,19, 39], and should be able to further the value of biochemical QTL as targeted breeding tools.

2.6 Tables

Table 2.1: Allele frequencies of *ccd1*-5p polymorphism in the Goodman-Buckler Diversity Lines

	All lines with data		White lines excluded		White lines only	
ccd1-5p Genotype	No.	Freq.	No.	Freq.	No.	Freq.
b73	103	0.57	92	0.51	11	0.26
wt	8	0.04	6	0.03	2	0.05
b73+Wc	5	0.03	1	0.01	4	0.10
Wc	43	0.24	18	0.10	25	0.60
Total No.	159		117		42	

Table 2.2: Correlation of ccd1-5p genotype with indel 388 of phytoene synthase (*psy1*) and color phenotype

Variables		N	Correlation (*r*)	P
psy1	*ccd1*	164	0.5	<0.0001
psy1	color	165	0.76	<0.0001
ccd1	color	179	0.56	<0.0001

Variables are categorical where *psy1* reflects state of indel-388 (insertion, deletion), *ccd1* reflects state of promoter insertion (Wc, wt, b73), and color indicates visually white or pigmented endosperm

Table 2.3: Statistical associations for *ccd1*-5p in 2005 and combined year trait panel using GLM

	Combined Years n=241			Year 2005 n=98		
Carotenoid Trait	p_Mkr	Rsq_Model	Rsq_Marker	p_Mkr	Rsq_Model	Rsq_Marker
lutein	0.0076	0.906	0.0549	0.006	0.2134	0.094
zeaxanthin	0.067	0.9306	0.0343	ns	ns	ns
total	0.0036	0.9258	0.0661	0.0709	0.0887	0.0691
β-carotene/total	0.0115	0.9458	0.0684	ns	ns	ns
lutein/total	0.069	0.9643	0.0326	ns	ns	ns
α-carotene/lutein	0.0025	0.7973	0.0721	ns	ns	ns
β-carotene /zeaxanthin	0.0035	0.6879	0.0664	ns	ns	ns

Statistical tests of association of trait variation with allelic variation in ccd1 promoter polymorphism evaluated using probability value of marker (p_Mkr), coefficient of determination for the model consisting of covariates and marker main effect (Rsq_Model) and coefficient of determination for the marker effect (Rsq_Marker).

Table 2.4: Least squares trait means in combined years and in 2005 for allelic classes of *ccd1*-5p in association panel

	Adjusted trait means, combined years (μg g^{-1})						
Allele	lutein	zeaxanthin	total	β-carotene/ total	lutein/ total	α-carotene/ lutein	β-carotene/ zeaxanthin
ccd1-b73	10.31	6.14	19.6	0.07	0.52	0.06	0.08
ccd1-wt	10.79	4.21	17.26	0.09	0.53	0.07	-0.09
ccd1-Wc	6.41	4.06	13	0.12	0.43	0.11	1.38

	Adjusted trait means, 2005 (μg g^{-1})	
Allele	lutein	zeaxanthin
ccd1-b73	7.09	4.9
ccd1-wt	5.51	3.58
ccd1-Wc	4.25	4.08

Table 2.5: Carotenoid traits affected by *lcyε, ccd1* and *crtRB1* QTL in A619 x SC55 $F_{2:3}$ population grown in 2005, El Batan, Mexico

QTL	Chr. Bin	Contig (ctg)	Interval	Trait	LOD	Add. (μg g^{-1})	R2 (%)
lcye	8.04	-	umc1343-LCYe-umc1340	α-Carotene	5.69	0.03	9.5
	8.05	354		total	3.87	0.98	8.9
				lutein	5.47	0.64	11.3
				α: β branch	6.54	0.09	14.6
ccd1	9.07	391	CCD1-zct128	total	21.21	-3.65	53.1
				β-Cryptoxanthin	4.08	-0.11	12.4
				zeaxanthin	4.22	-0.86	29.6
				zeinoxanthin	3.12	-0.07	2.2
				lutein	12.85	-1.79	28.8
				α: β branch	11.41	-0.21	40.1
	9.08	391	zct128-umc1505	phytoene	10.08	-1.35	19.8
				phytofluene	9.61	-0.24	18.4
crtRB1	10.05	414	umc1506-CrtR-B1	β-Cryptoxanthin	9.93	-0.2	17.6
				Zeaxanthin	4.32	-0.5	9.7
				β-Carotene: β-Cryptoxanthin	10.36	2.25	3.4
	10.06	417	CrtR-B1-umc1993	β-Carotene	6.42	0.37	10
				β-Carotene: β-Cryptoxanthin	7.63	2.74	4.8
				α-Carotene: Zeinoxanthin	4.53	0.08	12.2

Summary of QTL results for functional markers from Chapter 1 analysis, Table 1.12

Table 2.6: Contrasts of additive and dominance effects on selected carotenoid traits for *lcyε, ccd1* and *crtRB1* in population A619 x SC55

Trait	Gene	Regression	P
lutein	*lcyε*	linear	0.002
	lcyε	quadratic	0.2519
	ccd1	linear	<.0001
	ccd1	quadratic	0.0001
zeaxanthin	*crtRB1*	linear	0.0072
	crtRB1	quadratic	0.1034
	ccd1	linear	0.0112
	ccd1	quadratic	0.0275
β-Carotene	*crtRB1*	linear	0.0077
	crtRB1	quadratic	0.2434
total	*lcyε*	linear	0.012
	lcyε	quadratic	0.4572
	ccd1	linear	<.0001
	ccd1	quadratic	<.0001

Additive (linear) and dominance (quadratic) relationships tested using indicated contrasts.

2.7 Figures

Figure 2.1: Carotenoid biosynthetic pathway and pleiotropic effects observed in this study

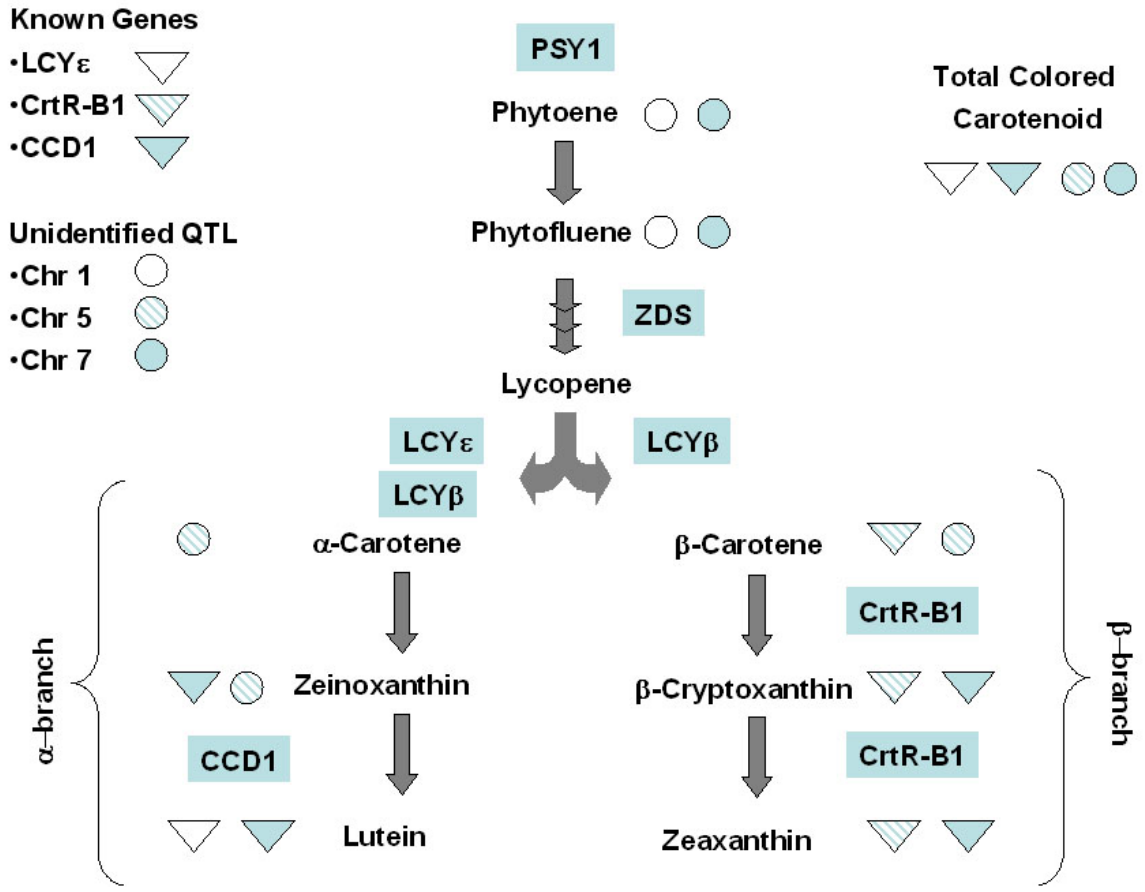
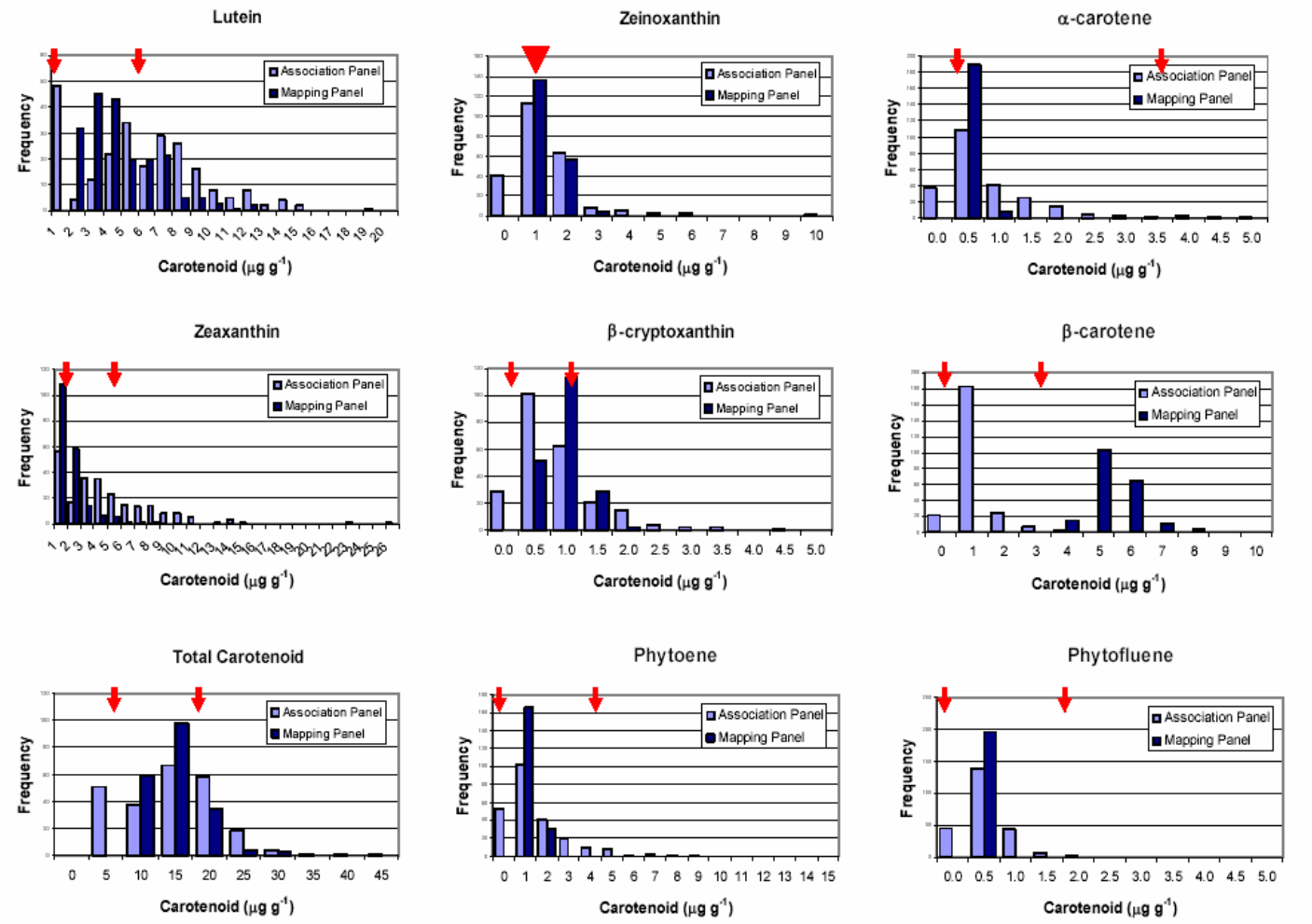


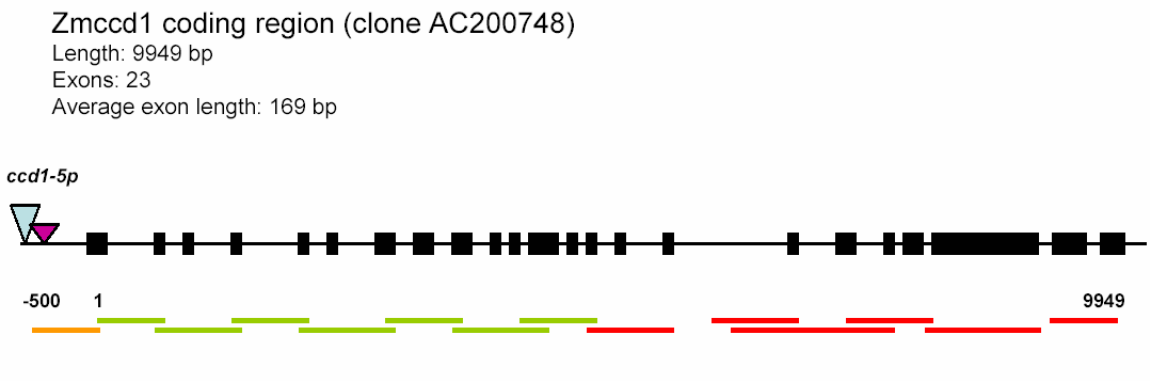
Figure 2.2: Comparison of trait distributions between association mapping panel and A619 x SC55 $F_{2:3}$ mapping population, grown in 2005 El Batan, Mexico



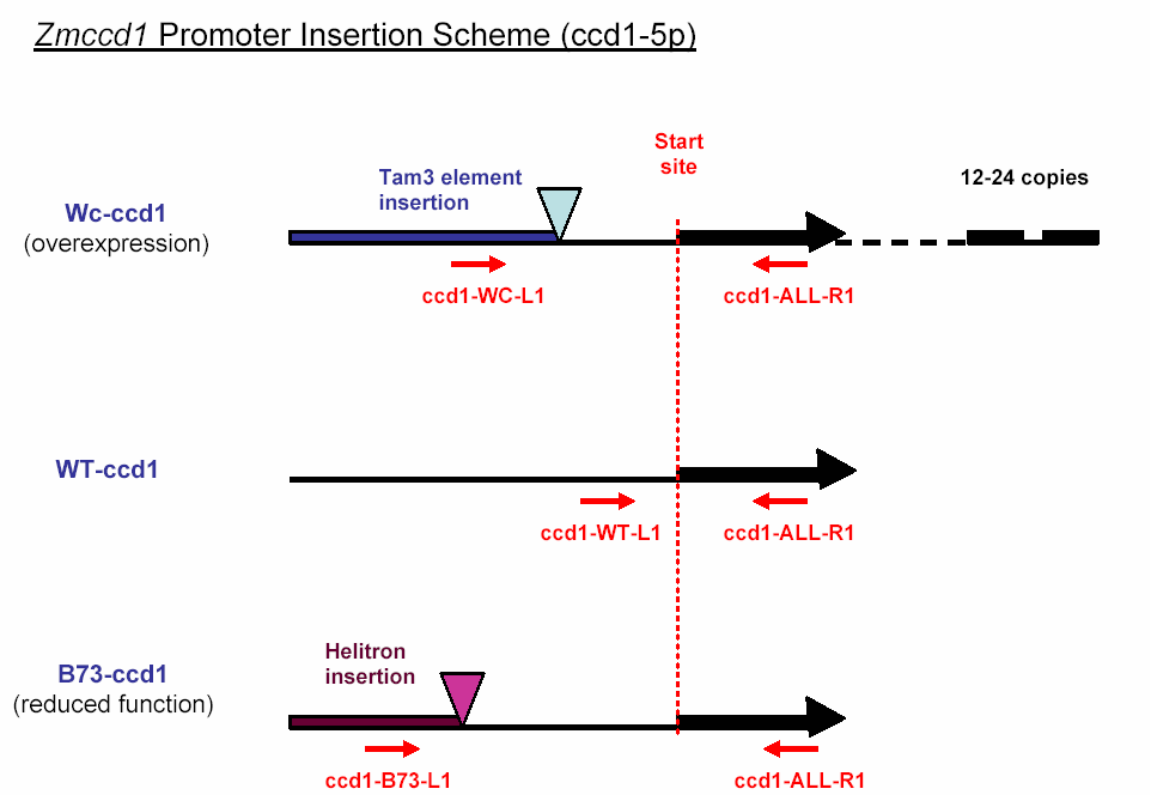
Trait distributions are shown for association panel (light blue) and mapping population (dark blue). Red arrows indicate parent values of mapping population.

Figure 2.3: Marker assay for *Zmccd1* promoter polymorphism

a.



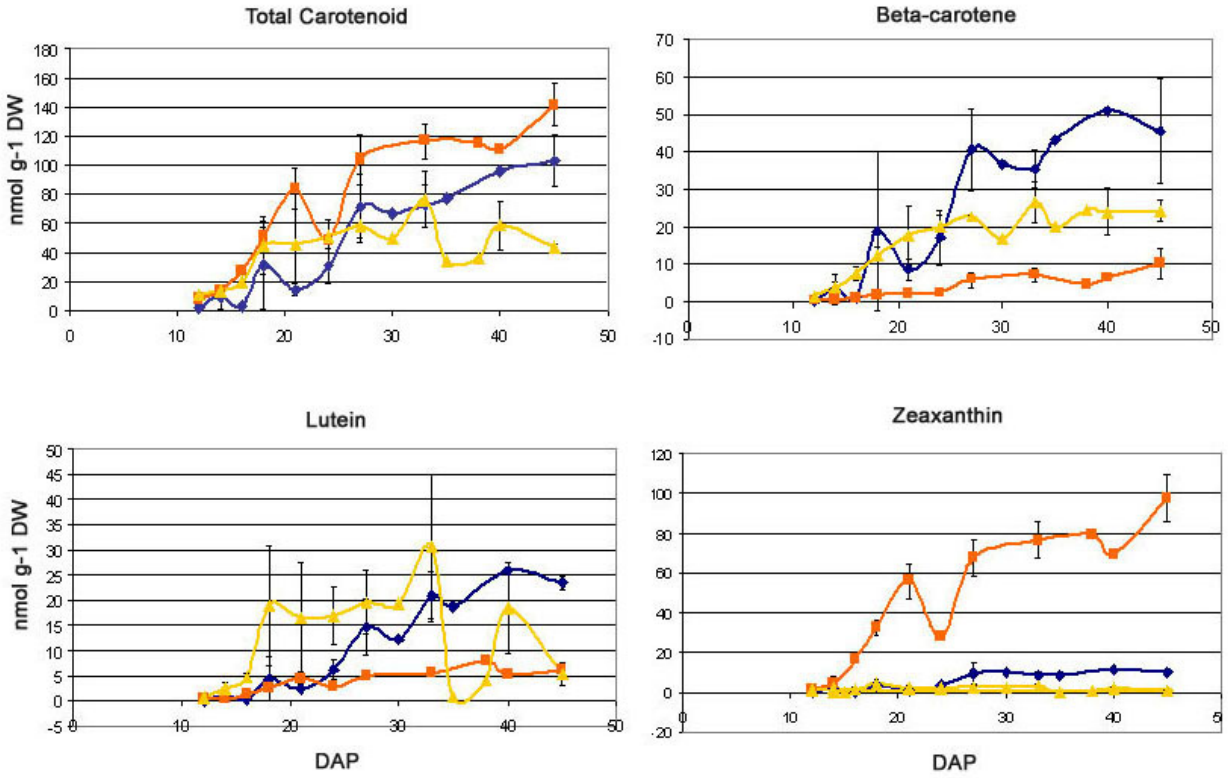
b.



a: Gene model for Zm*ccd1*. Exon boundaries are shown in black. Polymorphism described here is indicated by triangles in promoter region, and amplified by orange primer assay. Other PCR assays for full length association mapping are indicated in green (both primer annealing sites are ccd1 specific) and red (only one primer site is ccd1 specific).

b. Schematic of sequence variation and primer annealing sites in ccd1 promoter

Figure 2.4: Carotenogenesis profile of selected traits for three inbred lines varying in *ccd1* alleles



	LCYε	CrtR-B1	CCD1
CI7	.	weak allele	weak allele
KUI3	weak allele	.	weak allele
SC55	.	weak allele	.

Carotenoid profiles for inbreds CI7 (blue), KUI3 (orange) and SC55 (yellow) are plotted across days after pollination (DAP). Inbred haplotypes differ for *lcyε*, *crtRB1* and *ccd1*, where weak allele is designated as less efficient condition; actual allelic conditions for these lines are noted in Appendix A. For ccd1, weak allele is designated as b73, and non-weak allele is designated as Wc.

CHAPTER 3

Toward Efficient Use of ProVitamin A QTL in Germplasm Enhancement: An Evaluation of Allelic Selection in Segregating Progenies

3.1 Abstract

Allele-specific variation in pathway biosynthesis genes underlying quantitative trait loci and directly affecting β-carotene and provitamin A (proVA) concentrations have been identified. We investigated the effect of allelic variation at *beta-carotene hydroxylase* (*crtRB1*), *lycopene epsilon cyclase* (*lcyε*) and *carotenoid cleavage dioxygenase1* (*ccd1*) on maize grain carotenoid composition, and more specifically proVA carotenoids in the context of varying genetic backgrounds and environmental effects. Genes underlying carotenoid QTL were found to contribute to trait phenotypes through both additive and dominance effects. *crtRB1* allelic combinations previously associated with desirable endosperm β-carotene concentrations among inbred lines [29] were found to substantially increase β-carotene concentrations across four segregating populations (1.11-3.11 fold increase) and three advanced breeding populations (3.45-11.33 fold increase). Accounting for all measured carotenoids with proVA activity including β-carotene, β-cryptoxanthin and α-carotene, proVA concentrations were markedly improved when the favorable *crtRB1* variation was bred to homozygosity (1.45-3.22 fold increase). A more detailed investigation of the effects on substrate allocation to carotenoid pathway branches through *lcyε* effects [19] revealed that polymorphisms within this gene modified α/β branch ratio through different modes of action. Polymorphisms *lcyε*-5'TE and *lcyε*-3'TE were found to primarily affect the quantity of α-branch carotenoids, whereas *lcyε*-SNP216 was associated with the expected change in both α– and β–branch carotenoids. Results provide guidelines for use of allelic variation at these carotenoid QTL in ongoing and future selection programs.

3.2 Introduction

Improvement of maize grain quality traits by phenotype selection has had a long, successful history [106-108]. Breeding efforts have capitalized on recombination of superior pedigrees to attain improved progeny with predictable performance. The observation of heritable and enhanced phenotypes in subsequent generations highlights the favorable

contribution of parental gametes, and permits the association of phenotypic performance with genotypic composition. Our understanding of the genetic basis underlying trait variation has expanded recently as maize genomics, model and crop species physiology, biotechnology, and statistical methods have evolved [103, 109] (www.maizesequence.org). Varieties with improved nutritional qualities including oil composition, essential amino acid content, and starch type have been successfully developed, leading to many products that are commercially available today [110]. Enhanced grain micronutrient content for both vitamins and minerals is rapidly emerging as the next suite of seed quality traits to be improved by breeding and biotechnology [73, 111].

The biofortification strategy, which focuses on the nutritional enhancement of staple crops to combat dietary deficiencies using multi-disciplinary approaches in breeding and genetics, biotechnology, nutrition, and community health [14], has been widely adopted by many academic, non-profit and international agriculture centers. Grain provitamin A (proVA) carotenoid concentration, which has been noted to have enormous potential as a sustaintable dietary vitamin supply to at-risk populations in developing countries, is a biofortification target compound [12, 111]. ProVitamin A carotenoids such as β-carotene, α-carotene and β-cryptoxanthin are synthesized by plant secondary metabolism. Through animal metabolism, these carotenoids are cleaved into Vitamin A, or retinol units [10, 11], which are essential to many biological processes including vision, immunity, reproduction and growth and development [112]. Daily dietary requirements are dependent upon age and gender, but an average range of approximately 275-500 μg is estimated for individuals consuming a diverse diet of plant proVA carotenoids and animal-based Vitamin A [14].

Unlike typical diets found in industrialized nations which are composed of diverse and fortified foods, diets in developing countries are based primarily on cereal staple crops which tend to be less nutrient dense than other plant foods [113]. In 2003, over 50% of the daily caloric intake in sub-Saharan Africa was derived from maize, where individuals consumed an average of 101.3 g/capita/day as compared to consumption across all developing countries (54.8 g/capita/day) and all industrialized nations (32.9 g/capita/day; http://faostat.fao.org/site/609/default.aspx). Lack of dietary diversity coincides with high rates of Vitamin A deficiency (VAD) which affects 40% of children under five years of age in developing countries and leads to 1 million infant deaths each year (UNICEF REPORT 2004).

Improving the micronutrient balance of staple crops like maize through biofortification is therefore a reasonable alternative to address nutrient deficiencies, including VAD.

Plant carotenoid profiles can also be used to reduce the risk of age-related macular degeneration (AMD), a vision impairment common in aging populations of industrialized nations which ultimately results in blindness [114]. As of 2004, 1.75 million Americans were diagnosed with AMD; this number is projected to rise to 3 million by 2020. Lutein and zeaxanthin, carotenoids which predominate maize carotenoid profiles and are collectively known as xanthophylls, have been shown to help prevent oxidation and degeneration of the macula [17] and are being further tested in the Age Related Eye Disease Studies (AREDS2) sponsored by the National Institutes of Health which follows a successful initial study [115]. The average consumption of xanthophylls and specifically lutein in the US and Europe (2.0 mg day $^{-1}$) is below USDA recommended dietary guidelines (3.8 mg day $^{-1}$), which is far less than the xanthophyll amounts projected to prevent AMD (7.3 mg day $^{-1}$). Lutein can be supplied through low-dose vitamins and is added to most processed foods in crystalline form [116]; however, plant based carotenoids can provide an alternative xanthophyll source.

Based on measured concentrations of proVA carotenoids in fresh tissues, many fruits and vegetables appear to be superior proVA alternatives to grain based proVA supplementation; however, high cost, limited availability and reduced vitamin bioavailability from complex food matrices prohibit the widespread adoption of fruit and vegetable foods as major sources of proVA [13]. Compared to bioconversion rates of proVA measured at 19 units β-carotene to 1 unit retinol for carrot or 29:1 for spinach, bioconversion rates for maize grain β-carotene in model animal systems is substantially higher at a 3:1 ratio [117]. Uptake rates for lutein are estimated to be similar to those measured for β-carotene [116]. As maize consumption ranges from 30 to 100 g/ capita/ day worldwide, the availability of proVA and xanthophylls from maize grain provides an ample phytonutrient supply. Accounting for effects of post-harvest/ post cooking carotenoid stability and retention, vitamin bioavailability from the grain matrix, and conversion efficiency of proVA carotenoids to vitamin A, a breeding goal of 15 μg g^{-1} proVA grain carotenoids would satisfy recommended nutritional targets for consumption of 100 g of maize grain per day [14].

Breeding for grain carotenoids has taken advantage of extensive phenotypic diversity in maize. In a survey of antioxidant levels in sweet and dent corn using high throughput liquid chromatography (HPLC), Kurilich and Juvik [118] found a wide range of phenotypic variability in

the major grain carotenoids, lutein and zeaxanthin, but less variation in the proVA carotenoids of α-carotene, β-carotene and β-cryptoxanthin. Reports of carotenoid variation in Chinese [69] and Brazilian [119] dent germplasm found similar trends for major grain carotenoids and proVA carotenoids. The extent of phenotypic variation found depends on that which was sampled, as genetic variation for certain carotenoid traits is likely geographically restricted or rare. One example of this is demonstrated by orange flint germplasm of South America and the Caribbean which contains some of the highest zeaxanthin levels observed [120]. Similarly, surveys of genetically diverse inbreds have shown that favorable carotenoid profiles are found in distinctly unrelated lines [19, 25]; sample carotenoid profiles for a diverse set of germplasm collectively called the Goodman-Buckler Diversity Panel are presented in Chapter 2. It can be concluded, therefore, while genetic sources with superior carotenoid profiles do exist, the genetic variation contributing to high proVA phenotypes likely have never been pyramided due to genetic frequencies or mutation.

To date, hybrid development and population improvement for desired carotenoid composition has been achieved by relying on midparent-offspring covariance. Thus, parental selection has largely been driven by per se performance. Indeed, general combining ability among a set of historically elite inbreds was found to contribute most of the variation in hybrid carotenoid profiles, indicating that genetic effects on trait concentrations are largely additive [45]. A study of the comparative effects of pollen and female parents on carotenoid concentrations from hybrid-derived seed demonstrated that progeny profiles more often reflected the additive contribution of the female genotype, which donates two copies of a haplotype to the endosperm as opposed to one from the pollen parent [121]. Reports of moderate to high heritabilities for these traits [26, 69, 70] confirm that additive effects are highly important in population development. Non-additive effects represented by significant specific combining abilities [45] have been seen to contribute to lutein, β-carotene and β-cryptoxanthin hybrid profiles, suggesting that existing allelic variation contributing to dominance effects should be further explored. Population development through phenotypic selection of carotenoid profiles conducted with advanced breeding lines in the CIMMYT proVA breeding program (Pixley and Palacios-Rojas, personal communication) as well as in synthetic lines developed by the Rocheford laboratory at the University of Illinois has been successful for improvement of grain proVA levels.

Evidence for polygenic variation as the basis of inheritance for carotenoids has been confirmed by several QTL studies performed in recombinant inbred lines (Chander map) and $F_{2:3}$ mapping populations [25, 26, 44]. On average, these studies report 3-9 loci to affect carotenoid compositional traits, with up to 49% of the phenotypic variation in a given carotenoid trait explained by a single QTL. Pleiotropic effects at these QTL and positive trait correlations have been observed. Consistent with the understanding that traits are derived from a single metabolic pathway, several QTL map to chromosomal regions where classical maize mutations or known carotenoid biosynthesis genes are known to reside. These include: phytoene synthase 1 [21, 33] (PSY1, locus *y1,* bin 6.02), phytoene desaturase [34, 35] (PDS, locus *vp5*, bin 1.02), ζ-carotene desaturase [36] (ZDS, locus *vp9,* bin 7.02), lycopene β-cyclase [37] (LCYβ, locus *ps1/vp7*, bin 5.04), carotenoid Z-isomerase [38] (Z-ISO, locus, *y9/y12*, bin 10.02), carotenoid cleavage dioxygenase 1 (CCD1, locus *wc,* bin 9.07). Using comparative homologies with *Arabidopsis thaliana* and association mapping, two additional biosynthesis genes, lycopene ε-cyclase [19] (*lcyε*, bin 8.05) and β-carotene hydroxylase 1 [29, 40] (*crtRB1/hyd3*, bin 10.06), have recently been discovered. Elucidation of other genes in the maize carotenoid pathway will be greatly facilitated by drawing upon biochemical genetics from model species to guide targeted searches in the completed maize genome sequence (www.maizesequence.org) and will likely provide functional information for as yet unidentified QTL.

The majority of carotenoid traits appear to be affected by relatively few loci, as compared to traits such as flowering time [102] and oil composition [122] which are governed 50-60 major and minor QTL. Heritable quantitative variation for carotenoid QTL and many other traits arises from the enormous allelic diversity in the maize genome, as measured at both neutral markers and expressed gene sequences [123, 124]. For example, a gain-of-function insertion in the 5' regulatory region of phytoene synthase 1 (*psy1*), has been observed to enhance *psy1* gene expression and activate the carotenoid pathway in white endosperm to produce yellow pigmented grain [21]. Though much of the allelic variation at this locus is in linkage disequilibrium with insertion of the mobile element which may be monomorphic among yellow maize lines, sequence variants have appeared in yellow germplasm since the domestication event, and are likely contributors to further modification of the color phenotype [2]. Similarly, a functional allelic series created by helitron insertions into the 5' UTR of maize carotenoid cleavage dioxygenase 1 (*ccd1*) encoded by the *white cap* (*wc*) locus induced formation of a multi-copy tandem repeat of

the entire locus [84]. This high-copy allele presumably increases the degradation activity of the locus, and is associated with a large reduction in xanthophylls, but not β-carotene (Kandianis, unpublished).

Association mapping has been used as a new strategy for the dissection of complex traits in maize [125]. Sequence variation within the carotenoid biosynthetic pathway at *lcyε* [19] and *crtRB1* [29] (Figure 3.1) has been tested for statistical associations with carotenoid concentration and composition traits. Under the hypothesis that allelic variation at these genes would alter the pathway to cause a compositional change, three polymorphisms at each locus were found to modify pathway branch ratio and conversion of β-carotene, respectively [19, 29]. Among genetically diverse inbred lines, allelic variation at the two loci was found to explain a large proportion of the phenotypic variation in β-carotene proportion, and thus proVA concentration. Considering that these loci encode enzymes that draw from the same biochemical substrates, it is conceivable that selected allelic combinations from both genes could be used to further modify existing pathway fluxes and engineer more desirable carotenoid profiles. In the case of *lcyε*, a weak allele would allow accumulation of β-branch carotenoids at the expense of α-branch carotenoids, while at *crtRB1*, a weak allele is presumed to accumulate β-carotene at the expense of zeaxanthin.

Genetic studies of *crtRB1*, *lcyε*, and *ccd1* have identified allelic variation at seven intragenic polymorphisms as potential targets for manipulation of β-carotene and provitamin A concentrations through conventional breeding. Associated trait effects for these sequence variants have been estimated using highly inbred germplasm, where experimental tests are based primarily upon haplotype comparisons [19, 29]. This approach has great merit in resolving the molecular basis of trait variation [126], and facilitates application of the predicted effects to derived progeny, especially when the trait is largely controlled by additive inheritance. However, evidence for specific combining effects [45] and pathway based epistasis [31, 127] suggests that extension of estimated haplotype effects to predict progeny or genotype performance through an additive model may not be sufficient. Further investigation of dominance and epistasis within *crtRB1*, *lcyε*, and *ccd1* allelic series, as well as their contribution to combining abilities for carotenoid traits, would be highly useful in achieving desired modifications to grain carotenoid profiles through breeding.

Alleles favoring higher β-carotene through proportionally higher β-branch carotenoids or lower β-branch conversion to β-cryptoxanthin and zeaxanthin have been found to be very rare [29], and have not yet been found to co-occur in any line. The favorable alleles at these loci were found in geographically distinct germplasm pools, suggesting that these sequence variants were likely never pyramided and were thus under-utilized. Combinations of these rare and geographically separated sequence variants warrant testing to evaluate their potential use in enhancing micronutrient density of maize grain. To this end, a series of studies are presented where the effects of parental allelic combinations at a QTL (defined in methods) are measured in hybrid and segregating progenies. The objectives were to 1) determine how known alleles contribute genetic effects to proVA variation and total carotenoids in the developed progenies and 2) evaluate the use of these alleles in applied manner through allele-specific marker assisted selection.

3.3 Materials and Methods

3.3.1 Study Definitions and Assumptions

The three loci under investigation in this study, *crtRB1*, *lcyε*, and *ccd1,* are referred to as QTL (Quantitative Trait Locus) and encode the *Zea mays* proteins *beta-carotene hydroxylase 1, lycopene epsilon cyclase*, *carotenoid cleavage dioxygenase 1*, respectively. Regions of sequence found to vary across pedigrees within a QTL are denoted as polymorphisms; variation in the state of a polymorphism denotes an allele. The qualification of "favorable" is assigned to allelic states that promote increases in β-carotene (concentration or proportion) or decreases in α/β branch ratio. Haplotypes describe the linear combination of alleles (QTN, or Quantitative Trait Nucleotides) found at each polymorphism within a single QTL (Quantitative Trait Locus).

Previous association mapping studies [19, 29] have demonstrated that sequence variation at these QTN are associated with functional changes in a given carotenoid. As full length gene sequencing has been used to identify most of these polymorphisms, is it less likely, but still possible, that unidentified sequence variation in complete linkage with the haplotype could be the cause of the trait effects. Limited intragenic recombination is expected to have occurred during the limited number of meiosis (≤ 3) undergone in this experiment; considering the small population sizes under analysis, it is assumed that the chance for such recombination events to occur is close to zero. Therefore, QTN are assumed to be inherited as intact parental haplotype

blocks within a locus. All genes map to different chromosomes (as indicated in Chapter 1) and thus assort independently.

3.3.2 Single Cross and Population Development

3.3.2.1 Reciprocal Hybrids

The Goodman-Buckler Diversity Panel, a set of 281 genotypically and geographically distinct maize inbred lines used for association mapping [43, 86], was grown in multiple years (2001-2003, 2005) 2003 and 2005 at the University of Illinois in Urbana, Champaign. Phenotype data from 2003 and 2005 seasons are listed in Islam 2004 and Appendix A, respectively. A subset of lines with high yields in β-carotene concentration, β -cryptoxanthin concentration, total carotenoid concentration or β -carotene/total carotenoid ratio across 2003 and 2005 seasons were selected as parents for dosage and hybrid studies (listed in Table 3.1).

Crosses were made from a set of 11 inbreds; of 55 possible hybrids, only 18 were created, including 15 one-way crosses and 3 reciprocal crosses (Table 3.2). The allelic composition at QTL reported to influence endosperm carotenoid concentrations [19, 29] was obtained for all inbred lines used in this study (Table 3.3). Considering that parent inbreds ranged in physiological maturity and germination rates, it was not possible to obtain a more complete mating design that would facilitate estimation of combining abilities. F_1 seed was produced at the University of Illinois Crop Sciences Research and Education Center (Urbana, Illinois, latitude 40° 6' N) in 2007. Inbreds and F_1 hybrids were grown in single row plots (5 m long rows, 76 cm between rows) in one replication in an alpha (0,1) incomplete block design during winter 2007-2008 in Puerto Vallarta, Mexico (latitude 20° 40' N). The trial was replicated during summer 2008 in Urbana, Illinois in single row plots (5 m long rows, 76 cm between rows) in one replication in a completely randomized design. During both seasons, fertilizer, irrigation water, and herbicides were added according to standard agronomic practices. At both locations, ears were hand harvested at maturity (ranging 12-20% moisture, and were dried by forced hot air (approximately 100°F) to final moisture content of 4-5%.

3.3.2.2 $F_{2:3}$ Mapping Populations

Five $F_{2:3}$ populations segregating in at least one of two carotenoid QTL (*lcyε* and *crtR-B1*) were used in this study. DEexpxCI7 (n=103) and A619xSC55 (n=227) $F_{2:3}$ mapping

populations were created and grown as described by Stevens [44]. Two replicates of these populations were grown in alpha (0, 1) incomplete block design across two environments (El Batan, Mexico and Urbana, Illinois); only BLUPSs from the Mexico environment will be discussed. KUI3xB77 (n=110) and KUI3xSC55 (n=66) $F_{2:3}$ populations, developed with the intent to recombine the most favorable *crtRB1* and *lcyε* alleles, were derived from F_2 seed produced in Puerto Vallarta, Mexico 2008. F_2 seeds were single kernel genotyped as described below, and a genotyped subset was grown at the CIMMYT experimental station at Tlaltizapan, Morelos (latitude 18° 40' N) in a completely randomized design. W64axA632 (n=200) $F_{2:3}$ mapping population was created and grown as described by Wong et al, 2004. All populations are listed in Table 3.4 along with haplotype information for the QTL of interest.

3.3.2.3 Breeding Populations for Allele-Specific Marker Assisted Selection (MAS)

The following breeding program is outlined in Figure 3.3. Advanced breeding lines derived from tropical and subtropical material, and selected for high grain provitamin A concentration were donated by CIMMYT-Harvest Plus (K. Pixley and N. Palacios-Rojas). These tropical inbred lines have been bred for higher provitamin A (6-8 μg g^{-1}) and carotenoid levels using visual selection of orange kernel phenotype and HPLC profiles. Of 20 lines received from CIMMYT in 2007, only six were used in generating MAS populations. Carotenoid Syn3 FS-8-4-2-B and Carotenoid Syn3 FS-8-4-6-B, derived from University of Illinois Syn1376-1379, originated from a composite of KUI3, KUI11, KUI43, KUI2007 and SC55. KUI Carotenoid Syn-FS17-3-2-B, derived from University of Illinois Syn1388-1397, originated from a composite of KUI3, KUI11, KUI43 and KUI2007. FloridaA+Syn-FS6-3-1-B, derived from Illini Orange Synthetic, originated from a composite of A619, R30, PI18619, SAPhoto876TR1042, and FLA949161-2. CIMMYT inbreds CML297 and CML 324 were also used. Lines were grown in single row nursery plots during 2007 in Urbana, Illinois, and crosses were made to produce F_1 seed, hereafter noted as HP donors. Temperate inbred lines used as donors of carotenoid QTL (*lcyε* and *crtR-B1*) from the Goodman-Buckler Diversity Panel were used to produce F_1 seed at Urbana, Illinois in 2007, as described above; these hybrids are hereafter called GDL donor hybrids.

Double cross hybrids (synthetic populations, S_0 generation) were produced during winter 2008-2009 using HP donor hybrids as the female pedigree, and GDL donor hybrids as the male

pedigree. To pyramid desired QTL and background genetics in the male pedigree and recombine tropical materials in the female pedigree, hybrids were used as parents rather than inbreds. This also resulted in larger F_1 and S_0 seed yields. Populations listed in Table 3.5 were generated. Representative of most tropical lines tested, inbreds constituting hybrids of the female pedigree did not contain favorable alleles for any *crtRB1* polymorphism, but all had favorable alleles for *lcyε* for at least one of three polymorphisms. Hybrids of the male pedigree had at least one favorable *crtRB1* allele at each of three polymorphisms, but were lacking favorable alleles for the three polymorphisms at *lcyε*.

Ten S_0 populations were planted ear to row in summer 2008 in Urbana, Illinois in two row plots. Two additional S_0 populations were genotyped during 2008 (PopA-GH and PopB-GH), and were advanced to the S_1 generation in the greenhouse. S_0 plants were genotyped, and only those maintaining at least one copy of each favorable QTL were self-pollinated to create the S_1 generation. *lcyε*-5'TE and *crtRB1*-InDel4 polymorphisms were chosen to guide marker assisted selection (MAS) as they were noted in Harjes [19] and Yan and Kandianis [29] to contribute substantial additive effects to α/β branch ratio and β-carotene concentration respectively. In addition, PCR based assays for these polymorphisms were more reliable than other allele-specific markers available at that time. The proposed aim of genotyping in the S_0 generation was to increase the frequency of favorable alleles in each population, such that in the S_1 generation, subsequent phenotyping would permit selection of individuals homozygous for favorable alleles at both loci. Genotyping of the S_0 generation occurred during the field season, and plants with at least one copy of the favorable alleles at both loci were selected for self pollination.

Sampling from several S_1 ears, 8-16 seeds per ear were screened by single kernel genotyping with allele-specific markers. From available genotypic information, all S_1 seeds with favorable homozygous *lcyε* and *crtRB1* alleles were selected for advance. In order to evaluate genetic contrasts, additional genotypes with heterozygous or homozygous unfavorable allelic combinations were selected. S_1 plants were grown in greenhouse hills in Urbana, Illinois (20 cm within row spacing, 50 cm between row spacing; 15 hour day cycle, 80°C day/ 70°C night temperature cycle) in winter 2008-2009, and S_0 parent populations were recreated to establish parental checks. Three $S_{1:2}$ populations were grown in replicated field trials (two locations, two replicates per location) in a randomized complete block design at the Purdue University

Agronomy Center for Research and Education (West Lafayette, Indiana). In each replicate, S_0 parent populations were included as checks.

3.3.3 Carotenoid Extraction and Quantification

3.3.3.1 Overview of Methods Used and External Standard Quantification

Carotenoid profiles for germplasm sets used in this study were obtained through several extraction and separation methods as protocols evolved to achieve higher recovery and throughput with equivalent labor. Observations suggest that methods may vary in the recovery of particular compounds. Therefore, our studies have been separated by extraction and separation method as indicated below.

Quantification of compounds for all extractions was accomplished by standard linear regression with external standards. Lutein, zeaxanthin, β-cryptoxanthin and α-carotene standards were purchased from Carotenature (Lupsingen, Switzerland), and β-carotene was purchased from Sigma-Aldrich (St. Louis, MO). External standards were dissolved in organic solvent, and absorbance of stock solutions was measured spectrophotometrically at 450 nm. Stock concentrations were calculated using the Beer-Lambert law and published extinction coefficients [32]. Five dilutions for each compound were used in triplicate to perform linear regressions of peak area on external standard concentration. δ-tocopherol, purchased from Sigma-Aldrich (St. Louis, MO) was used as an internal standard (at 290 nm) to adjust for recovery of carotenoid extracts. Although present in very low levels, δ-tocopherol is produced in corn grain. Therefore, extracts were spiked with large amounts of internal standard, such that endogenous δ-tocopherol would account for less than 5% of the total amount.

3.3.3.2 Carotenoid Extraction Protocol: Kurilich and Juvik[118] and *Wong* [26]

Carotenoid extractions were performed as in Wong [26]. This protocol is almost identical to that proposed by Kurilich and Juvik [118]. Briefly, the method uses an intermediate scale (7 mL) organic extraction (hexanes or petroleum ether) and a heated saponification step to liberate and partition esterified carotenoids. Carotenoid separation was performed on a C-30 YMC Waters Carotenoid column, with isocratic separation and photodiode array detection as described in Wong [26]

Germplasm from the hybrid evaluation carried out in Puerto Vallarta, Mexico (two biological replicates and two analytical replicates) was analyzed with this method. $F_{2:3}$ mapping population W64a x A632, phenotyped in Wong et al, 2003, was also profiled as described. $F_{2:3}$ mapping population KUI3 x SC55 (one biological replicate, one analytical replicate) was phenotyped by X. Yang at China Agricultural University with the same method.

3.3.3.3 Carotenoid Extraction Protocol: Modification from Tian [128]

Carotenoid extraction and separation protocols were modified from Tian[128]. A bulk of 50 kernels was finely ground with a Stein Mill. Further homogenization was performed on 30 mg (DW) of sample using a QIAGEN Tissuelyser II homogenizer and two-4 mm steel ball bearings. Carotenoids were extracted by further homogenization in microtubes using 600 µl of 2:1 methanol:chloroform containing BHT (1 mg/mL) and δ-tocopherol as an internal standard. After addition of 400 µl water and 200 µl chloroform, the samples were vortexed for 15 minutes and spun at 12 000 g for 10 minutes. The bottom fraction was collected, dried, and resuspended in 200 µl injection buffer (95:5 acetonitrile: ethyl acetate). HPLC analysis was carried out on 50 µl of the final extract.

Carotenoids were separated by HPLC on a C18 column (Spherisorb ODS2 5 micron, 150 × 2.1 mm, Column Engineering, Ontario) with a Waters 2690/2695 Alliance Separations Module at variable flow rates with solvent A (acetonitrile: water [9:1 v/v]) and solvent B (ethyl acetate) on the following gradient: 0–20 min, 5% to 77% B, 1.0 mL/min; 20-20.2 min, 77 to 100% B, 1.0 mL/min; 20.2-22.2 min, 100% B, 1.5 mL/min, 22.2-22.4 min, 100-5% B, 1.5 mL/min; 22.4-25 min, 5% B, 1.0 mL/min. A Waters 996 Photodiode Array Detector was used for peak detection and HPLC peak areas were integrated at 450 nm using Waters Empower software.

Samples for the hybrid evaluations grown in Urbana, IL (three biological replicates, one analytical replicate) were evaluated with this protocol. $F_{2:3}$ mapping population KUI3 x B77 (one biological replicate, one analytical replicate) and all samples from the marker assisted selection experiment (one biological replicate, three analytical replicates) were also profiled with this method.

3.3.3.4 Carotenoid Extraction Protocol: Modification from Granado [46]

$F_{2:3}$ mapping populations DEexp x CI7 and A619 x SC55 (two environments, two biological replicates per environment, one analytical replicate) were profiled at Iowa State University by W. White and W. Liu and described in Stevens [44]. Briefly, this method uses an intermediate scale (5 mL) organic extraction (hexane: methylene chloride) and a heated saponification step. Carotenoid separation was performed on a C-30 YMC Waters Carotenoid column, with gradient separation (methanol/ methyl-tert-butyl ether) and photodiode array detection.

3.3.4 DNA Extraction and Genotyping

DNA micro-scale extractions from leaf tissue were modified from laboratory protocols originally based on a large scale CTAB method. Approximately 50 mg of finely ground, frozen leaf tissue was mixed with 500 µl CTAB extraction buffer (Tris-HCl, 100 mM, pH 8.0; EDTA, 20 mM, pH 8.0; NaCl, 1.4 M; CTAB, 2% w/v; β-mercaptoethanol, 28.6 mM), and was heated to 70°C in a water bath with intermittent vortexing. The sample was vigorously mixed with 500 µl chloroform/ isoamyl alcohol solution (24:1 v/v), and centrifuged for 10 minutes, 7 000 RPM. To reduce RNA contamination, Ribonuclease A (30 µg, Sigma-Aldrich) was added and the sample was incubated for 20 minutes at 37°C. Isopropanol (300 µl) was added and the sample was incubated for 30 minutes at -20°C to precipitate genomic DNA. The sample was centrifuged for 15 minutes, 12 000 RPM, and the resultant pellet was washed with 70% ethanol to remove salts. After removing residual ethanol, the pellet was re-hydrated in 1x TE (Tris-HCl, 10mM, pH 8.0; EDTA, 1 mM, pH 8.0). Endosperm DNA was prepared using a seed-based extraction method from Gao et al, 2008 with an SDS-Sarkosyl extraction buffer.

Genotype data for polymorphisms in carotenoid QTL including *lcyε* [19] , *crtR-B1* [29] and *ccd1*, (Kandianis, unpublished) were obtained for maize inbred lines in the Goodman-Buckler Diversity Panel including all inbreds used in dosage studies, hybrid evaluation, segregating populations and MAS donor pools (Table 3.3). Allele-specific PCR based assays distinguishing haplotypes for these carotenoid QTL include: *lcyε*-5'TE, *lcyε* -SNP216 and *lcyε* -3'TE as reported in Harjes.[19]; *crtR-B1*-5'TE, *crtR-B1*-InDel4; *crtR-B1*-3'TE as reported in Yan and Kandianis [29] (summarized in Figure 3.1) and ccd1-5p as designed in Chapter 2. Genotype information specific for the inbreds in this study is listed in Table 3.3. Primer information,

cycling conditions and genotypes for the entire Diversity Panel are listed in Appendices A, B and C.

Using genotype specific sequence provided in kind by Pioneer, an additional codominant PCR based marker for *lcyε* was designed against a pair of triplet SNPs in lcyε exon 8 which differ between A619 and SC55. A three primer set yielding two variable length amplicons (SC55 allele: 391 bp, A619 allele: 112 bp) was created. Primer sequences are as follows: LCYe-MZA-SC55-L ATT TTT CTG GTA TTT ATT CAG C, LCYe-MZA-A619-L AAG GCT ACT ACC TCC ATG AAA, LCYe-MZA-ALL-R1 AAT GAG AAT AGT ATG AGA TCG. Cycling conditions for LCYe-MZA assay were performed as follows: 1 cycle, 95°C, 2'; 8 cycles, 92°C, 30 s; 58°C, 30 s (reduce 2°C/ cycle), 68°C, 45 s; 30 cycles, 92°C, 30 s; 44°C, 30 s, 68°C, 45 s; 1 cycle, 68°C, 2'.

3.3.5 Statistical Analyses

All analyses were conducted with SAS software (SAS Institute, Version 9.2). Descriptive statistics for hybrid studies were obtained through Proc MEANS. Midparent heterosis was calculated as the difference between hybrid and midparent carotenoid concentrations using entry means from the hybrid evaluation as described in Falconer and Mackay [129].

Analyses of variance for segregating populations and MAS selections were performed in Proc GLM, in which QTL terms were considered as fixed effects, and Type I error rates of α=0.01 and 0.05 were used. Normality, heterogeneity of variances and correlated errors for model residuals were evaluated using Proc UNIVARIATE and graphical techniques. To satisfy ANOVA assumptions, phenotype data was transformed when necessary with power transformations selected from Proc TRANSREG; β-carotene was transformed using a square root transform, whereas all other traits were transformed using the natural logarithm . The significance of *crtRB1* and *lcyε* effects on carotenoid concentrations was evaluated across several genetic populations using two-way analysis of variance with fixed effects. Traits including β-carotene (βcar), proportion of β-carotene (βcar/ALL), ratio of β-carotene to downstream conversion products (βcar/ (βcry+zea), ratio of pathway branches (α/β branch) and total carotenoids (Total) were evaluated. Phenotypic variation was modeled as:

$y_{ijk} = \mu + a_i + b_j + ab_{ij} + e_{ijk}$, where y is the trait value (μg g^{-1}), μ is the population mean, a_i is the effect of *crtRB1* with the i^{th} allele, b_j is the effect of *lcyε* with the j^{th} allele, ab_{ij} is the interaction of *crtRB1* and *lcyε* at alleles i and j, and e_{ijk} is the experimental error of the k^{th} replicate.

Estimates for the effect of haplotype substitution were obtained through linear regression carotenoid concentration on QTL using Proc REG. Only polymorphic QTL *(lcyε* and/or *crtRB1*) were tested. Haplotype substitution was estimated only if the tested QTL was found to be statistically significant for the explanation of variation in β-carotene in a two-gene model, and linear regression was performed with one or both QTL as predictor variables. and models were evaluated on the basis of overall significance at a Type 1 error rate of α=0.01. As previously mentioned, the favorable haplotype at a carotenoid QTL is defined here as the condition which will increase β-carotene concentration. Accordingly, three haplotype states can be considered in a diploid plant: homozygous favorable (2 copies), heterozygous (1 copy), homozygous unfavorable (0 copies). Adjusted trait least square means at all three haplotype states of one given QTL were estimated using one-way ANOVA with the following model $y_{ijk} = \mu + a_i + B(x_j - \bar{x}) + e_{ijk}$ where y is the trait value (μg g^{-1}), μ is the population mean, a_i is the effect of *crtRB1* with the i^{th} allele, $B(x_j - \bar{x})$ accounts for the effect of *lcyε* as a covariate with all haplotypes yielding the common slope B, and e_{ijk} is the experimental error of the k^{th} replicate. A model with *lcyε* as the main effect and *crtRB1* as the covariate was also tested.

Analyzed traits are abbreviated in text as follows: β-carotene, βcar; β-cryptoxanthin, βcry,; zeaxanthin, Zea,; lutein, Lut,; total carotenoids, ALL or Total which is the summation of lutein, zeaxanthin, zeinoxanthin (if measured), β-cryptoxanthin, α-carotene (if detected), and β-carotene. Ratios of the absolute concentrations are designated by '/'. ProVA concentrations are calculated as β-carotene + 0.5*β-cryptoxanthin + 0.5*α-carotene unless otherwise indicated.

3.4 Results

3.4.1 Hybrid Evaluation for High Carotenoid Concentrations

Grain carotenoid concentrations were compared in a select set of diverse inbreds and hybrids grown in Puerto Vallarta, Mexico (winter 2007-2008) and Urbana Illinois (summer 2008). Means for lutein, zeaxanthin and β-carotene were found to significantly differ between environments (α=0.01, Table 3.6), whereas those for β-cryptoxanthin and total carotenoids did not. Correlations for absolute concentrations and for ranks between environments were found to be positive and significant (Table 3.6). Entry means for all measured carotenoids by environment are listed in Appendix D.

As different extraction and separation methods (Kurilich versus Tian, see methods for details) were used for samples from Mexico and Illinois, the environmental effect was confounded with the phenotyping method in this experiment. A test for bias using the same grain samples with different extraction/ separation methods (n=22) revealed that Tian's method has a bias against lutein (data not shown). Comparison of the Mexico and Illinois environments show that lutein concentrations are, on average, higher in the Illinois set which was prepared with Tian's method. This suggests that lutein concentrations could be even higher than reported here.

Several hybrids and inbreds consistently yielded high carotenoid concentrations based on comparison of ranks for entry means between the two environments. Inbred lines that ranked highest in certain carotenoid concentrations were frequently observed to contribute to hybrids with high levels of the corresponding carotenoid concentrations (Table 3.7), consistent with the high heritabilities reported for these traits. Lutein concentrations in hybrids were most positively affected by inbreds A619 and DEexp. Hybrids with A272, HI27 or CML328 parentage yielded the highest zeaxanthin and β-cryptoxanthin concentrations. β-carotene concentrations were highest in hybrids with CI7 or DEexp parentage, consistent with *per se* performance. Although it did not exhibit high *per se* β-carotene concentration, inbred SC55 gave rise to multiple high yielding β-carotene hybrids, suggesting that SC55 may be an excellent source of favorable β-carotene QTL. Rough mid-parent heterosis estimates across both the Illinois and Mexico environments were found to be positive for several of these high ranking hybrids (Appendix E) including CI7xDEexp (lutein and β-carotene), NC354xHI27 (zeaxanthin and β-cryptoxanthin), KUI43xB77 (lutein), B77xA272 (zeaxanthin) and HI27xCML328 (zeaxanthin). The results are

suggestive of favorable general and specific combining abilities, and support observations of significant additive effects found in studies of endosperm dose (Appendix F).

It was also observed that simultaneous achievement of high total carotenoids and β-carotene may be complex. Based on its *per se* profile, inbred A272 would be an excellent donor for total carotenoids; however, this overall increase in accumulation of pathway intermediates would not necessarily increase β-carotene, as marked by the absence of an A272 hybrid among the high ranking β-carotene hybrids.

3.4.2 Genetic Effects of Carotenoid QTL in Segregating Populations

An investigation of QTL significance and effect was performed with *crtRB1* and *lcyε* to assess QTL significance and effect in five different segregating populations. Two-gene models for β-carotene (βcar), proportion of β-carotene (βcar/ALL), proportion of β-carotene to other β-branch carotenoids (βcar/(βcry+zea)), α/β branch and total carotenoids were highly significant at α=0.01, and explained a large proportion of trait variation in each population (Table 3.8), with few exceptions. Main effects for both genes were not necessarily significant for a given trait. Variation for βcar and βcar/ (βcry+zea) was consistently attributed to *crtRB1* in all populations segregating at this locus, indicating that the *crtRB1* effect is robust across various genetic backgrounds. βcar/ALL, an indicator of proportional substrate flux to β–carotene, was also significantly affected by changes in *crtRB1* allelic complement. While it is possible that this effect is derived from changes in βcar alone, it is likely that the significant effect of *crtRB1* on total carotenoids observed in populations KUI3xB77 and KUI3xSC55 also contributed to the significance of the ratio in these crosses (Table 3.8).

lcyε was found to explain a large proportion of the variation found for α/β branch in all tested populations (Table 3.8), which was consistent with association analyses in Harjes [19]. In contrast to this work, however, *lcyε* was not found to consistently explain variation in βcar/ALL, suggesting that original detection of the effect on this ratio trait was enabled by use of association analysis where more diverse phenotypes were under evaluation than in narrow-based segregating populations. Although both *crtRB1* and *lcyε* were found to specifically affect βcar, or traits that would presumably affect βcar (i.e. α/β branch), no two-gene interactions were

detected for β-carotene traits. This indicated that alleles at both QTL were acting in an additive manner to produce the final carotenoid phenotype.

The direction and magnitude of trait effects attributed to allelic doses of *crtRB1* and *lcyε* was surveyed across populations. Models of carotenoid concentration with *crtRB1* as the main effect adjusted by *lcyε* covariate were all highly significant (α=0.01) and accounted for a large percentage of the variation in βcar (R^2 range: 7.8-43.5%), βcar/βcry (R^2 range: 21-69.3%), βcar/Zea (R^2 range: 15-63.1%), and βcar/ALL (R^2 range: 17-59.6%), across four separate populations (Table 3.9). A comparison of the fold change between homozygotes suggested that presence of two copies of the favorable *crtRB1* haplotype lead to moderate increases in βcar, and even larger increases for all ratio traits. Tests of mean separation revealed that heterozygote and unfavorable homozygote classes (haplotype copies 1 and 0) were often not significantly different (α=0.05), whereas the favorable homozygote class (haplotype copy 2) was different from both.

Effects of *lcyε* on α/β branch ratio were large (R^2 range: 10-56.8%), but fold change between the favorable and unfavorable homozygote classes was less variable (Table 3.10) than that found with *crtRB1* (Table 3.9) across all five populations. The lowest α/β branch ratio, and thus the more favorable phenotype, always occurred in the favorable homozygote (haplotype copy 2). Means for haplotype copies 0 and 1 of *lcyε* were often not significantly different (α=0.05), suggesting that the homozygous favorable condition of *lcyε* may be required to significantly alter carotenoid allocation through pathway branches.

Using the assumption of jointly inherited allelic states through parental haplotype blocks, β-carotene concentration was regressed on *crtRB1* and *lcyε* haplotype copy number. Populations were highly variable in the average amount of β-carotene produced as indicated by the intercept (Table 3.11). The range in concentrations among populations was likely attributed to other QTL affecting precursor synthesis, including the committed step at phytoene synthase (*psy1*). The largest effects of *crtRB1* haplotype substitution, found in KUI3xB77 (0.82 μg g^{-1} per copy) and KUI3xSC55 (1.25 μg g^{-1} per copy), correspond to segregation of the most extreme allelic contrasts at all three *crtRB1* polymorphisms (Tables 3.3 and 3.4). In addition, a substantial difference in effect was seen between populations that segregated at *crtRB1*-3'TE (KUI3xB77, 0.82 μg g^{-1}, R^2=39%, p<0.001; KUI3xSC55, 1.25 μg g^{-1}, R^2=27%, p<0.001) and those that were monomorphic for the favorable alleles at *crtRB1*-3'TE (A619xSC55, 0.3 μg g^{-1}, R^2=6.9%,

$p<0.001$; DEexpxCI7, 0.3 μg g^{-1}, R^2=8.4%, $p<0.001$; Table 3.11). Populations with the highest mean β-carotene concentrations (intercept) were also monomorphic at the favorable *crtRB1*-3'TE (A619xSC55 and DEexpxCI7) and *crtRB1*-5'TE (DEexpxCI7) polymorphisms. These results indicated that the greatest increase in β-carotene could be achieved by the inheritance of favorable alleles at all three *crtRB1* polymorphisms, and that *crtRB1*-3'TE could have the largest effect.

Although all populations were segregating for at least one *lcyε* polymorphism, this QTL had an effect on β-carotene concentration only in populations DEexpxCI7 (0.44 μg g^{-1} per copy, R^2=20%, $p<0.001$) and W64axA632 (0.07 μg g^{-1} per copy, R^2=10%, $p<0.001$), both of which had moderate allelic contrasts at *lcyε*-5'TE and contrasts at *lcyε* -SNP216 (Table 3.12). Interestingly, the KUI3xB77 and KUI3xSC55 populations, designed to have an extreme allelic contrast at *lcyε* -5'TE (which reportedly has the largest effect on α/β branch ratio [19]), did not reveal any *lcyε* specific effects on β-carotene. This suggests that significance for α/β branch ratio (Table 3.8), does not necessarily lead to a greater chance of increasing β-carotene in segregating populations.

A comparison of the effect of *lcyε* on branch traits was made across all populations to determine if *lcyε* contributed to modifications in both pathway branches. Varying allelic composition at *lcyε* was found to affect α branch amounts for all populations (Table 3.12), corresponding to significance of *lcyε* effect for α/β branch ratio. However, a change in α branch was not necessarily accompanied by a similar change in β branch, as β branch variation was found to co-vary with *lcyε* haplotype in only two populations (DEexpxCI7 and W64axA632). Variation in branch concentrations was found to associate with *lcyε* only when segregation occurred at SNP216. Therefore, although all three *lcyε* polymorphisms affect α/β branch ratio, only one polymorphism (SNP216) appears to do so by causing fluctuations in both α (decrease) and β (increase) branches. The coincidental overlap of dual branch effect and β-carotene concentration by variation at *lcyε*-SNP216 indicates that of all identified *lcyε* polymorphisms, only modification at SNP216 appears to be associated with the intended directional effect on β-carotene concentration.

At *lcyε*, accumulation of β -branch carotenoids through the favorable allele of SNP216 was found to occur with a reduction in total carotenoids (Table 3.12). A greater effect on total carotenoid was found at *crtRB1* where increases in β-carotene were found to be much smaller

than associated reduction in zeaxanthin and β-cryptoxanthin suggesting that an overall reduction in the carotenoid pathway is tied to the inheritance of favorable *crtRB1* alleles (Figure 3.2). The effect of favorable *crtRB1* haplotypes in three of four populations segregating in at least one *crtRB1* polymorphism was observed to significantly decrease the amount of total carotenoid (α=0.05) by as much as 23% per dose (Table 3.14). Concentrations of colorless carotenoids phytoene and phytofluene were also affected by *crtRB1* haplotype in A619xSC55 ($p<0.0001$) and DEexpxCI7 ($p=0.0192$); the homozygous classes were associated with a 42% and 25% reduction of these carotenoids (respectively) when the favorable *crtRB1* haplotype was present (data not shown). These results suggest that the favorable alleles of *crtRB1* may down-regulate the carotenoid pathway at or near the committed step of *psy1*.

3.4.3 Phenotypic Response to *crtRB1* and *lcyε* Marker Based Selection

Recombination and selection of desirable genetic variation at *lcyε* and *crtRB1* with the objective to improve grain β-carotene concentration were evaluated in breeding germplasm high in total carotenoids. Synthetic populations were generated from double cross hybrids as depicted in Figure 3.3. The first cycle of selection was designed to increase the frequency of desired alleles in each population at *lcyε*-5'TE (insertion allele 4 or insertion allele 1) and at *crtRB1*-InDel4 (12 bp insertion allele 1). As shown in Table 3.14, expected probabilities for the favorable *lcyε* alleles differed from those for *crtRB1*. For the S_0 generation, the observed percentage of selected genotypes (at least one copy of the favorable alleles at *lcyε*-5'TE and *crtRB1*-InDel4) were quite often less than the predicted percentage of selected genotypes. This suggests sampling error as number of individuals sampled from each population was very small (Table 3.14).

Genotyping of the S_1 generation occurred prior to planting using DNA from chipped endosperm, focused on fewer populations, permitting larger test populations (Table 3.15). Favorable alleles in selected S_0 plants maintained heterozygosity at both *lcyε* and *crtRB1,* or homozygosity at *lcyε* and heterozygosity at *crtRB1*, thus probability of an S_1 progeny homozygous for the favorable alleles at both loci ranged from 6.25 - 25%. Recovery of double homozygotes in all but population Q/R fell at the lower bound or below (3.1-8.5%). Since not all S_0 ears had desirable S_1 genotypes, it is highly likely that the percent recovery was reduced due

to sampling from misgenotyped or errant S_0 selections. Single kernel genotyping, while efficient in enabling the prioritization of resources to individuals of greater genotypic value in the greenhouse, led to poor germination rates (<80%) for most genotypes (Table 3.15). This is mainly attributed to inexperience with chipping endosperm tissue which may have compromised embryo tissues, but also suggests the method works best for genotypes with larger seeds. Recovery of $S_{1:2}$ ears was also reduced by the number of S_1 plants that reached reproductive maturity in the greenhouse.

Population sizes for the $S_{1:2}$ generation were small but still provided the ability to evaluate the effects of different *crtRB1* and *lcyε* genotype classes on specific carotenoid traits (Table 3.16). In most populations, inspection of the trait means for MAS *crtRB1* classes revealed that favorable *crtRB1* genotypes are significantly (α=0.05) associated with increasing β-carotene, but decreasing total carotenoids (Table 3.17). The direction and magnitude of effect for favorable *crtRB1* haplotypes on total carotenoids in the MAS populations was similar to that for the segregating populations (Table 3.18), where reductions of up to 16% total carotenoids per *crtRB1* dose were observed. In contrast, *lcyε* was found to affect α/β branch variation, but not necessarily the intended target trait, β-carotene (Table 3.16). Inspection of *lcyε*-5'TE polymorphism revealed significant effects on variation in α–branch carotenoids for all populations (α=0.05), but inconsistent effects for β-branch carotenoids.

Original HP donors had proVA concentrations (6-8 μg g^{-1}) similar to those for the homozygous unfavorable *crtRB1* genotype class among advanced $S_{1:2}$ families (2.8-9.5 μg g^{-1}), but the homozygous favorable *crtRB1* class had extremely high proVA concentrations (9.21-17.88 μg g^{-1}). Selection was found to have a significant effect on β-carotene and proVA concentrations in four of five populations as observed by the selection differentials (Table 3.17). Comparison of the homozygous classes in populations A-GH, B-GH and MN revealed a 3.5-11 fold difference in β-carotene. A smaller 1.5-3 fold improvement in proVA levels was attributed to the large amount of β-cryptoxanthin observed in the *crtRB1* heterozygotes and unfavorable homozygotes that was not present in the favorable homozygotes (Table 3.17). A substantial amount of variation in proVA (R^2 range 16-74%) was explained by *crtRB1* (data not shown). Of all GDL hybrids used in this study, those consisting of combinations of DEexp, CI7 and B77 were found to have very high β-carotene and β-cryptoxanthin concentrations, similar to the

hybrids evaluated in this study (Appendix D). This suggests that combinations of these three inbreds likely contributed additional favorable genetics for increased β-carotene and provitamin A other than QTL at *crtRB1*. On average, progeny in population M/N exhibited less total carotenoids than other sampled populations. As the SC55 inbred was found to contain the deleterious Wc-allele of *ccd1*, it is possible that an overall reduction in all carotenoids except b-carotene was incurred through the inheritance of this allele.

Lines producing proVA levels at or above target Harvest Plus amounts (15 μg g^{-1}) have been identified from evaluation of the $S_{1:2}$ generation. Replicated field tests of $S_{1:2}$ lines from populations A-GH, B-GH and M/N grown in summer 2009 will be assessed for carotenoid profiles in Fall 2009. The lines highest in both β-carotene and proVA concentrations are derived from different ears in population A-GH (09GH-70, S_1 ear A8 and 09-GH76, S_1 ear A23), and population M/N (09GH-40, S_1 ear M4), suggesting that further selections could be made from the original S_1 populations if additional sampling of recombinants from the same favorable genetic pool was desired. Parentage among the three selected populations is shared through one inbred on the male pedigree and one in the female pedigree (Table 3.5). Although population B was not genotyped in the S_1 generation and was therefore planted blindly, enough information from the S_0 generation permitted selection of this line for advancement to $S_{1:2}$ generation and it produced two of the highest provitamin A yielding lines 09GH-113 and 09GH-115, both from ear S_1 B31.

Inbreeding appeared to reduce total carotenoids, as observed in the contrast of the S_0 and $S_{1:2}$ generations (Table 3.16). Decrease in total carotenoids between generations did not appear to be attributed to *crtRB1*-associated effects as the S_0 parental checks, which are at most heterozygous for the favorable *crtRB1* alleles, and have more total carotenoids than both heterozygous and unfavorable homozygous classes of the $S_{1:2}$ generation in several of the populations. $S_{1:2}$ lines of the unfavorable *crtRB1* genotype, which should not be affected by *crtRB1*-related carotenoid reduction, yield lower amounts of total carotenoids than S_0 lines, which are segregating at *crtRB1* (Figure 3.4). Heterosis for total carotenoids was not observed between inbred and F_2 seed in the hybrid evaluation presented here (Appendices D and E), suggesting that the effects of inbreeding on total carotenoids will warrant future study.

3.5 Discussion

3.5.1 Genetic Characteristics of Carotenoid QTL *lcyε* and *crtRB1* in Non-Inbred Germplasm

Quantitative variation of carotenoid traits was investigated taking advantage of previously identified QTL affecting β-carotene conversion [29] (*crtRB1*), pathway branch proportion [19] (*lcyε*) and carotenoid degradation [84] (*ccd1*). Special emphasis was placed on *crtRB1* and *lcyε*, as the two QTL have been shown to affect β-carotene and proVA concentrations in maize grain, whereas *ccd1* affects all carotenoids except β-carotene (Kandianis, Chapter 2), and is therefore of less importance to the above breeding goal. We hypothesized that β-carotene could be increased through selection of weak alleles at both *lcyε* and *crtRB1* loci, which would reduce epsilon ring cyclization and shift α-branch carotenoids to a β-branch with less efficient intra-branch hydroxylation. With knowledge of the *crtRB1* and *lcyε* allelic series, we evaluated the correlation of carotenoid phenotypes in segregating populations with QTL effects predicted from an association panel comprised of diverse inbreds.

Dosage and hybrid evaluations revealed that variation in lutein, zeaxanthin and β-cryptoxanthin is attributed primarily to significant additive effects and less to dominance effects. These results are in agreement with combining ability estimates measured by Egesel [45]. Less variation was noted for β-carotene, and was most often attributed solely to additive effects. Mid-parent heterosis for β–carotene was observed in some hybrid combinations, suggesting that dominance at carotenoid QTL can positively affect the trait. However, it seems likely that with the various QTL affecting β-carotene, recombination to accumulate favorable alleles at different loci followed by fixation is necessary to make improvements in this trait. *lycε* and *crtRB1* QTL affecting β-carotene proportion in association populations [19, 29] were found to also have additive effects on the trait in segregating populations. In comparison to the effects of dominance noted for β-carotene in some hybrids, only additive effects of *crtRB1* and *lcyε* were significant for β-carotene in the segregating populations. In cases where both QTL had additive effects on β-carotene, a non-additive interaction was not statistically detected.

Using a similar assessment, effects for *lcyε* haplotypes were significant for variation in β-carotene only in populations segregating at SNP216. This was surprising, as prior estimation of QTL effects in association panels noted that *lcyε*–5'TE had a much larger effect than SNP216 on β-carotene proportion and on α/β branch ratio [19]. Since modification of the α/β branch ratio by

way of *lcyε* did not lead to equal effects on β-carotene by all *lcyε* polymorphisms, we tested if each branch was being equally affected by the allelic substitution. Of all *lcyε* polymorphisms, only SNP216 was found to increase β-branch accumulation. All lcye polymorphisms, however, were found to alter the concentration of α–branch carotenoids. These observations led us to conclude that only SNP216 was functioning in segregating populations according to the hypothesized mechanism to reallocate α-branch carotenoids to the β-branch which was based on association panel data. This effect was also noticed in the MAS experiments, where *lcyε*-5'TE was shown to be highly effective in changing the branch ratio, but did not contribute to the variation at β-carotene. The polymorphism at *lcyε*-SNP216 leads to an amino acid substitution at residue 70, at which the high α/β branch phenotype is associated with a glutamic acid residue and the low α/β branch phenotype is associated with an aspartic acid residue. The substitution maintains the acidic character at this residue; however, it is possible for a change in protein conformation to arise from amino acid size differences. A comparison with published Arabidopsis *lycopene cyclase* motifs and cofactor binding sites did not suggest that residue 70 in *Zea mays* was involved in known catalytic sites [130]. Although this information does not indicate what the basis of the SNP216 effect could be, it is still evident that tandem selection at both *crtRB1* and *lcyε* should be employed, using favorable *crtRB1* alleles at 5'TE and 3'TE, and the favorable *lcyε* allele at SNP216 for improved grain β-carotene concentration.

Information for *crtRB1* and *lcyε* allelic effects reported in inbreds could have correctly predicted some of the highest yielding entries. For example, CI7xDEexp is one of the few crosses with complementary haplotypes at the *crtRB1*-InDel4 and 5'TE polymorphisms, both of which are predicted to have large positive increases in β-carotene with the favorable allelic states. In contrast, haplotype predictions do not agree with the high zeaxanthin concentration of B77xA272, which would have been expected to have less zeaxanthin due to the B77 allele for *crtRB1*-5'TE and InDel4. Comparison of this hybrid with the dosage experiment of the same parentage shows that the A272 haplotype has a partly dominant effect on zeaxanthin and yields more total carotenoid than any other line. Given this observation, it is surprising that the A272 effect does not contribute to high yielding β-carotene hybrids, and illustrates that allelic information for higher order pathway steps (such as those at *psy1*) is needed to have better predictive capacities.

3.5.2 Epistatic Effects in Metabolic Pathways: Exception or Rule?

The molecular basis of higher β-carotene concentration or proportion through *crtRB1* and *lcyε* allelic effects has been shown to be largely dependent upon weakened transcript expression attributed to the allelic state of the 5' and 3' ends [19, 29]. Conversion of isoprenoid precursors or synthesis of downstream carotenoid pathway intermediates in maize is greatly affected by transcript levels of the biosynthesis genes themselves [31, 127, 131], suggesting that gene interaction through differential expression provides regulatory control of the pathway. Many interactions thus far appear to affect, or be affected by, the rate limiting step at *psy1* [72]. Inhibition of transcript levels for *lycopene β-cyclase* (*lcy*β) in a null mutant was shown to increase endosperm tissue transcript levels for *lcyε*, *psy1* and several isoprenoid biosynthesis genes as compared to wild type *lcy*β lines [127]. Similarly, increasing phytoene synthase activity through transgenic means was shown to upregulate endogenous *lcyβ* but reduce *crtRB1* (*hydb*) expression [132]. The success of Golden Rice and Golden Rice 2, where introduction of a suite of pathway transgenes was used to provide the genetic variation in carotenoid biosynthesis which was lacking in rice was based on increased biochemical activity and transcript upregulation of endogenous pathway components [133, 134]. It is necessary to note that in studies using null mutations or transgenic variation dominant to endogenous pathway activities, the integrated processes of synthesis, conversion and degradation may vary greatly in comparison to what would be observed with less extreme alleles typical of natural variants. The combined results indicate that transcriptional differences arising from genetic variation within the pathway can provide feedforward or feedback regulation on other pathway steps, providing examples of genetic epistasis through the biochemical manifestation of multigene interactions.

Based on the hypothesized mechanism of β-carotene accumulation through combined modifications to carotenoid pathway branch allocation and branch conversion, we hypothesized that the effect of *crtRB1* could be dependent upon *lcyε* genotype. Epistasis between *crtRB1* and *lcyε*, however, was not statistically detected in any of the segregating populations for β-carotene concentration or proportion. Considering the detection of transcriptional regulation in the maize carotenoid pathway with mutant and transgene alleles, we found the absence of this two gene

interaction to be surprising. The absence of an interaction in the case of *crtRB1* and *lcyε* could indicate that hydroxylation activities are not restricted by cyclization activities.

Assessment for statistical epistasis as measured by the significance of a two gene interaction across different backgrounds with small sample sizes in this study may not have had the power to detect biological epistasis [135]. In this context biological epistasis is defined as the molecular manifestation of a multi-component interaction rather than a departure from additivity [136]. Even with excellent mating designs that consider allele effects across several genetic backgrounds, and have large sample and replicate sizes in maize, the detection of epistasis can still be uncommon [102]. Therefore, another obvious alternative is that epistasis does not contribute greatly to the variation in β-carotene. Considering that biochemical epistasis has been found to be less heritable, and slightly less detectable, than transcriptional epistasis [137], an investigation of transcriptional expression across individuals with existing allelic variation and null mutations at *crtRB1* and *lcyε* could help to determine the level of interaction between these two loci.

An epistatic interaction between *crtRB1* and *psy1* may occur, as a substantial reduction in total carotenoids was found to result from *crtRB1* alleles that increased β-carotene. The interaction of β-carotene hydroxylases and carotenoid pathway precursor reactions has been observed in Solanaceous species. This effect can occur in either direction, as diminished hydroxylation of β-carotene has been shown to reduce total carotenoid in tomato [104], but has also been associated with an increase in total carotenoid in potato [138]. This may be an obstacle for concurrent improvement in β-carotene and total carotenoid concentrations while using favorable *crtRB1* alleles.

3.5.3 Influence of Allele Specific Selection on Pathway Flux

Correlations for carotenoid traits within segregating populations were largely positive and significant (Appendix G). While this is indicative that all traits could be controlled by one shared pathway step at *psy1*, the results do not indicate that there is competitive allocation within the pathway. However, if trait correlations are measured among individuals homozygous for the favorable state of *crtRB1*, competitive relationships are observed between β-carotene and β-cryptoxanthin concentrations, as well as between β-carotene and total carotenoid concentrations. Reductions in total carotenoids derived from all intermediates except β-carotene were observed

to be as high as 23% per copy of favorable *crtRB1* haplotype, meaning that nearly half of the total carotenoid may be lost to make gains in β-carotene concentration. Negative relationships could present challenges when simultaneous genetic gain for the competitive traits is the desired response to selection.

The carotenoid profiles of MAS S_1:2 progeny showed gain in β-carotene with the selected *crtRB1* haplotypes was much larger than that of proVA, which is calculated as β-carotene plus half the amount of α-carotene and β-cryptoxanthin. While the primary goal for the proVA breeding program is to attain the target amount of 15 μg g^{-1} proVA, the nature of the proVA composition must also be considered. Nutrition studies comparing the bioavailability of proVA carotenoids have concluded that β-cryptoxanthin is at least as bioavailable as β-carotene from the maize food matrix and possibly more so [139], potentially putting a slightly more favorable weight on the nutritional value of β-cryptoxanthin. Similarly, advancement of lines on the basis of proVA levels alone may ignore the need to maintain sufficient carotenoid precursors for greater proVA carotenoid compositional gains in the future. Therefore, an effort must be made to consider antagonistic but intrinsically related effects between β-carotene, β-cryptoxanthin and total carotenoids, in order to avoid concurrent reductions. Combining allele-specific selection at *crtRB1* with color-based visual selection of total carotenoids to ensure accumulation of precursor substrates appears to be a useful approach. Introgression of favorable *crtRB1* alleles in germplasm visually selected for higher total carotenoid (range 31-52 μg g^{-1}) which is largely derived from higher zeaxanthin concentrations (12-29 μg g^{-1}) has resulted in β-carotene gains larger than those seen in segregating populations with substantially less total carotenoid. Large gains in β-carotene such as those seen in the MAS experiments are largely due to the creation of high total carotenoid phenotypes in lines recipient to the favorable hydroxylase. Therefore, dark yellow to orange color phenotype can be used as a simple selectable marker to ensure that adequate precursor supply is entering the pathway. As useful allelic variants for precursor pathway steps such as *psy1,dxr* and others become available, marker assisted selection for these upstream pathway QTL can be readily integrated into this scheme.

Another alternative to select for competing traits is to weight the value of each metabolite according to a selection metric (economic, nutritional, etc.), determine the relationship between the metabolites by their covariance, and combine the information into a common index. Such

selection indices have been shown to enhance selection efficiency by incorporating data from multiple ancillary traits that are synergistic with the main trait, but are not necessarily of primary importance to selection [140]. Improvements in varietal development for negatively correlated traits such as protein and grain yield have also been observed [141]. Many of the existing selection indices have been based on the work of Smith [142] and Hazel [143] and account for the covariances between traits under selection. Such an index would be useful for the selection of lines with improved proVA content, as selection at *crtRB1* causes antagonistic shifts between β-carotene and β-cryptoxanthin accumulation, and leads to a smaller fold improvement in proVA as noted in the MAS experiment. The results from proVA index developed for a specific MAS populations (Appendix H) indicates the top ranked lines generally agree with those selected by the standard proVA calculation, but that the two methods diverge as ranks decrease. Efficiency of these two methods can be compared in the $S_{1:2}$ families.

3.5.4 Use of MAS to Increase Efficiency of ProVA Breeding Strategies

The merit of incorporating a marker assisted (or based) selection strategy into a breeding program must be justified by increased selection efficiency over that of phenotypic selection alone [144]. Metrics of efficiency include cost, time, technical effort and labor, and ease of integration with the existing program. While phenotypic selection alone has considerably advanced population development in proVA breeding programs (UIUC, CIMMYT-Mexico, IITA), efficiency is hampered by the technical challenge, high cost and limited throughput of current HPLC profiling methods. Alternative methods using near infrared reflectance spectroscopy [145, 146] or spectrophotometry with column separation [147] have been proposed, but are either in the early stages of development or require more labor than return. Grain carotenoid profiles are also dependent upon growth environment [68, 69], seed development, and post-harvest treatment (Kandianis, unpublished data), implying that the heritability of the trait, although high, decreases as these external factors vary and accumulate.

As more efficient alternative, molecular tools and methodologies have been developed to track and screen alleles that are associated with increased β-carotene, which constitutes most of the proVA concentration. Existing codominant molecular markers for *lcyε* [19] and *crtRB1* [29] generated to survey allelic frequencies in various association panels, are in complete linkage with the QTL ensuring absolute predictability of the marker-phenotype association (barring

genotyping errors). QTL studies and evaluation of genetic effects have confirmed that each of these loci explain a large proportion of the trait variation, permitting the inference that selection for a given allele would lead to a marked phenotypic response. Effectiveness of MAS has been reported to be most effective when tight linkage exists between markers and QTL [148, 149], and heritability is high [150], implying that allele specific MAS for increased β-carotene would have a high success rate.

Significant and substantial fold differences in β-carotene and proVA were observed between MAS genotype classes in $S_{1:2}$ progeny indicating that the marker-phenotype association was very strong. Moreover, selection of the same allele in different genetic backgrounds had the same directional effect, although the magnitude of change was different. Evaluation of MAS in different genetic backgrounds has demonstrated that the extent of response associated with the introgressed QTL is highly dependent upon the recipient parents, and is often derived from dominant and epistatic effects that are imposed by the recipient genome [151, 152]. In this study, several genetic backgrounds were combined not only to pyramid the favorable alleles, but also to take advantage of genetics leading to high total carotenoids in the female pedigree and elevated β-carotene genetics in the male pedigree. $S_{1:2}$ populations shown to have the highest trait means were those derived from DEexp, CI7 and B77 parentage, which were all found to have the highest per se and hybrid β-carotene levels. This suggests that these three lines are sources of additional QTL that can enhance the β-carotene trait, and supports future efforts towards the identification of these regions in the $S_{1:2}$ progeny. Further development of these populations to reconstitute the agronomic qualities of the original CIMMYT lines while maintaining high proVA should continue foreground selection of desired alleles and incorporate background selection of the tropical parent genome.

3.5.5 Final Remarks

Candidate gene association mapping has provided increased resolution to the association of phenotype with genotype in maize for a wide range of biochemical, developmental and physiological traits [99, 100, 102]. This approach should only gain more appeal when genome-wide association mapping becomes available in maize, as multiple testing of putative QTL/QTN may allow complete genetic networks to be identified. The question remains if haplotype effects identified in a genetically diverse population of inbreds are predictive of genotype effects in the

context of narrow based biparental segregating populations. For carotenoid traits, much of the variation measured in this study is attributed to additive effects, which would indicate that consistency in the allelic hierarchy should exist between haplotype and genotype effects. Dominance, which is not transmitted through gametic contribution, was also found to contribute to carotenoid variation. The effects of dominance, or epistasis, can be exploited in planned crosses.

Interestingly, certain genetic effects were detected in only one population type, association panel or segregating population. This included detection of joint additive effects for *crtRB1* and *lcyε* on β-carotene in association panels but not segregating populations [19, 29], and detection of *crtRB1* effects on total carotenoids for segregating populations but not association panels. Statistical detection of these genetic effects likely results from several factors including better mean separation across genotypes with more extreme allele pairings in association panels, and smaller background complexity in bi-allelic segregating populations that may pinpoint specific QTL effects on a composite trait. Inferences drawn from other comparisons of association and QTL populations [105] reveal that intragenic polymorphisms found in association studies typically explain QTL effects, or at least uncover polymorphisms in LD with causal lesions. This study affirms that association and segregating populations can successfully cross-validate QTL effects. Additionally, this study highlights the usefulness of testing haplotype effects identified in association populations in segregating progeny genotypes. Breeders can draw upon results from both association and segregating populations to make more informed selection decisions.

Predictive models for β-carotene and proVA concentration to assist with parental selection may have value in breeding applications. This study confirms that the directional effects of alleles within *lcyε* and *crtRB1* polymorphisms are similar across genetic backgrounds. Prediction of the magnitude of effect, however, will require consideration of the range of additive effects at *psy1* as well as epistatic interactions with downstream pathway steps. The large β-carotene responses resulting from introgression into genetic backgrounds with extremely high total carotenoid production in the MAS S_0 parental population contrast with much smaller β-carotene gains and total carotenoid values in narrow-based segregating populations, illustrating the need to focus on favorable genetic variation of pathway precursors for future MAS experiments.

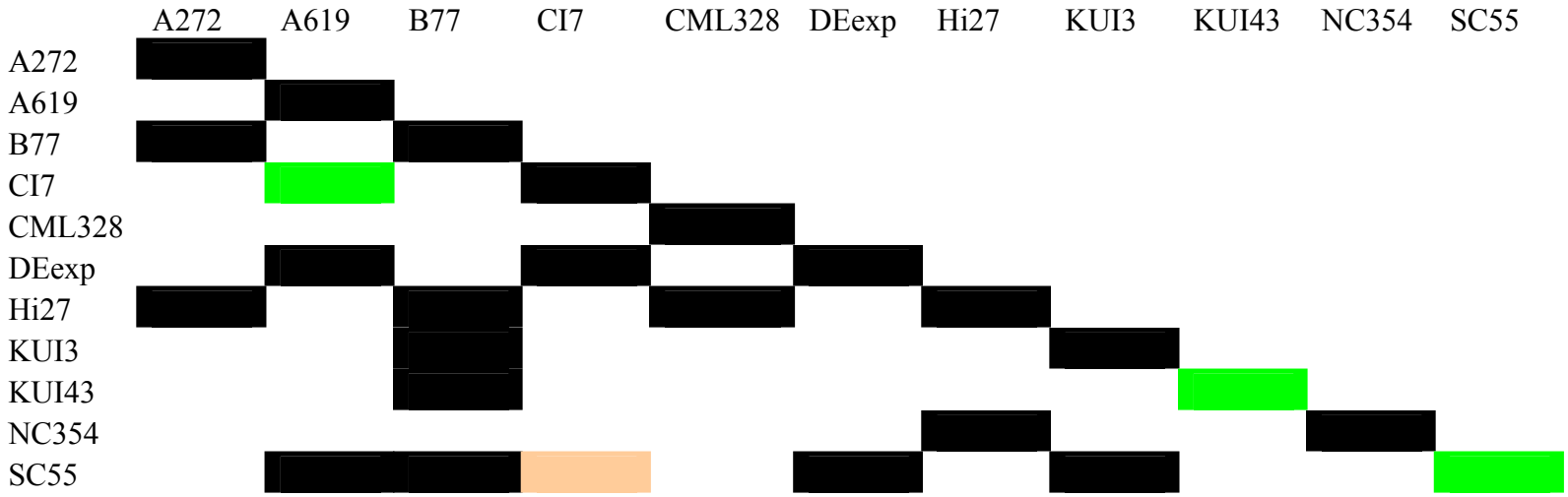
3.6 Tables

Table 3.1: Inbreds from Goodman-Buckler Diversity Panel used in Chapter 3 studies

Inbred	Pedigree	Maturity[a]	Pop Structure[b]		Notes on Carotenoid Traits
			Group[c]	SubGroup	
A272	KwaZulu-Natal University, South Africa, Yellow Boesman Variety	Early	Mixed		Highest for all measured carotenoids in association panel
A619	[(A171 x Oh43) Oh43]	Early	NSS	M14:Oh43	High lutein
A632	[(Mt42 X B14) B14(3)]	Early	SS	B14A	High zeaxanthin
B77	BS11(FR)CO-Q51-3-2-1-2-1 (Pioneer Two-Ear Composite)	Early	NSS	NSS-mixed	High β-carotene
CI7	(L317 x 33-16) 33-16[2]	Mid	NSS	K64W	Highest β-carotene
CML328	Recycled yellow lines.: 89[HTSG29-1/TEYFDMR]#-96-1-2-B*3	Late	Mixed		High β-cryptoxanthin, high zeaxanthin
DEexp	Delaware Experimental Inbred, self from Pioneer Hi Bred F1	Early	Mixed		Highest β-carotene
Hi27	(CM104(India) x MV source)BC6	Late	Mixed		High β-cryptoxanthin
KUI3	Suwan 1(S)C4-S8-5-3 (2007)	Late	TS	Suwan	High zeaxanthin, orange germplasm
KUI43	Suwan 3(S)C3-S7-138, (Kei 9101)	Mid	Mixed		High zeaxanthin, orange germplasm
NC354	Of NC296 and Pioneer X304A x H101 origin	Late	TS	NC	High β-cryptoxanthin
SC55	(L501 x L503) X (L548 x L569)	Late	Mixed		Highest β-carotene: total carotenoid ratio
W64a	WF9 x C.I.187-2	Early	NSS	Hy:T8: Wf9	High β-carotene, β-cryptoxanthin, lutein

a: Maturity relative to flowering dates of all lines in association panel [43]
b: Population structure as delimited in Liu [43]
c: Groups are given as NSS: Non-Stiff Stalk; SS: Stiff Stalk; TS: Tropical or Semitropical; Mixed: < 80% identity with major groups

Table 3.2: Crosses generated for hybrid seed comparison grown over two environments in 2008



Shown are self pollinations (diagonal) and F2 seed derived from hybrid plants (off diagonal)
Box codes indicate samples from Urbana, Illinois, 2008 only (green), Puerto Vallarta, Mexico, winter 2007-2008 only (orange) both locations (black)

Table 3.3: Genotypes for *crtRB1*, *lcyε* and *ccd1* carotenoid QTL in selected inbreds used in Chapter 3 studies

QTL	*crtRB1*			*lcyε*			*ccd1*
Polymorphism	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	5p
Allelic series[a]	1/**2**/3	**1**/2	**1**/2/3	**1**/2/3/**4**	**1**/2	1/**2**	WC/**b73**
A272	1	2	3	.	.	.	
A619	1	2	**1**	2	2	**2**	**b73**
A632	**2**	2	3	2	2	1	**b73**
B77	**2**	**1**	**1**	2	2	**2**	**b73**
CI7	**2**	**1**	**1**	2	2	1	**b73**
CML328	.	2	3	**1**	2	1	**b73**
DEexp	**2**	2	**1**	3	**1**	1	**b73**
HI27	1	2	3	2	2	1	**b73**
KI3	1	2	3	**4**	2	**2**	**b73**
KI43	1	2	3	2	2	**2**	**b73**
NC354	1	2	3	**4**	2	1	**b73**
SC55	**2**	**1**	**1**	2	2	1	WC
W64A	**2**	2	3	3	**1**	1	**b73**

a: Allele associated with higher β-carotene is considered favorable and is denoted in bold

Allele codes	1=397 bp ins 2=200 bp ins 3=0 bp ins	1=12 bp ins 2=0 bp ins	1=0 bp ins 2=325 bp ins 3=1250 bp ins	1=150+280 bp 2=250 bp 3=250+380 bp 4=993 bp	1=G 2=T	1=0 bp del 2=8 bp del	b73=small ins WC=large ins, multicopy

Table 3.4: $F_{2:3}$ populations used for segregation analysis of *crtRB1* and *lcyε* carotenoid QTL effects

			Allele-specific marker genotypes[a]					
Population	n	Genotype Information	crtRB1 5'TE	crtRB1 InDel4	crtRB1 3'TE	lcyE 5'TE	lcyE SNP216	lcyE 3'TE
A619xSC55	227	Linkage map[b]	**1,2**	**2,1**	1,1	2,2	2,2	**2,1**
DE3xCI7	103	Linkage map[b]	2,2	**2,1**	1,1	**3,2**	**1,2**	1,1
KI3xB77	110	Allele-specific markers[c]	**1,2**	**2,1**	**3,1**	**4,2**	2,2	**2,2**
KI3xSC55	66	Allele-specific markers[c]	**1,2**	**2,1**	**3,1**	**4,2**	2,2	**2,1**
W64axA632	200	Linkage map[d]	2,2	2,2	3,3	**3,2**	**1,2**	1,1
		Allele codes	1/**2**/3	**1**/2	**1**/2/3	**1**/2/3/**4**	**1**/2	1/**2**

a: Allele-specific marker genotypes are listed by population and polymorphism in terms of "parent 1 allele, parent 2 allele"; favorable alleles are in bold.
b: Linkage map constructed by Stevens [44] and further modified by CBK
c: Allele-specific markers for *lcyε* [19] and *crtRB1* [29]
d: Linkage map constructed by Wong [26] and further modified by CBK

Table 3.5: Germplasm and segregating *crtRB1* and *lcyε* carotenoid QTL genotypes in marker assisted selection study

Pop.	**Source**	**Pedigree - S_0 generation**
A/B	UI_PV08-860/923	[(CML297-BxCML324-B) x (DEexpxCI7)]
C/D	UI_PV08-868/900	[(Carotenoid Syn3-FS8-4-6-BxCML324-B) x (B77xHi27)]
E/F	UI_PV08-868/907-1	[(Carotenoid Syn3-FS8-4-6-BxCML324-B) x (KUI43xB77)]
G/H[a]	UI_PV08-849/898-1	[(Carotenoid Syn3-FS8-4-6-BxCML324-B) x (KUI3 xB77)]
I/J	UI_PV08-852/884	[(Carotenoid Syn3-FS8-4-2-BxFlorida A + Syn-FS6-3-1-B) x (DEexpxCI7)]
K/L	UI_PV08-852/887	[(Carotenoid Syn3-FS8-4-2-BxFlorida A + Syn-FS6-3-1-B) x (A619xSC55)]
M/N	UI_PV08-881/896	[(CML297-BxCML324-B) x (KUI3xSC55)]
O/P	UI_PV08-872/898	[(Carotenoid Syn3-FS8-4-6-BxKUI carotenoid syn-FS17-3-2-B) x (KUI3 xB77)]
Q/R	UI_PV08-872/923	[(Carotenoid Syn3-FS8-4-6-BxKUI carotenoid syn-FS17-3-2-B) x (KUI3 xB77)]
S/T	UI_PV08-872/924-1	[(Carotenoid Syn3-FS8-4-6-BxKUI carotenoid syn-FS17-3-2-B) x (A619xSC55)]
A-GH	UI_PV08-853/910	[(Carotenoid Syn3-FS8-4-6-BxFlorida A + Syn-FS6-3-1-B) x (DEexp/CI7)xB77)]
B-GH[a]	UI_PV08-849/898-1	[(Carotenoid Syn3-FS8-4-6-BxCML324-B) x (KUI3 xB77)]

Table 3.5:, continued

	female hybrid genotypes[b]						**male hybrid genotypes**[b]					
Pop.	***lcye* 5'TE**[c]	***lcye* SNP216**	***lcye* 3'TE**	***crtRB1* 5'TE**	***crtRB1* InDel4**[c]	***crtRB1* 3'TE**	***lcye* 5'TE**[c]	***lcye* SNP216**	***lcye* 3'TE**	***crtRB1* 5'TE**	***crtRB1* InDel4**[c]	***crtRB1* 3'TE**
A/B	**4**/2	**1**	na	1	2	3	3/2	**1**/2	**1/1**	**2/2**	2/**1**	**1/1**
C/D	**4**/2	**1**	na	1	2	3	2/2	2/2	2/**1**	**2**/1	**1**/2	**1**/3
E/F	**4**/2	**1**	na	1	2	3	2/2	2/2	2/2	1/**2**	2/**1**	3/**1**
G/H[a]	**4**/2	**1**	na	1	2	3	**4**/2	2/2	2/2	1/**2**	2/**1**	3/**1**
I/J	2/**1**	**1**	na	1	2	3	3/2	**1**/2	**1/1**	**2/2**	2/**1**	**1/1**
K/L	2/**1**	**1**	na	1	2	3	2/2	2/2	**1/1**	1/**2**	2/**1**	**1/1**
M/N	**4**/2	**1**	na	1	2	3	**4**/2	2/2	2/2	1/**2**	2/**1**	3/**1**
O/P	**4**/2	**1**	na	1	2	3	**4**/2	2/2	2/2	1/**2**	2/**1**	3/**1**
Q/R	**4**/2	**1**	na	1	2	3	3/2	2/2	2/2	1/**2**	2/**1**	**1/1**
S/T	**4**/2	**1**	na	1	2	3	2/2	2/2	**1/1**	1/**2**	2/**1**	**1/1**
A-GH	**4**/1	**1**	na	1	2	3	3/2-2/2	**1**/2-2/2	**1/1**-2/2	**2/2-2/2**	**1**/2-**1/1**	**1/1-1/1**
B-GH[a]	**4**/2	**1**	na	1	2	3	**4**/2	2/2	2/2	1/**2**	2/**1**	3/**1**

a: Same population sampled twice (in both UIUC field and Greenhouse)
b: Genotypes are assigned according to screening or inferred based on Table 3.3. Allele codes are listed in Table 3.3, and favorable alleles are shown here in bold.
c: Genotype selection of progeny based on *lcye*-5'TE and *crtRB1*-InDel4 marker data

Table 3.6: Comparison of environment effects on carotenoid composition for hybrids and inbreds grown in Puerto Vallarta, Mexico in winter 2008-2009 and Urbana, IL in 2008

	Means (stdev)		Trait Correlation between Environments[a]	
Traits	Mexico n=100	Illinois n=82	Pearson	Spearman
Lutein	10.54 (5.52)	5.34 (3.90)	0.72**	0.60**
Zeaxanthin	8.14 (5.86)	9.86 (8.97)	0.96**	0.96**
βcryptoxanthin	3.81 (3.22)	2.52 (2.66)	0.96**	0.90**
βcarotene	7.03 (4.35)	2.55 (1.52)	0.86**	0.77**

a. Trait correlations between environments are indicated by product-moment correlations (Pearson) and rank correlations (Spearman). Statistical significance is given by ** (α=0.01), * (α=0.05)

Table 3.7: Lines with highest carotenoid concentrations across Urbana, Illinois in 2008 and Puerto Vallarta, Mexico 2008-09

Set	Total tested	Lut	Zea	βcry	βcar	Total
Hybrids	18	A619xDEexp	B77xA272	B77xA272	CI7xDEexp	B77xA272
		CI7xDEexp	HI27xA272	HI27xA272	DEexpxSC55	DEexpxCI7
		DEexpxA619	HI27xCML328	HI27xCML328	DEexpxCI7	HI27xA272
		KUI43xB77	NC354xHI27	NC354xHI27	SC55xDEexp	HI27xCML328
Inbreds	11	A619	A272	A272	A272	A272
		DEexp	HI27	CML328	CI7	A619
		KUI43	KUI3	HI27	DEexp	KUI43

Table 3.8: Significance of two-gene model (*lcyε* and *crtRB1*) for carotenoid composition traits among four segregating populations

$F_{2:3}$ Population	Trait	Transform[a]	Model[b]	R^2	Main Effects[b]		Two-way Interaction[b]
					crtRB1	*lcye*	
A619 x SC55	βcar	yes	**	10.0%	**	ns	ns
	βcar/ALL	no	**	21.1%	**	ns	*
	βcar/(βcry+zea)	yes	**	25.9%	**	ns	ns
	α/β branch	yes	**	16.0%	ns	**	*
	Total	yes	**	13.6%	ns	ns	**
DEexp x CI7	βcar	no	**	41.6%	**	**	ns
	βcar/ALL	no	**	41.2%	ns	**	ns
	βcar/(βcry+zea)	no	*	27.2%	**	ns	ns
	α/β branch	no	**	39.6%	ns	**	ns
	Total	no	ns	-	ns	ns	ns
KUI3 x B77	βcar	yes	**	51.1%	**	*	ns
	βcar/ALL	yes	**	62.2%	**	ns	ns
	βcar/(βcry+zea)	yes	**	66.0%	**	*	ns
	α/β branch	no	**	31.1%	**	**	ns
	Total	no	**	50.7%	**	ns	ns
KUI3 x SC55	βcar	no	**	33.6%	**	ns	ns
	βcar/ALL	yes	**	53.4%	**	ns	ns
	βcar/(βcry+zea)	yes	**	64.9%	**	ns	ns
	α/β branch	yes	**	42.9%	ns	**	ns
	Total	yes	ns	18.6%	*	ns	ns
Single gene model							
W64a x A632[c]	βcar	yes	**	10.7%		**	
	βcar/ALL	yes	**	34.7%		**	
	βcar/(βcry+zea)	no	ns	-		ns	
	α/β branch	yes	**	58.7%		**	
	Total	no	**	7.9%		**	

a: Requirement of power transformation to satisfy normality assumptions for analysis of variance.
b: Model significance is indicated α=0.01 (**), α=0.05 (*), or non-significant (ns)
c: W64a x A632 population only segregates at *lcyε*, therefore a one gene model was used

Table 3.9: Effects of *crtRB1* on β–carotene related traits across genetic backgrounds

Copies of favorable haplotype	N				Traits															
					βCar				βCar/ALL				βCar/Zea				βCar/βCry			
	A	B	C	D	A	B	C	D*	A	B	C	D*	A	B	C	D*	A	B	C	D*
2	32	33	12	14	5.15[a]	2.37	5.57	6.27	0.47	0.17	0.46[a]	0.31[a]	8.98	0.33	9.65	2.5	12.33	2.25	5.8	19.49
1	114	42	26	33	5.02[a]	1.49	3.39[a]	5.79[a]	0.41	0.07	0.28[a]	0.3[a]	4.83	0.1	1.50	2.16[a]	7.28	0.63	1.63	18[a]
0	50	35	28	13	4.62	0.76	2.79[a]	5.67[a]	0.38	0.03	0.17	0.29[a]	3.6	0.04	0.62	2.06[a]	5.4	0.26	0.77	17.31[a]
R^2 (%)					7.8	43.5	25.4	31.8	17	59.6	49	35.2	15.0	63.1	46.8	21	36.8	58.1	69.3	21
Actual difference between homozygotes					0.53	1.61	2.78	0.60	0.09	0.14	0.29	0.02	5.38	0.29	9.03	0.44	6.93	1.99	5.03	0.44
Fold change between homozygotes					1.11	3.11	2	1.11	1.23	5.67	2.70	1.08	2.49	8.25	15.56	1.22	2.28	8.65	7.53	1.22

crtRB1 genotypic class means for βCar (µg/g), βCar/ALL, βC/Zea and βC/βCry traits. Effects are adjusted by *lcyε* covariates, all the adjusted means are significantly different at α=0.05, except where noted[a].

Populations are coded as: A: A619 x SC55; B: KUI3 x B77; C: KUI3 x SC55; D: DEexp x CI7

Origin of favorable haplotype for *crtRB1* is listed in Table 3.11

Table 3.10: Effect of *lcyε* on α/β branch ratio across five segregating populations

Copies of favorable haplotype	n					Trait: α/β branch ratio				
	A	B	C	D	E	A	B	C	D	E
2	32	33	12	14	25	0.65	0.09	0.31	1.19	0.21
1	114	42	26	33	81	0.78[a]	0.18[a]	0.56[a]	1.62	0.33
0	50	35	28	13	38	0.84[a]	0.21[a]	0.59[a]	2.02	0.57
R^2 (%)						10	16.8	24.1	31.8	56.8
Actual difference between homozygotes						0.19	0.12	0.28	0.83	0.36
Fold change between homozygotes						1.29	2.33	1.90	1.70	2.71

lcyε genotypic class means for α/β branch ratio; all least square means are significantly different at α=0.05, except where noted[a].

Populations are coded as: A: A619 x SC55; B: KUI3 x B77; C: KUI3 x SC55; D: DEexp x CI7; E: W64a x A632

Origin of favorable haplotype for *lcyε* is listed in Table 3.14

Table 3.11: Estimated effects of crtRB1 haplotype substitutions on β-carotene in five segregating populations

		Donor of Favorable Haplotype[a]		Average Effect of Haplotype Substitution[b]			
Population	n	*crtRB1*	*lcyε*	Intercept	$\beta_{crtR\text{-}B1}$	β_{lcye}	R^2_{adj} (P)[c]
A619 x SC55	193	SC55	A619	4.68 (0.09)	0.30 (0.08)	-	6% **
DEexp x CI7	60	DEexp	CI7	6.18 (0.17)	0.30 (0.10)	0.44 (0.11)	27% **
KUI3 x B77	110	B77	KUI3	0.8 (0.12)	0.82 (0.10)	-	39% **
KUI3 x SC55	66	SC55	KUI3	2.76 (0.27)	1.25 (0.25)	-	26% **
W64a x A632	154	-	W64a	0.62 (0.07)	-	0.07 (0.02)	10% **

a: Based on predicted haplotype performance in association panel
b. Phenotypic gain (μg g^{-1}) as a copy of the favorable haplotype is added; assumption of no intragenic recombination; β indicates regression coefficient
c: R^2 adjusted for number of terms in model; p-value for model, where ** indicates significant at α=0.01

Table 3.12: Comparison of *lcyε* polymorphism effects on carotenoid ratios and composition across five segregating populations

Population	Segregating polymorphisms		*lcye*	Significance of *lcye* for trait[b]					Estimated Effect for Total
				α/β branch	α branch	β branch	βcar	Total	
A619 x SC55			3'TE[a]	**	**	ns	ns	†	least total with 2 favorable copies
DEexp x CI7	5'TE	SNP216[a]		**	**	**	**	**	least total with 2 favorable copies
KUI3 x B77	5'TE[a]			**	**	ns	*	ns	
KUI3 x SC55	5'TE[a]			**	**	ns	ns	ns	
W64a x A632	5'TE	SNP216[a]		**	**	**	**	**	least total with 2 favorable copies

a: Segregating at extreme allelic contrasts
b: Significance is indicated at α=0.01 (**), α=0.10 (†), or non-significant (ns)

Table 3.13: Estimated effects of *crtRB1* haplotype substitutions on total carotenoids in four segregating populations

		Average Effect of Haplotype Substitution[a]			
Population	N	Intercept	$\beta_{crtR\text{-}B1}$	R^2_{adj} (P)[b]	% decrease per copy
A619 x SC55	193	13.02 (0.43)	-0.80 (0.39)	2% *	6%
DEexp x CI7	60	21.55 (0.52)	0.03 (0.43)	-	-
KUI3 x B77	110	26.74 (0.85)	-6.24 (0.67)	43.9% **	23%
KUI3 x SC55	66	16.71 (1.13)	-2.90 (1.07)	8.9% **	17%

a. Phenotypic gain ($\mu g\ g^{-1}$) as a copy of the favorable haplotype is added; assumption of no intragenic recombination
b: R^2 adjusted for number of terms in model; p-value for model, where ** indicates significant at α=0.01, * significant at α=0.05, - is non-significant

Table 3.14: Genotype selection probabilities in the S_0 generation

Population	Probability of encountering favorable alleles in S_0[a]		Exp. % selected genotypes	No. Genotyped S_0 plants[c]	Obs % selected genotypes[d]
	lcye-5'TE	*crtRB1*-InDel4			
A/B	50% "4" het	50% het	25.00%	26	0.00%
C/D	50% "4" het	50% het	25.00%	37	24.30%
E/F	50% "4" het	50% het	25.00%	35	0.00%
G/H[a]	25% "4" hom; 50% "4" het	50% het	37.50%	37	0.00%
I/J	50% "1" het	50% het	25.00%	35	31.40%
K/L	50% "1" het	50% het	25.00%	35	0.00%
M/N	25% "4" hom; 50% "4" het	50% het	37.50%	35	34.30%
O/P	25% "4" hom; 50% "4" het	50% het	37.50%	37	0.00%
Q/R	50% "4" het	50% het	25.00%	33	12.10%
S/T	50% "4" het	50% het	25.00%	32	12.50%
A-GH	50% "4" het; 50% "1" het	37.5% het	37.50%	54	40.00%
B-GH[b]	25% "4" hom; 50% "4" het	50% het	25.00%	40	20.00%
			29.17%		**14.55%**

Probability of encountering favorable alleles in S0 generation depends on parental allelic composition (listed in Table 3.5 for female and male hybrid genotypes). Desirable alleles are dependent upon parents used for population development; generally, desired *lcyε*-5'TE allele is "4" or "1", and favorable *crtRB1* allele is "ins"

a: Genotypes containing the favorable alleles are heterozygous (het) or homozygous (hom)

b: Probability is the outcome of progeny containing at least one copy of the favorable allele at both polymorphisms

c: Using *lcyε*-5'TE and *crtRB1*-InDel4 codominant markers

d: Selection of any plant with at least one favorable allele at both loci

Table 3.15: Selection percentage and survival probabilities in S_1 generation

	Genotyped S_1 seeds		Selected S_1 seeds[a]		Advanced S_1 seeds[b]		
Population	S_1 Kernels Sampled (#)	S_0 Ears Tested	S_1 Kernels Selected (%)	Derived from S_0 Ears	Total Kernels	Germinated S_1 seeds (%)	Mature $S_{1:2}$ ears[c] (%)
A/B	-	-	-	-	-	-	-
C/D	88	6	5.7%	2	13	61.5%	75.0%
E/F	-	-	-	-	-	-	-
G/H[a]	-	-	-	-	-	-	-
I/J	88	6	4.5%	4	13	76.9%	70.0%
K/L	-	-	-	-	-	-	-
M/N	116	8	6.0%	7	30	50.0%	40.0%
O/P	-	-	-	-	-	-	-
Q/R	28	2	28.6%	1	10	0.0%	-
S/T	32	2	3.1%	1	5	100.0%	80.0%
A-GH	176	12	8.5%	8	23	60.9%	57.1%
B-GH[d]	not tested	not tested	-	-	22	100.0%[d]	63.6%

a: Seeds selected if favorable alleles were homozygous at both *lcyε*-5'TE and *crtRB1*-InDel4
b: Reflects the number of S_1 planted in greenhouse from selected set in addition to unfavorable contrasts for further genetic study.
c: Percentage of surviving S_1 plants reaching reproductive stage
d: S_1 seeds from this population were planted without screening kernel genotypes

Table 3.16: Effects of *crtRB1* genotype, *lcyε* genotype and generation on selected traits in MAS study

			$S_{1:2}$ Genotype Classes[a]		*crtRB1* effect in $S_{1:2}$[b]		*lcye* effect in $S_{1:2}$[b]		Contrast: S_0 vs $S_{1:2}$ *crtRB1* het[b]	
Pop.	No. S_0	No. S_2	*crtRB1*	*lcye*	βcar	Total	βcar	α/β branch	βcar	Total
C/D	2	6	2	2	ns	**	ns	**	ns	*
I/J	2	7	2	3	*	**	**	**	ns	**
M/N	2	6	3	2	**	*	ns	**	ns	**
A-GH	2	8	3	3	**	**	ns	**	*	**
B-GH	2	14	3	1	**	*	no seg[c]	no seg[c]	**	ns

a: Genotype classes include homozygous favorable, heterozygous, homozygous unfavorable
b: Significance is indicated at α=0.01 (**), α=0.05 (*), or non-significant (ns)
c: Only one *lcyε*-5'TE genotype class was observed in this population

Table 3.17: Response to *crtRB1* selection in MAS populations, $S_{1:2}$ generation

Pop.	MAS *crtRB1* Class	N	Average concentrations μg g^{-1} (Std Dev) βcar	βcry	Total	proVA	Selection Differential[a] βcar	proVA	Fold Improvement[b] βcar	proVA
C/D	homozygous favorable	2	6.93 (1.00)	0.64 (0.10)	11.09 (1.24)	7.25 (1.09)	2.23	-1.91*	-	-
	heterozygous	4	4.70 (2.73)	8.92 (3.90)	31.80 (8.52)	9.16 (1.82)				
I/J	homozygous favorable	4	9.83 (4.11)	1.82 (1.27)	26.90 (8.38)	10.94 (4.21)	5.31*	1.44	-	-
	heterozygous	2	4.52 (1.75)	9.97 (4.63)	47.27 (13.07)	9.50 (3.75)				
M/N	homozygous favorable	3	7.82 (3.01)	2.47 (1.62)	18.53 (4.42)	9.21 (2.72)	4.09**	2.54*	11.33	3.22
	heterozygous	2	3.73 (1.54)	5.88 (1.07)	23.09 (4.17)	6.67 (1.73)				
	homozygous unfavorable	1	0.69 (0.11)	4.35 (0.20)	16.07 (0.74)	2.86 (0.15)	-3.04	-3.81*		
A-GH	homozygous favorable	3	17.18 (4.11)	1.19 (0.38)	27.15 (5.17)	17.88 (4.31)[c]	11.48**	9.58**	4.75	2.02
	heterozygous	4	5.70 (1.97)	5.15 (0.51)	26.09 (4.92)	8.30 (1.70)				
	homozygous unfavorable	1	3.62 (0.27)	10.41 (0.15)	43.02 (3.31)	8.83 (0.65)	-2.08	-0.53		
B-GH	homozygous favorable	2	11.92 (2.78)	3.33 (0.95)	25.15 (2.24)	13.61 (2.34)[c]	5.83**	3.04**	3.45	1.45
	heterozygous	9	6.09 (2.60)	8.96 (3.62)	38.04 (12.73)	10.57 (3.12)				
	homozygous unfavorable	3	3.46 (1.26)	11.84 (4.06)	40.51 (7.11)	9.39 (3.00)	-2.63**	-1.18		

a: Difference from heterozygote (μg g^{-1}), significance denoted with a t-test at α=0.01 (**), or α=0.05 (*); S_0 genotype (heterozygote) used as reference
b: Improvement in trait (fold change) between homozygous classes
c: Indicates selection cohort that has reached target proVA levels of 15 μg g^{-1}

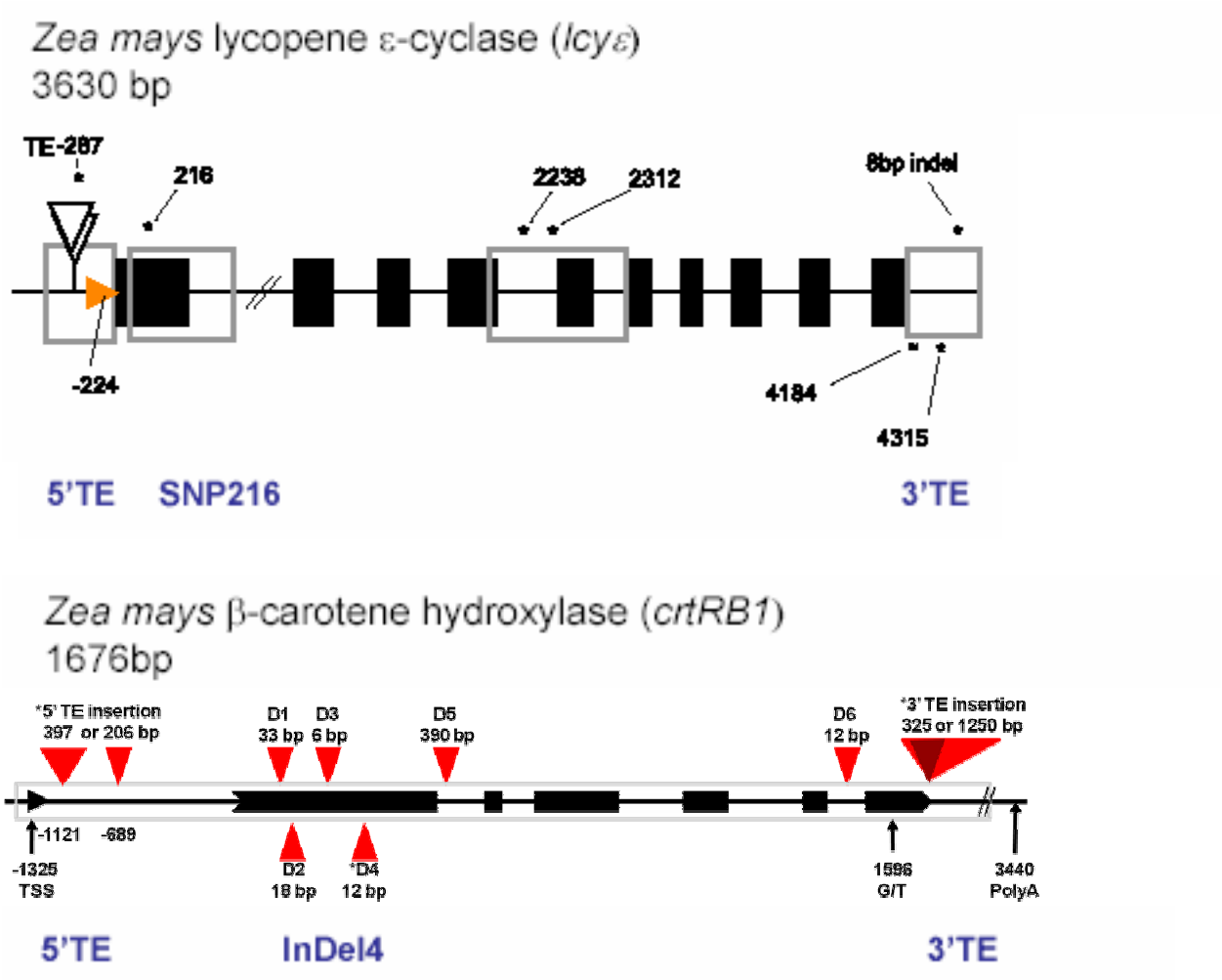
Table 3.18: Estimated effects of carotenoid QTL haplotype substitutions on total carotenoids in MAS populations

		Average Effect of Haplotype Substitution[a]			
Population	N	Intercept	$\beta_{crtR\text{-}B1}$	R^2_{adj} (P)[b]	% decrease per copy
A-GH	24	35.31 (2.82)	-5.36 (1.99)	21.3% **	15%
B-GH	42	43.24 (3.17)	-7.01 (2.88)	10.7% *	16%
MN	18	19.55 (2.31)	-0.07 (1.51)	-	-

a. Phenotypic gain ($\mu g\ g^{-1}$) as a copy of the favorable haplotype is added; assumption of no intragenic recombination
b: R^2 adjusted for number of terms in model; p-value for model, where ** indicates significant at α=0.01, * significant at α=0.05, - is non-significant

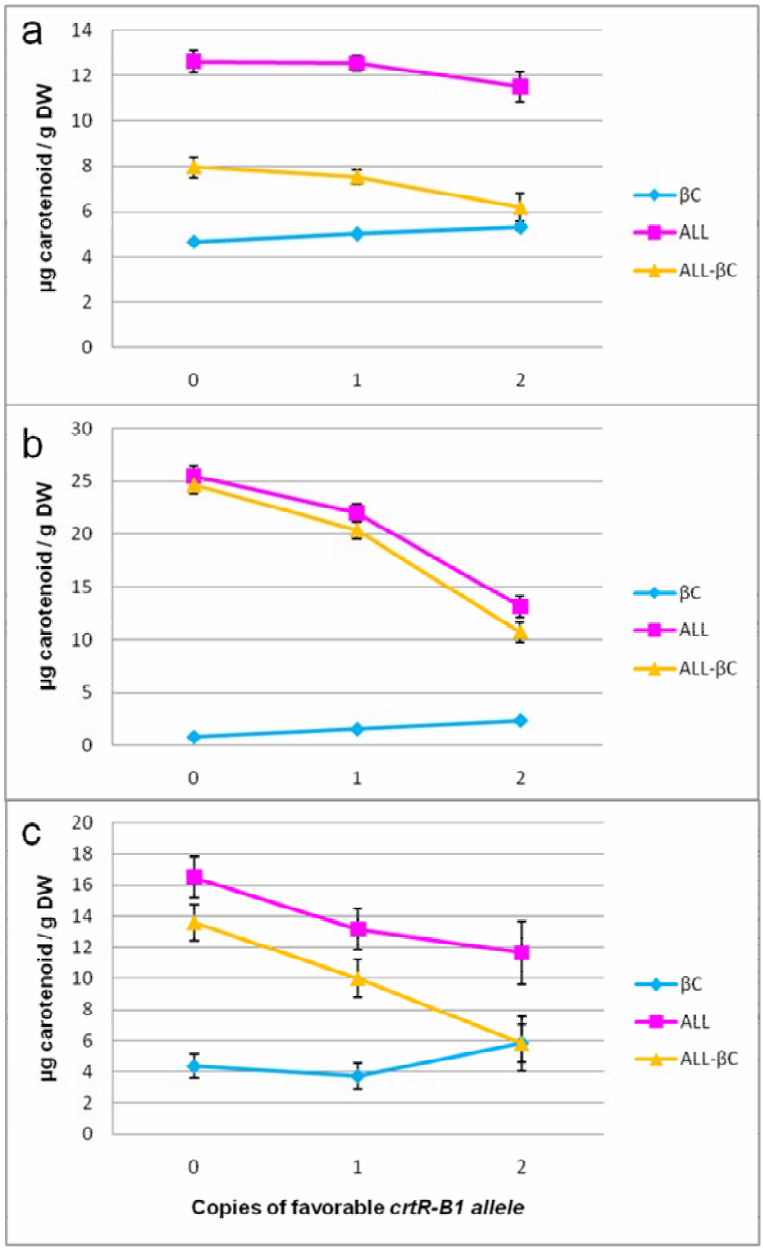
3.7 Figures

Figure 3.1: Schematic of polymorphisms within *Zea mays lycopene ε-cyclase* and *β-carotene hydroxylase*



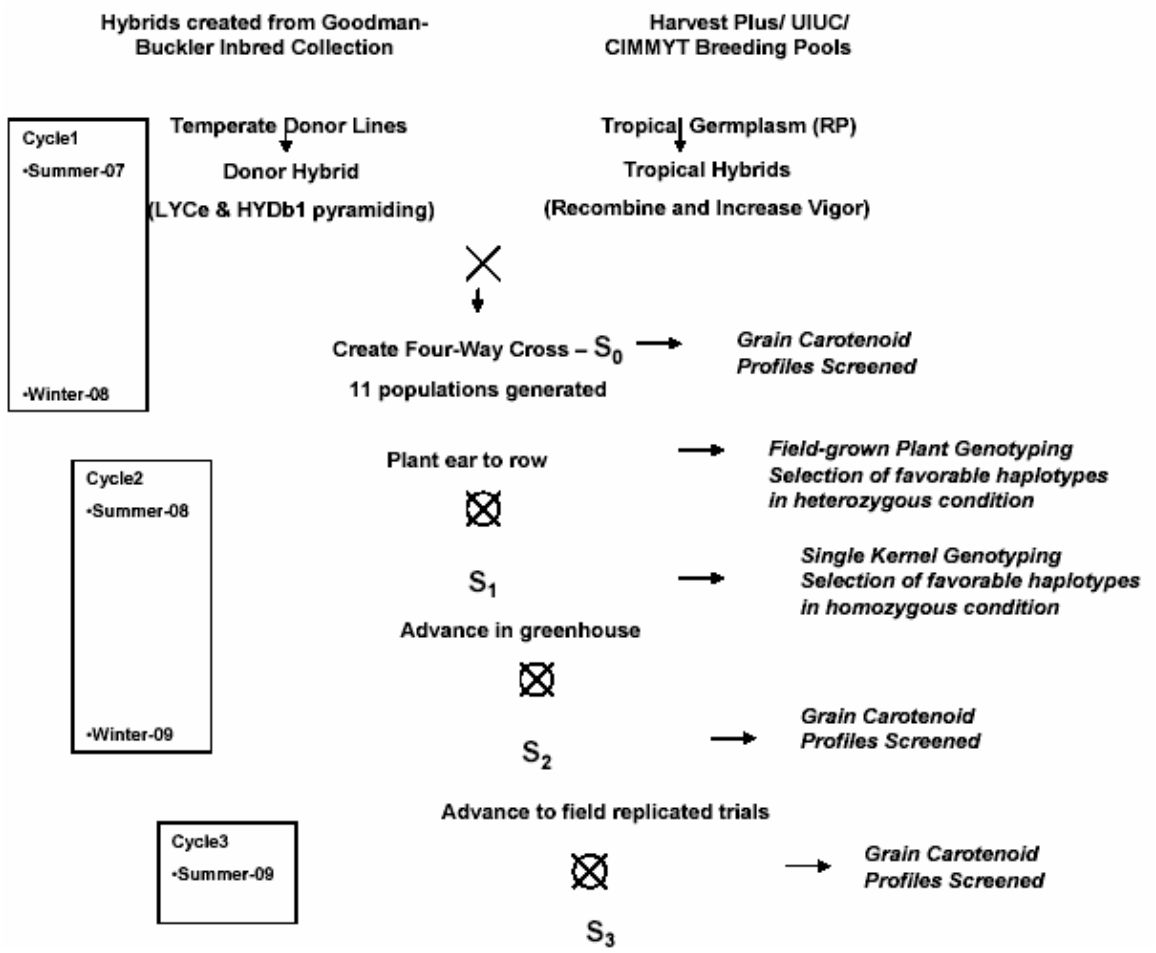
Modified from Harjes [19] and Yan and Kandianis [29]

Figure 3.2: Mean β-carotene, total and total minus β-carotene concentrations for *crtR-B1* allele classes across three $F_{2:3}$ populations



Means from the following populations are shown: **a**: A619xSC55, **b**: KUI3xB77, **c**: KUI3xSC55. Error bars are standard error. Trait legend indicates β-carotene (βC), total carotenoid (ALL) and total minus β-carotene (ALL-βC).

Figure 3.3: Marker assisted selection strategy for increased β-carotene concentration over three cycles



Note: Grain profile screens conducted in S_0, S_2 and S_3 generations were not used for selection, only to gauge response to recombination and advance.

Figure 3.4: Relationship between β-carotene and total carotenoids among selected $S_{1:2}$ individuals in five MAS experiments

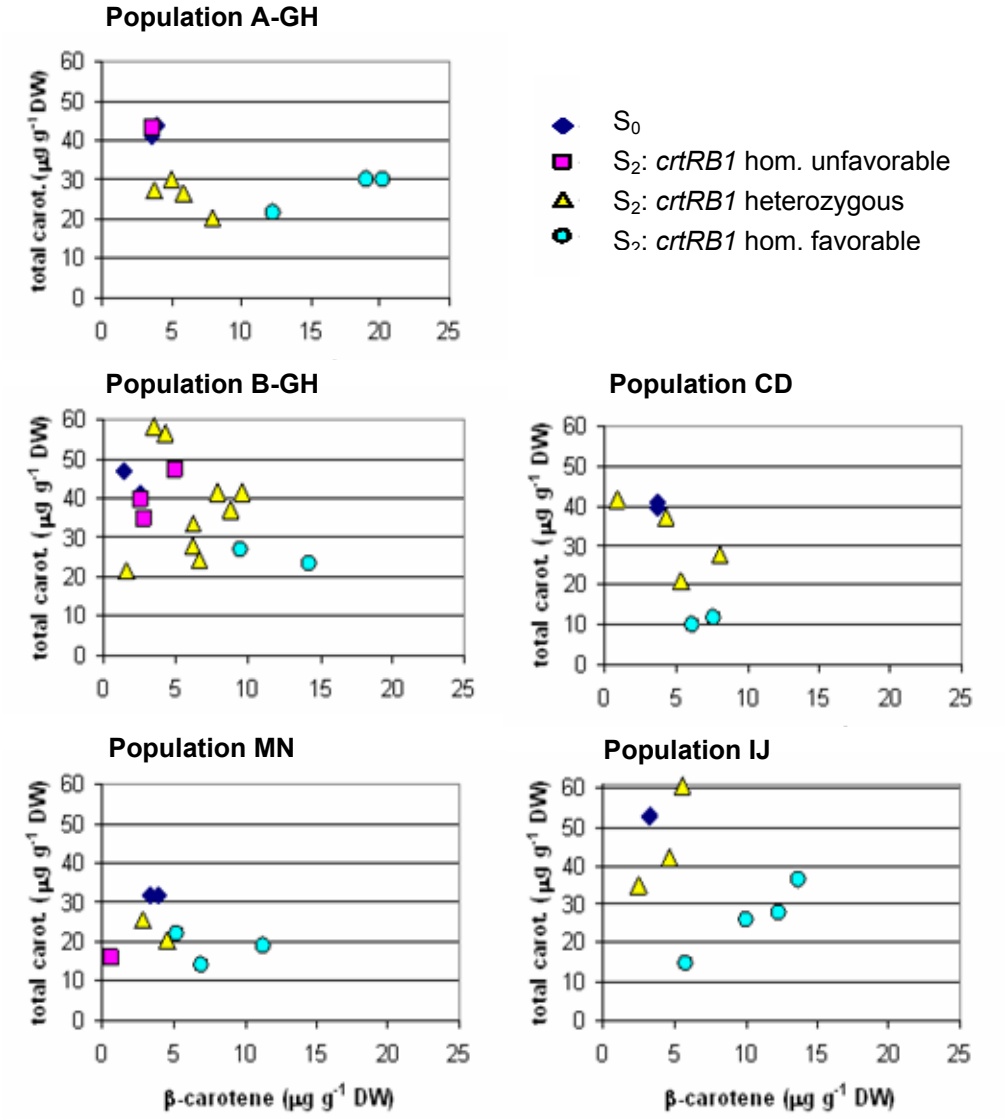
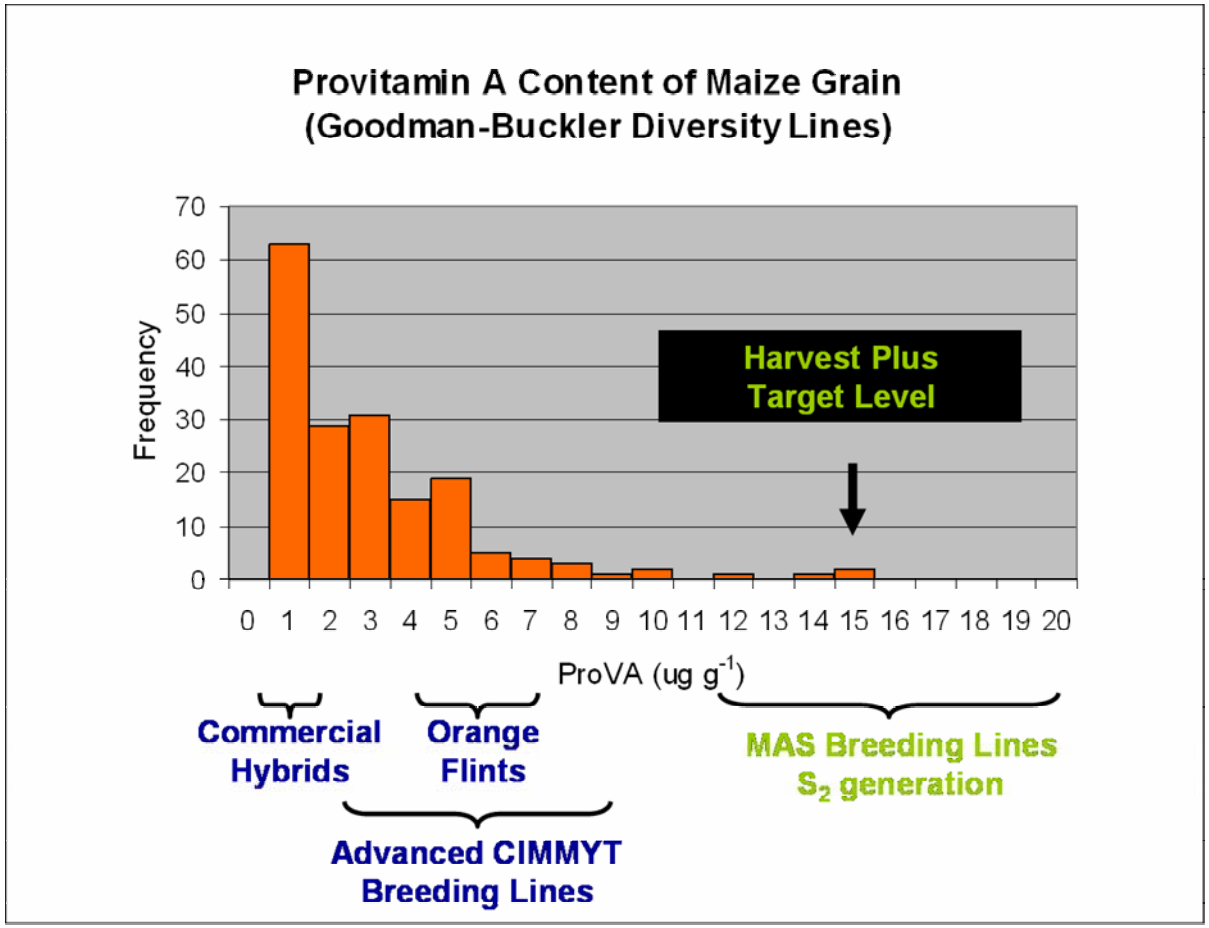


Figure 3.5: ProVitamin A levels of maize varieties used in breeding program



Marker Assisted Selection (MAS) breeding lines used in this evaluation were greenhouse grown

EPILOGUE

Through the projects detailed here, we evaluated the hypothesis of metabolic QTL underlying carotenoid traits with linkage mapping studies, and directly tested this hypothesis through three genes involved in carotenogenesis by association mapping of *ccd1*, and further characterization of *lcyε* and *crtRB1* in a variety of genetic backgrounds. Furthermore, we leveraged haplotype information from basic genetic studies to develop improved carotenoid and proVA profiles in genetic populations. From these investigations, it is evident that additive effects are most important in contributing to progeny performance, that dominance exists within several allelic series, and that epistasis, though rare, can occur between pathway genes. Dominance was more often observed with alleles that comparatively yield reduced β-carotene and proVA, indicating that desired alleles should be bred to homozygosity, and will likely need to be introduced to both sides of the pedigree if hybrids are produced.

Improvement of carotenoid profiles through selection of specific QTL should consider the manipulation of genetic factors controlling synthesis, conversion and degradation processes as all three influence the final harvest phenotype. Results from QTL analyses of two populations with strikingly different carotenoid profiles jointly conclude that increases to any specific carotenoid pool can be facilitated through the selection of favorable genetic variation within the MEP and isoprenoid pathways which provide substrate precursor to the carotenoid pathway. Not much is known about the allelic variation at these loci; therefore, characterization of loci such as DXR and GGPPS with association mapping could help to identify the next genetic targets to be evaluated in marker assisted selection projects.

Consideration of degradation effects was found to be equally important in generating lines with favorable carotenoid profiles. Degradation processes associated with the 5p promoter polymorphism at *ccd1* do not appear to affect β-carotene concentrations. However, lines which have a large proportion of their proVA concentration attributed to β–cryptoxanthin can still be negatively affected by this allelic variation. Selection of allelic variation at *ccd1* will be most useful for MAS in breeding total carotenoid traits, since increased degradation through the Wc allele can be maintained in yellow lines. Additionally, a QTL potentially encoding ZEP was found to reduce total carotenoids through a reduction in zeaxanthin. Variation at this locus could potentially be used to keep total carotenoids high in the endosperm. However, it is critical to

note that ZEP is necessary for abscisic acid biosynthesis to occur in the embryo, and modification carotenoid profiles through this locus may lead to physiological complexities.

The effect of degradation during synthesis and conversion is particularly interesting as selection for lines with potential for enhanced accumulation could be confounded by degradation activity at CCD1. The initial rate of carotenoid accumulation in a representative Wc line appears to be similar to that of representative b73 lines until late in kernel development, confirming that the effect of *ccd1* can cause lines with highly favorable biosynthesis genetics to be discarded upon selection. This also suggests that the biological role of *ccd1* in seed endosperm may be tied to developmental processes that specifically occur at that time, such as the termination of starch fill, reduction of water content or plateau in seed volume.

Modification of the carotenoid profile through intra-pathway variation at both *crtRB1* and *lcyε* can be largely effective in increasing β-carotene concentration. To successfully use this genetic variation, however, an understanding of the effect of each polymorphism on the entire carotenoid trait profile is required. Results from chapter 3 note that unanticipated trait effects resulting from these genetic polymorphisms were observed. The effect of *crtRB1* on total carotenoids would not have been predicted from known carotenoid biochemistry, and may have resulted from an epistatic interaction on upstream pathway components through transcriptional or metabolic signals rather than from *crtRB1* itself. Similarly, upon noticing that use of the more "favorable" alleles at *lcyε* - 5'TE polymorphism did not significantly increase β-carotene in any population, a more detailed investigation of the *lcyε* effect at this polymorphism revealed that this sequence variation largely affected only the α branch. These results modify our understanding of substrate allocation within the carotenoid pathway branches. Firstly, down stream pathway reactions can feedback onto governing pathway control points, meaning that selection of genetic variation below phytoene synthase should always be evaluated on the effect it has on precursor pools. Secondly, allocation of substrate to the α– and β– branches might also be subject to feedback, meaning that reduced allocation to one branch by selected genetic variation may not necessarily provide more substrate to the other branch. These two examples support exploration of pathway effects using systematic, multivariate approaches as well as derived trait ratios to fully elucidate the nature of genetic effects, and also suggest that rare intragenic or intergenic recombinants at both QTL should be identified in order to more fully explore the genetic effects of the favorable combinations.

β-carotene is a unique trait as it is inversely proportional in magnitude to all other carotenoids, was not significantly affected by additive or complementary responses in dosage studies, and appears to have an interesting relationship with *ccd1*-related degradation processes as it is the only carotenoid not affected by *ccd1*. In addition, there is a unique relationship between β-carotene and total carotenoid through *crtRB1* mediated processes revealing that β-carotene levels may govern carotenoid accumulation in both branches. These observations and the larger number of QTL that seem to regulate β–carotene concentration in comparison to other carotenoids suggest that this metabolite pool is under tight regulation and therefore may be a physiological control point. It is highly possible that the role of β–carotene as a precursor to abscisic acid synthesis and thus germination processes could be a reason for this regulation. As this could be the case, germination and vivipary should be closely monitored through germination tests for developed high carotenoid lines.

Interestingly, reports on the stability of genetic effects on carotenoid production across temperate, subtropical and tropical climates environments have not yet been conducted. Carotenoid compounds were found to be highly heritable, despite the differences in climate, which encourages the possibility of performing both advance and selection of lines in tropical or subtropical winter nurseries. Some differences in carotenoid concentrations were observed across environments, yet trait ratios appeared to stay constant, indicating that an environmental effect may be primarily due to increased carotenoid pathway activity, rather than shifts in allocation of substrate within the pathway. These differences could have arisen due to temperature, UV, precipitation and cultural practices.

Results from the marker assisted selection of *lcyε* and *crtRB1* suggested that in high total carotenoid backgrounds, selection of the desired alleles could make a large improvement over genotypes carrying the undesirable alleles. The ideal experiment for evaluating genetic gain and realized heritability from allele-specific selection is yet to be done. To estimate genetic gain from the selection, we consider the following relationship: $R = h^2 S$ where R is the response to selection as measured by the phenotype difference of the selected and original population means, S is the selection differential which reflects the phenotype difference between the original population mean and the mean of selected individuals from the original population, and h^2 is the realized heritability which indicates the transmissibility of favorable phenotype in terms of additive genetic effects. In our experiment, individuals from the S_0 generation were not

phenotyped and rather were selected on the basis of their marker profiles. As such, we have little indication of how the population in general would have fared, and have difficulty making the comparison of relative gain in the $S_{1:2}$ generation over the parent population. Considering that the current marker assisted selection approach has enabled impressive phenotype gains in proVA and has successfully reached a realistic nutritional target using sustainable breeding approaches (Figure 3.5), it is certain that our limited (but evolving!) fundamental understanding of the genetics and biochemistry affecting carotenoid accumulation is achieving its intended application.

REFERENCES

1. Doebley, J. F., Gaut, B. S. & Smith, B. D. The molecular genetics of crop domestication. *Cell* **127**, 1309-1321 (2006).

2. Palaisa, K. A., Morgante, M., Williams, M. & Rafalski, A. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell* **15**, 1795-1806 (2003).

3. Poneleit, C. G. in *Specialty Corns* (ed Hallauer, A. R.) 225-262 (CRC Press, Boca Raton, 1994).

4. Mangelsdorf, P. C. & Fraps, G. S. A direct quantitative relationship between vitamin A in corn and the number of genes for yellow pigmentation. *Science* **73**, 241-242 (1931).

5. Bauerfeind, J. C., Adams, C. R. & Marusich, W. L. in *Carotenoids as Colorants and Vitamin A Precursors* (ed Bauernfeind, J. C.) p563 (Academic Press, New York, 1981).

6. Weber, E. J. Carotenoids and Tocols of Corn Grain Determined by Hplc. *Journal of the American Oil Chemists Society* **64**, 1129-1134 (1987).

7. Johnson, E. J. The role of carotenoids in human health. *Nutrition in Clinical Care* **5**, 56-65 (2002).

8. Stephensen, C. B. Vitamin A, infection, and immune function. *Annu. Rev. Nutr.* **21**, 167-192 (2001).

9. Whitcher, J. P., Srinivasan, M. & Upadhyay, M. P. Corneal blindness: a global perspective. *Bull. World Health Organ.* **79**, 214-221 (2001).

10. Leuenberger, M. G., Engeloch-Jarret, C. & Woggon, W. D. The reaction mechanism of the enzyme-catalyzed central cleavage of beta-carotene to retinal. *Angewandte Chemie-International Edition* **40**, 2614-2617 (2001).

11. Fierce, Y. *et al*. In vitro and in vivo characterization of retinoid synthesis from beta-carotene. *Arch. Biochem. Biophys.* **472**, 126-138 (2008).

12. Tanumihardjo, S. A., Bouis, H., Hotz, C., Meenakshi, J. V. & McClafferty, B. Biofortification of staple crops: An emerging strategy to combat hidden hunger. *Comprehensive Reviews in Food Science and Food Safety* **7**, 329-334 (2008).

13. Tang, G. W., Qin, J., Dolnikowski, G. G., Russell, R. M. & Grusak, M. A. Spinach or carrots can supply significant amounts of vitamin A as assessed by feeding with intrinsically deuterated vegetables. *Am. J. Clin. Nutr.* **82**, 821-828 (2005).

14. Pfeiffer, W. H. & McClafferty, B. HarvestPlus: Breeding crops for better nutrition. *Crop Sci.* **47**, S88-S105 (2007).

15. Tanumihardjo, S. A. Food-based approaches for ensuring adequate vitamin A nutrition. *Comprehensive Reviews in Food Science and Food Safety* **7**, 373-381 (2008).

16. Chucair, A. J. *et al*. Lutein and zeaxanthin protect photoreceptors from apoptosis induced by oxidative stress: Relation with docosahexaenoic acid. *Invest. Ophthalmol. Vis. Sci.* **48**, 5168-5177 (2007).

17. Seddon, J. M. *et al*. Dietary Carotenoids, Vitamin-A, Vitamin-C, and Vitamin-E, and Advanced Age-Related Macular Degeneration. *Jama-Journal of the American Medical Association* **272**, 1413-1420 (1994).

18. Howitt, C. A. & Pogson, B. J. Carotenoid accumulation and function in seeds and non-green tissues. *Plant Cell and Environment* **29**, 435-445 (2006).

19. Harjes, C. E. *et al*. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **319**, 330-333 (2008).

20. Blessin, C. W., Brecher, J. D., Grogan, C. O., Campbell, C. M. & Dimler, R. J. Carotenoids of Corn and Sorghum .3. Variation in Xanthophylls and Carotenes in Hybrid, Inbred, and Exotic Corn Lines. *Cereal Chem.* **40**, 436-& (1963).

21. Buckner, B., Miguel, P. S., JanickBuckner, D. & Bennetzen, J. L. The Y1 gene of maize codes for phytoene synthase. *Genetics* **143**, 479-488 (1996).

22. Rodriguez-Concepcion, M. Early steps in isoprenoid biosynthesis: Multilevel regulation of the supply of common precursors in plant cells. *Phytochemistry Reviews* **5**, 1-15 (2006).

23. Estevez, J. M., Cantero, A., Reindl, A., Reichler, S. & Leon, P. 1-deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J. Biol. Chem.* **276**, 22901-22909 (2001).

24. Robertson, D. S. Survey of Albino and White-Endosperm Mutants of Maize - their Phenotypes and Gene Symbols. *J. Hered.* **66**, 67-74 (1975).

25. Islam, S. N. Survey of Carotenoid Variation and Quantitative Trait Loci Mapping for Carotenoid and Tocopherol Variation in Maize. (2004).

26. Wong, J. C., Lambert, R. J., Wurtzel, E. T. & Rocheford, T. R. QTL and candidate genes phytoene synthase and zeta-carotene desaturase associated with the accumulation of carotenoids in maize. *Theor. Appl. Genet.* **108**, 349-359 (2004).

27. Stevens, R. L. Genetic and QTL analysis of carotenoid variation in two mapping populations of maize. (2007).

28. Chander, S. *et al*. Using molecular markers to identify two major loci controlling carotenoid contents in maize grain. *Theor. Appl. Genet.* **116**, 223-233 (2008).

29. Yan, J. *et al*. *Rare Genetic Variation at Zea mays crtRB1 Increases β-carotene in Maize Grain*. *Nat.Genet.* (submitted).

30. Lichtenthaler, H. K., Rohmer, M. & Schwender, J. Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. *Physiol. Plantarum* **101**, 643-652 (1997).

31. Vallabhaneni, R. & Wurtzel, E. T. Timing and Biosynthetic Potential for Carotenoid Accumulation in Genetically Diverse Germplasm of Maize. *Plant Physiol.* **150**, 562-572 (2009).

32. Rodriguez-Amaya, D. B. & Kimura, M. Harvest Plus Handbook for Carotenoid Analysis. (2004).

33. Buckner, B., Kelson, T. L. & Robertson, D. S. Cloning of the Y1 Locus of Maize, a Gene Involved in the Biosynthesis of Carotenoids. *Plant Cell* **2**, 867-876 (1990).

34. Li, F., Murillo, C. & Wurtzel, E. T. Maize Y9 encodes a product essential for 15-cis-zeta-carotene isomerization. *Plant Physiol.* **144**, 1181-1189 (2007).

35. Hable, W. E., Oishi, K. K. & Schumaker, K. S. Viviparous-5 encodes phytoene desaturase, an enzyme essential for abscisic acid (ABA) accumulation and seed development in maize. *Molecular and General Genetics* **257**, 167-176 (1998).

36. Matthews, P. D., Luo, R. B. & Wurtzel, E. T. Maize phytoene desaturase and zeta-carotene desaturase catalyse a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. *J. Exp. Bot.* **54**, 2215-2230 (2003).

37. Singh, M. *et al*. Activator mutagenesis of the pink scutellum1/viviparous7 locus of maize. *Plant Cell* **15**, 874-884 (2003).

38. Li, F., Murillo, C. & Wurtzel, E. T. Maize Y9 encodes a product essential for 15-cis-zeta-carotene isomerization. *Plant Physiol.* **144**, 1181-1189 (2007).

39. Harjes, C. E. *et al*. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **319**, 330-333 (2008).

40. Vallabhaneni, R. *et al*. Metabolite sorting of a germplasm collection reveals the *Hydroxylase3* locus as a new target for maize provitamin A biofortification. *Plant Physiology* (2009).

41. Sandmann, G., Roemer, S. & Fraser, P. D. Understanding carotenoid metabolism as a necessity for genetic engineering of crop plants. *Metab. Eng.* **8**, 291-302 (2006).

42. Lois, L. M., Rodriguez-Concepcion, M., Gallego, F., Campos, N. & Boronat, A. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant Journal* **22**, 503-513 (2000).

43. Liu, K. J. *et al*. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics* **165**, 2117-2128 (2003).

44. Stevens, R. L. Genetic and QTL Analysis of Carotenoid Variation in Two Mapping Populations of Maize. (2007).

45. Egesel, C. O., Wong, J. C., Lambert, R. J. & Rocheford, T. R. Combining ability of maize inbreds for carotenoids and tocopherols. *Crop Sci.* **43**, 818-823 (2003).

46. Granado, F., Olmedilla, B., Gil-Martinez, E. & Blanco, I. A fast, reliable and low-cost saponification protocol for analysis of carotenoids in vegetables. *Journal of Food Composition and Analysis* **14**, 479-489 (2001).

47. Ooijen, J. W. V. & Voorrips, R. E. *JOINMAP version 3.0, Software for the Calculation of Genetic Linkage Maps*. Plant Research International, Wageningen. (2001).

48. SAS Institute. Version 9.2 Cary, North Carolina

49. Hallauer, A. R. & Miranda, J. B. in *Quantitative Genetics in Maize Breeding* 468 (Iowa State University Press, Ames, Iowa, 1988).

50. Utz, H. F. & Melchinger, A. E. PLABQTL: A computer program to map QTL. **1.2** (1993).

51. Haley, C. S. & Knott, S. A. A Simple Regression Method for Mapping Quantitative Trait Loci in Line Crosses using Flanking Markers. *Heredity* **69**, 315-324 (1992).

52. Jansen, R. C. & Stam, P. High-Resolution of Quantitative Traits into Multiple Loci Via Interval Mapping. *Genetics* **136**, 1447-1455 (1994).

53. Zeng, Z. B. Precision Mapping of Quantitative Trait Loci. *Genetics* **136**, 1457-1468 (1994).

54. Lander, E. S. & Botstein, D. Mapping Mendelian Factors Underlying Quantitative Traits using Rflp Linkage Maps. *Genetics* **121**, 185-199 (1989).

55. Doerge, R. W. & Churchill, G. A. Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285-294 (1996).

56. Lopez, A. B. *et al*. Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.* **59**, 213-223 (2008).

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

58. Phelps, T. L., Hall, A. E. & Buckner, B. Microsatellite repeat variation within the y1 gene of maize and teosinte. *J. Hered.* **87**, 396-399 (1996).

59. Andersen, J. R. & Lubberstedt, T. Functional markers in plants. *Trends Plant Sci.* **8**, 554-560 (2003).

60. Jiang, C. J. & Zeng, Z. B. Multiple-Trait Analysis of Genetic-Mapping for Quantitative Trait Loci. *Genetics* **140**, 1111-1127 (1995).

61. Zeng, Z. B. Theoretical Basis for Separation of Multiple Linked Gene Effects in Mapping Quantitative Trait Loci. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 10972-10976 (1993).

62. Korol, A. B., Ronin, Y. I. & Kirzhner, V. M. Interval Mapping of Quantitative Trait Loci Employing Correlated Trait Complexes. *Genetics* **140**, 1137-1147 (1995).

63. Mangin, B., Thoquet, P. & Grimsley, N. Pleiotropic QTL analysis. *Biometrics* **54**, 88-99 (1998).

64. Gilbert, H. & Le Roy, P. Power of three multitrait methods for QTL detection in crossbred populations. *Genetics Selection Evolution* **36**, 347-361 (2004).

65. Upadyayula, N., da Silva, H. S., Bohn, M. O. & Rocheford, T. R. Genetic and QTL analysis of maize tassel and ear inflorescence architecture. *Theor. Appl. Genet.* **112**, 592-606 (2006).

66. Brewer, M. T., Moyseenko, J. B., Monforte, A. J. & van der Knaap, E. Morphological variation in tomato: a comprehensive study of quantitative trait loci controlling fruit shape and development. *J. Exp. Bot.* **58**, 1339-1349 (2007).

67. Ebrahimi, A. *et al*. QTL mapping of seed-quality traits in sunflower recombinant inbred lines under different water regimes. *Genome* **51**, 599-615 (2008).

68. Menkir, A. & Maziya-Dixon, B. Influence of genotype and environment on beta-carotene content of tropical yellow-endosperm maize genotypes. *Maydica* **49**, 313-318 (2004).

69. Chander, S., Meng, Y., Zhang, Y., Yan, J. & Li, J. Comparison of nutritional traits variability in selected eighty-seven inbreds from Chinese maize (Zea mays L.) germplasm. *J. Agric. Food Chem.* **56**, 6506-6511 (2008).

70. Menkir, A., Liu, W., White, W. S., Mazlya-Dixon, B. & Rocheford, T. Carotenoid diversity in tropical-adapted yellow maize inbred lines. *Food Chem.* **109**, 521-529 (2008).

71. Li, F., Vallabhaneni, R., Yu, J., Rocheford, T. & Wurtzel, E. T. The maize phytoene synthase gene family: Overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. *Plant Physiol.* **147**, 1334-1346 (2008).

72. Gallagher, C. E., Matthews, P. D., Li, F. Q. & Wurtzel, E. T. Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. *Plant Physiol.* **135**, 1776-1783 (2004).

73. Ortiz-Monasterio, J. I. *et al*. Enhancing the mineral and vitamin content of wheat and maize through plant breeding. *J. Cereal Sci.* **46**, 293-307 (2007).

74. Penuelas, J. & Munne-Bosch, S. Isoprenoids: an evolutionary pool for photoprotection. *Trends Plant Sci.* **10**, 166-169 (2005).

75. Sun, Z. *et al*. 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 (2008).

76. Robertson, D. S. Some thoughts about the nature of Mu-induced Y1 mutants. *Maize News Letters* **61**, 9-10 (1987).

77. Li, F., Vallabhaneni, R., Yu, J., Rocheford, T. & Wurtzel, E. T. The maize phytoene synthase gene family: Overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. *Plant Physiol.* **147**, 1334-1346 (2008).

78. DellaPenna, D. & Pogson, B. J. Vitamin synthesis in plants: Tocopherols and carotenoids. *Annual Review of Plant Biology* **57**, 711-738 (2006).

79. Matthews, P. D. & Wurtzel, E. T. in *Food Colorants: Chemical and Functional Properties* (ed Sociaciu, C.) (CRC Press, Boca Raton, 2007).

80. Kean, E. G., Ejeta, G., Hamaker, B. R. & Ferruzzi, M. G. Characterization of carotenoid pigments in mature and developing kernels of selected yellow-endosperm sorghum varieties. *J. Agric. Food Chem.* **55**, 2619-2626 (2007).

81. Auldridge, M. E. *et al*. Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. *Plant Journal* **45**, 982-993 (2006).

82. Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A. D. & McCarty, D. R. VP14 of maize catalyzes the carotenoid cleavage reaction of abscisic acid biosynthesis. *Plant Physiol.* **114**, 798-798 (1997).

83. Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A. D. & McCarty, D. R. Specific oxidative cleavage of carotenoids by VP14 of maize. *Science* **276**, 1872-1874 (1997).

84. Tan, B. C. et al. *The Dominant White Endosperm Factor White Cap Encodes the ZmCCD1 Carotenoid Dioxygenase in a Large Multiple Copy Gene Array* (Maize Genetics Conference Abstracts Ser. 46, 2004).

85. Vogel, J. T., Tan, B., McCarty, D. R. & Klee, H. J. The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *J. Biol. Chem.* **283**, 11364-11373 (2008).

86. Flint-Garcia, S. A. *et al*. Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant Journal* **44**, 1054-1064 (2005).

87. Haley, C. S. & Knott, S. A. A Simple Regression Method for Mapping Quantitative Trait Loci in Line Crosses using Flanking Markers. *Heredity* **69**, 315-324 (1992).

88. Bradbury, P. J. *et al*. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633-2635 (2007).

89. Schwartz, S. H., Qin, X. Q. & Zeevaart, J. A. D. Characterization of a novel carotenoid cleavage dioxygenase from plants. *J. Biol. Chem.* **276**, 25208-25211 (2001).

90. Vogel, J. T., Tan, B., McCarty, D. R. & Klee, H. J. The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *J. Biol. Chem.* **283**, 11364-11373 (2008).

91. Kapitonov, V. V. & Jurka, J. Rolling-circle transposons in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8714-8719 (2001).

92. Lal, S. K., Giroux, M. J., Brendel, V., Vallejos, C. E. & Hannah, L. C. The maize genome contains a Helitron insertion. *Plant Cell* **15**, 381-391 (2003).

93. Morgante, M. *et al*. Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.* **37**, 997-1002 (2005).

94. Lal, S. K. & Hannah, L. C. Plant genomes - Massive changes of the maize genome are caused by Helitrons. *Heredity* **95**, 421-422 (2005).

95. Yang, L. & Bennetzen, J. L. Structure-based discovery and description of plant and animal Helitrons. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12832-12837 (2009).

96. Troyer, A. F. Background of US hybrid corn. *Crop Sci.* **39**, 601-626 (1999).

97. Rodriguez-Amaya, D. B. Changes in carotenoids during processing and storage of foods. *Arch. Latinoam. Nutr.* **49**, 38S-47S (1999).

98. Coe, E. H., Neuffer, M. G. & Hoisington, D. A. in *Corn and Corn Improvement* (eds Sprague, G. F. & Dudley, J. W.) 81-258 (American Sociery of Agronomy, Boca Raton, 1988).

99. Wilson, L. M. *et al*. Dissection of maize kernel composition and starch production by candidate gene association. *Plant Cell* **16**, 2719-2733 (2004).

100. Weber, A. L. *et al*. The Genetic Architecture of Complex Traits in Teosinte (Zea mays ssp parviglumis): New Evidence From Association Mapping. *Genetics* **180**, 1221-1232 (2008).

101. Salvi, S. *et al*. Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus m maize. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11376-11381 (2007).

102. Buckler, E. S. *et al*. The Genetic Architecture of Maize Flowering Time. *Science* **325**, 714-718 (2009).

103. Buckler, E. & Gore, M. An Arabidopsis haplotype map takes root. *Nat. Genet.* **39**, 1056-1057 (2007).

104. Galpaz, N., Ronen, G., Khalfa, Z., Zamir, D. & Hirschberg, J. A chromoplast-specific carotenoid biosynthesis pathway is revealed by cloning of the tomato white-flower locus. *Plant Cell* **18**, 1947-1960 (2006).

105. Ducrocq, S. *et al*. Key impact of Vgt1 on flowering time adaptation in maize: Evidence from association mapping and ecogeographical information. *Genetics* **178**, 2433-2437 (2008).

106. Dudley, J. W. & Lambert, R. J. 90-Generations of Selection for Oil and Protein in Maize. *Maydica* **37**, 81-87 (1992).

107. Duvick, D. N. & Cassman, K. G. Post-green revolution trends in yield potential of temperate maize in the north-central United States. *Crop Sci.* **39**, 1622-1630 (1999).

108. Duvick, D. N. Biotechnology in the 1930s: the development of hybrid maize. *Nature Reviews Genetics* **2**, 69-74 (2001).

109. Moose, S. P. & Mumm, R. H. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* **147**, 969-977 (2008).

110. Mazur, B., Krebbers, E. & Tingey, S. Gene discovery and product development for grain quality traits. *Science* **285**, 372-375 (1999).

111. Graham, R. D., Welch, R. M. & Bouis, H. E. Addressing micronutrient malnutrition through enhancing the nutritional quality of staple foods: Principles, perspectives and knowledge gaps. *Advances in Agronomy, Vol 70* **70**, 77-142 (2001).

112. Debier, C. & Larondelle, Y. Vitamins A and E: metabolism, roles and transfer to offspring. *Br. J. Nutr.* **93**, 153-174 (2005).

113. Graham, R., Senadhira, D., Beebe, S., Iglesias, C. & Monasterio, I. Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crops Res.* **60**, 57-80 (1999).

114. Friedman, D. S. *et al*. Prevalence of age-related macular degeneration in the United States. *Arch. Ophthalmol.* **122**, 564-572 (2004).

115. Kassoff, A. *et al*. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss - AREDS Report No. 8. *Arch. Ophthalmol.* **119**, 1417-1436 (2001).

116. Alves-Rodrigues, A. & Shao, A. The science behind lutein. *Toxicol. Lett.* **150**, 57-83 (2004).

117. Howe, J. A. & Tanumihardjo, S. A. Carotenoid-biofortified maize maintains adequate vitamin A status in Mongolian gerbils. *J. Nutr.* **136**, 2562-2567 (2006).

118. Kurilich, A. C. & Juvik, J. A. Quantification of carotenoid and tocopherol antioxidants in Zea mays. *J. Agric. Food Chem.* **47**, 1948-1955 (1999).

119. Cardoso, W. S. *et al*. Variability of maize genotypes for grain carotenoid composition. *Pesquisa Agropecuaria Brasileira* **44**, 164-172 (2009).

120. Goodman, M. M. & Brown, W. L. in *Corn and Corn Improvement* (eds Sprague, G. F. & Dudley, J. W.) 33-79 (American Society of Agronomy, Madison, WI, 1988).

121. Egesel, C. O., Wong, J. C., Lambert, R. J. & Rocheford, T. R. Gene dosage effects on carotenoid concentration in maize grain. *Maydica* **48**, 183-190 (2003).

122. Laurie, C. C. *et al*. The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics* **168**, 2141-2155 (2004).

123. Remington, D. L. *et al*. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 11479-11484 (2001).

124. Tenaillon, M. I. *et al*. Patterns of DNA sequence polymorphism along chromosome 1 of maize (Zea mays ssp mays L.). *Proc. Natl. Acad. Sci. U. S. A.* **98**, 9161-9166 (2001).

125. Zhu, C., Gore, M., Buckler, E. & Yu, J. M. Status and prospects of association mapping in plants. *Plant Genome* **1**, 5 (2008).

126. Myles, S. *et al*. Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. *Plant Cell* **21**, 2194-2202 (2009).

127. Bai, L., Kim, E., DellaPenna, D. & Brutnell, T. P. Novel lycopene epsilon cyclase activities in maize revealed through perturbation of carotenoid biosynthesis. *Plant Journal* **59**, 588-599 (2009).

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

129. FALCONER, D. S. & Mackay, T. F. C. in *Introduction to Quantitative Genetics* (Longmans Green, Harlow, Essex, UK, 1996).

130. Cunningham, F. X. *et al*. Functional analysis of the beta and epsilon lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell* **8**, 1613-1626 (1996).

131. Li, F., Vallabhaneni, R., Yu, J., Rocheford, T. & Wurtzel, E. T. The maize phytoene synthase gene family: Overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. *Plant Physiol.* **147**, 1334-1346 (2008).

132. Aluru, M. *et al*. Generation of transgenic maize with enhanced provitamin A content. *J. Exp. Bot.* **59**, 3551-3562 (2008).

133. Ye, X. D. *et al*. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**, 303-305 (2000).

134. Paine, J. A. *et al*. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat. Biotechnol.* **23**, 482-487 (2005).

135. Moore, J. H. & Williams, S. M. Traversing the conceptual divide between biological and statistical epistasis: Systems biology and a more modern synthesis. *Bioessays* **27**, 637-646 (2005).

136. Fisher, R. A. The correlation between relatives on the supposition of mendelian inheritance. *royal society of edinburgh transactions of the society* **52**, 399-433 (1918).

137. Kliebenstein, D. J. Advancing Genetic Theory and Application by Metabolic Quantitative Trait Loci Analysis. *Plant Cell* **21**, 1637-1646 (2009).

138. Diretto, G. *et al*. Silencing of beta-carotene hydroxylase increases total carotenoid and beta-carotene levels in potato tubers. *Bmc Plant Biology* **7**, 11 (2007).

139. Davis, C., Jing, H., Howe, J. A., Rocheford, T. & Tanumihardjo, S. A. beta-Cryptoxanthin from supplements or carotenoid-enhanced maize maintains liver vitamin A in Mongolian gerbils (Meriones unguiculatus) better than or equal to beta-carotene supplements. *Br. J. Nutr.* **100**, 786-793 (2008).

140. Banziger, M. & Lafitte, H. R. Efficiency of secondary traits for improving maize for low-nitrogen target environments. *Crop Sci.* **37**, 1110-1117 (1997).

141. Iqbal, M., Navabi, A., Salmon, D. F., Yang, R. -. & Spaner, D. Simultaneous selection for early maturity, increased grain yield and elevated grain protein content in spring wheat. *Plant Breeding* **126**, 244-250 (2007).

142. Smith, H. F. A discriminant function for plant selection. *Ann. Eugenics* **7**, 240-250 (1936).

143. Hazel, L. N. The genetic basis for constructing selection indices. *Genetics* **28**, 476-490 (1943).

144. Hospital, F. Challenges for effective marker-assisted selection in plants. *Genetica* **136**, 303-310 (2009).

145. Brenna, O. V. & Berardo, N. Application of near-infrared reflectance spectroscopy (NIRS) to the evaluation of carotenoids content in maize. *J. Agric. Food Chem.* **52**, 5577-5582 (2004).

146. Rathore, S. S. S. et al. *Corn Carotenoid Measurements Using Near-Infrared Spectroscopy* (International Grain Quality & Technology Congress Proceedings, 2008).

147. Hulshof, P. J. M., Kosmeijer-Schuil, T., West, C. E. & Hollman, P. C. H. Quick screening of maize kernels for provitamin A content. *Journal of Food Composition and Analysis* **20**, 655-661 (2007).

148. Lande, R. & Thompson, R. Efficiency of Marker-Assisted Selection in the Improvement of Quantitative Traits. *Genetics* **124**, 743-756 (1990).

149. Zhang, W. & Smith, C. Simulation of Marker-Assisted Selection Utilizing Linkage Disequilibrium - the Effects of several Additional Factors. *Theor. Appl. Genet.* **86**, 492-496 (1993).

150. Moerkerke, B., Goetghebeur, E., De Riek, J. & Roldan-Ruiz, I. Significance and impotence: towards a balanced view of the null and the alternative hypotheses in marker selection for plant breeding. *Journal of the Royal Statistical Society Series A-Statistics in Society* **169**, 61-79 (2006).

151. Lecomte, L. *et al*. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *Theor. Appl. Genet.* **109**, 658-668 (2004).

152. Chaib, J., Lecomte, L., Buret, M. & Causse, M. Stability over genetic backgrounds, generations and years of quantitative trait locus (QTLs) for organoleptic quality in tomato. *Theor. Appl. Genet.* **112**, 934-944 (2006).

APPENDICES

Appendix A: Carotenoid QTL genotypes and carotenoid concentrations for Goodman-Buckler Diversity Panel grown in Urbana, IL 2005

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
3316	2	1	1	2	2	1	3	3	.	.	.	.	.	.	.	.	.
3811	.	2	3	.	2	2	3	1	4.97	10.23	0.95	1.34	0.14	1.82	18.51	3.05	0.91
4226	.	2	3	3	1	1	1	.	8.80	5.44	1.61	0.39	0.42	0.19	15.24	4.80	0.94
4722	3	2	3	2	2	1	1	.	4.28	8.35	1.17	1.13	0.13	1.36	15.25	0.30	0.09
A188	3	2	2	2	2	1	1	3	.	.	.	.	.	.	.	.	.
A214N	3	2	2	3	1	1	1	.	.	.	.	.	.	.	.	.	.
A239	3	2	3	.	.	.	1	.	8.01	7.75	3.29	1.78	1.89	0.38	19.82	3.88	0.82
A272	3	2	3	.	.	.	1	.	4.72	25.24	1.30	4.18	4.66	0.49	39.29	2.83	0.81
A4415	3	2	3	4	2	1	1	3	3.40	9.05	0.90	1.97	0.11	1.63	16.16	0.41	0.13
A554	3	2	3	2	2	1	1	1	5.16	7.73	1.14	1.76	1.43	0.19	16.27	1.23	0.33
A556	3	2	1	3	1	1	1	.	7.37	2.36	0.65	0.39	1.15	0.43	11.70	2.86	0.56
A6	3	2	3	4	2	1	1	1	.	.	.	.	.	.	.	.	.
A619	3	2	1	2	2	2	1	.	6.90	5.78	0.94	1.12	3.54	0.21	17.54	4.29	1.57
A632	3	2	3	2	2	1	1	1	4.13	7.04	0.97	0.69	0.58	0.08	12.51	0.05	0.04
A634	3	2	3	3	1	1	1	1	4.89	2.09	0.74	0.29	0.24	0.12	7.63	0.03	0.04
A635	3	2	3	3	1	1	1	.	11.74	4.64	1.22	0.42	0.57	0.17	17.54	0.78	0.17
A641	3	2	3	3	1	1	1	.	3.86	2.84	1.20	1.74	0.69	0.17	9.30	1.49	0.33
A654	3	2	1	3	1	1	1	1	2.83	0.64	0.31	0.21	0.21	0.07	3.95	0.69	0.12
A659	3	2	3	3	1	1	1	1	7.84	6.36	2.57	0.61	0.97	0.35	16.12	1.86	0.62
A661	3	2	2	3	1	1	1	.	4.90	2.80	0.51	0.23	0.13	0.08	8.15	0.52	0.13
A679	3	2	1	3	1	1	1	1	4.12	1.49	1.01	0.51	0.56	0.63	7.30	0.85	0.04
A680	3	2	3	3	1	1	1	1	8.21	2.92	2.23	0.56	0.45	0.47	12.62	0.76	0.18
A682	3	2	2	3	1	1	1	1	7.94	3.75	0.79	0.33	0.23	0.14	12.40	2.54	0.52
AB28A	3	2	3	2	.	1	1	1	.	.	.	.	.	.	.	.	.
B10	3	2	3	3	1	1	1	1	.	.	.	.	.	.	.	.	.
B103	3	2	2	3	1	1	1	1	4.24	4.42	0.67	0.69	0.47	0.09	9.91	0.76	0.16

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
B104	3	2	1	.	.	.	1	1	11.81	3.94	1.13	0.80	1.26	0.24	18.04	3.64	0.78
B105	3	2	1	3	1	1	1	1	11.10	3.49	1.74	0.40	1.29	0.34	16.61	4.19	1.21
B109	3	2	3	3	1	1	1	1	8.05	3.95	2.40	0.74	0.71	0.72	14.16	0.77	0.15
B115	3	2	1	3	1	1	1	1	11.45	4.03	1.12	0.45	0.40	2.15	18.48	0.53	0.23
B14A	3	2	3	3	1	1	1	1	6.18	4.55	1.24	0.68	0.57	0.17	12.15	0.17	0.08
B164	3	2	3	3	1	1	1	1	7.41	3.56	0.61	0.29	0.78	0.14	12.17	2.88	1.00
B2	3	2	3	3	1	1	1	.	7.45	3.96	0.78	0.46	0.20	0.52	12.58	0.35	0.18
B37	.	2	3	2	2	1	1	1	13.81	5.36	1.24	1.30	0.64	1.80	22.91	3.22	0.77
B46	3	2	3	3	1	1	1	.	5.84	4.98	2.34	0.64	0.84	0.29	12.59	3.98	1.00
B52	3	2	3	3	1	1	2	3	3.53	2.58	1.15	0.22	0.10	0.20	6.63	0.62	0.22
B57	3	2	2	3	1	1	1	1	11.88	2.06	5.10	0.68	0.88	0.75	16.24	2.76	0.73
B64	3	2	3	2	2	2	1	.	1.95	8.04	0.83	0.94	0.13	1.46	12.51	0.58	0.19
B68	3	2	3	3	1	1	1	.	4.94	3.75	0.88	0.39	0.31	0.13	9.52	0.16	0.09
B73	3	2	3	3	1	1	1	1	7.22	3.27	1.68	0.84	0.47	0.61	12.40	0.49	0.12
B73HTRHM	3	2	3	3	1	1	1	1	6.59	2.65	1.68	0.62	0.49	0.51	10.85	0.63	0.11
B75	3	2	1	3	1	1	1	1	9.64	2.29	2.52	0.27	1.85	0.49	14.54	1.15	0.33
B76	3	2	3	.	.	.	1	1	13.49	2.98	1.05	0.49	1.02	0.44	18.42	2.73	0.74
B77	2	1	1	2	2	2	1	1	4.62	3.36	0.74	0.47	4.00	0.22	12.67	3.56	0.73
B79	3	2	1	3	1	1	1	.	11.25	1.83	0.65	0.23	0.92	0.22	14.44	0.71	0.24
B84	3	2	3	.	.	.	1	1	.	.	.	.	.	.	.	.	.
B97	3	2	3	.	.	.	1	.	5.00	2.89	0.61	0.27	0.28	0.14	8.58	2.15	0.71
C103	3	2	3	2	2	1	1	.	.	.	.	.	.	.	.	.	.
C123	3	2	3	2	2	1	1	.	4.15	5.41	0.74	0.91	0.39	0.08	10.94	1.03	0.36
C49A	3	2	1	3	1	1	1	.	9.31	1.08	2.11	0.56	1.27	0.79	13.01	0.75	0.15
CH70130	3	2	1	2	2	2	1	1	3.43	3.72	0.65	0.76	1.01	0.16	9.08	0.81	0.18
CH9	3	2	1	3	1	1	1	.	4.68	3.19	1.70	0.76	1.16	0.18	9.96	1.48	0.45
CI1872	3	2	3	2	2	2	1	1	.	.	.	.	.	.	.	.	.
CI21E	.	2	3	2	2	2	1	1	2.99	3.53	0.67	0.46	0.12	0.38	7.47	0.11	0.02

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
CI28A	3	2	3	3	1	1	1	.	.	.	.	.	.	.	.	.	.
CI31A	.	2	3	2	2	2	1	1	.	.	.	.	.	.	.	.	.
CI3A	3	2	3	2	2	1	1	.	4.20	6.65	1.05	0.94	0.25	1.91	13.95	2.07	0.52
CI64	3	2	1	4	2	2	3	3	.	.	.	.	.	.	.	.	.
CI66	3	2	3	2	2	1	3	3	.	.	.	.	.	.	.	.	.
CI7	2	1	1	2	2	1	1	1	3.43	1.52	0.44	0.29	0.22	3.72	9.18	2.94	0.78
CI90C	3	2	3	2	2	1	1	1	6.45	1.94	0.46	0.22	0.12	0.57	9.29	0.30	0.15
CI91B	.	2	3	3	1	1	1	.	8.19	5.28	0.92	0.78	0.22	0.87	15.34	1.52	0.30
CM105	3	1	.	3	1	1	1	1	3.67	1.69	0.65	0.26	1.15	0.23	7.00	0.12	0.05
CM174	3	1	1	3	1	1	1	1	6.50	1.70	0.89	0.26	1.22	0.18	9.86	0.13	0.04
CM37	3	1	1	3	1	1	1	.	12.06	2.94	0.87	0.61	1.73	0.29	17.63	0.64	0.24
CM7	3	2	1	3	1	1	1	.	7.03	1.78	0.69	0.34	0.83	0.34	10.32	1.12	0.24
CML10	.	2	3	.	2	1	.	3	.	.	.	.	.	.	.	.	.
CML103	3	2	2	.	2	1	3	3	.	.	.	.	.	.	.	.	.
CML108	.	2	3	2	2	1	3	1	.	.	.	.	.	.	.	.	.
CML11	3	2	3	4	2	1	3	3	.	.	.	.	.	.	.	.	.
CML14	.	2	3	.	.	.	.	3	.	.	.	.	.	.	.	.	.
CML154Q	3	2	3	4	2	1	3	.	.	.	.	.	.	.	.	.	.
CML157Q	3	2	3	4	2	1	3	.	.	.	.	.	.	.	.	.	.
CML158Q	.	2	3	4	2	1	.	3	.	.	.	.	.	.	.	.	.
CML218	3	2	3	2	2	2	3	3	.	.	.	.	.	.	.	.	.
CML220	3	2	3	.	.	.	3	.	.	.	.	.	.	.	.	.	.
CML228	.	2	3	1	2	1	1	.	.	.	.	.	.	.	.	.	.
CML238	3	2	2	2	2	2	.	.	.	.	.	.	.	.	.	.	.
CML247	3	2	3	2	2	2	3	.	.	.	.	.	.	.	.	.	.
CML254	.	2	3	2	2	2	.	3	.	.	.	.	.	.	.	.	.
CML258	3	2	3	.	.	.	3	3	.	.	.	.	.	.	.	.	.
CML261	.	2	3	2	2	1	.	3	.	.	.	.	.	.	.	.	.

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
CML264	.	2	3	4	2	1	.	3	.	.	.	.	.	.	.	.	.
CML277	3	2	3	2	2	2	3	3	.	.	.	.	.	.	.	.	.
CML281	.	2	3	2	2	1	.	.	.	.	.	.	.	.	.	.	.
CML287	.	2	3	4	2	1	.	.	.	.	.	.	.	.	.	.	.
CML311	.	2	3	2	2	1	.	1	.	.	.	.	.	.	.	.	.
CML314	3	2	3	2	1	1	3	3	.	.	.	.	.	.	.	.	.
CML321	3	2	3	4	2	1	3	3	.	.	.	.	.	.	.	.	.
CML322	.	2	3	4	2	1	3	3	.	.	.	.	.	.	.	.	.
CML323	3	2	3	2	2	2	1	.	2.66	9.27	0.75	3.12	0.09	2.89	18.03	0.78	0.34
CML328	.	2	3	1	2	1	3	1	.	.	.	.	.	.	.	.	.
CML331	.	2	3	.	2	1	.	2	.	.	.	.	.	.	.	.	.
CML332	3	2	3	2	2	1	1	2	.	.	.	.	.	.	.	.	.
CML333	3	2	3	2	2	2	3	.	.	.	.	.	.	.	.	.	.
CML341	.	2	3	4	2	2	.	3	.	.	.	.	.	.	.	.	.
CML38	.	2	3	2	2	2	_	1	.	.	.	.	.	.	.	.	.
CML45	.	2	2	2	2	2	.	3	.	.	.	.	.	.	.	.	.
CML5	.	2	3	4	2	1	.	1	.	.	.	.	.	.	.	.	.
CML52	3	2	3	2	2	1	2	.	7.00	14.88	1.45	3.47	0.24	3.22	28.81	0.37	0.10
CML61	.	2	3	2	2	2	.	.	.	.	.	.	.	.	.	.	.
CML69	.	2	3	2	2	1	.	3	.	.	.	.	.	.	.	.	.
CML77	.	2	3	2	2	2	3	3	.	.	.	.	.	.	.	.	.
CML91	3	2	3	2	2	2	3	3	.	.	.	.	.	.	.	.	.
CML92	3	2	3	2	2	2	3	.	.	.	.	.	.	.	.	.	.
CO125	3	2	1	2	2	1	1	2	.	.	.	.	.	.	.	.	.
CO255	3	2	3	2	2	1	1	1	5.03	0.42	0.42	0.24	0.31	0.24	6.23	0.16	0.05
D940Y	3	2	3	3	2	1	1	1	7.26	4.57	1.83	0.53	0.14	0.84	13.33	1.60	0.56
DE1	3	2	2	3	1	1	1	.	7.71	5.40	0.61	0.63	0.21	1.03	14.97	1.89	0.41
DE2	3	2	2	.	.	.	1	.	7.11	3.89	0.56	0.38	0.16	0.59	12.12	1.44	0.37

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
DEexp	2	1	1	3	1	1	1	.	.	.	.	.	.	.	.	.	.
DE811	3	2	3	2	2	1	1	1	4.75	13.31	1.11	1.43	1.33	0.09	20.90	2.60	0.83
E2558W	3	1	1	3	1	1	3	3	.	.	.	.	.	.	.	.	.
EP1	3	2	2	2	2	1	1	2	6.13	6.56	0.86	0.98	0.61	0.08	14.37	1.18	0.33
F2	.	.	.	.	.	.	1	.	3.87	1.82	0.48	0.25	0.28	0.11	6.33	0.46	0.08
F2834T	3	2	2	2	2	1	1	3	2.79	4.74	1.11	0.69	0.87	0.07	9.15	0.28	0.07
F44	3	2	3	2	2	1	1	1	3.68	7.35	1.06	0.79	0.16	1.00	12.98	0.40	0.19
F6	3	2	2	2	2	1	1	.	.	.	.	.	.	.	.	.	.
F7	3	2	3	2	2	1	1	2	9.34	2.70	0.50	0.25	0.51	0.25	13.04	1.25	0.34
GA209	3	2	3	2	2	1	3	1	.	.	.	.	.	.	.	.	.
GT112	3	2	3	2	2	1	1	1	.	.	.	.	.	.	.	.	.
H105W	3	2	3	2	2	1	3	1	.	.	.	.	.	.	.	.	.
H49	3	2	3	3	1	1	1	1	10.53	4.99	0.88	0.32	0.77	0.25	16.85	2.17	0.63
H84	2	1	1	2	2	1	1	1	5.60	1.54	0.51	0.25	0.28	2.47	10.15	0.14	0.08
H91	3	2	3	3	1	1	1	1	4.46	2.38	0.53	0.25	0.22	0.09	7.39	0.04	0.01
H95	.	2	1	3	1	1	1	1	8.53	2.06	0.62	0.22	1.58	0.25	12.64	3.61	0.63
H99	3	2	3	2	2	1	1	1	14.26	5.56	1.75	0.43	1.55	0.81	22.61	1.92	0.65
HI27	3	2	3	2	2	1	1	3	.	.	.	.	.	.	.	.	.
HP301	.	2	2	2	2	1	.	.	.	.	.	.	.	.	.	.	.
HY	3	2	3	3	1	1	1	1	4.41	2.68	1.39	0.37	0.19	0.21	7.87	0.38	0.03
I137TN	3	2	1	4	2	1	1	2	2.51	1.56	0.57	0.27	0.18	1.31	5.83	0.13	0.04
I205	.	2	3	3	1	1	.	.	.	.	.	.	.	.	.	.	.
I29	3	2	1	2	2	1	3	2	.	.	.	.	.	.	.	.	.
IA2132	3	2	2	3	1	1	1	1	5.82	10.61	1.13	0.76	0.52	0.13	17.84	0.41	0.14
IA5125	3	2	2	3	1	1	1	3	7.34	7.48	1.28	0.67	0.41	0.11	16.00	0.08	0.04
IDS28	3	2	3	3	1	1	1	.	6.61	3.31	0.55	0.38	0.20	0.11	10.60	0.31	0.09
IDS69	3	2	2	2	2	1	1	.	3.05	9.92	0.89	1.74	0.87	0.15	15.73	0.58	0.17
IDS91	3	2	2	2	2	1	1	3	3.84	10.69	0.84	1.94	1.82	0.27	18.56	1.91	0.74

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
IL101	3	2	2	2	2	1	1	2	7.41	7.40	1.63	0.87	0.49	0.24	16.41	1.21	0.22
IL14H	3	2	3	3	1	1	3	1	.	.	.	.	.	.	.	.	.
IL677A	.	2	2	4	2	1	1	.	1.48	3.06	0.44	0.38	0.08	0.61	5.60	0.10	0.06
K148	3	1	1	2	2	1	1	1	9.34	0.48	0.51	0.40	3.00	0.29	13.50	0.18	0.06
K4	3	2	1	2	2	1	2	1	4.83	2.52	0.57	0.65	0.13	2.77	10.90	6.49	1.26
K55	3	2	3	.	.	.	1	3	.	.	.	.	.	.	.	.	.
K64	3	2	1	4	2	2	3	3	.	.	.	.	.	.	.	.	.
KI11	.	2	3	2	2	1	1	.	.	.	.	.	.	.	.	.	.
KI14	3	2	3	2	2	1	3	1	.	.	.	.	.	.	.	.	.
KI2021	3	2	3	2	2	1	1	1	.	.	.	.	.	.	.	.	.
KI21	.	2	3	2	2	1	1	1	.	.	.	.	.	.	.	.	.
KI3	3	2	3	4	2	2	1	1	3.04	12.42	1.12	2.24	0.09	1.39	19.19	0.44	0.12
KI43	3	2	3	2	2	2	1	1	.	.	.	.	.	.	.	.	.
KI44	3	2	3	2	2	2	1	.	5.37	9.79	1.64	1.01	0.10	0.76	17.02	0.94	0.24
KUI2007	.	.	.	.	.	.	1	.	6.63	3.47	0.82	0.41	0.16	0.36	11.03	0.43	0.18
KY21	3	2	2	4	2	2	3	3	.	.	.	.	.	.	.	.	.
KY226	.	2	3	2	2	1	.	.	.	.	.	.	.	.	.	.	.
KY228	3	2	1	.	2	1	3	3	.	.	.	.	.	.	.	.	.
L317	3	2	2	3	1	1	1	.	5.84	3.62	0.68	0.49	0.43	0.10	10.47	1.40	0.38
L578	3	2	3	2	2	1	1	.	3.04	2.88	0.51	0.40	0.10	0.48	6.90	0.13	0.05
M14	.	2	3	.	.	.	.	.	13.18	3.98	1.83	0.34	1.30	0.42	19.22	3.00	0.90
M162W	.	1	1	3	1	1	.	.	.	.	.	.	.	.	.	.	.
M37W	3	2	.	2	2	2	3	1	.	.	.	.	.	.	.	.	.
MEF156552	3	2	3	3	1	1	1	.	5.39	4.32	0.65	1.12	0.47	0.22	11.51	0.51	0.15
MO17	3	2	.	2	2	1	1	1	8.77	7.43	0.83	0.66	0.61	0.11	17.57	4.16	0.71
MO18W	.	2	3	3	1	1	3	3	.	.	.	.	.	.	.	.	.
MO1W	3	2	3	3	1	1	3	3	.	.	.	.	.	.	.	.	.
MO24W	.	1	.	2	2	1	3	3	.	.	.	.	.	.	.	.	.

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
MO44	.	2	3	2	2	1	.	1	5.51	2.93	0.55	0.33	0.43	0.13	9.33	1.47	0.33
MO45	3	2	3	3	1	1	1	1	10.36	7.61	1.65	1.80	0.19	1.48	21.45	1.54	0.42
MO46	3	2	3	3	1	1	1	.	7.41	3.06	0.86	0.43	0.55	0.30	11.76	0.79	0.19
MO47	3	2	3	2	1	1	1	.	7.79	6.65	0.75	1.10	0.86	0.29	16.68	1.24	0.39
MOG	3	2	3	.	2	2	1	1	4.49	8.71	1.02	1.62	0.12	2.70	17.63	1.29	0.57
MP339	2	1	1	2	2	2	3	1	.	.	.	.	.	.	.	.	.
MS1334	3	2	2	3	1	1	1	.	14.12	2.43	1.20	0.47	0.50	0.24	17.75	1.86	0.42
MS153	3	1	1	3	1	1	1	1	8.57	2.49	0.83	0.38	1.53	0.15	13.11	2.00	0.60
MS71	.	2	3	3	1	1	1	.	6.09	2.33	0.59	0.24	0.33	0.10	9.09	0.71	0.17
MT42	1	2	3	2	1	1	1	.	4.66	3.26	1.00	0.46	1.44	0.38	10.19	0.77	0.13
N192	3	2	3	3	1	1	1	1	7.14	4.03	1.92	0.54	0.43	0.25	12.39	0.30	0.10
N28HT	3	2	3	3	1	1	1	1	11.80	4.78	1.74	0.51	0.22	0.70	18.00	0.03	0.07
N6	3	2	1	2	2	1	1	3	5.39	3.86	1.14	0.88	2.35	0.37	12.85	0.60	0.12
N7A	3	2	3	3	1	1	1	.	13.22	4.80	4.82	0.44	0.95	0.77	20.17	0.89	0.30
NC196A	.	.	.	.	.	.	3	3	.	.	.	.	.	.	.	.	.
NC222	1	2	3	2	2	1	1	.	.	.	.	.	.	.	.	.	.
NC230	3	2	3	.	.	.	?3	1	4.68	7.85	0.88	1.07	0.18	1.76	15.54	1.10	0.34
NC232	3	2	1	4	2	1	1	1	2.92	4.11	0.79	0.58	0.09	1.35	9.04	0.54	0.19
NC236	3	2	1	2	2	1	1	3	3.60	4.08	1.18	0.46	0.20	1.79	10.14	1.29	0.68
NC238	.	2	3	2	2	1	1	1	.	.	.	.	.	.	.	.	.
NC250	.	2	1	2	2	1	1	.	6.29	2.27	0.45	0.27	0.14	0.48	9.46	0.15	0.08
NC258	3	2	3	2	2	1	_	.	6.73	4.61	0.52	0.25	0.06	0.30	11.95	6.90	0.82
NC260	3	2	3	2	2	1	.	1	3.35	5.09	0.63	0.77	0.07	0.86	10.14	7.02	0.81
NC262	3	2	3	2	2	2	1	1	3.77	7.49	0.69	0.70	0.90	0.07	12.92	4.97	0.74
NC264	3	2	3	3	1	1	1	.	8.58	3.93	0.54	0.29	0.15	0.62	13.57	0.32	0.14
NC268	.	.	.	.	.	.	1	.	7.24	6.94	1.92	2.20	1.53	0.24	18.16	0.60	0.19
NC290A	3	2	3	.	.	.	1	1	4.60	8.46	0.75	1.28	1.64	0.14	16.13	8.14	1.08
NC292	.	.	.	.	.	.	1	.	6.91	4.07	2.76	1.18	0.66	0.56	13.39	0.91	0.19

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
NC294	3	2	1	3	1	1	1	1	.	.	.	.	.	.	.	.	.
NC296	3	2	3	4	2	2	3	3	.	.	.	.	.	.	.	.	.
NC296A	.	2	3	4	2	2	.	.	.	.	.	.	.	.	.	.	.
NC298	3	2	3	2	2	2	1	1	9.93	13.98	1.27	1.74	1.80	0.13	27.58	0.03	0.05
NC300	3	2	3	2	2	2	1	3	0.50	0.60	1.24	1.44	0.25	0.96	3.73	1.04	0.26
NC302	.	2	3	2	2	2	.	.	3.25	10.31	0.87	1.56	1.05	0.11	16.27	0.34	0.07
NC304	3	2	3	4	2	1	1	.	4.47	8.40	0.88	1.46	0.14	1.48	15.95	0.76	0.25
NC306	3	2	3	2	2	1	1	1	5.55	6.05	1.31	1.98	1.41	0.21	15.20	0.57	0.15
NC308	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.
NC310	3	2	3	3	1	1	1	.	2.14	1.34	0.52	0.27	0.13	0.08	3.96	0.08	0.04
NC312	.	.	.	.	.	.	1	.	3.33	3.32	0.68	0.48	0.08	0.29	7.50	0.11	0.05
NC314	3	2	3	2	2	1	1	1	4.41	4.37	0.65	0.68	0.09	0.36	9.91	0.31	0.10
NC316	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.
NC318	3	2	3	2	.	2	1	.	.	.	.	.	.	.	.	.	.
NC320	3	2	3	4	2	2	1	.	2.16	5.91	0.53	0.42	0.12	0.71	9.30	0.43	0.27
NC322	.	.	.	.	.	.	1	3	2.98	2.78	0.44	0.25	0.09	0.24	6.34	0.10	0.03
NC324	3	2	3	2	2	1	1	1	4.16	3.91	0.52	0.42	0.11	0.31	8.91	0.30	0.10
NC326	3	2	3	3	1	1	1	1	6.01	3.06	1.50	0.72	0.44	0.43	10.67	0.51	0.10
NC328	3	2	1	3	1	1	1	1	2.79	0.83	0.44	0.24	0.37	0.10	4.33	0.13	0.05
NC33	3	2	2	1	2	1	3	1	.	.	.	.	.	.	.	.	.
NC330	.	.	.	.	.	.	1	.	6.15	2.73	1.90	0.71	0.44	0.45	10.47	0.97	0.16
NC332	.	.	.	.	.	.	1	3	3.12	3.23	0.52	0.26	0.07	0.29	6.96	0.09	0.05
NC334	.	.	.	.	.	.	1	.	3.40	3.55	0.56	0.29	0.07	0.27	7.58	0.09	0.04
NC336	3	2	3	4	2	2	3	3	.	.	.	.	.	.	.	.	.
NC338	.	2	3	.	.	.	1	3	1.20	0.30	0.91	0.41	0.29	0.44	2.63	0.60	0.17
NC340	3	2	3	2	2	2	3	3	.	.	.	.	.	.	.	.	.
NC342	3	2	3	2	2	1	1	1	6.36	5.32	0.73	0.66	0.80	0.13	13.26	4.23	0.91
NC344	3	2	3	.	2	1	1	.	.	.	.	.	.	.	.	.	.

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
NC346	3	2	3	4	2	2	3	3	.	.	.	.	.	.	.	.	.
NC348	3	2	3	2	2	2	3	1	.	.	.	.	.	.	.	.	.
NC350	3	2	3	2	2	2	1	1	1.26	0.67	0.86	2.00	0.35	1.93	6.21	2.23	0.67
NC352	3	2	3	4	2	2	3	3	.	.	.	.	.	.	.	.	.
NC354	3	2	3	4	2	1	1	3	3.93	9.13	1.95	2.29	0.28	2.75	18.38	0.61	0.13
NC356	.	2	3	4	2	2	.	.	7.72	22.45	3.83	2.30	0.22	2.45	35.14	0.54	0.19
NC358	3	2	3	3	1	1	1	1	8.08	5.92	9.12	0.72	0.91	0.57	16.21	1.96	0.55
NC360	3	2	2	2	2	2	1	1	4.34	9.40	1.12	0.66	0.14	0.62	15.16	0.69	0.28
NC362	3	2	3	2	2	2	1	1	6.62	5.32	0.84	0.94	1.05	0.20	14.13	3.20	0.62
NC364	3	2	3	2	2	2	1	1	5.63	5.05	0.83	0.95	1.15	0.22	12.98	2.87	0.54
NC366	3	2	2	2	2	1	1	.	.	.	.	.	.	.	.	.	.
NC368	.	2	3	3	1	1	.	1	.	.	.	.	.	.	.	.	.
NC370	.	.	.	.	.	.	1	.	2.90	6.54	0.66	0.45	0.14	0.83	10.85	0.59	0.36
NC372	.	.	.	.	.	.	3	.	7.56	2.79	2.05	0.78	0.50	0.58	12.22	1.10	0.20
ND246	3	2	3	3	1	1	1	1	9.16	2.56	0.57	0.25	0.65	0.21	12.82	0.33	0.13
OH40B	1	2	3	3	1	1	1	.	18.66	2.65	1.46	0.44	2.45	0.48	24.69	1.28	0.35
OH43	3	2	1	3	1	1	1	.	12.75	3.83	1.06	0.56	2.62	0.31	20.07	0.66	0.27
OH43E	3	2	1	2	2	2	.	.	5.32	5.22	0.79	1.20	3.24	0.20	15.19	2.00	0.62
OH603	.	2	2	3	1	1	.	1	.	.	.	.	.	.	.	.	.
OH7B	3	2	3	3	1	1	1	.	10.80	7.27	3.99	0.83	0.42	1.67	20.98	3.49	0.73
OS420	3	2	1	2	2	1	1	.	7.28	4.40	1.57	0.68	1.76	0.47	14.59	2.97	0.47
P39	.	2	2	3	1	1	1	2	7.03	7.24	2.00	0.90	0.71	0.34	16.20	0.84	0.17
PA762	3	2	1	3	1	1	.	.	6.09	3.99	5.69	0.43	0.49	1.99	12.99	5.51	1.32
PA875	3	2	1	3	1	1	1	1	7.40	2.10	0.72	0.24	0.97	0.19	10.90	0.68	0.21
PA880	3	2	3	3	1	1	3	1	8.95	4.11	0.64	0.30	0.30	0.13	13.79	2.41	0.75
PA91	.	2	3	4	2	2	1	1	8.42	8.78	1.23	1.35	0.49	1.12	20.16	1.09	0.25
R109B	3	2	3	3	1	1	1	.	.	.	.	.	.	.	.	.	.
R168	3	2	3	.	.	.	1	.	6.99	3.54	0.80	0.41	0.34	0.15	11.44	0.43	0.19

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
R177	3	2	2	3	1	1	1	1	11.57	3.58	1.13	1.01	0.79	0.42	17.37	0.28	0.16
R229	3	2	3	3	1	1	1	1	8.71	4.32	4.27	0.97	0.84	0.91	15.75	0.65	0.25
R4	.	2	3	3	1	1	.	2	.	.	.	.	.	.	.	.	.
SA24	3	2	3	2	2	1	1	.	6.81	6.98	0.94	1.13	0.14	1.38	16.43	0.19	0.09
SC213R	1	2	3	.	.	.	1	1	.	.	.	.	.	.	.	.	.
SC357	3	2	3	2	2	1	1	3	6.19	6.49	1.41	0.67	0.15	1.03	14.53	0.21	0.07
SC55	2	1	1	2	2	1	2	3	0.38	0.29	0.90	0.28	0.39	3.09	4.43	0.39	0.16
SD40	3	2	3	2	2	.	1	1	7.64	7.23	0.72	0.87	1.23	0.12	17.08	4.63	0.83
SD44	3	2	1	3	1	1	1	1	6.98	2.66	0.51	0.35	1.30	0.17	11.47	0.86	0.30
SG1533	3	2	3	2	2	1	1	.	4.32	6.88	0.82	0.96	0.66	0.22	13.05	1.32	0.55
SG18	3	2	2	2	2	1	1	.	3.99	13.96	1.21	2.51	0.28	1.41	22.15	1.11	0.52
T232	3	2	3	2	2	2	1	.	2.84	6.90	0.93	1.94	0.18	0.81	12.67	0.80	0.17
T234	3	1	1	3	1	1	1	.	.	.	.	.	.	.	.	.	.
T8	3	2	3	2	2	1	1	.	4.25	5.95	0.87	0.63	0.11	0.64	11.58	0.03	0.04
TX303	3	2	1	3	1	1	1	3	.	.	.	.	.	.	.	.	.
TX601	3	2	3	2	2	2	_	3	.	.	.	.	.	.	.	.	.
TZI10	1	2	3	2	2	2	3	.	.	.	.	.	.	.	.	.	.
TZI11	3	2	1	2	2	1	3	3	.	.	.	.	.	.	.	.	.
TZI16	3	2	3	4	2	1	3	3	.	.	.	.	.	.	.	.	.
TZI18	3	2	3	4	2	2	1	.	.	.	.	.	.	.	.	.	.
TZI25	.	2	3	1	1	1	1	1	.	.	.	.	.	.	.	.	.
TZI8	.	2	3	2	2	1	3	3	.	.	.	.	.	.	.	.	.
TZI9	3	2	3	2	2	2	3	.	.	.	.	.	.	.	.	.	.
VA102	3	2	2	2	2	1	1	1	7.56	10.63	4.01	1.86	1.87	0.41	22.33	1.12	0.30
VA14	3	2	3	3	1	1	2	.	9.79	2.80	0.61	0.36	0.50	0.19	13.65	0.03	0.05
VA17	3	2	3	3	1	1	1	.	8.70	4.06	0.72	0.57	0.53	0.16	14.03	0.03	0.04
VA22	3	2	3	3	1	1	1	.	4.51	3.05	0.55	0.44	0.12	0.29	8.42	0.09	0.04
VA26	3	2	1	3	1	1	1	.	6.75	4.02	0.94	0.95	2.01	0.23	13.96	1.05	0.35

Appendix A:, continued

	--------crtRB1--------			--------lcyE--------			psy1	CCD1	----------Carotenoid Traits (ug g^{-1})----------								
Entry	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	InDel388	5p	Lut	Zea	Zeino	βcry	αcar	βcar	Total	Phyen	Phyflu
VA35	3	2	3	2	2	1	1	.	7.85	8.27	1.17	1.27	1.21	0.19	18.78	0.34	0.16
VA59	3	2	2	2	2	1	1	1	6.64	6.74	1.13	0.77	0.45	0.17	14.76	0.13	0.09
VA85	3	2	3	3	1	1	1	.	10.33	1.98	5.87	1.22	1.38	0.65	15.55	4.20	1.45
VA99	3	2	3	.	.	.	1	.	4.63	9.37	0.88	1.24	0.13	1.89	17.27	1.40	0.52
VAW6	2	1	1	2	2	1	1	.	3.76	0.85	0.42	0.20	1.39	0.13	6.33	0.37	0.10
W117HT	3	2	1	2	1	1	1	.	4.38	2.74	0.60	0.53	1.21	0.15	9.01	2.12	0.29
W153R	3	2	1	3	1	1	1	.	10.79	3.90	1.87	0.57	2.21	0.40	17.87	1.59	0.48
W182B	3	2	3	2	2	1	1	.	8.80	6.75	1.67	1.02	0.67	0.12	17.35	1.11	0.38
W22	3	2	3	.	.	.	1	.	4.96	1.67	1.09	0.23	0.24	0.11	7.21	0.60	0.18
W22 RR:STD	.	2	2	3	1	1	1	.	4.57	0.70	0.66	0.38	0.66	0.63	6.93	0.41	0.12
W64A	3	2	3	3	1	1	1	.	6.30	2.40	1.84	0.28	0.34	0.16	9.47	0.44	0.24
WD	3	2	3	2	2	1	1	1	6.49	9.70	3.41	2.57	1.89	0.53	21.18	2.90	0.76
WF9	3	2	3	3	1	1	1	1	9.44	3.60	0.99	0.72	0.73	0.25	14.74	1.26	0.56
YU796_NS	3	1	1	3	1	1	1	.	6.59	1.38	0.77	0.39	1.17	0.24	9.76	0.36	0.10

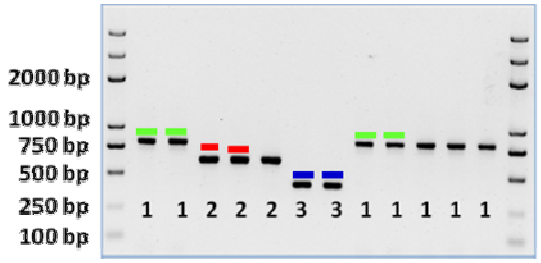
Carotenoid QTL genotypes are listed for beta-carotene hydroxylase (crtRB1), lycopene epsilon cyclase (lcyE), phytoene synthase 1 (psy1) and carotenoid cleavage dioxygenase 1 (ccd1). Traits are lutein, zeaxanthin, zeinoxanthin, β-cryptoxanthin, α-carotene, β-carotene, total carotenoids, phytoene and phytofluene. Trait measurements were taken from a single bulk of seed..

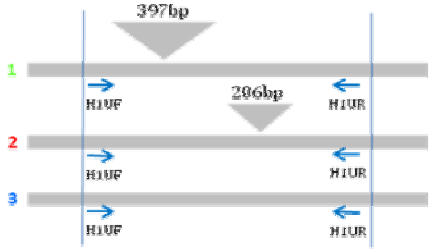
Polymorphisms are listed below QTL, and allele codes are as follows (favorable defined in methods):

	crtRB1-5'TE	crtRB1-InDel4	crtRB1-3'TE	lcyE-5'TE	lcye-SNP216	lcye-3'TE	psy1-InDel388	ccd1-5p
Allele codes	1=397 bp ins 2=200 bp ins 3=0 bp ins	1=12 bp ins 2=0 bp ins	1=0 bp ins 2=325 bp ins 3=1250 bp ins	1=150+280 bp 2=250 bp 3=250+380 bp 4=993 bp	1=G 2=T	1=0 bp del 2=8 bp del	1=ins 2=het 3=del	b73=small ins WC=large ins, multicopy
Favorable Alleles	2	1	1	1,4	G	2	1	b73

Appendix B: Allele-specific marker assays for *crtRB1* and *lcyε* polymorphisms

crtRB1 5'TE PCR assay

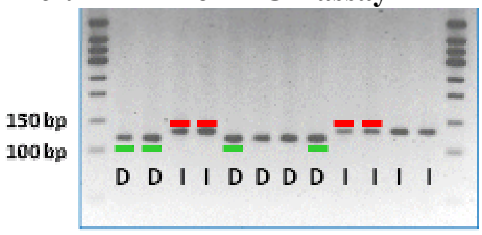


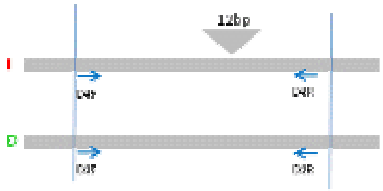


Allele hierarchy according to β-carotene concentrations: 2>3>1
1: 800 bp (with 397bp insertion)
2: 600 bp (with 206bp insertion)
3: 400 bp (without insertion)

crtRB1 H1UF TTAGAGCCTCGACCCTCTGTG
crtRB1 H1UR AATCCCTTTCCATGTTACGC

crtRB1 InDel4 PCR assay

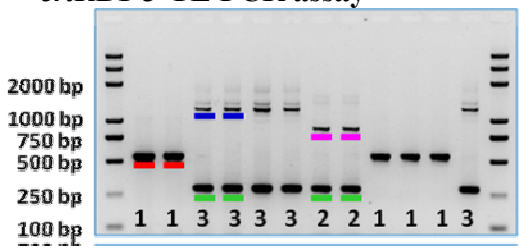


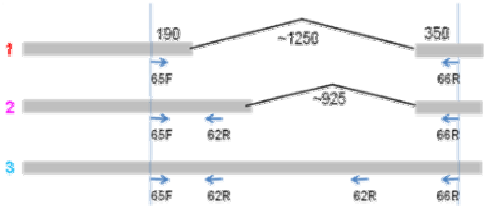


Allele hierarchy according to β-carotene concentrations: I>D
I: 129 bp (with 12 bp insertion at InDel4)
D: 117 bp (with 12 bp deletion at InDel4)

Note: highly rare and non-functional 6bp insertion at InDel3 may interfere with banding pattern
crtRB1 D4F ACCGTCACGTGCTTCGTGCC
crtRB1 D4R CTTCCGCGCCTCCTTCTC

crtRB1 3'TE PCR assay



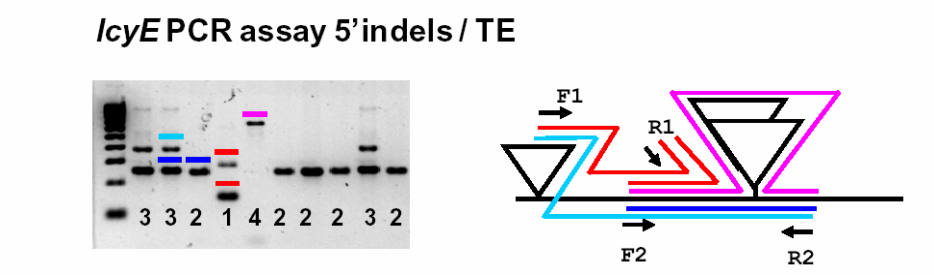


Allele hierarchy according to β-carotene concentrations: 1>3>2
1: 543 bp (with 1250bp deletion)
2: 296+875 bp (with 925bp deletion)
3: 296+1221+1800 bp (without deletion)

Note: the largest fragment (1800 bp) by primers 65F and 66R was usually weak or not amplified.
crtRB1 65F ACACCACATGGACAAGTTCG
crtRB1 62R ACACTCTGGCCCATGAACAC
crtRB1 66R ACAGCAATACAGGGGACCAG

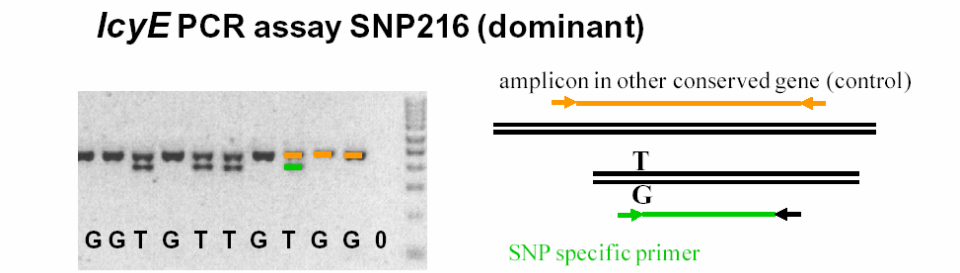
(Assay schema published in Yan and Kandianis [29])

Appendix B:, continued



Allele hierarchy according to β-carotene concentrations: 4=1>2>3
1: 150 + 280 bp; 2: 250 bp; 3: 250 + 280 bp; 4 993 bp

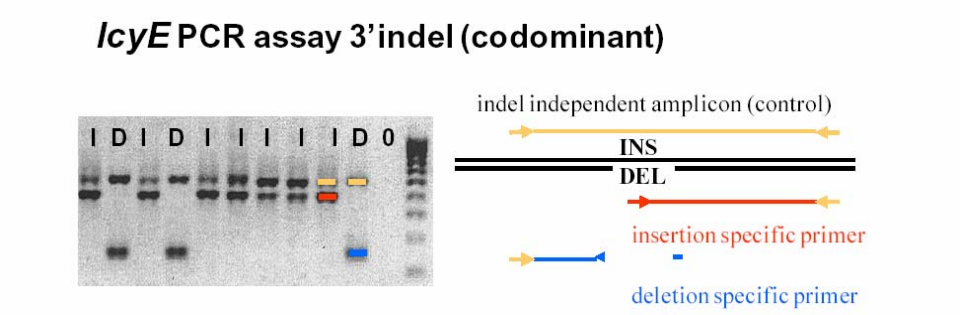
LYCe-TE103PF-F1	CGC TAG CAA GCC CAT TAT TTT TA
LYCe-TE103PF-R1	CGG TAT GGT TTT TGG TAT ACG G
LYCe-ZGt111R1-F1	AAG CAT CCG ACC AAA ATA ACA G
LYCe-TE105PR-R1	GAG AGG GAG ACG ACG AGA CAC



Allele hierarchy according to β-carotene concentrations: G>T
G: absence of band; T/C: presence of band

LYCe-SNP216-F1	GCG GCA GTG GGC GTG GAT
LYCe-SNP216-R1	TGA AGT ACG GCT GCA GGA CAA CG
umc1178-F	CTG TCG TAA GAG CGC CAA CAG
umc1178-R	GTC TGA ACG ATG AAC AGT ACA CGC

Note: umc primer set is used as the control amplicon – any other compatible control can be used



Allele hierarchy according to β-carotene concentrations: D>I
D: 144 bp; I: 399 bp

LYCe-3'indl-F1	GTA CGT CGT TCA TCT CCC GTA CCC
LYCe-3'indl-R1	CTT GGT GAA CGC ATT TCT GTT GG
LYCe-3'indl-F2	GGA CCG GAA CAG CCA ACT G
LYCe-3'indl-R2	GGC GAA ATG GGT ACG GCC

(Assay schema published in Harjes [19])

Appendix C: Temperature cycling conditions for allele-specific assays

TD65-54

step	temp	time (min)
1	94	5
2	94	1
3	64	1
4	reduce temp by 0.5C/cyc	
5	72	1
6	repeat from step 2 19x	
7	95	1
8	54	1
9	72	1.5
10	repeat from step 7 19x	
11	72	10
12	4	forever

Assays: lcyE-5'TE, lcyE-3'TE

FL_NT58Touchdown

step	temp	time (min)
1	94	3
2	94	1
3	64	1
4	reduce temp by 1C/cyc	
5	72	1.5
6	goto step 2 5x	
7	94	1
8	58	1
9	72	1.5
10	goto step 7 28x	
11	72	15

Assay: lcyE-SNP216

A60

step	temp	time (min)
1	94	3
2	94	0.66
3	60	0.66
4	72	1
5	repeat from step 2 39x	
6	72	10
7	4	forever

Assay: ccd1-5p

TD64-54

step	temp	time (min)
1	95	5
2	94	1
3	64	1
4	reduce temp by 1C/cyc	
5	72	1
6	repeat from step 2 9x	
7	95	1
8	54	1
9	72	1.5
10	repeat from step 7 26x	
11	72	10
12	4	forever

Assay: crtRB1-5'TE

TD68-58

step	temp	time (min)
1	95	3
2	94	1
3	68	1
4	reduce temp by 1C/cyc	
5	72	1.5
6	repeat from step 2 10x	
7	94	1
8	58	1
9	72	1.5
10	repeat from step 7 30x	
11	72	10
12	4	forever

Assay: crtRB1-InDel4

PZB57

step	temp	time (min)
1	95	5
2	95	1
3	57	1
4	72	4
5	repeat from step 2 34 x	
6	72	10
7	4	forever

Assay: crtRB1-3'TE

Appendix C:, continued

UIUC-M

step	temp	time (min)
1	92	0.5
2	68	0.75
3	reduce temp by 1C/cyc	
4	72	0.75
5	repeat from step 1 10x	
6	92	0.5
7	58	0.75
8	72	0.75
9	repeat from step 6 25x	
10	72	5

Assay: BMC markers

UIUC-H

step	temp	time (min)
1	93	0.5
2	70	0.75
3	reduce temp by 1C/cyc	
4	72	0.75
5	repeat from step 1 10x	
6	92	0.5
7	60	0.75
8	72	0.75
9	repeat from step 6 25x	

Assay: UMC, BNLG, ZCT markers

MMC

step	temp	time (min)
1	94	0.5
2	65	1
3	reduce temp by 1C/cyc	
4	72	1.5
5	repeat from step 1 10x	
6	92	0.5
7	55	1
8	72	1.5
9	repeat from step 6 30x	

Assay: MMC markers

Appendix D: Carotenoid concentrations by entry mean for hybrids and inbreds grown in Mexico and Illinois during 2008

		Puerto Vallarta, Mexico (winter 2007-2008)						Urbana, Illinois (summer 2008)					
Gen.	Entry	N	Lut	Zea	βcry	βcar	Total	N	Lut	Zea	βcry	βcar	Total
Hybrids	A619xDE3	4	14.16	4.55	1.54	6.21	26.46	3	10.46	4.58	0.96	3.01	19.89
	A619xSC55	8	5.89	2.44	1.28	6.53	16.14	3	2.98	0.39	0.19	2.74	7.30
	B77xA272	4	11.25	15.40	5.83	9.78	42.26	3	2.77	21.38	3.43	2.71	30.44
	B77xHI27	4	5.96	9.43	3.85	5.20	24.43	3	2.11	13.55	3.14	2.45	21.29
	CI7xA619		-	-	-	-	-	3	9.07	4.30	0.58	2.79	17.18
	CI7xSC55	4	5.61	2.38	1.02	10.51	19.51		-	-	-	-	-
	CI7xDE3		-	-	-	-	-	3	11.19	2.32	0.06	5.46	19.46
	DE3xA619		-	-	-	-	-	3	12.20	4.34	0.84	2.62	21.02
	DE3xCI7	4	17.61	3.09	1.32	15.46	37.48		-	-	-	-	-
	DE3xSC55	4	7.78	2.43	3.45	9.58	23.23	3	3.98	0.42	0.03	3.01	8.56
	HI27xA272	4	9.28	17.60	8.83	5.03	40.74	3	2.67	29.75	7.10	3.19	42.79
	HI27xCML328	4	9.25	13.97	8.48	4.49	36.18	3	4.27	20.01	7.26	1.34	33.49
	KUI3xB77	4	8.98	9.98	4.09	4.69	27.73	3	5.33	13.29	2.32	1.11	22.59
	KUI3xSC55	4	4.27	3.20	2.03	3.91	13.41	3	2.46	2.56	1.34	1.50	8.86
	KUI43xB77	4	17.84	8.85	2.42	4.31	33.41	3	11.45	10.24	1.35	1.89	26.92
	NC354xHI27	4	9.36	11.64	5.71	3.20	29.90	3	4.78	16.64	4.81	1.58	29.34
	SC55xB77	4	4.38	2.30	1.12	6.99	14.79	3	2.63	0.66	0.08	2.98	7.01
	SC55xDE3	4	7.43	2.29	0.93	9.19	19.83	3	5.05	0.48	0.03	4.23	11.08
Inbreds	A272	4	22.31	24.29	9.76	8.92	65.27	3	3.93	30.40	7.79	3.53	47.34
	A619	4	21.66	7.16	1.80	5.60	36.22	3	10.77	8.11	1.53	2.90	24.06
	B77	6	8.48	4.86	1.60	10.05	24.98	3	4.04	3.23	0.28	1.89	9.88
	CI7	4	15.63	3.59	1.22	11.86	32.29	3	4.35	1.11	0.04	3.12	8.78
	CML328	4	10.38	13.49	9.32	6.04	39.22	3	1.50	11.42	5.99	1.86	21.31
	DE3	2	16.67	3.23	1.42	23.50	44.81	3	10.65	1.67	0.00	6.46	19.16
	HI27	4	5.52	11.47	9.33	1.13	27.45	3	1.03	14.99	7.28	1.42	25.50
	KUI3	4	7.08	12.73	4.73	1.20	25.73	6	1.54	17.70	4.22	0.94	24.65
	KUI43	2	20.98	10.82	2.48	2.75	37.02	3	10.68	14.52	1.90	1.29	32.14
	NC354	2	9.20	9.05	3.44	2.30	23.97	3	4.07	5.61	3.25	1.57	17.54
	SC55		-	-	-	-	-	3	0.84	0.08	0.21	2.68	4.99

Appendix E: Midparent heterosis (MPH) observed for carotenoid concentrations across Puerto Vallarta, Mexico winter 2008-09 and Urbana, IL, 2008

	Puerto Vallarta, MX Environment					Urbana, IL Environment			
Entry	Lut	Zea	βcry	βcar		Lut	Zea	βcry	βcar
A619xDE3	-5.01	-0.65	-0.06	-8.34		-0.25	-0.31	0.20	-1.68
A619xSC55[a]	-	-	-	-		-2.83	-3.71	-0.68	-0.05
B77xA272	-4.14	**0.82**	0.15	0.30		-1.21	**4.57**	-0.61	0.00
B77xHI27	-1.04	**1.27**	-1.62	-0.39		-0.43	**4.44**	-0.64	0.79
CI7xA619	-	-	-	-		1.51	-0.31	-0.21	-0.22
CI7xSC55[a]	-	-	-	-		-2.60	-0.59	-0.13	-2.90
CI7xDE3	-	-	-	-		3.69	0.93	0.03	0.67
DE3xA619	-	-	-	-		1.49	-0.54	0.08	-2.06
DE3xCI7	1.46	-0.32	0.01	-2.22		-	-	-	-
DE3xSC55[a]	-	-	-	-		-1.76	-0.45	-0.07	-1.56
HI27xA272	-4.63	-0.29	-0.71	**0.01**		0.19	7.06	-0.44	**0.72**
HI27xCML328	**1.30**	**1.49**	-0.84	0.91		**3.00**	**6.81**	0.62	-0.30
KUI3xB77	**1.20**	**1.19**	**0.92**	-0.93		**2.54**	**2.82**	**0.07**	-0.31
KUI3xSC55[a]	-	-	-	-		1.27	-6.34	-0.87	-0.30
KUI43xB77	**3.11**	**1.01**	**0.38**	-2.09		**4.09**	**1.36**	**0.25**	0.30
NC354xHI27	**2.01**	**1.38**	-0.67	**1.49**		**2.23**	**6.34**	-0.46	**0.09**
SC55xB77[a]	-	-	-	-		0.19	-0.99	-0.16	0.70
SC55xDE3[a]	-	-	-	-		-0.70	-0.39	-0.08	-0.34

MPH derived from: $F_1 - (P_1+P_2)/2$
Positive MPH in both environments is coded by gray box; positive MPH in a single environment is coded by green box
a: MPH not calculated in Puerto Vallarta environment since SC55 data was not available

Appendix F: Endosperm Dosage Effects on Carotenoid Concentration

Problem:

Allelic variation has been identified within several genes encoding enzymes within the carotenoid biosynthesis pathway. These alleles have a pronounced effect on carotenoid concentrations in inbred lines. Information based on haplotype effects may not give much indication of intragenic interactions (dominance effects) and intergeneic interactions (epistasis effects) that could heavily account for phenotypic differences in segregating populations. As one of the uses of haplotype information is to selectively use the favorable genetic variation in breeding germplasm, it is necessary to understand how haplotypes function in combination.

Objectives:

1. To determine if the effect of endosperm dosage is significant on carotenoid concentrations in per se and reciprocally crossed seed

2. To associate any detected dosage effects with known allelic variation at QTL known to affect carotenoid concentrations, specifically lycopene epsilon cyclase (*lcyε*), beta carotene hydroxylase (*crtRB1*) and carotenoid cleavage dioxygenase 1 (*ccd1*)

3. To evaluate detected effects in the context of additive and dominance gene action

Methods:

F_1 seed for the dosage study was produced during summer 2008, during which inbred lines were selected to make reciprocal crosses and self pollinations to create a complete array of haplotype dose in triploid endosperm. Reciprocal crosses and parental self pollinations were evaluated to determine the effect of haplotype copies, or doses, on the carotenoid profile of the endosperm tissue. Parent-of-origin effects can be measured in endosperm of F_1 seed since the female contribution of two haploid chromosomes and the male contribution of one haploid chromosome is asymmetrical. Inbreds from the core set of carotenoid lines (Table 3.1) were used to create seven experimental sets consisting of reciprocal F_1 and self pollinated seed. Inbred carotenoid QTL genotypes for *crtRB1*, *lcyε* and *ccd1* identified in previous studies [19, 29]

are listed in Table 3.3 and Appendix A. Using this information, allelic contrasts between inbreds were pre-determined; accordingly, these contrasts would segregate in a dose-dependent manner among reciprocal crosses (Table G-1).

Carotenoid concentrations for lutein, zeaxanthin, β-cryptoxanthin, β-carotene and the summation of these carotenoids (defined as Total) were measured for triplicate biological samples on a dry weight basis using Tian's modified protocol as indicated in Chapter 3 Methods. Metabolic data was analyzed with SAS software [48]. Descriptive statistics were obtained through Proc MEANS. Multiple comparisons for dose classes were drawn using Tukey's t-tests in Proc GLM with an experiment-wise error rate of α_{EWE}=0.05, and Honestly Significant Differences ($HSD_{0.05}$) are reported. Normality, heterogeneity of variances and correlated errors for model residuals were evaluated using Proc UNIVARIATE and graphical techniques. Models of carotenoid concentrations by dose were performed in Proc GLM by one way ANOVA; dose was measured as 0, 1, 2, or 3 copies of the female parent, leading to four possible genotype classes.

Results:

Seven experimental sets consisting of self pollinated inbred parents and reciprocal hybrids were used to evaluate the combinatorial effect of differing haplotype doses. Segregation at polymorphisms for *crtRB1*, *lcyε* and *ccd1* was known a priori and is outlined in Table F-1. Haplotype dose was highly significant (α=0.05) in explaining variation for six carotenoid traits across most experimental sets (Table F-2 indicating that genetic effects contributing to carotenoid phenotypes differed between the parents. The direction of these effects was evaluated through linear contrasts which were found to be largely significant for most carotenoid traits. This indicated that haplotype contributions to most carotenoid traits were highly additive. Tests of quadratic contrasts showed some crosses (sets 1, 3 and 4) exhibited dominance. β-carotene was found to be least affected by haplotype dosage, suggesting either that QTL affecting this trait did not vary between the parents, or that QTL affecting this trait were in repulsion. Considering that several of these crosses were segregating at QTL known to affect β-carotene, it was presumed that dosage effects would have been significant for this trait. As this was not the case, it is possible that an antagonistic effect between genetic backgrounds precluded observation of an effect on β-carotene. Interestingly, even in a cross segregating at only one known polymorphism (Set 6, *lcyε*-5'TE), several traits were found to be significantly affected.

Variation among doses for lutein occurred in Sets 1 and 7, which coincides with segregation at *ccd1*-5p. A dominance effect for the *ccd1*-Wc allele was observed in both sets. Classes with one, two or three doses of the dominant allele through SC55 parentage did not significantly differ from one another for lutein, and had significantly reduced lutein concentrations from that of the other parent, which had 0 doses. Lutein was found to be highest in the reciprocal crosses for sets with common B77 parentage (Sets 2, 3, 5), similar to that in the reciprocal crosses for DEexp and CI7 (Set 4). Similar reductions in zeaxanthin and β-cryptoxanthin were observed with SC55 contribution.

Effect of *lcyε* polymorphisms on α/β branch ratio were significant in almost all crosses and were largely found to follow haplotype predictions. Notably, in two sets segregating only at the 3'TE allele for lcyε (Sets 1 and 2), α/β branch ratios were found to be higher with the weaker, or less efficient, allele, the opposite of what was expected. A619 (Set 1), predicted to have the more favorable, less efficient allele at *lcyε*-5'TE, was observed to yield a higher α/β branch ratio (Table 3.4). In this cross, the ratio is likely affected by *ccd1*effects in SC55 due to a negative epistatic interaction between *ccd1* (Wc allele) and *lcyε* (allele 4, more efficient). This allelic combination is associated with less lutein, and thus a smaller α/β ratio.

Dose dependent increases in zeaxanthin were positively correlated with β-cryptoxanthin in all experimental sets suggesting that both compounds are regulated by similar controls. Zeaxanthin and β-cryptoxanthin concentrations were most variable in sets 2, 3, and 5 in which B77 contribution was associated with a decrease in these β-branch carotenoids. Comparison with the predicted effect of B77 haplotype points to reduced conversion of β-carotene through *crtRB1* polymorphisms. Interestingly, an increase in favorable *crtRB1* haplotype doses reduced zeaxanthin and β-cryptoxanthin but did not augment β-carotene concentrations in the reciprocal hybrids. A substantial reduction in both zeaxanthin and β-cryptoxanthin concentrations was associated with the presence of the SC55 haplotype in Sets 1 and 7, which coincided with the overexpression form of *ccd1*-5p.

No complementation in haplotypes was found to increase β-carotene through dosage, as observed in Table F-2. Given that transgressive segregation has been observed in $F_{2:3}$ progeny from A619 x SC55 and DEexp x CI7 mapping populations, these results suggest that parental haplotype blocks need to be be further recombined to achieve β-carotene increases.

Table F.1: Crosses generated for haplotype dosage comparisons in Urbana, Illinois in 2008

			Polymorphisms Segregating in Cross						
Experiment	Parent 1	Parent 2	crtRB1 5'TE	crtRB1 InDel4	crtRB1 3'TE	lcyE 5'TE	lcyE SNP216	lcyE 3'TE	ccd 5p
1	A619	SC55	**x**	**x**	-	-	-	**x**	**x**
			1,2	**2,1**	1,1	2,2	2,2	**2,1**	**b,W**
2	B77	HI27	**x**	**x**	**x**	**-**	-	**x**	**-**
			2,1	**1,2**	**1,3**	2,2	2,2	**2,1**	b,b
3	B77	KUI3	**x**	**x**	**x**	**x**	-	-	-
			2,1	**1,2**	**1,3**	**2,4**	2,2	2,2	b,b
4	DE3	CI7	-	**x**	-	**x**	**x**	-	-
			2,2	2,1	1,1	**3,2**	**1,2**	1,1	b,b
5	A272	B77	**x**	**x**	**x**	na	na	na	-
			1,2	**2,1**	**3,1**	-,2	-,2	-,2	b,b
6	CML328	HI27	-	-	-	**x**	-	-	-
			-,1	2,2	3,3	**1,2**	2,2	1,1	b,b
7	KUI3	SC55	**x**	**x**	**x**	**x**	-	**x**	**x**
			1,2	**2,1**	**3,1**	**4,2**	2,2	**2,1**	**b,W**

Allele code	1/**2**/3	**1**/2	**1**/2/3	**1**/2/3/**4**	**1**/2	1/**2**	WC/**b73**

Polymorphism in a cross is denoted by '**x**'.
Lack of genotype information is denoted by 'na'.

Table F.2: Endoperm carotenoid concentrations for reciprocal crosses of selected high carotenoid lines grown in Urbana, Illinois in 2008

Set	Entry	Lut	Zea	βcry	βcar	Total	α/β Ratio
1	A619	10.77	8.11	1.53	2.90	24.06	0.92
	A619xSC55	0.43	0.04	0.27	1.29	2.91	0.81
	SC55xA619	0.10	0.03	0.23	1.69	2.51	0.34
	SC55	0.84	0.08	0.21	2.68	4.99	0.45
	$HSD_{(0.05)}$	0.71	0.82	0.54	0.94	3.06	0.29
	ANOVA (p)	**	**	**	**	**	**
	Linear Effect	**	**	**	ns	**	**
	Quadratic Effect	**	**	**	**	**	ns
2	B77	4.04	3.23	0.28	1.89	9.88	0.79
	B77xHI27	2.70	9.62	1.66	2.27	16.55	0.22
	HI27xB77	3.52	14.73	4.92	1.66	25.59	0.2
	Hi27	1.03	14.99	7.28	1.42	25.50	0.08
	$HSD_{(0.05)}$	1.41	2.99	1.07	1.10	6.12	0.1
	ANOVA (p)	**	**	**	ns	**	**
	Linear Effect	**	**	**	ns	**	**
	Quadratic Effect	ns	**	ns	ns	*	**
3	B77	4.04	3.23	0.28	1.89	9.88	0.79
	B77xKUI3	6.89	14.77	2.11	0.83	25.17	0.42
	KUI3xB77	2.94	15.46	3.79	1.40	24.14	0.17
	KUI3	1.54	17.70	4.22	0.94	24.65	0.08
	$HSD_{(0.05)}$	1.38	3.70	1.54	1.11	6.90	0.11
	ANOVA (p)	**	**	**	ns	**	**
	Linear Effect	**	**	**	ns	**	**
	Quadratic Effect	**	**	*	ns	**	**
4	CI7	4.35	1.11	0.04	3.12	8.78	1.07
	CI7xDEexp	10.42	1.19	0.14	2.31	14.83	3.07
	DEexpxCI7	21.01	2.01	0.23	2.35	27.30	4.89
	DEexp	10.65	1.67	0.00	6.46	19.16	1.57
	$HSD_{(0.05)}$	5.85	0.71	0.06	5.29	7.69	1.61
	ANOVA (p)	**	**	**	0.1055	**	**
	Linear Effect	**	*	ns	ns	**	*
	Quadratic Effect	**	ns	**	ns	**	**

Table F.2:, continued

Set	Entry	Lut	Zea	βcry	βcar	Total	α/β Ratio
5	A272	3.93	30.40	7.79	3.53	47.34	0.79
	A272xB77	5.20	20.92	4.97	2.38	34.48	0.31
	B77xA272	5.48	15.02	2.35	2.13	25.74	0.21
	B77	4.04	3.23	0.28	1.89	9.88	0.12
	$HSD_{(0.05)}$	1.35	5.99	3.37	2.08	12.87	0.11
	ANOVA (p)	*	**	**	*	**	**
	Linear Effect	ns	**	**	*	**	**
	Quadratic Effect	**	ns	ns	ns	ns	**
6	CML328	1.50	11.42	5.99	1.86	21.31	0.10
	CML328xHI27	2.13	14.41	6.34	2.07	25.53	0.12
	HI27xCML328	1.74	16.91	6.61	1.64	27.56	0.09
	Hi27	1.03	14.99	7.28	1.42	25.50	0.08
	$HSD_{(0.05)}$	1.61	3.75	1.40	0.56	5.74	0.07
	ANOVA (p)	ns	*	*	*	*	ns
	Linear Effect	ns	**	*	*	*	ns
	Quadratic Effect	ns	*	ns	ns	*	ns
7	KUI3	1.54	17.70	4.22	0.94	24.65	0.08
	KUI3xSC55	0.62	2.98	2.36	1.28	7.82	0.18
	SC55xKUI3	0.16	0.13	0.83	1.45	3.93	0.45
	SC55	0.84	0.08	0.21	2.68	4.99	0.56
	$HSD_{(0.05)}$	0.81	2.41	1.67	1.10	6.06	0.29
	ANOVA (p)	**	**	**	**	**	*
	Linear Effect	*	**	**	**	**	**
	Quadratic Effect	**	**	ns	ns	**	ns

Lines grown in single environment; means computed from carotenoid profiles from 3 separate ears.
Dose models and contrasts of linear effects (additive) and of quadratic effects (dominance) are significant at α=0.01 (**), α=0.05 (*), or non-significant (ns).

Appendix G: Selected carotenoid trait correlations for five $F_{2:3}$ populations

Population: A619xSC55

	total	abranch	bbranch	total-BC	precursors
bcar	0.29**	0.15*	0.39**	0.14*	-0.07
total		0.76**	0.75**	0.99**	0.55**
abranch			0.69**	0.76**	0.49**
bbranch				0.71**	0.71**
total-BC					0.57**

Population: DE3xCI7

	total	abranch	bbranch	total-BC	precursor
bcar	-0.04	-0.30**	0.89**	-0.24**	0.16
total		0.95	0.13	0.98**	-0.23*
abranch			-0.18	0.99**	-0.23*
bbranch				-0.06	0.01
total-BC					-0.26**

Population: KUI3xB77

	total	abranch	bbranch	total-BC
bcar	-0.40**	-0.10	-0.42**	-0.51**
total		0.55**	0.98**	0.99**
abranch			0.35**	0.53**
bbranch				0.97**

Population: KUI3xSC55

	total	abranch	bbranch	total-BC
bcar	0.29*	0.11	0.25*	0.03
total		0.83**	0.96**	0.96**
abranch			0.65**	0.81**
bbranch				0.92**

Population: W64axA632

	total	abranch	bbranch	total-BC
bcar	0.53**	0.09	0.76**	0.49**
total		0.79**	0.73**	0.99**
abranch			0.15*	0.81**
bbranch				0.70**

Appendix H: Derivation of ProVA selection index and comparison with commonly used weighting schemes

An index coefficient (b) for a given trait can be calculated on the basis of trait covariances among genotypes (G) and weights (d), such that **b=G'd**. The index, I, which summarizes the weighted sum of the trait values for each entry or genotype would be equivalent to: $I = \sum_{i=1}^{n} \hat{b}_i P_i$, where $\hat{b}_i$ is the index coefficient for the ith trait, and P is the phenotypic value of the ith trait. The basis of weight selection is critical, yet is often reported to be obscured (as reviewed in Solkner, 2008) In the case of breeding for increased proVA, selection could be performed on an index using weights of proVA activity, proVA quality (bioavailability) or predicted tradeoff of increased β-carotene to reduction in total carotenoids. Each individual would yield an index value on the basis of the trait phenotypes and the index coefficient.

As an example case, an index for proVA considering only β-carotene and β-cryptoxanthin was constructed from population B-GH S_2 progeny (n=14). This population was selected for index selection as uncommonly superior b-cryptoxanthin profiles were expected to complicate selection of the best progeny. Using weights of 1 and 0.5 for the β-carotene and β-cryptoxanthin, respectively, and the $S_{1:2}$ population variance-covariance matrix, index coefficients were found to be $b_{\beta car}$=8.8398 and $b_{\beta cry}$=2.7865. Index scores were compared to progeny means for β-carotene and proVA (which was calculated as β-carotene + 0.5*β-cryptoxanthin). As expected, rank correlations among β-carotene rank, proVA rank, and index-based proVA rank were all positive, high and significant (α=0.01). While ranks between absolute concentrations had a significant correlation of r=0.75, correlations between the proVA index ranks and absolute concentrations were much higher (β-carotene, r=0.91 and proVA, r=0.93). Comparison of the rank data for absolute proVA concentration versus index proVA indicates that the index accommodates the superior b-cryptoxanthin profiles more than an absolute weighting index would allow. This suggests that use of a least squares approach will maximize the amount of trait covariation used to make selection decisions.

To evaluate selection efficiency by both ranking methods, carotenoid profile data from the $S_{1:3}$ progeny grown in replicated trials during 2009 can be compared to the expected ranks. Use of index selection for more complex trait relationships should facilitate (1) input of multiple

traits into a single index value (2) handle complex weighting systems, and (3) adjust indices to reflect the magnitude and direction of trait covariation.

Population B-GH $S_{1:2}$ progeny ranks by β-carotene, proVA, and index concentrations.

	Carotenoid Conc. (μg g^{-1})			Ranks		
$S_{1:2}$ Progeny	βcar	βcry	ProVA	βcar	ProVA	Index ProVA
09GH_115	14.31	2.48	15.54	1	1	1
09GH_129	9.61	9.56	14.39	2	2	2
09GH_118	7.94	12.61	14.24	5	3	3
09GH_124	5.00	15.71	12.85	9	4	6
09GH_128	4.33	15.49	12.07	10	5	7
09GH_126	8.77	6.07	11.81	4	6	5
09GH_113	9.54	4.19	11.64	3	7	4
09GH_130	3.56	12.35	9.74	11	8	11
09GH_122	6.63	5.97	9.61	6	9	8
09GH_121	6.19	6.71	9.54	8	10	9
09GH_116	6.20	6.27	9.34	7	11	10
09GH_117	2.56	12.99	9.06	13	12	12
09GH_127	2.83	6.82	6.24	12	13	13
09GH_114	1.63	5.55	4.41	14	14	14
Parents						
09GH_P13	2.59	13.02	9.11			
09GH_P14	1.45	14.00	8.45			

Rank correlations between indices

	PV_rank	IndexPV_rank
BC_rank	0.75**	0.91**
PV_rank		0.93**

SAS Code

```
title 'obtaining variance-covariance matrix - progeny only';
proc corr cov;
var bcar bcry;
run; quit;

title 'obtaining index coefficients - progeny only';
proc iml;
g={11.95451307 -6.22940533, -6.22940533 18.03175339};
d={1, 0.5};
transposg=g`;
b=transposg*d;
print g, d, transposg, b;
quit;

title 'measuring correlation btwn indices - progeny only';
proc corr data=indices spearman;
var bcrank pvrank idpvrank;
run; quit
```

Appendix I: Modeling carotenogenesis in developing maize kernels as a dynamic system

Research Background and Objectives:

Enhanced production of carotenoid compounds in seeds and grains through plant secondary metabolism has the potential to facilitate increased dietary intake of proVitamin A. Genetic breeding and biochemistry efforts to biofortify dent and sweet corn with these compounds require the ability to screen and identify crop varieties that differ in carotenoid accumulation [1-3]. Conventional practices for screening involve chemical profiling experiments of carotenoids on mature, dry seed. As most food products derived from maize grain use fully mature seed, the relevance of these measurements is indisputable. Grain carotenoid concentrations are sensitive to endogenous degradation processes (Chapter 2, [4]) and post-harvest storage conditions (M. Grusak, personal communication), suggesting that the biosynthetic potential for carotenoid concentration of screened inbred lines may be confounded with external environmental factors. This suggests that lines superior in carotenoid accumulation could be inadvertently selected against due to this confound. Additionally, carotenoid accumulation in grains such as sweet corn are consumed very early in the process of seed development, during which time total carotenoids are beginning to increase in synthesis and deposition [5]. Therefore, for the selection of sweet corn varieties that have increased carotenogenesis (or synthesis of carotenoids), it is necessary to perform concurrent selection for premature carotenogenesis.

Kernel development involves the growth and differentiation of cells that constitute cereal endosperm (seed nutrient reserves), embryo (new generation) and maternally derived tissues including the seed pericarp (outer covering). During this process, the dynamics of cell division and expansion are accompanied by starch deposition and proportional water loss, as described in Figure I-1. To facilitate classification of kernels undergoing development, the peak and termination of these physiological processes have been marked by reproductive stages in maize, also indicated in Figure I-1. Carotenoid accumulation, which can be visually noted by the transition of a colorless to yellow pigmented tissue, is regulated by the activity of phytoene synthase (*psy1*), the first committed step of the carotenoid pathway. Comparison of total carotenoid accumulation across kernel development for maize inbred line B73 revealed that the increase of carotenoids as well as their plateau was correlated with the presence of *psy1* transcript and PSY1 enzyme [6]. The accumulation of carotenoids including lutein, zeaxanthin,

zeinoxanthin, β-cryptoxanthin, α-carotene and β-carotene during kernel development is likely similar that of total carotenoids through a pattern limited growth, which is described by an initial rate, tapering to a steady state or maximum value with increased time. Considering the phenotypic diversity of maize grain carotenoid profiles, it is also likely that these carotenoids may not accumulate at the same rate, or perhaps, even through the same biochemical mechanisms.

Knowledge of the genetics controlling carotenoid biosynthesis has been shown to enhance predictability for selection of lines with desirable carotenoid profiles, as demonstrated through marker assisted selection (Chapter 3). As carotenoid concentration and composition traits are quantitatively inherited and highly heritable (Chapter 1), genetic regulation of carotenoid accumulation is arguably highly important at both the final harvest timepoint and throughout development. Metabolic conversion rates can be compared with what is known about the genetics of carotenoid pathway regulation (allelic composition, combination of haplotype effects in hybrids) to test if rates cluster with allelic composition. In addition, differences in the final carotenoid phenotype could be due to the timing and coupling of reactions involved in carotenogenesis; this hypothesis could also be tested with carotenoid profiles of developing seeds.

Using biochemical profiles of carotenoid accumulation over kernel development in six inbred lines, this study was developed to test: (1) if accumulation of single carotenoids over time could be described using models for limited growth; (2) if a pathway model describing the simultaneous change of six colored carotenoids could explain the observed data; (3) if this model could estimate genotype-specific parameters that were relatable to known genetics governing the carotenoid biosynthesis pathway; 4) if estimated parameters provided any new insight to pathway dynamics. Specific objectives of this work will include the evaluation of a first order model for carotenoid inter-conversion (please see pathway in Chapter 1 for details on carotenoid intermediates) and the determination of rate constants for conversion steps within the pathway.

Methods

Experimental Design and Sample Collection

A set of six maize inbreds contrasting for haplotypes in at least one polymorphism for *crtRB1*, *lcyε* and *ccd1* (Tables I-1 and I-2), were planted in a randomized split plot design during

summer 2008 in Urbana, Illinois. Three replicates each consisting of five whole plots, or blocks, were planted. The five contrasting blocks were comprised of the following inbred lines: (1) A619, SC55; (2) KUI3, B77; (3) CI7, KUI3; (4) CI7, DE3; (5) KUI3, SC55. Each block contained four sub plots which upon harvest, yielded seed of the following: plot 1, inbred 1 *per se*; plot 2, inbred 2 *per se*; plot 3, F_1 seed of inbred 1 x inbred 2; plot 4, F_1 seed of inbred 2 x inbred 1. Up to 15 self or cross pollinations were made in each row; however, germination rates were poor for several plots, and therefore the number of plants/ pollinations was greatly reduced in some plots.

Sampling of pollinated ears was performed according to a time-course on every subplot in the experiment. Developing ears were harvested on the basis of days after pollination (DAP) according to a 15 point time-course including 12, 14, 16, 18, 21, 24, 27, 30, 33, 35, 38, 40, 45 DAP, a fresh final harvest point (55-65 DAP) and a final harvest point with heat treatment of approximately 37° F in a dryer for five days. Whole ears were flash frozen in liquid nitrogen and stored at -80° C. Frozen kernels were removed and divided into four portions to 1) stage the reproductive maturity of the kernels (R1 through R4), 2) obtain fresh and dry five-kernel weights (3 replicates), 3) measure kernel volume (7 replicates), and 4) reserve a bulk seed sample for carotenoid and transcript profiling.

Physiological Measurements

Percent moisture was measured as the fraction of fresh kernel weight unaccounted by dry kernel weight. Five fresh frozen kernels were placed in a weighed borosilicate glass vial and weighed. Care was taken to remove any moisture along the pericarp cell layer resulting from condensate with a laboratory wipe, prior to weighing. The vials were incubated at approximately 67°C in a drying oven for 48 hours, upon which they were quickly weighed. After subtracting initial vial weight, adjusted fresh and dry weights were used to calculate percent moisture as: $(fresh\_wt - dry\_wt) \big/ fresh\_wt$. Percent moisture loss over time is depicted in Figure I-2.

Seed volume was estimated as a function of length (distance from tip to crown), width (germ width at crown) and thickness (crown width) of the seed. Care was taken to sample frozen kernels from the middle third of the ear. Seven kernels were sampled and measured with a pair

of calibrated electronic calipers to the closest tenth of a millimeter. Seed volume increase over time is depicted in Figure I-2.

Maturity was roughly assessed on a four class scale defined by maize reproductive stages: R1, blister, R2, milk; R3, dough, R4, dent. Kernel samples from developing B73 ears were used as a comparison for this rough estimation.

Carotenoid Extraction and Quantification

Fresh frozen bulk kernel samples from were used for carotenoid profiling. Embryos from 4-6 kernels were removed, and the remaining frozen endosperm was coarsely homogenized. Further homogenization was performed on 30-50 mg of sample using a QIAGEN Tissuelyser II homogenizer and 4 mm steel ball bearings. Carotenoids were extracted in microtubes using 600 µl of 2:1 methanol:chloroform containing BHT (1 mg/mL) and tocol as an internal standard by further homogenization. After addition of 400 µl water and 200 chloroform, the samples were vortexed for 15 minutes and spun at 12 000 g for 10 minutes. The bottom fraction was collected, dried, and resuspended in 200 µl injection buffer (95:5 acetonitrile: ethyl acetate). HPLC analysis was carried out on 50 µl of the final extract.

Carotenoids were separated by HPLC on a C18 column (Spherisorb ODS2 5 micron, 150 × 2.1 mm, Column Engineering, Ontario) with a Shimadzu LC-20AD HPLC at variable flow rates with solvent A (acetonitrile: water [9:1 v/v]) and solvent B (ethyl acetate) and the following gradient: 0–20 min, 5% to 77% B, 1.0 mL/min; 20-20.2 min, 77 to 100% B, 1.0 mL/min; 20.2-22.2 min, 100% B, 1.5 mL/min, 22.2-22.4 min, 100-5% B, 1.5 mL/min; 22.4-25 min, 5% B, 1.0 mL/min. HPLC peak areas were integrated at 450 nm for the carotenoid compounds and 290 nm for the tocol standard. Data is represented on a fresh weight basis.

Statistical Analysis and Model Testing

Only phenotypic data from sampled self pollinations was analyzed, as the full set of materials was not profiled for carotenoids. All descriptive statistics and analysis of variance was performed with SAS v9.2 [7]. Integration of derived differential equations describing growth rates for individual carotenoid pools not based on Richards Family growth models was performed by hand and later validated with Derive 5 (Texas Instruments) software. Parameter estimates, model variance and goodness of fit measures for non-linear models of single carotenoid pools

were obtained using the NLREG non-linear regression software package [8]. Parameter estimates for a system of differential equations describing changes in more than one carotenoid pool (as in the derived pathway models) were obtained using Berkley Madonna version 8.0 software which provides numerical solutions to systems of ordinary differential equations. Model fits were generated using the Runge-Kutta 4 integration method, with dt=0.001. STELLA version 7.0 was used to interpret the data from MADONNA.

Description of Pathway Model and Assumptions

Construction of the Pathway Chain Model was based on a framework of the relationships in the carotenoid pathway, as shown in Chapter 1, Figure 1.2. A system of differential equations was created to describe :(1) the input of precursor from the main pathway branch to downstream branches represented as Chains I and II; (2) the flow of substrate within Chain I, also known as the α-branch; (3) the flow of substrate within Chain II, also known as the β-branch; (4) efflux out of each chain. Parameter "input" is defined as the amount of precursor carotenoid cycled into the pathway at each time step, and is assumed to be constant. Change in reservoir concentration over time is governed by the difference between input rate from upstream reservoirs and output rate to downstream reservoirs. These rates are linearly dependent on the concentration of carotenoid as it changes over time.

The first order rate constants of growth for component carotenoids within the pathway, k, are expressed on a per day basis. Proportion factor "p" is constant and designates the amount of total carotenoid allocated to Chain I; (1-p) and Chain II (p).

A pictoral representation of the model is shown in Figure I-3. Within Chain I, metabolite stocks are designated as A (α-carotene), B (zeinoxanthin), and C (lutein). Within Chain II, metabolite stocks are designated as D (β-carotene), E (β-cryptoxanthin) and F (zeaxanthin).

Results and Interpretation

Behavior of carotenogenesis trends in maize

Profiles of the six major colored carotenoids in maize endosperm were measured across a developmental window spanning 12 to 65 days after pollination (DAP). Of all carotenoids measured, lutein, zeaxanthin, β-carotene and β-cryptoxanthin increased most through the sampled developmental window (Figure I-4). This increase in magnitude was accompanied by

increased variance with increasing time, suggesting that it could be more difficult to determine the true character of the growth curve at the latter end of the time course. Carotenoid concentrations were 10% or less than their maximal values when measured at 12 DAP, indicating that the carotenoid pathway is not highly active during the initial stages of kernel development. By the R2 stage, carotenoid accumulation is occurring at its fastest rate. Most of the highly varying carotenoid pools demonstrate high rates of growth through the R2 stage. These observations follow the reported activity of PSY1 during kernel development.

Striking genotype-specific differences in carotenoid accumulation were noted across the six inbreds (Figure I-4). Some of these differences appeared to be associated with changes in the allelic state of known genes that affect the pathway including: *lcyε*, which is reported to govern allocation between Chains I and II; *crtRB1*, which affects the conversion of intermediates within Chain II; *ccd1*, reported to deplete most carotenoid intermediates. β-carotene had the largest maximal accumulation of all carotenoids for inbreds SC55, CI7 and B77, all of which had alleles at *crtRB1* polymorphisms which are associated with reduced β-carotene conversion to β-cryptoxanthin. Comparison of the α/β (or Chain I /Chain II) branch ratio at the maximal timepoints showed that DEexp had the highest α/β ratio, whereas KUI3 had the lowest. Allelic conditions for *lcye* do not appear to strongly associate with the phenotypic ratios. The deleterious Wc-ccd1 allele, reported to significantly reduce lutein levels (Chapter 2), appeared to have a large negative effect on SC55, the only line with the Wc allele in this set of germplasm. Interestingly, had the Wc-ccd1 effect not been present in this line, the general trend of the lutein curve suggests that it would have surpassed β-carotene accumulation.

Evaluating models of single stock growth

Given that carotenoid accumulation over time generally appeared to follow a sigmoidal growth curve reaching a maximal level at or close to 50 DAP, models of limited growth were investigated to determine if model parameters could fit the growth trend of the observed data. We tested if carotenoid accumulation over kernel development is consistent with growth models from the Richards' family [10], which have been used to describe S-shaped growth with first order rate constants. The three Richards' models tested here (Table I-3) differ in the relationship between growth rate and concentration (linear versus logarithmic), the presence or absence of an inflection point, and the time it takes to attain steady state. Initial values for all Richards' Family

models are approximately zero, which facilitates the description of early growth unlike many other models. The three models share parameters of k, a rate constant; $W_{max,}$ the maximum size of a population; and a factor A.

Using β-carotene concentrations (pool D) from the time course of A619 developing kernels, monomolecular, Gompertz and logistic growth models were tested in NLREG. With three parameters each, 76.59% to 79.48% of the phenotypic data was explained with these models (Table I-4). Predicted values for all models were quite similar, falling within the observed values. The spread of residuals was also very similar among the three models, and demonstrated most deviation from the observed data at the most variable timestep of 65 DAP (Figure I-5). Estimated rate constants and maximal population size were similar among the tested models, and parameter estimates for A were found to vary widely in order to accommodate the structure of each model (Table I-4).

These models were satisfactory in explaining the amount of variation in β-carotene as it changed over time; however, the parameter estimates could not be interpreted in the context of changes within a pathway, as these models lacked parameters that relate growth to changes in other pools. As a result, it appeared that using any one of these models to describe a system of six reservoirs would be insufficient to describe the biochemical pathway for carotenogenesis. Therefore, the Pathway Chain model, composed of six component reservoirs, was used. Details of this model are described in the methods as well as in Table I-5. The differential equation describing growth of pool D over time is nearly equivalent to that of pool A in Table I-5, with the exceptions that (1-p) is replaced by p, and k_b is replaced by k_e yielding: $\frac{dD}{dt} = (p)(Input) - k_E D$.

This model explained slightly less variation in β-carotene accumulation over time (R^2_{adj} = 75.42%) than those of the Richards' family, and had an extra parameter to account for its integration within a pathway. However, parameter estimates were interpretable in terms of the pathway input (3.652 nmol μg^{-1} per day), allocation of input to Chain II through p (0.518) and rate constant describing the transition to E (k_E=0.066), and the initial concentration of D_0 (-36.01 nmol μg^{-1}). The modeled initial value for reservoir D (at DAP=0) was large and negative, suggesting that this model failed to account for the lag phase of carotenoid production from 0-12 DAP. However, this enabled the model to provide an accurate estimate of the steady state value at DAP ≥ 30. Most importantly, the maximal value for this model was interpretable through the

integrated form (Table I-5), where if t is large, only the first term remains, describing the maximal value as the ratio of input to efflux rate parameters. The predicted and residual plots largely express that this curve is similar to that of the monomolecular curve (Figure I-5).

Testing System Pathway Chain Model

To express the change in total carotenoid accumulation on a pathway basis, a system of differential equations describing changes in each reservoir was delineated (Figure I-6). The integrated form of each function is listed in Table I-5. Solving for the system of ordinary differential equations describing this model, model parameters including rate constants, the proportionality constant and input were obtained for each genotype (Table I-6).

Parameter estimates for the input value were within the same order of magnitude for all genotypes and were therefore fairly similar. The allocation of substrate to each chain dictated by the parameter p differed, indicating that the majority of these genotypes favored Chain II (β-branch) over Chain I (α-branch).

Conversion within Chain I, influenced by rate constants k_b, k_c and k_{1out}, generally reflected larger rate constants for k_b than k_c, and highly variable magnitude in k_{1out} across genotypes. From the integrated growth equations in Table I-5, the steady state conditions for each of these reservoirs are:

$$A(t) = \frac{Input(1-p)}{k_B},\ B(t) = \frac{Input(1-p)}{k_C},\ C(t) = \frac{Input(1-p)}{k_{1OUT}}.$$

If the magnitude of input is the same for each reservoir, as is the case in this model, the steady state concentration of a component is inversely proportional to the size of the rate constant. Accordingly, the steady-state value for A (α-carotene) is consistently larger than that of B (zeinoxanthin). In general, a large rate constant for k_b reduces the time it takes for any pool to reach its steady state value. Attainment of steady state does however differ between individual pools because of the influence of other rate constants. The small values of k_{1out} reflect that genotypes including B77, CI7 and DEexp are more likely to accumulate reservoir C (lutein) due to minimal efflux from the pathway, than other genotypes. Conversion within Chain II affected by rate constants k_e, k_f and k_{2out}, was marked by much smaller rate constants than those in Chain I, indicating that reservoirs of D (β-carotene), E (β-cryptoxanthin) and F (zeaxanthin) were more likely to accumulate than A, B, and C. Considering that selection of these six inbreds was

heavily based on their ability to produce β-carotene, this result is not surprising. Model predictions and trait observations are compared for all genotypes in Figures I-7.

Model adjustment: seed volume as a driver of carotenogenesis

As can be seen by the data plots of the predicted versus observed values, the system model does not describe the initial lag of carotenoid accumulation between anthesis (0 DAP) and the beginning of our collection period (12 DAP). Accordingly, we lack the resolution to differentiate the timing of accumulation rates early in kernel development. The components of the system model shown here are controlled by growth characteristic of reaction models which do not experience limited growth at the start of carotenoid accumulation. Gompertz models, however, do describe initial limited growth, and often are used to model biological systems where the size of the initial population is critical in allowing more growth to occur and/or where constraints on growth emerge as the reservoirs attain a critical size. It is unlikely that the presence of carotenoids is required for the production of carotenoids during the early stages of growth. It is more likely that the initial input carotenoid pool (phytoene) is affected by physiological growth, as both cell division (Figure I-1) and phytoene synthase protein levels [5] are increasing during the same developmental window. Cell sink size has also been reported to be linked to β-carotene and total carotenoid accumulation in cauliflower and tomato mutants [11,12].

Considering a possible relationship between carotenoid accumulation and cell growth, we hypothesized that the growth rate of the input carotenoid reservoir could be governed according to the growth of a physiological parameter such as seed volume size (Figure I-2). As seed volume is known to experience limited initial growth, we are testing the Gompertz and logistic growth models with seed volume data. Although both the seed volume models and the incorporation of this modified input into to the Pathway Chain Model is not yet complete, preliminary results suggest that the Gompertz model yields a very good fit to the volume data which is encouraging for further use in a combined seed physiology-metabolite model. Alternate physiological parameters of water content or dry biomass could be tested for their influences on carotenoidgenesis during kernel development.

References:

1. Ye, X. D. *et al*. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**, 303-305 (2000).

2. Harjes, C. E. *et al*. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **319**, 330-333 (2008).

3. Yan, J. *et al*. *Rare Genetic Variation at Zea mays crtRB1 Increases β-carotene in Maize Grain*. *Nat.Genet.* (submitted).

4. Kean, E. G., Ejeta, G., Hamaker, B. R. & Ferruzzi, M. G. Characterization of carotenoid pigments in mature and developing kernels of selected yellow-endosperm sorghum varieties. *J. Agric. Food Chem.* **55**, 2619-2626 (2007).

5. Vallabhaneni, R. & Wurtzel, E. T. Timing and Biosynthetic Potential for Carotenoid Accumulation in Genetically Diverse Germplasm of Maize. *Plant Physiol.* **150**, 562-572 (2009).

6. Li, F., Vallabhaneni, R., Yu, J., Rocheford, T. & Wurtzel, E. T. The maize phytoene synthase gene family: Overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. *Plant Physiol.* **147**, 1334-1346 (2008).

7. SAS Institute. Version 9.2 Cary, North Carolina

8. Dennis, J.E., Gay, D.M., & Welsch,R.E. An adaptive nonlinear least-squares algorithm.*ACM Transactions on Mathematical Software 7,3* (Sept. 1981)

9. Macey, R., Oster, G., & Zahnley, T. Berkeley Madonna User's Guide,
Version 5.0, University of California (2000). http://www.berkeleymadonna.com

10. Brown, D. & Rothery, P. in *Models in Biology: Mathematics, Statistics and Computing* (John Wiley & Sons Ltd, West Sussex, England, 1993).

11. Lopez, A. B. *et al*. Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.* **59**, 213-223 (2008).

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

Tables

Table I-1: Carotenoid QTL genotypes for inbred lines used in carotenogenesis modeling study

QTL	*crtRB1*			*lcyε*			*ccd1*
Polymorphism	5'TE	InDel4	3'TE	5'TE	SNP216	3'TE	5p
Allelic series[a]	1/**2**/3	**1**/2	**1**/2/3	**1**/2/3/**4**	**1**/2	1/**2**	**b73**/Wc
A619	1	2	**1**	2	2	**2**	**b73**
B77	**2**	**1**	**1**	2	2	**2**	**b73**
CI7	**2**	**1**	**1**	2	2	1	**b73**
DEexp	**2**	2	**1**	3	**1**	1	**b73**
KI3	1	2	3	**4**	2	**2**	**b73**
SC55	**2**	**1**	**1**	2	2	1	Wc

a: Allele associated with higher β-carotene and/ or larger β-branch carotenoid pool is considered favorable and is denoted in bold

Table I-2: Inbreds and genotypes comprising experimental blocks of carotenogenesis modeling study

Blocks			Allele-specific marker genotypes						
No.	Inbred Line 1	Inbred Line 2	crtRB1 5'TE	crtRB1 InDel4	crtRB1 3'TE	lcyE 5'TE	lcyE SNP216	lcyE 3'TE	ccd1 5p
1	A619	SC55	**1,2**	**2,1**	1,1	2,2	2,2	**2,1**	**b73, Wc**
2	KUI3	B77	**1,2**	**2,1**	**3,1**	**4,2**	2,2	**2,2**	b73, b73
3	KUI3	CI7	**1,2**	**2,1**	**3,1**	**4,2**	2,2	**2,1**	b73, b73
4	DEexp	CI7	2,2	**2,1**	1,1	**3,2**	**1,2**	1,1	b73, b73
5	KUI3	SC55	**1,2**	**2,1**	**3,1**	**4,2**	2,2	**2,1**	**b73, Wc**

Bold face indicates differing haplotype between two inbred lines in block

Table I-3: Richards' Family of Growth Models

General Richards Family Growth Rate Model: $\frac{dW}{dt} = \frac{kW}{(1-m)}\left[\left(\frac{W}{W_{\max}}\right)^{1-m} - 1\right]$

Functional/ Growth Equation: $W(t) = W_{\max}(1 - Ae^{-kt})^{1/(1-m)}$

Growth Model	m	Growth Equation	Growth Rate	Inflection Point
Monomolecular	0	$W(f) = W_{\max}(1 - Ae^{-kt})$	$k[W_{\max} - W(t)]$	None
Gompertz	1	$W(f) = W_{\max}e^{-Ae^{-kt}}$	$kW(t)\log_e\left(\frac{W_{\max}}{W(t)}\right)$	0.368 W_{max}
Logistic	2	$W(f) = \frac{W_{\max}}{(1 - Ae^{-kt})}$	$kW(t)\left(1 - \frac{W_{\max}}{W(t)}\right)$	0.5 W_{max}

Table I-4: Evaluation of growth model fit to β-carotene (compartment D) concentration in developing kernels of A619 genotype

Parameters Estimated	Initial Guess	Richards Family Growth Models			Parameters Estimated	Initial Guess	Reservoir Model D
		Monomolecular	Gompertz	Logistic			
k	0.5	0.067	0.139	0.222	k_E	0.5	0.066
W_{max}	30	28.561	26.639	25.907	p	0.5	0.518
A	0	2.261	16.864	180.042	D_0	0	-36.01
					Input	5	3.652
Model Fit							
R^2		0.7862	0.8086	0.8127			0.7862
R^2 adj		0.7659	0.7904	0.7948			0.7542
Durbin-Watson		2.473	2.676	2.677			2.473

Models used in non-linear regression are listed in Table I-3. Parameter estimates solved by non-linear regression in NLREG.

Table I-5: Growth Functions for Derived Pathway Chain Model

Growth Model	Stock	Type	Representative Equation[a]
Component 1	A	Differential Equation	$\frac{dA}{dt} = (1-p)(Input) - k_B A$
		Integrated Form	$A(t) = \frac{Input(1-p)}{k_B} - \frac{Input(1-p)}{k_B}(e^{-k_B t}) + A_0(e^{-k_B t})$
Component 2	B	Differential Equation	$\frac{dB}{dt} = k_B A - k_c B$
		Integrated Form	$B(t) = \frac{Input(1-p)}{k_C} - \frac{Input(1-p)}{(k_C - k_B)}(e^{-k_B t}) + \frac{k_B A_0}{(k_C - k_B)}(e^{-k_B t})$
Component 3	C	Differential Equation	$\frac{dC}{dt} = k_c B - k_{1hOUT} C$
		Integrated Form	$C(t) = \frac{Input(1-p)}{k_{1OUT}} - \frac{Input(1-p)K_C}{(k_C - k_B)(k_{1OUT} - k_B)}(e^{-k_B t}) + \frac{k_C k_B A_0}{(k_{1OUT} - k_B)(k_C - k_B)}(e^{-k_B t})$

a: Both differential and integrated forms of the equations can be adapted for stocks D, E, and F, there proportion factor equals p rather than 1-p

Table I-6: Genotype specific parameter estimates for pathway chain model

Parameter Estimates	Genotypes					
	A619	B77	CI7	DEexp	KUI3	SC55
k1OUT	0.024	3.20E-09	1.85E-09	4.58E-09	0.548	0.099
k2OUT	2.01E-07	6.51E-07	0.022	5.59E-09	7.98E-04	0.092
kb	7.596	1.917	1.453	1.401	3.774	0.828
kc	0.581	0.432	0.445	2.720	1.59E-08	0.299
ke	0.020	0.022	0.016	0.021	0.282	0.007
kf	0.099	0.556	0.426	0.940	0.258	0.097
Input	1.270	0.795	1.014	0.774	1.210	1.058
p	0.609	0.749	0.693	0.755	0.985	0.377
RMS	12.8447	13.1043	13.7278	13.4382	20.3651	11.6919

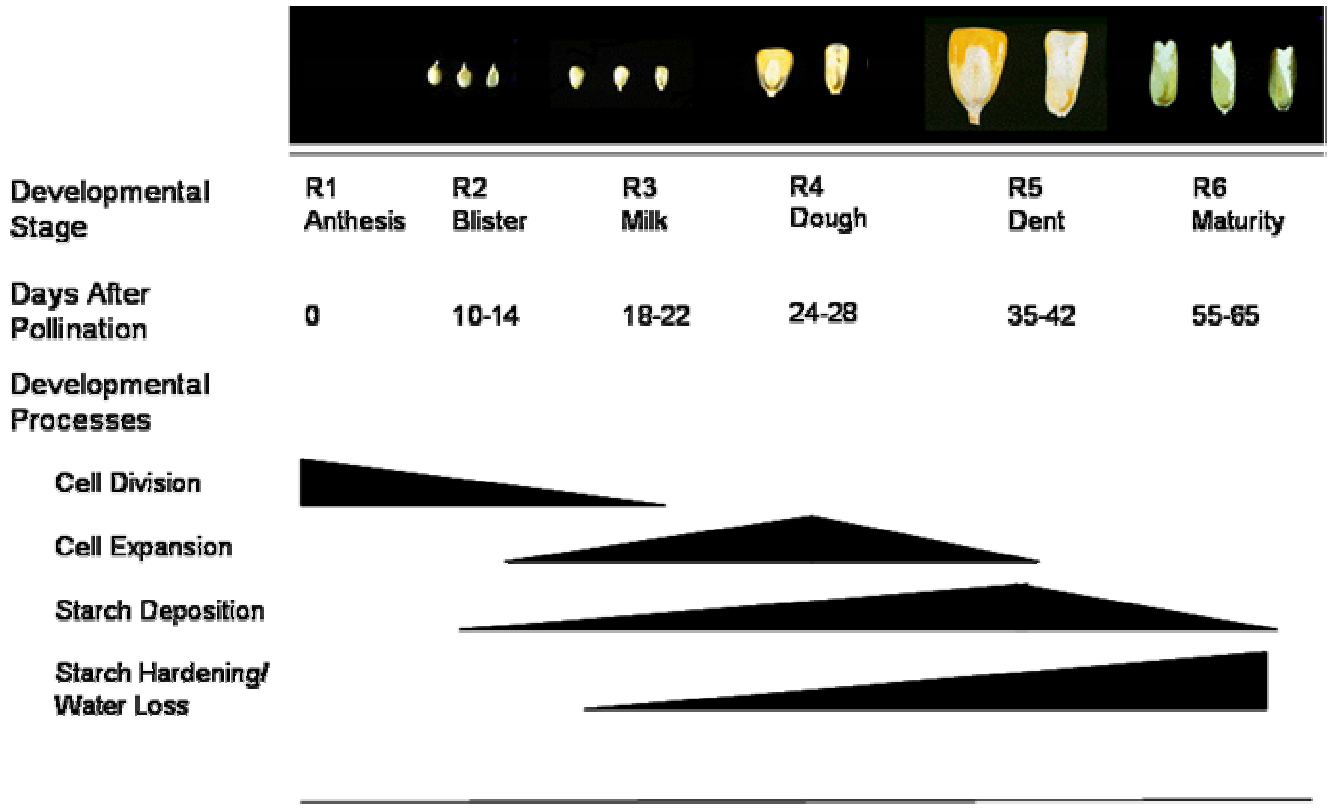
Minimal root mean square (RMS) achieved by numerical search for parameters is listed below parameter estimates. Parameter estimates obtained using Runge-Kutta 4 method of integration in MADONNA.

Table I-7: Code used to solve simultaneous ordinary differential equations in MADONNA

```
METHOD RK4

STARTTIME = 0
STOPTIME=70
DT = 0.001

ka = kprec*(1-p);
d/dt(A) = ka-kb*A
d/dt(B) = kb*A-kc*B
d/dt(C) = kc*B-k1OUT*C

kd = kprec*p
d/dt(D) = kd-ke*D
d/dt(E) = ke*D-kf*E
d/dt(F) = kf*E-k2OUT*F

k1OUT = .15
k2OUT = 0.085
kb = 5
kc = .9
ke = 0.06617
kf = 0.4
kprec = 4.0664
p = 0.46476

INIT A = 0
INIT B = 0
INIT C = 0
INIT D = 0
INIT E = 0
INIT F = 0
```

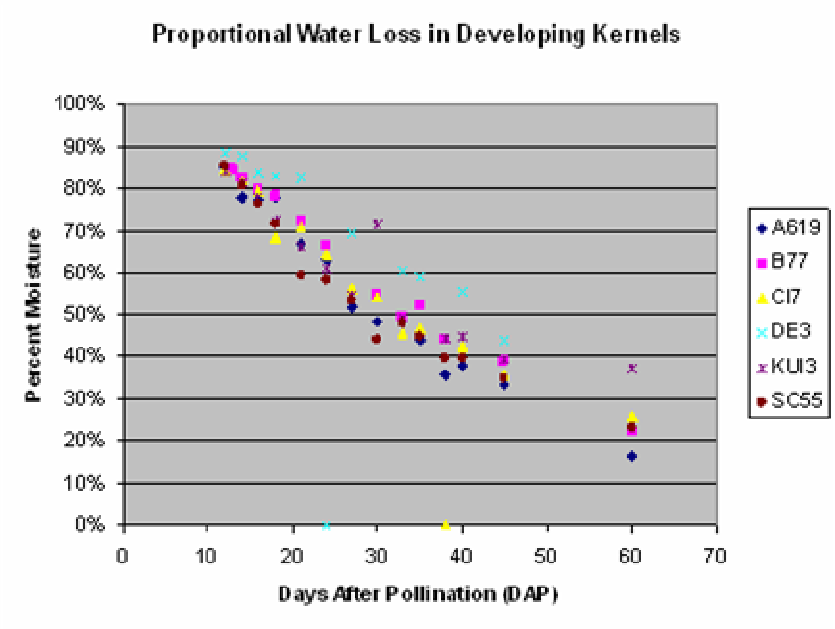
Figures

Figure I-1: Growth stages of maize and important developmental process during reproductive growth



Black bands indicate occurrence and relative magnitude of physiological processes during maize reproductive stages. Data obtained from Dr. Fred Below, CPSC518 Crop Growth and Development Course Notes

Figure I-2: Changes in seed moisture content and volume in developing kernels across six inbred genotypes



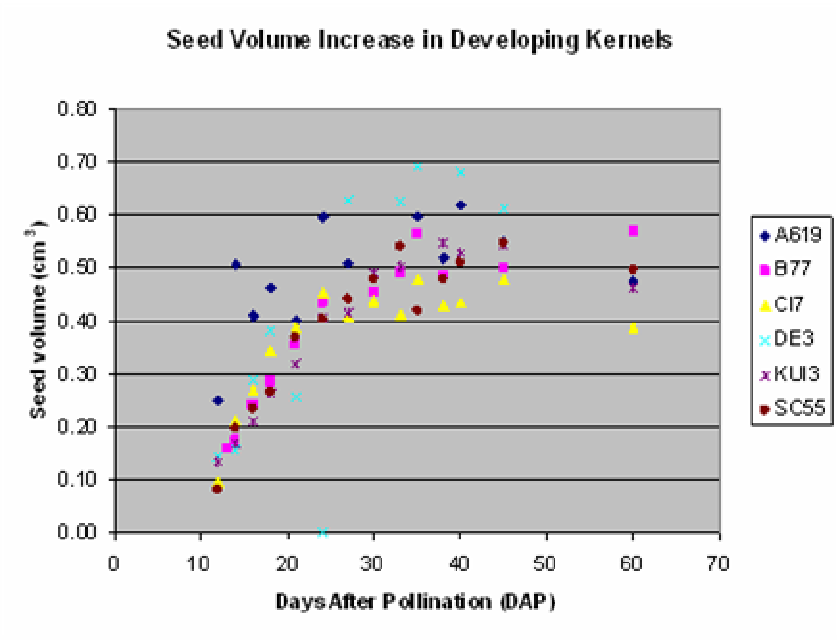
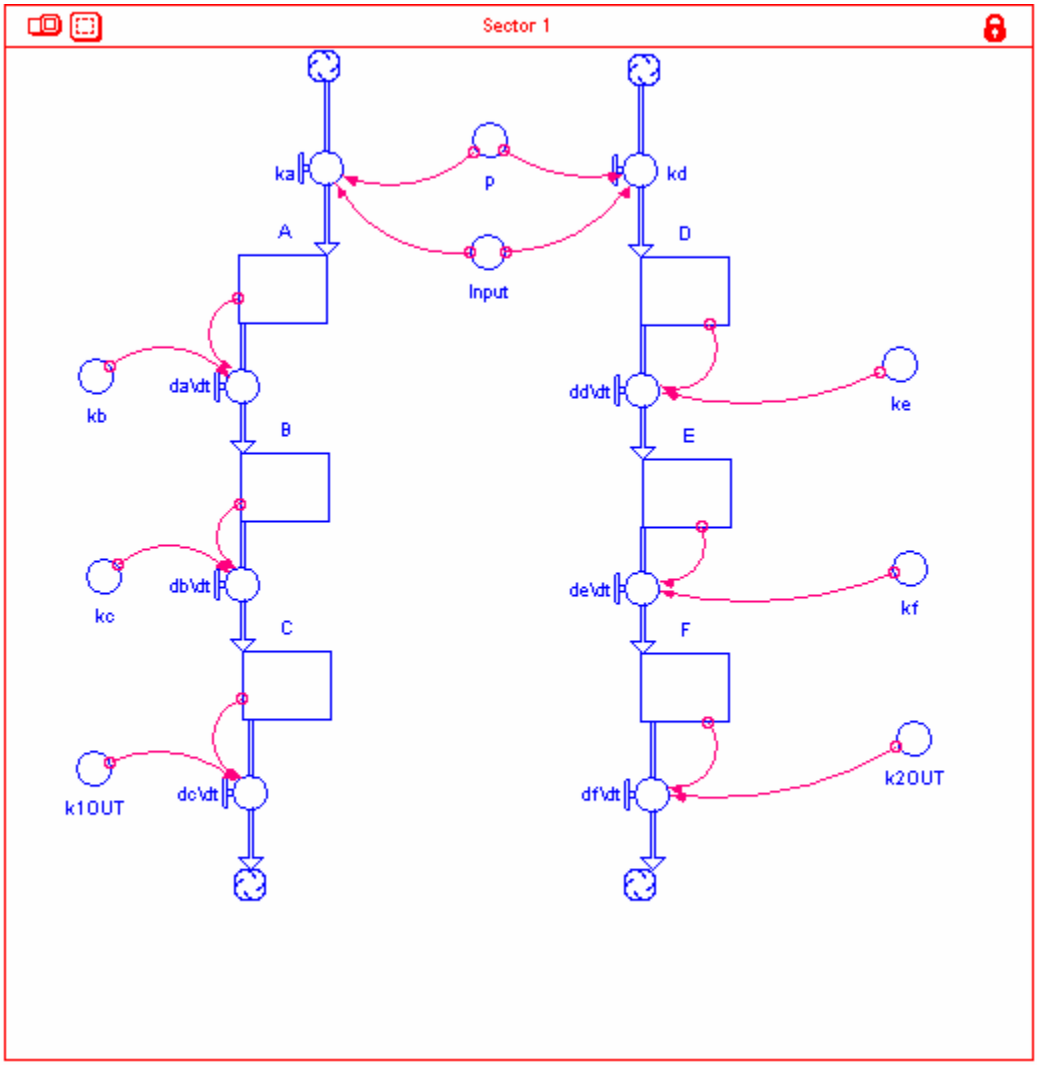


Figure I-3: Relationship of stock reservoirs in Pathway Chain Model



Main branch of carotenoid pathway provides carotenoid substrate to colored carotenoids (Chains I and II) through ka and kd fluxes respectively. Left hand branch is designated as Chain I with components A (α-carotene), B (zeinoxanthin), and C (lutein). Right hand branch is designated as D (β-carotene), E (β-cryptoxanthin) and F (zeaxanthin).

Figure I-4: Carotenoid profiles of developing kernels across six inbred genotypes

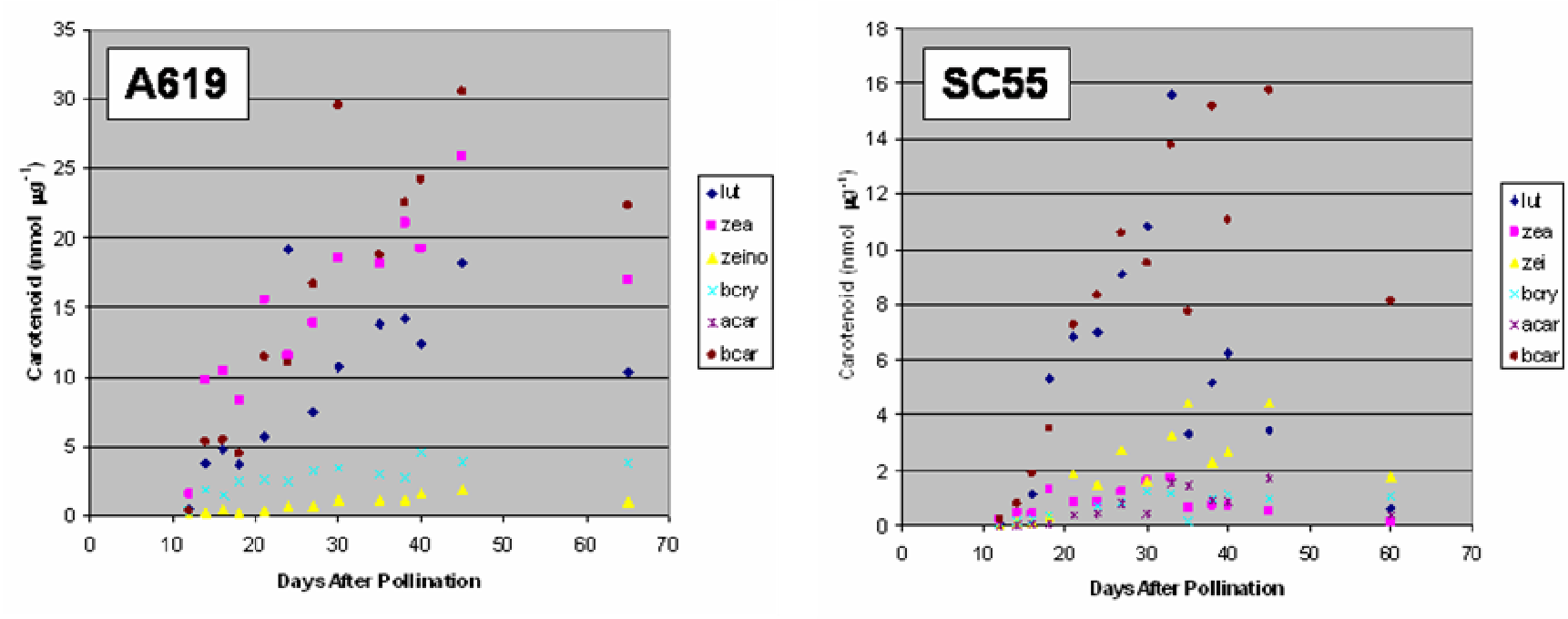
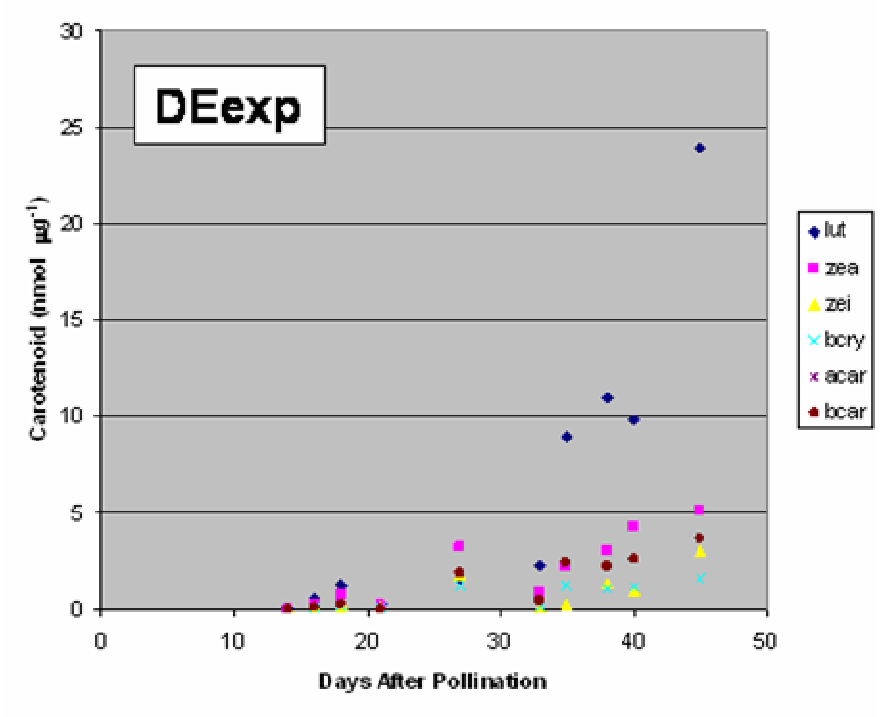


Figure I-4:, continued



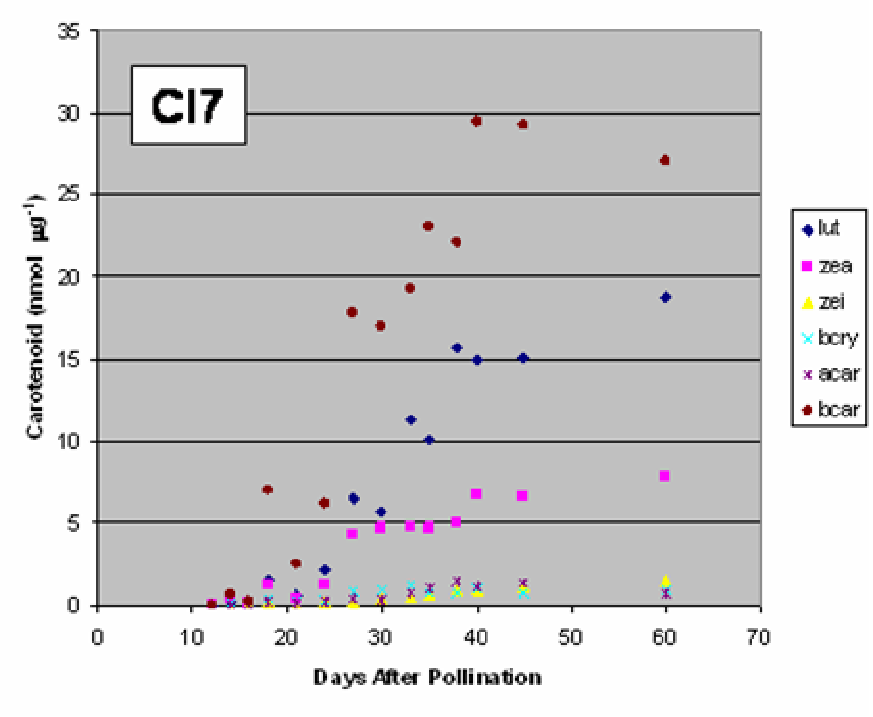
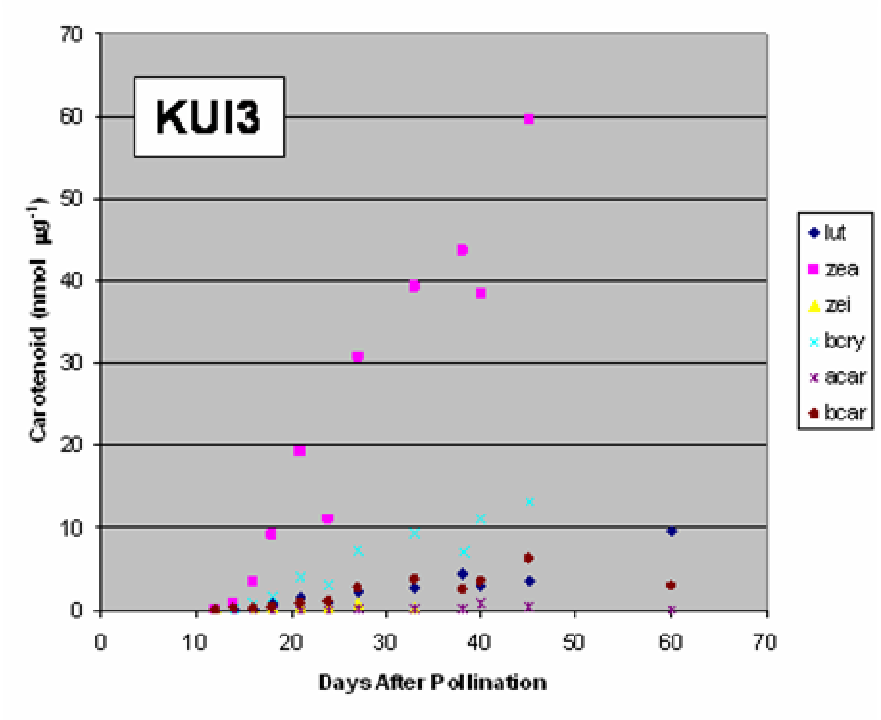


Figure I-4:, continued



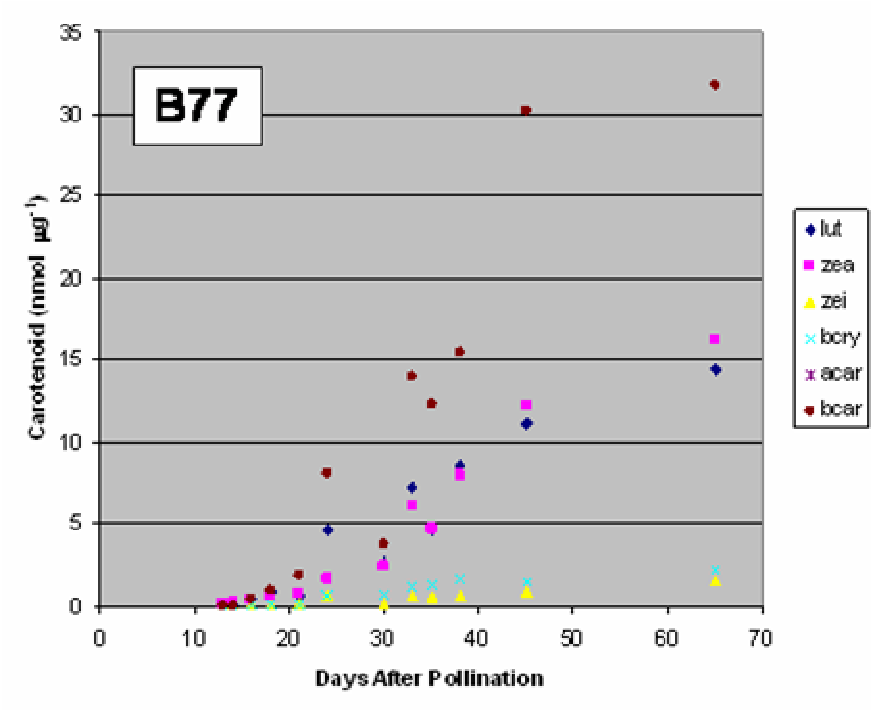
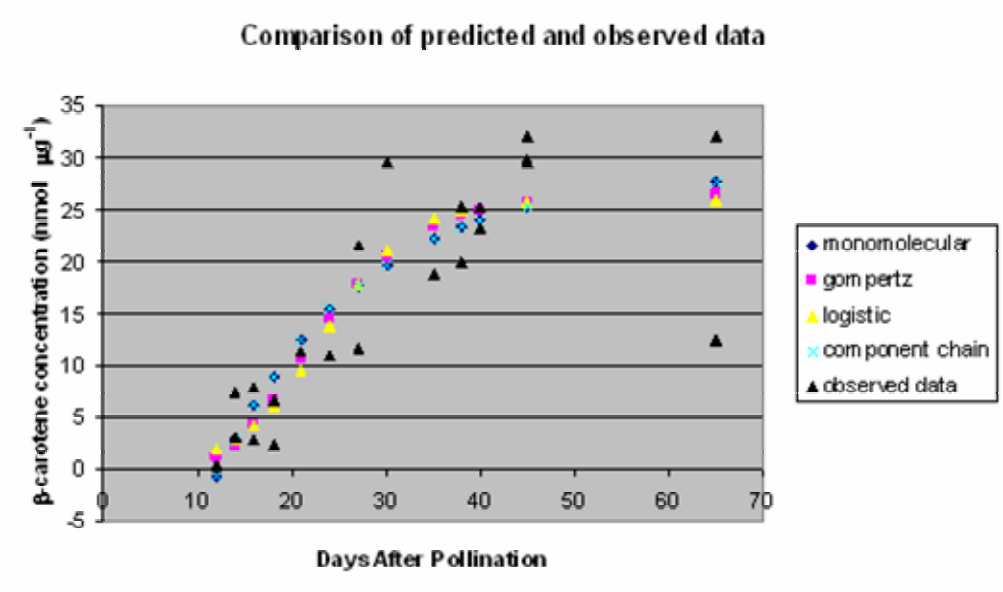


Figure I-5: Fit of Richards family and single component chain models on β-carotene concentration of developing A619 genotype kernels



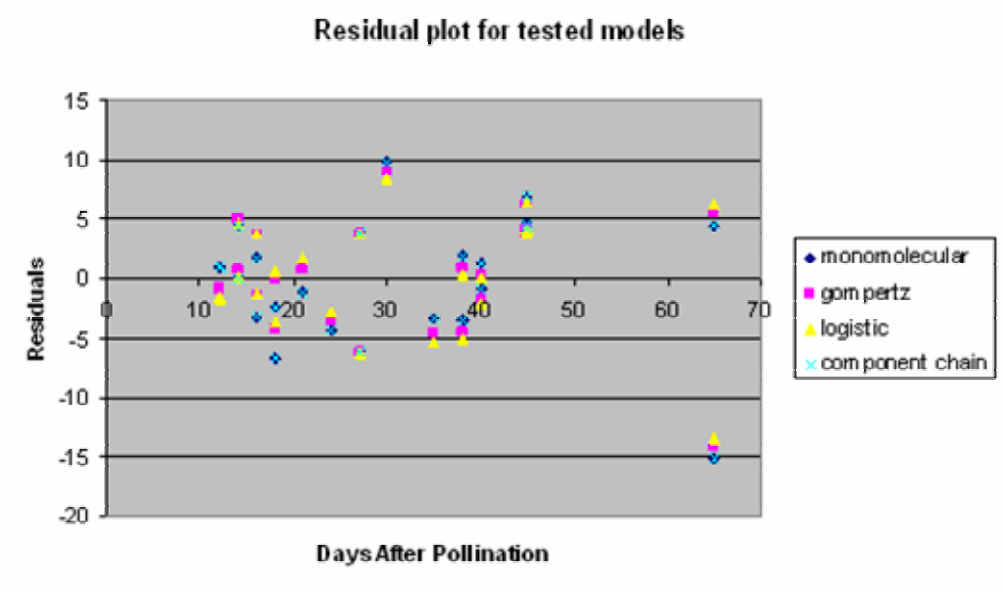


Figure I-6: Growth rates of stock components in Pathway Chain Model as defined in STELLA

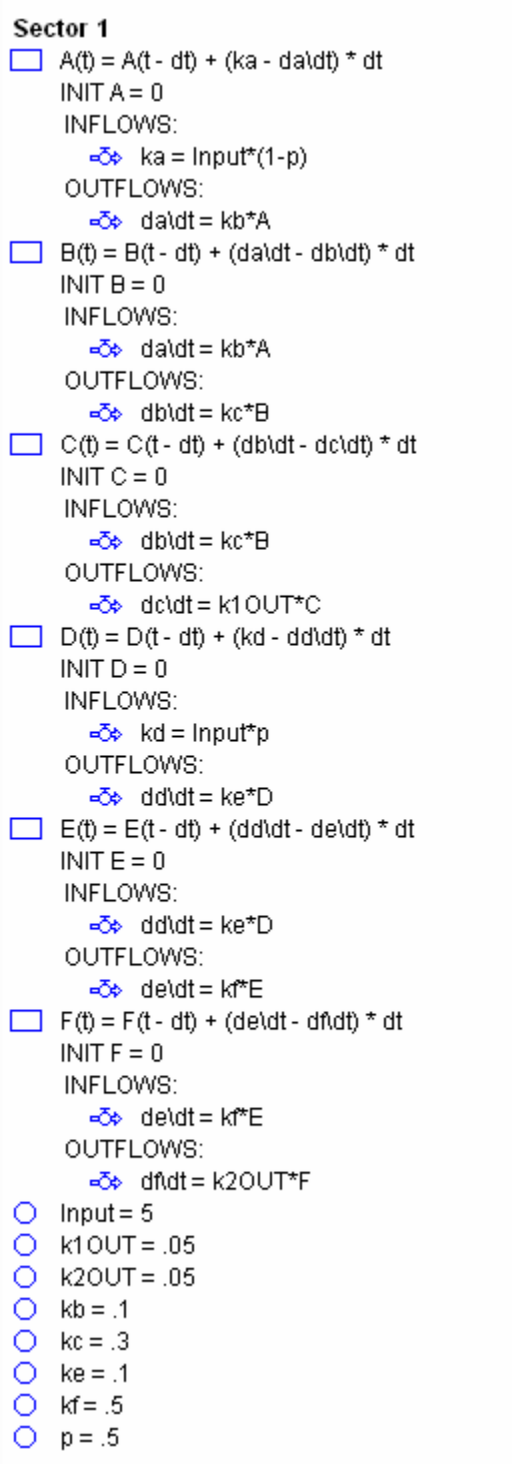
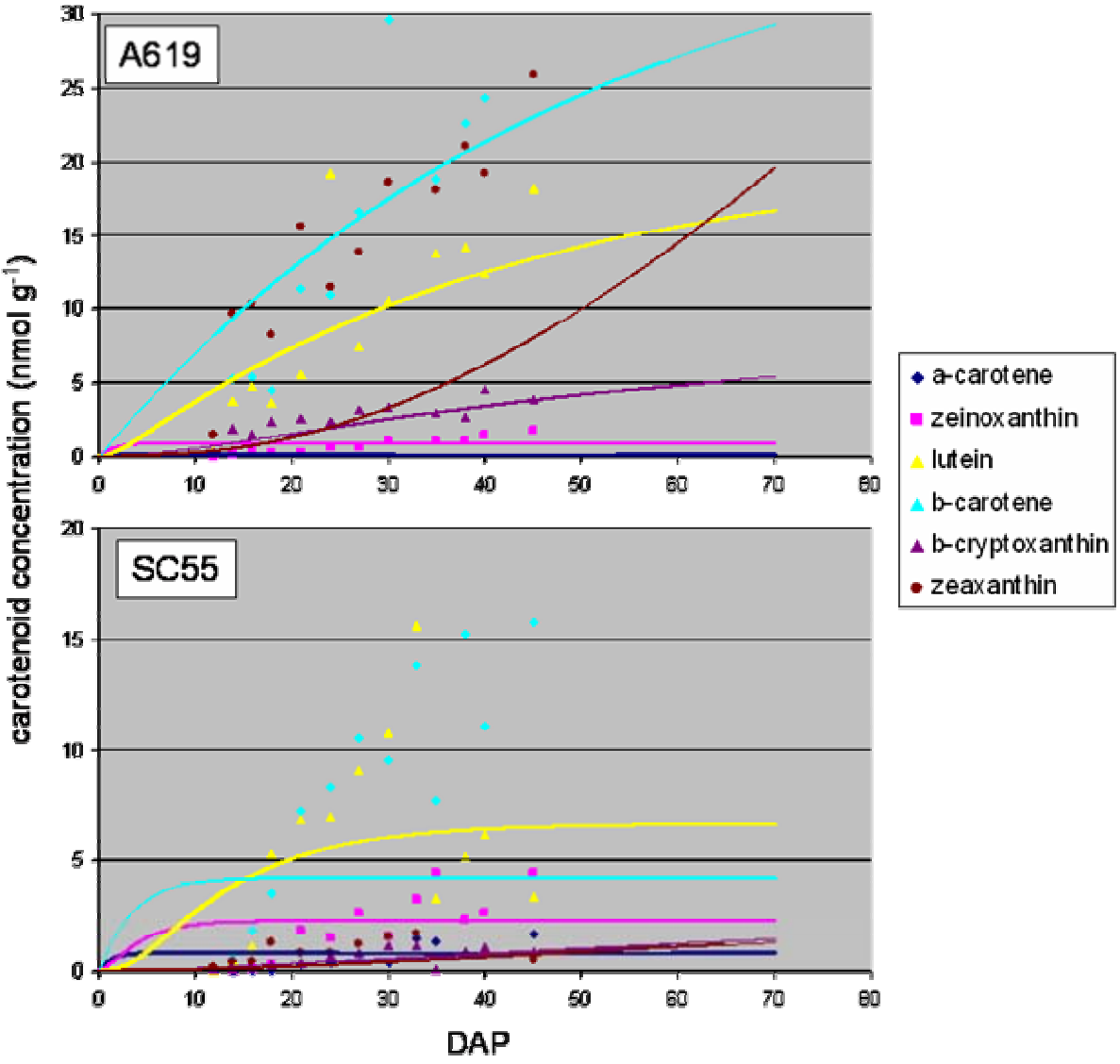
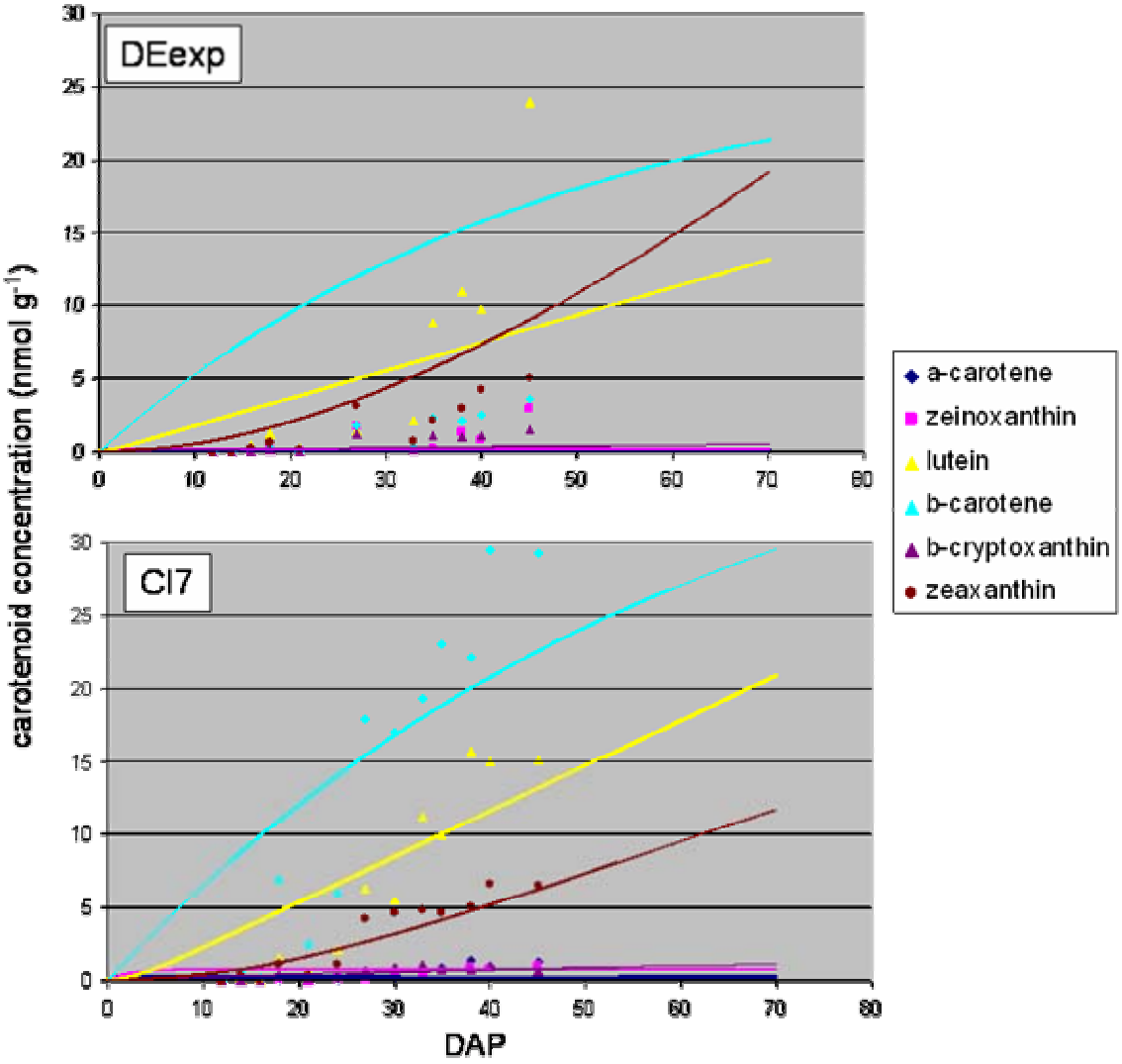


Figure I-7: Comparison of Pathway Chain Model predictions and observed carotenoid concentrations by genotype



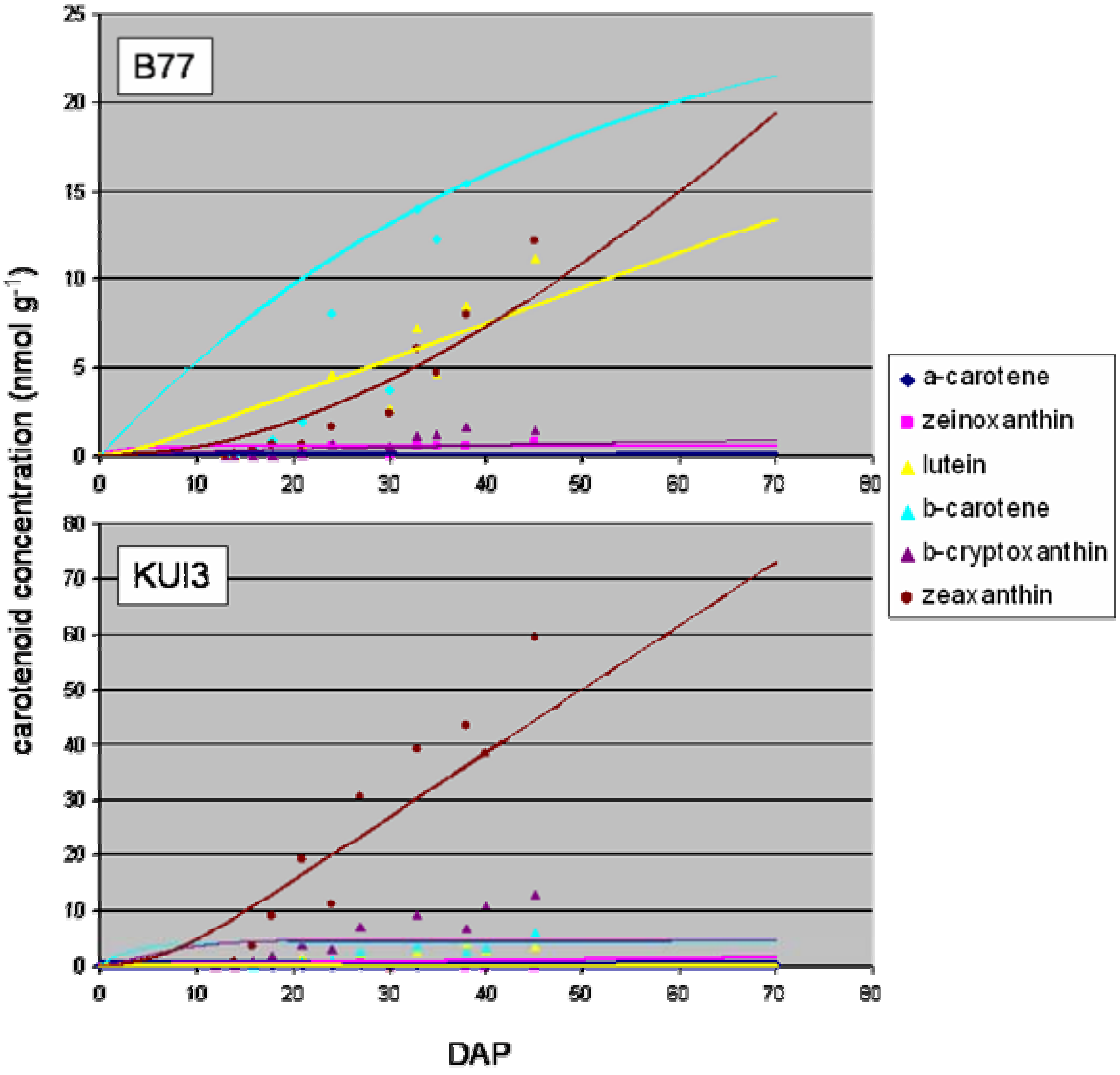
Predicted values are shown in lines; observed mean carotenoid concentrations are shown in scatter plot.

Figure I-7:, continued



Predicted values are shown in lines; observed mean carotenoid concentrations are shown in scatter plot.

Figure I-7: continued



Predicted values are shown in lines; observed mean carotenoid concentrations are shown in scatter plot.