ORIGINAL PAPER

Major QTL for carrot color are positionally associated with carotenoid biosynthetic genes and interact epistatically in a domesticated × wild carrot cross

**Brian J. Just · Carlos A. F. Santos ·
Brian S. Yandell · Philipp W. Simon**

Abstract We performed QTL analyses for pigment content on a carotenoid biosynthesis function map based on progeny of a wild white carrot (QAL) which accumulates no pigments × domesticated orange carrot (B493), one of the richest sources of carotenoid pigments—mainly provitamin A α- and β- carotenes. Two major interacting loci, Y and Y_2 on linkage groups 2 and 5, respectively, control much variation for carotenoid accumulation in carrot roots. They are associated with carotenoid biosynthetic genes zeaxanthin epoxidase and carotene hydroxylase and carotenoid dioxygenase gene family members as positional candidate genes. Dominant Y allele inhibits carotenoid accumulation. When Y is homozygous recessive, carotenoids that accumulate are either only xanthophylls in $Y_2\_\_$ plants, or both carotenes and xanthophylls, in $y_2 y_2$ plants. These two genes played a major role in carrot domestication and account for

Communicated by I. Paran.

B. J. Just · C. A. F. Santos · B. S. Yandell · P. W. Simon
Plant Breeding and Plant Genetics Program,
Department of Horticulture, University of Wisconsin,
1575 Linden Drive, Madison, WI 53706, USA

B. J. Just
Seminis Vegetable Seeds, PO Box 249, Felda, FL 33930, USA

C. A. F. Santos
Embrapa-Semi-Arido, CP 23, Petrolina, PE 56300-970, Brazil

B. S. Yandell
Department of Statistics, University of Wisconsin, Madison,
WI 53706, USA

P. W. Simon (✉)
USDA-Agricultural Research Service, Vegetable Crops Research
Unit, University of Wisconsin, Madison, WI 53706, USA
e-mail: Philipp.Simon@usda.ars.gov; psimon@wisc.edu

the significant role that modern carrot plays in vitamin A nutrition.

Introduction

One of the most striking differences between the modern cultivated carrot and the wild carrot (also known as Queen Anne's Lace, QAL) is that the cultivated carrot accumulates high levels of carotenoids in the storage roots giving it its characteristic orange color, while the wild carrot roots accumulate no carotenoid pigments, and consequently are white. Carrot historians believe that the first domesticated carotene carrot was yellow in color (accumulating mainly xanthophylls consisting primarily of lutein) and that the familiar modern orange carrot, that accumulates β- and α-carotene, was selected through this yellow carrot intermediate within the last 500 years (Banga 1957, 1963; Simon 2000).

The genetic control of carotenoid accumulation in carrot roots has been the subject of several research reports. The only quantitative trait loci (QTL) study reporting the inheritance and genetics of the carotenoid content in carrots utilized an intercross between a cultivated orange carrot, B493, and a wild white carrot, QAL (Santos 2001; Santos and Simon 2002). Heritability and number of genetic factors have also been estimated for total carotenes, β-carotene, α-carotene, ζ-carotene, lycopene, and phytoene (Santos and Simon 2006). Heritabilities for each trait were around 90%. The estimated number of factors was 4 for α-carotene, 1–2 for lycopene and total carotenes, and 1 each for β-carotene, ζ-carotene, and phytoene. Interestingly, while several factors account for variation of individual carotenes, most of them were clustered, so that discrete inheritance of as few as 2 major loci separate white from

orange carrot. Interesting too, a two-gene model for orange versus white color was also suggested by several studies involving cultivated white, yellow, and orange carrots (reviewed by Buishand and Gabelman 1979; Simon 1996).

QTL analysis on this data using a single marker stepwise regression approach demonstrated that the same markers were linked to the accumulation of several carotenes, especially in linkage groups 2 and 5 (Santos 2001; Santos and Simon 2002). This clustering was interpreted as due to linkage of genes for enzymes in the same biosynthetic pathway. These two regions interacted epistatically to reduce carotenoid accumulation so that many individuals in the population had white roots with 0 values for all traits. Because of this, all carotenes had non-normal distributions in the F_2 population. Normality of distributions for traits is an assumption in models used to detect QTL using parametric interval mapping (Kruglyak and Lander 1995). However, there are methods available for interval analysis of complex traits that do not fit a normal distribution and they are being incorporated into standard QTL mapping software. Kruglyak and Lander (1995) first proposed an interval analysis based on the Wilcoxon rank sum test. The technique they developed is considered to be robust regardless of the trait distribution. One drawback to this method of QTL mapping is that it does not provide a direct estimate of the phenotypic effect of the QTL; it only tests for the presence of a QTL. Another approach to nonparametric QTL analysis is an extension of the Kruskal–Wallis test statistic and is analogous to Kruglyak and Lander's (1995) extension of the Wilcoxon rank sum test (Broman 2003). This method does not provide means for estimating the effect of a QTL either. Zou et al. (2003), however, developed a rank-based nonparametric technique to estimate both positions and effects of QTL. Another specific type of non-normality is evident in binary traits. A binary trait has two possible values (e.g. susceptible and resistant) and is heritable but not necessarily in a simple single gene fashion. The basis for binary traits is presumed to be polygenic inheritance of genes that affect a physiological threshold that, once reached, will cause the phenotype to be expressed (Wright 1934). Techniques have been developed (Xu and Atchley 1996; Yi and Xu 1999; Broman 2003) that allow for QTL mapping of binary traits.

While QTL analysis attributes no particular function to loci identified, candidate gene analysis makes inferences about gene function. The identification of candidate genes facilitates the cloning of genes that may produce a marked phenotypic effect. It allows researchers to use information about biochemical pathways or information coming from other organisms to postulate specific underlying genes for a specific trait. Pflieger et al. (2000) classify two types of candidate genes (CGs). Functional CGs are cloned genes that are thought to affect a trait based on knowledge of the pathway or information from other organisms. Positional CGs are mapped genes that co-localize with a trait on a genetic map. This approach can be a powerful tool in establishing correlation but is not an end itself. Validation of positional and functional CGs is essential and can be done with physiological analyses, genetic transformation, and/or complementation studies. This technique has been useful in plants and was recently used in carrot to identify invertase isozyme II as the gene underlying the *Rs* locus that controls the accumulation of reducing sugars in carrot roots (Yau and Simon 2003).

Several crop species have tissue-specific carotenoid mutants and quantitative variants analogous to those of carrot. In many cases, especially for single gene mutations with large effects, the underlying genes have been demonstrated to be variants of some of the structural genes in the carotenoid pathway. For example, the *y1* locus in *Zea mays* is responsible for the difference between yellow and white kernel color. This locus is important in developing white, yellow, and bicolor sweet corn (Marshall and Tracy 2003). Map-based cloning, followed by sequence and expression analysis revealed that the *y1* locus in maize is phytoene synthase and that the differences in observed phenotypes are probably due to a transposable element insertion in the promoter of the gene that affects expression during development without causing complete dysfunction (Buckner et al. 1990; Buckner et al. 1996; Palaisa et al. 2003). Similarly, large effect genes and QTL in *Capsicum* and *Lycopersicon* have been demonstrated to be caused by polymorphisms in carotenoid biosynthetic structural gene loci. Table 1 summarizes associations between carotenoid structural gene loci and phenotypic loci for *Capsicum*, *Lycopersicon* and *Zea mays*. With varying degrees of evidence, it appears that it is quite common for carotenoid biosynthetic structural genes to be associated with drastic or moderate effects on pigment profiles in fruits of some species. However, it should be noted that there are also several examples of major genes and QTL affecting carotenoid accumulation in cauliflower, pepper and tomato that appear to not be due to polymorphisms in carotenoid structural genes or regulatory regions, and while genes have been cloned, their function in conditioning accumulation of carotenoids is not well understood (Ben Chaim et al. 2001; Li et al. 2001, 2006; Liu et al. 2004).

By extension, we hypothesize that some of the genetic control for root pigments can be associated with carotenoid biosynthetic structural genes in carrot. While a two-gene model for orange versus white root color has been proposed, naming these genes as Y and Y_2 (Buishand and Gabelman 1979; Simon 1996), there has been no analysis placing the carotenoid biosynthetic genes on the QTL map including Y and Y_2 before. Here we present a QTL analysis of carrot carotenoids performed on F_2 derivatives of an

Table 1 Major genes and QTL that affect pigmentation and are closely linked to carotenoid biosynthesis structural genes

Genus	Mutant	Phenotype	Gene	Evidence	Reference
Major genes					
Lycopersicon	*Del*	Orange fruit	Lycopene ε-cyclase	1, 5	Ronen et al. (1999)
Lycopersicon	*r*	Pale yellow fruit	Phytoene synthase	1, 3	Fray and Grierson (1993)
Lycopersicon	*Beta*	Orange fruit	Lycopene β-cyclase	2, 3	Ronen et al. (2000)
Lycopersicon	*og, og*[c]	Deep red fruit	Lycopene β-cyclase	1, 3	*Ibid*
Lycopersicon	*tangerine*	Orange fruit	Carotenoid isomerase	2, 4	Isaacson et al. (2002)
Capsicum	*y*	Red fruit	Capsorubin-capsanthin synthase	1	Lefebvre et al. (1998)
Capsicum	*c2*	Red fruit	Phytoene synthase	1	Huh et al. (2001)
Zea	*y1*	Yellow kernel	Phytoene synthase	2	Buckner et al. (1996)
Quantitative trait loci					
Capsicum	*rc3.1*	Increased red hue	β-Carotene hydroxylase	1	Thorup et al. (2000)
Capsicum	*rl3.1*	Increased red lightness	β-Carotene hydroxylase	1	*Ibid*
Capsicum	*pfc6.1*	Increased red chroma	Phytoene synthase	1	*Ibid*
Lycopersicon	*fc1.1*	Increased color	ζ-Carotene desaturase	6	*Ibid*
Lycopersicon	*fc2.2*	Increased color	Zeaxanthin epoxidase	6	*Ibid*
Lycopersicon	*fc4.2*	Increased color	Geranylgeranyl pyrophosphate synthase	6	*Ibid*
Lycopersicon	*fc10.1*	Increased color	Lycopene β-cyclase	6	*Ibid*

1 mapping, *2* map-based cloning, *3* transgenic assay, *4* bacterial complementation, *5* transcript quantification, *6* comparative mapping

intercross between B493 and Queen Anne's Lace (wild carrot, QAL) as described by Santos (2001) and Santos and Simon (2002) that combines QTL maps for carrot carotenoids with maps containing 24 genes putatively coding for enzymes in the carotenoid pathway (Just et al. 2007).

The purpose of the study is two-fold: (1) To test the hypothesis that genes coding for enzymes in the carotenoid biosynthetic pathway (functional CGs) are positional CGs for QTL. (2) To apply nonparametric, binary, and two dimensional genome scans to detect QTL in a population in which none of the traits are normally distributed, to gain a better understanding of the genetic basis of differences underlying pigment accumulation patterns in white-rooted Queen Anne's lace and a modern orange inbred (B493).

Materials and methods

Plant material, DNA and RNA extraction

The B493 × QAL F_2 population (Santos 2001; Santos and Simon 2002) was used for mapping STS (sequence tagged site) markers. B493 is a dark orange USDA inbred carrot (Simon et al. 1990) and QAL is a white wild carrot (*D. carota* var. *carota*). The population evaluated was derived from crossing a single B493 plant with a single QAL plant in Madison, WI, in 1989. A single F_1 plant was self-pollinated to produce the F_2 generation used for mapping. A total of 183 F_2 plants grown in field conditions in 1998 and 59 plants in 2000 were included in this study.

Additional plants of the same F_2 population and both parental stocks were also grown in 2001 (27 F_2 plants) and 2007 (21 F_2 plants). F_3 populations were derived from plants of the F_2 population evaluated in this study by self pollinating F_2 plants. Segregation patterns of AFLP and carotenoid gene STS markers with high LOD values in interval analysis and epistasis/joint analysis were scored in F_3 populations and also in two other mapping populations (Vivek and Simon 1999; Boiteux 2000; Boiteux et al. 2000) to confirm linkage relationships. Total genomic DNA was extracted from freeze-dried leaf tissue following the protocol of Doyle and Doyle (1990) with modifications by Boiteux et al. (1999).

Carotene extraction and analysis

Carrot storage root samples were lyophilized and major carotenoids were quantified by high-performance liquid chromatography (HPLC), as described by Simon and Wolff (1987). Lutein quantification was based upon modification of this procedure (Craft and Soares 1992; Surles et al. 2004) and concentrations were estimated.

QTL analysis

R/qtl software was used (Broman et al. 2003) for QTL analysis. For the single QTL model interval analyses, genotype probabilities were calculated with a step value of 1 over the entire linkage map. The "scanone" command was used with nonparametric (model = "np") and binary (model = "bin"). For the nonparametric analysis, the μg of pigment per gram

Table 2 Means and standard errors of each generation for total carotenoids, phytoene, ζ-carotene, lycopene, α-carotene, and β-carotene (μg/g fresh weight) in the carrot B493 × QAL cross, its parents, F_1 and F_2 (Santos and Simon 2006)

Trait	Generation					
	B493 (1998; $n = 48$)	B 493 (2007; $n = 17$)	QAL (1998; $n = 34$)	F_1 (1998; $n = 3$)	F_2 (1998; $n = 178$)	F_2 (2000; $n = 59$)
Total carotenoids	947 ± 34	1027 ± 41	13 ± 4	3 ± 1	64 ± 9	80 ± 23
Phytoene	500 ± 24	573 ± 44	0	0	16 ± 4	50 ± 17
ζ-Carotene	188 ± 12	199 ± 10	0	0	8 ± 3	19 ± 6
Lycopene	18 ± 2	21 ± 5	0	0	2 ± 1	2 ± 1
α-Carotene	266 ± 11	301 ± 17	0	0	5 ± 2	7 ± 5
β-Carotene	461 ± 20	520 ± 21	0	0	14 ± 4	39 ± 12

of fresh tissue for each of the pigments as calculated in β-carotene equivalents were used except for xanthophylls for which maximum absorbance units of the xanthophyll peak on the chromatogram was used. For the binary analyses the traits were re-coded as 0 for all samples for which the calculated pigment concentration was 0 and 1 for samples that had non-zero values for the trait. Since for total carotenes there were very few 0 values, and the application of binary trait analysis requires a relatively large number of 0 values, this analysis was not performed for that trait. For both binary and nonparametric analyses, approximate 95% support intervals were estimated by using the "one LOD rule" of Lander and Botstein (1989) ("lodint" command in R/qtl).

Two-dimensional scans were initially performed using a marker regression approach ("scantwo" with method = "mr-imp") with a single imputation for missing data. Some pairs of linkage groups were reexamined using an interval mapping technique based on the Haley–Knott regression (method = "hk"). The analysis was performed on a 1 cM grid (step = 1).

Post hoc analyses of variance

For each significant region detected in the interval analyses, the marker nearest to the LOD peak was used to perform a post hoc analysis of variance ("aov" in R) of marker class means to get an indication of allelic effects of the QTL (Ihaka and Gentleman 1996). When a codominant marker was investigated, pairwise t tests of the three marker classes were performed. Post hoc tests were also performed for each significant maximum pair of loci for each pair of linkage groups being compared. Pairwise t tests were performed for the different marker class combinations of all pairs of loci in joint and epistatic post hoc tests.

Genome wide significance determination

Genome wide significance thresholds for LOD scores were determined by permutation tests as suggested by Churchill

and Doerge (1994). For each analysis of both single QTL and two dimensional genome scans, 10,000 permutations were performed in R/qtl ("n.perm" = 10000). LOD scores significant at the 0.90, 0.95, and 0.99 levels were calculated empirically from this data.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under the accession nos. DQ192183 to DQ192205, DQ222429, and DQ222430.

Results

Carotene analysis of parents, F_1, and F_2 generations of B493 × QAL

Means for phytoene, ζ-carotene, lycopene, β-carotene, α-carotene, and total carotenoids of parental, F_1, and F_2 generations of B493 × QAL as determined by Santos and Simon (2006) and subsequent years are presented in Table 2. A high incidence of white roots in the F_2 generation resulted in a significant kurtosis and lack of normality (Santos 2001; Santos and Simon 2006). Ranges of carotenoid content in parental stocks and smaller F_2 populations in 2001 and 2007 were within 12% of values in Table 2 (data not presented).

Genome-wide LOD thresholds

Mean LOD thresholds for all traits corresponding to α levels of 0.10, 0.05, and 0.01 were 3.1, 3.5, and 4.4, respectively. Mean LOD thresholds corresponding to α levels of 0.10, 0.05, and 0.01 for epistasis and joint analyses are presented in Table 3. For the joint analyses LOD thresholds for the 0.10 level ranged from 13.5 to 36.2, for the 0.05 threshold they ranged from 15.1 to 38.3 and for the 0.01 level they ranged from 18.0 to 118.2. For the epistasis analyses LOD thresholds for the 0.10 level ranged from 11.0 to 31.0, for the 0.05 threshold they ranged from 12.4 to 34.5 and for the 0.01 level they ranged from 15.1 to 113.2. Values for parents B493 and QAL were very similar.

Table 3 Genome-wide LOD thresholds determined for each joint and interaction analysis for $\alpha = 0.10$, 0.05 and 0.01 in the carrot B493 × QAL cross

Trait	Joint analysis			Interaction analysis		
	LOD thresholds for α =			LOD thresholds for α =		
	0.10	0.05	0.01	0.10	0.05	0.01
QAL						
Total carotenes	13.6	15.1	18.1	11.0	12.4	15.1
Phytoene	16.5	18.2	22.8	14.0	15.6	20.2
ζ-Carotene	21.1	23.1	27.9	19.1	20.5	25.0
Lycopene	36.2	38.3	118.2	30.9	34.5	113.2
α-Carotene	29.3	35.0	45.7	26.9	31.8	42.4
β-Carotene	15.5	17.0	20.5	12.7	14.3	17.3
Xanthophyll	15.2	16.9	19.5	13.2	14.4	17.4
B493						
Total carotenes	13.5	15.2	18.0	11.2	12.6	15.2
Phytoene	16.6	18.2	22.9	14.2	15.8	20.3
ζ-Carotene	13.6	15.2	17.8	11.2	12.5	15.2
Lycopene	34.9	38.2	118.0	31.0	34.5	112.6
α-Carotene	29.2	33.9	45.9	26.9	31.0	42.8
β-Carotene	15.6	17.1	20.8	12.9	14.5	17.5
Xanthophyll	15.2	17.0	19.6	13.3	14.7	17.6

Nonparametric and binary interval mapping

Tables 4 and 5 list the QTL detected in this study using nonparametric and binary interval mapping, the calculated positions of their peaks with 95% support intervals, as well as the results of post hoc tests for the nonparametric and binary interval mapping. Four regions were identified as having significant QTL in binary and nonparametric interval analyses for total carotenes, phytoene, ζ-carotene, lycopene, α-carotene, β-carotene, and xanthophylls. QTL on linkage groups 2 (two separate QTL) and 5 were particularly common and the closest markers associated with carotenoid QTL are clustered (Fig. 1). In all cases where post hoc tests were significant, the A (QAL) allele was associated with a lesser amount of pigment than the B (B493 allele).

For all traits except xanthophylls, there were significant QTL ($p < 0.1$ after genome wide significance adjustment using permutation tests). QTL were detected in a fairly narrow region on linkage group 5 in both parents (Fig. 1) and in both the binary and nonparametric analyses. For each of these traits Y2mark, a co-dominant marker common to both parental maps (Bradeen and Simon 1998), was included in the approximate 95% support intervals (*sensu* Lander and Botstein 1989). On the QAL map all of the post hoc analyses were performed on a dominant marker for the linkage group 5 QTL, but it was clear that in all cases, the QAL allele (A-) had lower group means than homozygotes for the B493 allele (BB) (see Table 4). Post hoc tests for markers near QTL detected for phytoene in the binary and α-car-

otene in both the binary and nonparametric tests were nonsignificant ($p > 0.10$). Post hoc tests for total carotenes in the nonparametric analysis, ζ-carotene in the binary analysis, and β-carotene in both the nonparametric and binary analyses were performed on Y2mark, a codominant marker. In all cases the AA and AB genotype means were not significantly different from one another, but were significantly lower ($p < 0.05$) than the BB genotype in pairwise t tests (see Table 5), as is typical for dominant gene action. The closest STS marker for a carotenoid biosynthetic gene on linkage group 5 was zeaxanthin epoxidase (ZEP) which mapped just outside the 95% support intervals for most traits on both maps.

On linkage group 2 of the QAL map significant QTL were found for β-carotene in the binary ($p < 0.05$) and nonparametric tests ($p < 0.1$), α-carotene in the binary ($p < 0.05$) and nonparametric ($p < 0.05$) tests, total carotenes ($p < .01$), and xanthophylls for both the binary ($p < 0.01$) and nonparametric ($p < 0.01$) tests (Fig. 1). For each of these QTL, post hoc tests performed on a dominant linked marker were significant with the A- marker class having a lower mean than the BB marker class (see Table 4). Two of the STS markers for putative carotenoid biosynthesis genes were located in this region of high QTL density. Epsilon ring carotene hydroxylase (CHXE) and a copy of nine-*cis* epoxycarotenoid dioxygenase, a carotenoid cleavage enzyme, (NCED2) were both in the 95% support intervals for the α-carotene binary, β-carotene nonparametric and β-carotene binary analyses. CHXE was also in the 95% support interval for the lutein binary analysis.

Table 4 Interval analysis of the carrot QAL map displaying approximate 95% support intervals for linkage group maxima and corresponding post hoc statistical tests for marker means

Analysis	LG[a]	Peak maximum position (cM)[b]	LOD[c]	Approximate 95% support intervals[d]		Closest marker	Group means for nearest marker[e]				R^2	p
				Lower extreme (cM)	Upper extreme (cM)		AA	AB	A-	BB		
QAL												
Total carotenes nonparametric	2	45	11.6***	39	51	aggctg078			40.7	131.6	0.11	0.00
Total carotenes nonparametric	5	102	5.7**	96	115	aagcag233			28.5	151.7	0.23	0.00
Phytoene nonparametric	5	103	9.8***	98	108	aagcag233			0.5	57.5	0.19	0.00
Phytoene nonparametric	8	116	3.9**	94	143	gggcat322			9.9	34.2	0.03	0.02
Phytoene binary	5	104	8.3***	97	112	aagcag233	See phytoene nonparametric					
Phytoene binary	8	119	3.0*	97	157	gggcat322	See phytoene nonparametric					
ζ-Carotene nonparametric	5	102	7.4***	98	108	aagcag233			0.5	27.9	0.14	0.00
ζ-Carotene nonparametric	8	111	4.0**	93	123	gggcat322			5.2	16.2	0.02	0.06
ζ-Carotene binary	4	11	2.9*	0	25	gggcat119			7.30	10.9	0.00	0.53
ζ-Carotene binary	5	103	6.4***	96	112	aagcag233	See ζ-carotene nonparametric					
ζ-Carotene binary	8	115	2.9*	94	154	gggcat322	See ζ-carotene nonparametric					
Lycopene nonparametric	5	102	11.2***	98	107	aagcag233			0.0	6.2	0.07	0.00
Lycopene nonparametric	8	112	3.6*	90	124	gggcat322			2.1	0.9	0.00	0.56
Lycopene binary	5	103	9.9***	97	109	aagcag233	See lycopene nonparametric					
α-Carotene nonparametric	2	47	4.0**	29	57	aggctg078			0.1	18.18	0.07	0.00
α-Carotene nonparametric	5	104	6.4***	96	116	aagcag233			0.0	17.2	0.06	0.00
α-Carotene binary	2	45	3.0**	1	85	aggctg078	See α-carotene nonparametric					
α-Carotene binary	5	105	5.4***	94	115	aagcag233	See α-carotene nonparametric					
β-Carotene nonparametric	2	29	3.24*	1	85	ggacaa272			6.0	33.5	6	0.00
β-Carotene nonparametric	5	103	8.7***	98	114	aagcag233			0.7	48.8	0.19	0.00
β-Carotene nonparametric	8	115	3.4*	93	151	gggcat322			6.5	34.4	0.06	0.00
β-Carotene binary	2	29	3.4**	7	84	ggacaa272	See β-carotene nonparametric					
β-Carotene binary	5	104	7.5***	97	114	aagcag233	See β-carotene nonparametric					
Lutein nonparametric	2	45	26.0***	41	48	aggctg078			2.134	26.760	0.30	0.00
Lutein binary	2	48	16.9***	42	72	aggctg080			2.776	23.670	0.23	0.00
Lutein binary	2	61	16.9***	55	67	gggctg146	0.2	–	10.6	–	0.04	0.01

[a] Linkage group of significant peak

[b] Position on linkage group of significant peak displayed in cM

[c] LOD score and significance after determination by 10,000 permutation tests. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

[d] Approximate 95% support intervals for QTL location as determined by 'One LOD' rule. Displayed in terms of positions of ends of intervals

[e] Group means for marker classes of marker closest to LOD peak. *AA* homozygous QAL allele, *AB* heterozygous, *A-* homozygous QAL or heterozygous, *BB* homozygous B493 allele. Where more than two genotypes are considered, means marked with the same letter were not significantly different from one another in pairwise *t* tests

Linkage group 2 of the B493 map had significant QTL for phytoene and α-carotene in the binary analysis ($p < 0.1$) only, while β-carotene, total carotenes, and xanthophylls had significant QTL in both the binary ($p < 0.01$) and nonparametric ($p < 0.1$) analyses. Post hoc tests for phytoene, α-carotene, and β-carotene in the binary analysis and for β-carotene in the nonparametric tests were non-significant. For total carotenes and xanthophyll (nonparametric) post hoc tests were performed on NCED2. These tests were significant, with the AA and AB marker class means similar to each other and lower than the BB marker class, indicating dominance for this QTL region. For the binary xanthophyll analysis, the post hoc test was performed on a dominant marker with class AA significantly lower than B- (see Table 5). For all traits, there was at least one fairly large peak on linkage group 2. In some cases, however, the peak maximum did not meet the threshold set by our permutation tests. Two of the STS markers representing carotenoid biosynthetic genes were in or close to this region on linkage group 2. CHXE was in the 95% support interval in the α-carotene binary, β-carotene nonparametric, and lutein binary analyses. NCED2 was in the 95% support interval

Table 5 Interval analysis of the carrot B493 map displaying approximate 95% support intervals for linkage group maxima and corresponding post hoc statistical tests for marker means

Analysis	LG[a]	Peak maximum position (cM)[b]	LOD[c]	Approximate 95% support intervals[d]		Closest marker	Group means for nearest marker[e]				R^2	p
				Lower extreme (cM)	Upper extreme (cM)		AA	AB	B-	BB		
B493												
Total carotenes nonparametric	2	75	9.7***	66	82	NCED2	38.2 a	39.3 a	–	119.5 b	0.10	0.00
Total carotenes nonparametric	5	91	4.71**	71	105	Y2mark	26.9 a	39.7 a	–	175.8 b	0.23	0.00
Phytoene nonparametric	5	91	7.9***	72	105	Y2mark	0.0 a	5.4 a	–	67.4 b	0.2	0.0
Phytoene binary	2	0	3.2*	0	46	aaccta353	0.0	–	20.7	–	0.02	0.07
Phytoene binary	4	32	3.4**	0	37	aggcta380	0.0	–	22.0	–	0.03	0.02
Phytoene binary	5	78	6.8***	66	105	gggcaa229	0.0	–	18.8	–	0.02	0.07
ζ-Carotene nonparametric	5	78	5.2***	60	105	gggcaa229	0.0	–	9.3	–	0.01	0.12
ζ-Carotene binary	5	91	5.58***	64	105	Y2mark	0.0 a	1.5 a	–	36.7 b	0.18	0.00
Lycopene binary	4	32	3.0*	0	38	aggcta380	0.0		2.6		0.01	0.20
α-Carotene nonparametric	5	83	5.7***	63	105	acactt135	0.0	–	4.1	–	0.01	0.35
α-Carotene binary	2	2	2.8*	0	91	aaccta353	0.0	–	5.6	–	0.01	0.35
α-carotene binary	5	81	5.2***	61	105	gggcaa229	0.0	–	6.2	–	0.01	0.30
β-Carotene nonparametric	2	3	3.6*	0	87	aaccta353	0.0	–	18.3	–	0.02	0.06
β-Carotene nonparametric	5	98	7.2***	73	105	Y2mark	0.0 a	4.1 a		59.7 b	0.21	0.00
β-Carotene binary	2	0	4.0***	0	45	aaccta353	See β-carotene nonparametric					
β-Carotene binary	4	32	3.9**	27	37	aggcta380	0.0	–	20.1	–	0.03	0.02
β-Carotene binary	5	95	6.3***	66	105	Y2mark	See β-carotene nonparametric					
Lutein nonparametric	2	75	22.2***	65	80	NCED2	0.3 a	1.3 a	–	25.9 b	0.34	0.00

[a] Linkage group of significant peak

[b] Position on linkage group of significant peak displayed in cM

[c] LOD score and significance after determination by 10,000 permutation tests. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$

[d] Approximate 95% support intervals for QTL location as determined by 'One LOD' rule. Displayed in terms of positions of ends of intervals

[e] Group means for marker classes of marker closest to LOD peak. *AA* homozygous QAL allele, *AB* heterozygous, *B-* homozygous B493 or heterozygous, *BB* homozygous B493 allele. Where more than two genotypes are considered, means marked with the same letter were not significantly different from one another in pairwise t tests

for the linkage group 2 QTL in the total carotene nonparametric, α-carotene binary, β-carotene nonparametric and the lutein nonparametric analyses.

On QAL linkage group 4 a single QTL for ζ-carotene was detected ($p < 0.1$) using the binary method, though the post hoc test was non-significant. On B493 linkage group 4 we found significant QTL for lycopene ($p < 0.1$) in the nonparametric analysis, phytoene ($p < 0.05$) in the binary analysis, and β-carotene ($p < 0.05$) in the binary analysis. For lycopene, the post hoc test was non-significant. For the other two traits the tests were significant with the AA genotype means significantly lower than the BB means.

QTL were detected on QAL linkage group 8 (Table 4). They were for phytoene in the nonparametric ($p < 0.05$) and binary ($p < 0.1$) analyses, ζ-carotene in the nonparametric ($p < 0.05$) and binary ($p < 0.1$) analyses, lycopene in the nonparametric analysis ($p < 0.1$), and for β-carotene in the nonparametric analysis ($p < 0.1$). Post hoc tests were significant for the β-carotene nonparametric and phytoene

nonparametric analyses. For both of these, genotype means were lower for the A-genotype than the BB genotype.

Two-dimensional marker regression

Two dimensional genome scans of all pairs of markers revealed strong evidence for epistasis for total carotenes on both the QAL and B493 maps, and for β-carotene. Tables 6 and 7 display the results of the epistatic and joint (two-QTL model) tests along with the appropriate post hoc tests. Linkage relationships among markers on linkage groups 2 and 5 with significant LOD scores in the two-way QTL scan are noted in Fig. 1. On both maps the major interaction for total carotenes was between the CHXE region on linkage group 2 and the Y2mark region on linkage group 5. Since these are both codominant markers common to both maps, the post hoc analysis was the same: the BB/BB genotype (homozygous B493 allele) accumulated significantly higher carotenes than all the others, with levels nearly five fold
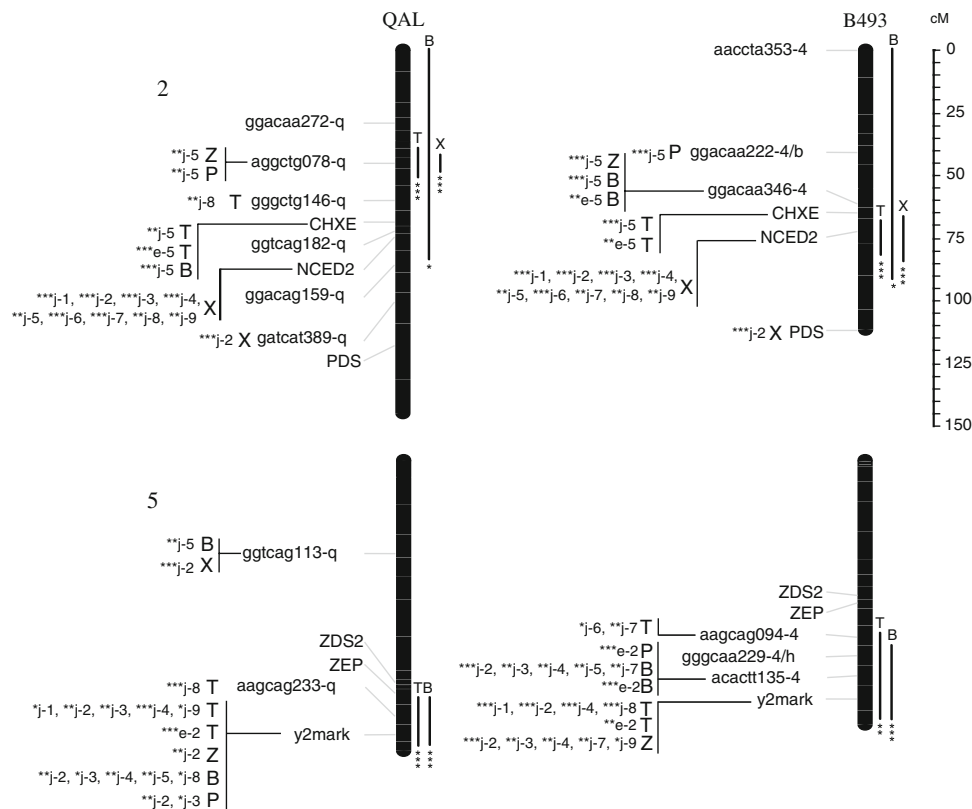
Fig. 1 Linkage maps with the positions of all QTLs detected on linkage groups 2 and 5 of carrot. QAL and B493 linkage maps are displayed side by side. The numbers 2 and 5 on the *left* are the linkage group number. Numbering assigned by Santos (2001) and Santos and Simon (2002) was retained. *Bars* to the *right* of the linkage group indicate the 1 LOD confidence intervals for QTLs detected in nonparametric analyses for total carotenes (*T*), β-carotene (*B*), and xanthophylls (*X*). *Stars* indicate the significance level after genome wide LOD score adjustment by permutation tests (*p < 0.10, **p < 0.05, ***p < 0.01). Letters "j" or "e" to the *left* of a marker name indicate that the marker was detected as a joint or epistatic QTL in the marker regression based two-QTL analysis, respectively, for phytoene (*P*), zeta- carotene (*Z*), total carotenes (*T*), β-carotene (*B*), and xanthophylls (*X*). The number next to the letter corresponds to the linkage group on which the other member of the pair is located

higher than the next marker class. Figure 2f graphically presents the nature of the interaction for total carotenes. This interaction explains 57% of the variation for total carotenes. Figure 2a–e displays the effect of these two markers on ζ-carotene, α-carotene, β-carotene, phytoene, and lycopene, respectively. While this pair of loci was not significant in the initial analyses for ζ-carotene, α-carotene, β-carotene, phytoene, and lycopene, it appears that the effects of these loci are similar to those on total carotenes. Additionally, the interaction term is highly significant ($p < 0.000$) for all these traits in a two-way analysis of variance on CHXE on linkage group 2 and Y2mark on linkage group 5 (data not shown). The LOD thresholds for significance in these two-dimensional scans are very high. Larger population sizes or segregation of these loci on a more uniform background will be needed to confirm the epistatic effects for these traits.

In the joint analyses, several other pairs of loci were significant on both maps. The significant pairs from the QAL map are presented on Table 6 with their corresponding

LOD scores, significance level (indicated by asterisk), and post hoc tests. For total carotenes Y2mark was significant in combination with five other markers including ZDS1, a STS marker for zeta-carotene desaturase on linkage group 4. A marker adjacent to Y2mark, aagcag233, was indicated in a joint QTL pair for total carotenes. Lastly, for ζ-carotene, Y2mark was jointly significant with one other marker. Y2mark was also jointly significant with five other markers for β-carotene (also including ZDS1 on linkage group 4) and two other markers for phytoene. For xanthophyll, a marker on linkage group 2, ggtcag182, was jointly significant with eight other markers and NCED2 on linkage group 2 (adjacent to ggtcag182) was significant with one other marker. With only two exceptions, all B allele markers were associated with higher levels of pigment than were A alleles. These exceptions were acccta508 (QAL linkage group1) and ggtcag113 (QAL linkage group 8) for which the A alleles were associated with more xanthophylls in the presence of BB alleles at ggtcag182 (on linkage group2).

Fig. 2 a–f Interaction plots for
each individual carotene and to-
tal carotenoid content at CHXE
(ε-ring carotene hydroxylase)
on linkage group 2 and Y2mark
on linkage group 5
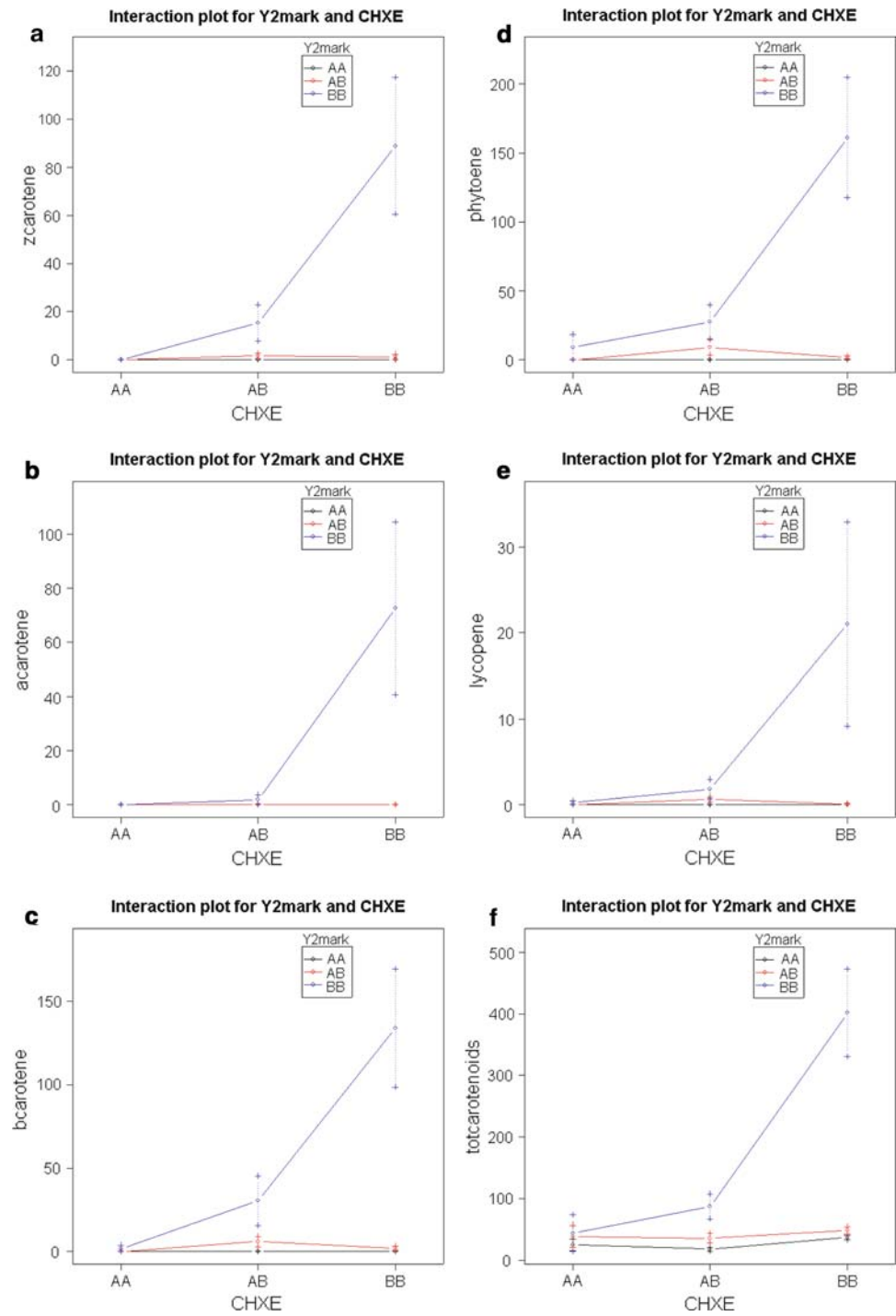
Significant pairs of joint markers on the B493 map are
presented on Table 7, with their LOD scores, significance
and appropriate post hoc test results. All of the pairs for
total carotenes, ζ-carotene, β-carotene, and phytoene
involve Y2mark or another nearby marker on linkage
group 5. ZDS1 on linkage group 4 was jointly significant
with Y2mark for total carotenes and ζ-carotene. It was
also jointly significant with acactt135 an AFLP marker
adjacent to Y2mark. Every significant pair for xantho-

phyll involves NCED2 (linkage group 2) as one of the
members.

To further confirm the interaction between QTL on link-
age groups 2 and 5 two way interval analysis was per-
formed in R/QTL. Results are shown in Fig. 3. This output
was obtained from running a two-way interval analysis on
the QAL map using total carotenes as the trait. Clear joint
and epistatic effects are observed between linkage groups 2
and 5 with LOD scores of >20 and >15, respectively.
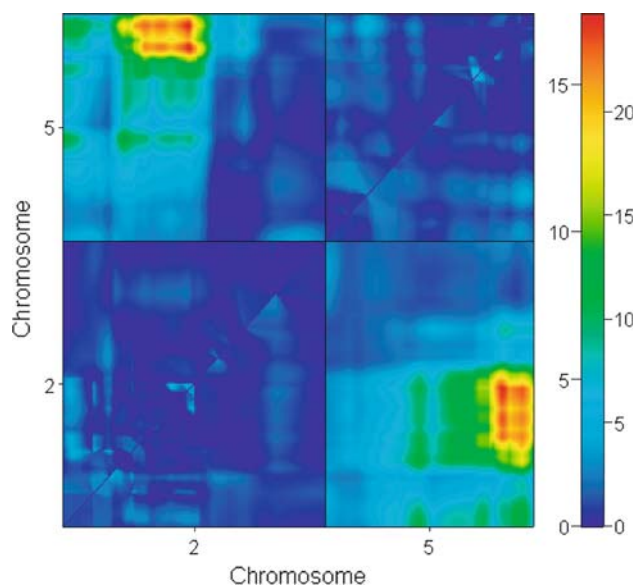
Fig. 3 R/qtl scantwo output of total carotenes data on linkage groups 2 and 5. The *lower right hand corner* is a graphical representation of joint LOD score for each pair of loci. The LOD scores to the *right* of the *scale bar* correspond to this portion of the graph. The *upper left hand corner* of the graph represents the interaction LOD score for each pair of loci. The LOD scores to the *left* of the *scale bar* correspond to this portion of the graph

Marker confirmation

AFLP and STS markers in Fig. 1 were confirmed to be linked in pairwise combination evaluations in at least one F_3 population derived from the B493 × QAL cross for all markers, and in at least one of the two additional mapping populations (Vivek and Simon 1999; Boiteux 2000; Boiteux et al. 2000), with the exception of AFLP marker gggctg146 which segregated in F_3 populations, but did not segregate in either population of these latter mapping populations. Linkages among markers on linkage group 2 included on Fig. 1 were confirmed, as were those on linkage group 5, and distances were similar to those noted (±8 cM) (data not presented).

Discussion

Phenotypic effects of major QTL conditioning carotenoid accumulation in carrots

While mutants affecting carotenoid accumulation in fruit and leaves have been studied in numerous plants, this study is the first in depth genetic analysis of carotenoid accumulation in roots. Our data indicate that there are two major genomic regions located on linkage groups 5 and 2 that control a large portion of the carotenoid profile differences between wild white and cultivated orange carrot. These

regions were implicated in several of the joint QTL and epistatic QTL tests. The interaction is such that high levels of carotenes only accumulate if the genotype is recessive (B493 allele) at each locus. Thus it appears that the genetics of the differences between wild white carrot and orange carrot is similar to the genetics of the differences between cultivated white carrot and orange carrot as studied by Buishand and Gabelman (1979). This study went further, to demonstrate the *Y* gene conditions xanthophyll accumulation without carotenes, when the Y_2 locus is homozygous dominant or heterozygous. Intercrosses between wild carrot and white cultivated carrot yield only white-rooted progeny (accumulating no carotenoids) in F_2 and F_3 generations (Simon, unpublished) to support the idea that the genetic constitution of these two white phenotypes is the same.

Santos and Simon (2002) suggested that clustering of several QTL for various carotenoid traits may be due to clustering of carotenoid biosynthetic genes. This clearly is not the case, since Just et al (2007) demonstrated that putative genes coding for known enzymes in the carotenoid pathway are distributed over all linkage groups and in most cases, tight linkage was not observed. Furthermore, while a few of the QTL we detected here mapped relatively close to carotenoid structural genes, others did not. An alternative explanation for the clustering of QTL into tight regions is pleiotropy combined with the inherent imprecision of QTL analyses with limited numbers of informative individuals (Beavis 1998). Since the traits in this study are all related, as they represent compounds produced at different steps in the same biosynthetic pathway, it is very possible that a single regulatory gene affects many components of the pathway.

Since we knew the allelic constitution of each individual in segregating populations, this study provides the first insights into the similarities and differences of phenotypic effects of the major QTL that control carotenoid accumulation in carrots. For the QTL on linkage group 5, the dominant allele reduces all carotenes (but not xanthophylls) in contrast to the QTL on linkage group 2, where the dominant allele reduces all carotenoids, including both carotenes and xanthophylls. Another way to describe the interaction between the two loci is that recessive alleles at the linkage group 2 QTL allow carotenoids to accumulate. When this QTL is homozygous recessive, the major type of carotenoids that accumulate are either xanthophylls only, when a dominant allele occurs at the QTL on linkage group 5, or both carotenes and xanthophylls, when the linkage group 5 QTL is homozygous recessive.

Role of major QTL for carrot color in carrot domestication and human health

Our results suggest that the major QTL region on linkage group 5 is the Y_2 gene and the major QTL on linkage group

Table 6 Epistasis and joint analysis of the carrot QAL map displaying significant pairs of markers for linkage group maxima and corresponding post hoc statistical tests for marker means

Trait	Marker1 (LG)[a]	Marker2 (LG)[a]	LOD[b]	Marker genotype means[c]														R^2[d]	p1[e]	p2[f]	pint[g]
				AA/AA	AA/AB	AA/BB	AB/AA	AB/AB	AB/BB	BB/AA	BB/AB	BB/BB	A-/AA	A-/AB	A-/BB	A-/A-	BB/A-				
Total	gatctc155(1)	Y2Mark(5)	14.2*	25.0 ab	38.2 ab	43.7 ab	17.8 a	35.2 a	87.1 b	33.5 a	52.9 a	165.0 b	26.8 a	36.2 a	162.7 b			0.21	0.58	0.00	0.95
Total	CHXE(2)	Y2Mark(5)	32.9**[h]							36.7 ab	48.6 ab	401.3 c			67.7 a	36.9 a	63.1 a	0.57	0.00	0.00	0.00
Total	gggctc146(2)	gggcat322(8)	16.2**									347.8 b						0.35	0.00	0.00	0.00
Total	gggcat136(3)	Y2Mark(5)	15.5**							29.0 a	44.2 a	206.2 b	26.3 a	38.9 a	172.4 b			0.24	0.87	0.00	0.82
Total	ZDS1(4)	Y2Mark(5)	22.2***	26.5 a	19.0 a	31.8 ab	25.16 a	35.32 a	293.0 c	32.64 ab	57.54 ab	112.6 b						0.43	0.01	0.00	0.00
Total	acactt079(8)	aagcag233(5)	18.9***							27.9 a	26.8 a	250.7 c	27.6 a	38.0 a	94.9 b	28.5 a	25.5 a	0.36	0.00	0.00	0.00
Total	ggacag428(9)	Y2Mark(5)	14.9*							0.0 a	0.5 a	97.3 b		1.6 a	9.7 a			0.35	0.04	0.00	0.00
ζ-Carotene	aggctg078(2)	Y2Mark(5)	23.1**							0.0 a	2.0 a	133.8 c						0.49	0.00	0.00	0.00
β-Carotene	CHXE(2)	Y2Mark(5)	21.2***	0.0 ab	0.0 a	1.9 ab	0.0 a	6.0 a	30.4 b	0.0 a	2.1 a	101.3 c						0.43	0.00	0.00	0.00
β-Carotene	aggcat116(3)	Y2Mark(5)	15.9*							0.0 a	8.2 a	38.7 b						0.29	0.01	0.00	0.00
β-Carotene	ZDS1(4)	Y2Mark(5)	18.2**	0.0 a	0.0 a	0.8 a	0.0 a	3.0 a	105.9 c	0.0 a				4.9 a	39.3 b			0.37	0.05	0.00	0.00
β-Carotene	ggtcag113(5)	Y2Mark(5)	18.1**							0.0 a	7.0 a	58.2 b	0.0 a	3.6 a	66.6 b			0.23	0.23	0.00	0.84
β-Carotene	gggcat322(8)	Y2Mark(5)	16.4*							0.0 ab	8.25 ab	132.4 c	0.0 ab	2.53 a	27.5 b			0.41	0.00	0.00	0.00
Phytoene	aggctg078(2)	Y2Mark(5)	21.7**							0.0 a	1.2 a	170.1 b	0.0 a	6.7 a	22.2 a			0.47	0.00	0.00	0.00
Phytoene	acactt105(3)	Y2Mark(5)	17.2*							0.0 a	1.1 a	97.8 c	0.0 a	7.1 a	50.5 b			0.23	0.50	0.00	0.06
Lutien	acccta508(1)	ggtcag182(2)	20.4***									13.4 b	1.2 a		33.2 c		1.9 ab	0.42	0.09	0.00	0.00
Lutien	gatcat389(2)	NCED2(2)	22.3***							0.2 a	1.3 a	35.7 c	0.2 a	1.3 a	14.4 b			0.44	0.00	0.00	0.00
Lutien	acccta260(3)	ggtcag182(2)	22.5***									43.4 c	1.2 a		21.1 b		1.7 a	0.46	0.00	0.00	0.00
Lutien	acccac112(4)	ggtcag182(2)	20.5***									25.0 b	1.3 a		25.1 b		1.8 a	0.32	0.59	0.00	0.93
Lutien	ggtcag113(5)	ggtcag182(2)	19.0**									15.5 b	1.2 a		30.0 c		2.2 a	0.37	0.18	0.00	0.02
Lutien	gatcat258(6)	ggtcag182(2)	25.1***									50.4 c	1.1 a		18.9 b		2.3 a	0.51	0.00	0.00	0.00
Lutien	gatcat196(7)	ggtcag182(2)	21.3***									34.4 b	1.1 a		24.7 b		2.8 a	0.38	0.00	0.00	0.21
Lutien	acacaa165(8)	ggtcag182(2)	19.3**									21.0 b	1.2 a		34.2 b		1.9 a	0.39	0.86	0.00	0.02
Lutien	ggacag428(9)	ggtcag182(2)	18.2**									31.8 b	1.5 a		23.7 b		1.5 a	0.33	0.28	0.00	0.21

[a] Marker names for pairs of markers with significant joint or epistatic LOD score in two-way QTL scan followed by their respective linkage groups in parentheses

[b] LOD score for the joint analysis in the two-way QTL scan. Significance after genome wide significance adjustment using permutation tests. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$

[c] Marker genotype group means for significant marker pairs. Genotype for marker 1 appears before "/" and genotype for marker 2 appears after "/". AA homozygous QAL allele, AB heterozygous or heterozygous (for dominant markers), BB homozygous for B493 allele, A- homozygous QAL allele or heterozygous (for dominant markers)

[d] R^2 for model using both marker genotypes as regressors on trait

[e] p value for marker 1

[f] p value for marker 2

[g] p value for the interaction term of marker 1 and marker 2

[h] Marker pair was detected as significant epistatic pair in two-way QTL scan (LOD = 17.5, $p < 0.01$)

Table 7 Epistasis and joint analysis of the carrot B493 map displaying significant pairs of markers for linkage group maxima and corresponding post hoc statistical tests for marker means

Trait	Marker1(LG)[a]	Marker2(LG)[a]	LOD[b]	Marker genotypes[c]														R^2[d]	p1[e]	p2[f]	pint[g]
				AA/AA	AA/AB	AA/BB	AB/AA	AB/AB	AB/BB	BB/AA	BB/AB	BB/BB	B-/AA	B-/AB	B-/BB	B-/B-	AA/B-				
Total	gggcaa477(1)	Y2Mark(5)	16.1**	30.4 a	34.2 a	48.8 a	17.8 a	35.2 a	87.1 b	36.7 ab	48.6 ab	401.3 c	205.1 b	42.6 a	26.0 a			0.29	0.05	0.00	0.01
Total	CHXE(2)	Y2Mark(5)	28.0***[h]	25.0 ab	38.2 ab	43.7 ab	25.16 a	35.32 a	293.0 c	32.64 ab	57.54 ab	112.6 b						0.57	0.00	0.00	0.00
Total	ZDS1(4)	Y2Mark(5)	18.7***	26.5 a	19.0 a	31.8 ab												0.43	0.01	0.00	0.00
Total	aagcag094(5)	ggtcag072-381(6)	15.3**	25.7 a	30.3 a	26.8 a							80.0 a	86.2 a	38.4 a			0.04	0.03	0.37	0.76
Total	gtgtct364(7)	aagcag094(5)	14.0*	20.8 ab		38.2 a							30.1 a			83.2 b	56.9 ab	0.03	0.29	0.03	0.74
Total	aagctc523(8)	Y2Mark(5)	21.6***	26.7 a	29.8 a	38.2 a							25.7 a	44.5 a	219.0 b			0.35	0.00	0.00	0.00
ζ-Carotene	ggcaa346(2)	Y2Mark(5)	20.3***	0.0 a	0.0 a	0.0 a							0.0 a	1.9 a	36.9 b			0.19	0.11	0.00	0.11
ζ-carotene	ggacag259(3)	Y2Mark(5)	15.3**	0.0 a	0.0 a	41.2 b							0.0 a	1.3 a	36.0 b			0.19	0.81	0.00	0.93
ζ-carotene	ZDS1(4)	Y2Mark(5)	17.2*	0.0 a	0.0 a	0.0 a	0.0 a	1.0 a	69.7 b	0.0 a	3.1 a	16.9 a						0.36	0.05	0.00	0.00
ζ-carotene	aggcta203(7)	Y2Mark(5)	15.9**	0.0 ab	1.2 a	30.1 bc							0.0 a	1.7 a	38.6 c			0.19	0.55	0.00	0.86
ζ-Carotene	aaccat106(9)	Y2Mark(5)	14.9*	0.0 a	2.0 a	22.6 ab							0.0 a	0.6 a	41.7 b			0.2	0.26	0.00	0.38
β-Carotene	ggacaa346(2)	acactt135(5)	35.3***[i]	0.0 ab	0.0 a	0.0 ab							0.3 a		18.8 b			0.04	0.11	0.06	0.52
β-Carotene	Gggcaa229(5)	CCD1(3)	17.2**	0.0 ab	0.0 a	0.4 a							36.2 b	13.0 a	11.5 ab			0.05	0.07	0.15	0.72
β-Carotene	acactt135(5)	ZDS1(4)	19.8**	0.0 ab	0.5 ab	0.0 ab							0.3 a	23.0 b	15.6 ab			0.05	0.07	0.16	0.60
β-Carotene	aggctg066(5)	gggcaa229(5)	18.5**	0.0 a									0.3 a		14.0 a	42.4 b		0.05	0.29	0.02	0.18
β-Carotene	aagcag176(7)	acactt135(5)	19.3*	0.9 a										0.0 a		18.9 a	9.6 a	0.03	0.34	0.08	0.60
Phytoene	ggacaa222(2)	gggcaa229(5)	23.9***	0.0 a									0.0 ab			25.0 b	0.0 a	0.05	0.07	0.05	0.36
Lutien	acactt195(1)	NCED2(2)	20.3***	0.3 a	1.1 a	22.7 b							0.2 a	1.4 a	28.2 b			0.36	0.5	0.00	0.74
Lutien	PDS(2)	NCED2(2)	21.1***	0.2 a	0.4 a	5.0 ab	0.18	1.7 a	16.6 b	1.0 ab	1.8 a	34.6 c						0.41	1.00	0.00	0.04
Lutien	gggcat096(3)	NCED2(2)	21.7***	0.0 a	0.8 a	10.6 a							0.4 a	1.4 a	31.0 b			0.4	0.01	0.00	0.00
Lutien	aggcta084(4)	NCED2(2)	22.1***	0.4 a	0.5 a	10.1 a							0.2 a	1.7 a	31.5 b			0.4	0	0.00	0.01
Lutien	acacaa079(5)	NCED2(2)	18.2**	0.0 a	0.6 a	18.5 b							0.3 a	1.5 a	30.6 c			0.39	0.29	0.00	0.23
Lutien	scar-opk9c-co(6)	NCED2(2)	22.6***	0.5 ab	0.6 a	14.9 bc	0.3 a	1.4 a	20.2 c	0 ab	1.8 a	44.3 d						0.45	0.01	0.00	0.00
Lutien	SSR-w93(7)	NCED2(2)	17.6***	0.1 a	0.7 a	23.7 bc	0.1 a	1.3 a	22 b	0.3 a	2.2 a	34.8 c						0.37	0.14	0.00	0.45
Lutien	ggacag464(8)	NCED2(2)	18.0**	0.4 a	1.1 a	11.1 a							0.2 a	1.6 a	28.3 b			0.35	0.02	0.00	0.07
Lutien	aaccat106(9)	NCED2(2)	19.0**	0.0 a	1.1 a	17.8 b							0.4 a	1.5 a	32 c			0.39	0.42	0.00	0.05

[a] Marker names for pairs of markers with significant joint or epistatic LOD score in two-way QTL scan followed by their respective linkage groups in parentheses

[b] LOD score for the joint analysis in the two-way QTL scan. Significance after genome wide significance adjustment using permutation tests. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$

[c] Marker genotype group means for significant marker pairs. Genotype for marker 1 appears before "/" and genotype for marker 2 appears after "/". AA homozygous for QAL allele, AB heterozygous, BB homozygous for B493 allele, B- homozygous B493 allele or heterozygous (for dominant markers)

[d] R^2 for model using both marker genotypes as regressors on trait

[e] p value for marker 1

[f] p value for marker 2

[g] p value for the interaction term of marker 1 and marker 2

[h] Marker pair was detected as significant epistatic pair in two-way QTL scan (LOD = 17.5, $p < 0.01$)

[i] Marker pair was detected as significant epistatic pair in two-way QTL scan (LOD = 22.5, $p < 0.05$)

2 is the Y gene as described by Buishand and Gabelman (1979). There are several lines of evidence to support this. First, the major QTL locus on linkage group 5 maps near Y2mark, a SCAR developed from an AFLP fragment linked to the Y_2 gene (Bradeen and Simon 1998). Second, the differences between cultivated white carrot and orange carrot have been demonstrated to be due in part to allelic differences at the Y_2 gene. In segregating crosses from cultivated white × orange crosses the Y_2 gene is involved in a similar epistatic interaction: only roots with recessive alleles from the orange parent at both the Y_2 and the Y loci will accumulate orange α- and β-carotene pigments (Buishand and Gabelman 1979). This is the same pattern of inheritance suggested by the epistatic interaction between the linkage group 2 and linkage group 5 alleles. The fact that the linkage group 2 QTL region had highly significant effects on xanthophyll accumulation and the linkage group 5 QTL had no effects on xanthophylls, also supports this idea. These biological observations have interesting implications about genetic changes that occurred during carrot domestication. Historical evidence suggests that the orange carrot was developed through a yellow carrot intermediate which was grown for hundreds of years before the first reported orange carrots (Banga 1957, 1963). Thus wild carrots probably harbor YYY_2Y_2 alleles and the development of the yellow carrot from white carrot as long as 1,000 years ago involved, in part, a mutation at the Y locus to yy alleles, which conditioned the accumulation of xanthophylls in yyY_2Y_2 roots. The development of the orange carrot from yellow progenitors probably involved a second mutation at the Y_2 locus since the recessive alleles coming from B493 appear to condition the accumulation of carotenes in addition to xanthophylls.

Root pigment composition clearly has played an important role in the domestication and subsequent breeding of carrot. It is interesting, and perhaps not surprising, to note that of the two major loci that determine a large proportion of the differences in carotenoid accumulation between wild and cultivated carrot are recessive, given the pattern that emerges from other domestication studies. Like our studies, a small number of genes with large effect, often with the cultivated alleles as recessive or partially recessive to the wild alleles, confer most of the cultivated phenotype for a given trait (Doebley et al. 1990; Grandillo and Tanksley 1996; Xiong et al. 1999; Poncet et al. 2002; reviewed by Gepts 2002).

While there is no indication that the nutritional value of carotenoids was understood as the first orange carrots were developed (Simon 2000), this was a fortunate development for modern nutritional health. Orange is by far the most predominant color of most carrots cultivated today, accounting for 65% of the β-carotene and 95% of the α-carotene in the U.S. diet, and accounting for approximately 1/4 of the dietary vitamin A (Simon et al. 2009).

Candidate genes for Y and Y_2

Based upon our conclusions about the identity of these genes, it is interesting to note that the Y and Y_2 loci on chromosome region on linkage group 2 and 5, respectively, were linked to several carotenoid biosynthetic enzyme sequence tagged sites. The Y locus on linkage group 2 is closely linked to the STS marker for CHXE gene, the STS marker for NCED2, and more distantly linked to the STS marker for the PDS gene. Other carotenoid biosynthetic genes demonstrated significant values in epistasis estimates (Tables 6 and 7). All of these carotenoid biosynthetic genes may be considered positional candidate genes. Considering the placement of CHXE and NCED2 in the carotenoid biosynthetic pathway beyond carotene biosynthesis, a mechanism to account for their role controlling accumulation of both carotenes and xanthophylls is difficult to explain. A possible role for PDS in accounting for Y gene phenotype is more readily explained with its position early in the pathway. The STS markers for ZEP and ZDS are not as closely linked to Y_2 as are the positional candidate genes for Y, and like CHXE and NCED2, a mechanism to account for their role in limiting carotene, but not xanthophyll, accumulation is not obvious. More research in carotenoid gene expression as it relates to carotenoid accumulation is necessary to advance any of these biosynthetic genes to candidate gene status.

Because these two loci in carrot control so much of the variability and the doubly recessive state was required for accumulation of large amounts of orange carotene pigments, the power to detect QTL with a smaller effect on carotenoid content is greatly reduced. QTL with smaller effects were found on linkage groups 4 and 8 and several loci were detected as significant in the joint (two locus model) analyses. Notably, the ZDS1 STS was jointly significant with Y2mark (or the adjacent AFLP marker) for total carotenes and β-carotene on both maps. Future studies of these loci in appropriate backcross populations are needed to confirm their presence and to estimate their effects more accurately.

References

Banga O (1957) Origin of the European cultivated carrot. Euphytica 6:54–63

Banga O (1963) Main types of the western carotene carrot an their origin. W. E. J. Tjeenk Willink, Zwolle, p 153

Beavis WD (1998) QTL analyses: power, precision and accuracy. In: Paterson AH (ed) Molecular analysis of complex traits. CRC Press, Boca Raton, pp 45–161

Ben Chaim A, Paran I, Grube R, Jahn M, van Wijk R, Peleman J (2001) QTL mapping of fruit related traits in pepper (*Capsicum annuum*). Theor Appl Genet 102:1016–1028

Boiteux LS (2000) Characterization of the *Meloidogyne javanica* resistance locus employing molecular markers and isolation of candidate disease resistance loci in the carrot (*Daucus carota* L.) genome. Ph.D. Dissertation, University of Wisconsin-Madison, 265 pp

Boiteux LS, Fonseca MEN, Simon PW (1999) Effects of plant tissue and DNA purification methods on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. J A Soc Hortic Sci 124:32–38

Boiteux LS, Belter JG, Roberts PA, Simon PW (2000) RAPD linkage map of the genomic region encompassing the root-knot nematode (*Meloidogyne javanica*) resistance locus in carrot. Theor Appl Genet 100:439–446

Bradeen JM, Simon PW (1998) Conversion of an AFLP fragment linked to the carrot Y_2 locus to a simple, codominant, PCR-based marker form. Theor Appl Genet 97:960–967

Broman KW (2003) Mapping quantitative trait loci in the case of a spike in the phenotype distribution. Genetics 163:1169–1175

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890

Buckner B, Kelson TL, Robertson DS (1990) Cloning of the *y1* locus of maize, a gene involved in the biosynthesis of carotenoids. Plant Cell 2:867–876

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

Buishand JG, Gabelman WH (1979) Investigations on the inheritance of color and carotenoid content in phloem and xylem of carrot roots (*Daucus carota* L.). Euphytica 28:611–632

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

Craft NE, Soares JH Jr (1992) Relative solubility, stability, and absorptivity of lutein and beta-carotene in organic solvents. J Agric Food Chem 40:431–434

Doebley J, Stec A, Wendel J, Edwards M (1990) Genetic and morphological analysis of a maize-teosinte F_2 population: implications for the origins of maize. Proc Natl Acad Sci (USA) 87:9888–9892

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Fray RJ, Grierson D (1993) Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression. Plant Molec Biol 22:589–602

Gepts P (2002) A comparison between crop domestication, classical plant breeding, and genetic engineering. Crop Sci 42:1780–1790

Grandillo S, Tanksley SD (1996) QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. Theor Appl Genet 92:935–951

Huh JH, Kang BC, Nahm SH, Kim S, Ha KS, Lee MH, Kim BD (2001) A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (*Capsicum* spp.). Theor Appl Genet 102:524–530

Ihaka R, Gentleman R (1996) R: A language for data analysis and graphics. J Comput Graph Stat 5:299–314

Isaacson T, Ronen G, Zamir D, Hirschberg J (2002) Cloning of *tangerine* from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants. Plant Cell 14:333–342

Just BJ, Santos CAF, Fonseca MEN, Boiteux LS, Oloizia BB, Simon PW (2007) Carotenoid biosynthesis structural genes in carrot (*Daucus carota*): isolation, sequence-characterization, single nucleotide polymorphism (SNP) markers and genome mapping. Theor Appl Genet 114:693–704

Kruglyak L, Lander ES (1995) A nonparametric approach for mapping quantitative trait loci. Genetics 139:1421–1428

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lefebvre V, Kuntz M, Camara B, Palloix A (1998) The capsanthin-capsorubin synthase gene: a candidate gene for the y locus controlling the red fruit colour in pepper. Plant Mol Biol 36:785–789

Li L, Paolillo DJ, Parthasarathy MV, DiMuzio EM, Garvin DF (2001) A novel gene mutation that confers abnormal patterns of β-carotene accumulation in cauliflower (*Brassica oleracea* var. *botrytis*. Plant J 26:59–67

Li L, Lu S, Van Eck J, O'Halloran D, Zhou X, Lopez AB, Cosman K, Conlin B, Paolillo D, Garvin DF, Vrebalov J, Kochian LV, Kupper H, Earle E, Cao J (2006) The cauliflower *or* gene encodes a cysteine-rich zinc finger domain-containing protein that induces high-level of b-carotene accumulation. Plant Cell 18:3594–3605

Liu Y-S, Gur A, Ronen G, Causse M, Damidaux R, Buret M, Hirschberg J, Zamir D (2004) There is more to tomato fruit colour than candidate carotenoid genes. Plant Biotech J 1:195–207

Marshall SW, Tracy WF (2003) Sweet corn. In: Ramstad PE, White P (eds) Corn chemistry and technology, 2nd edn. Amer Assoc Cereal Chem, Minneapolis, pp 537–569

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

Pflieger S, Lefebvre V, Causse M (2000) The candidate gene approach in plant genetics: a review. Mol Breed 7:275–291

Poncet V, Martel E, Allouis S, Devos KM, Lamy F, Sarr A, Robert T (2002) Comparative analysis of QTLs affecting domestication traits between two domesticated × wild pearl millet (*Pennisetum glaucum* L. *Poaceae*) crosses. Theor Appl Genet 104:965–975

Ronen G, Cohen M, Zamir D, Hirschberg J (1999) Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. Plant J 17:341–351

Ronen G, Carmel-Goren L, Zamir D, Hirschberg J (2000) An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning in *Beta* and *old-gold* color mutations in tomato. Proc Natl Acad Sci (USA) 97:11102–11107

Santos CAF (2001) Biometrical studies and quantitative trait loci associated with major products of the carotenoid pathway of carrot (*Daucus* carota L.). Dissertation, University of Wisconsin-Madison, 265 pp

Santos CAF, Simon PW (2002) QTL analyses reveal clustered loci for accumulation of major provitamin A carotenes and lycopene in carrot roots. Mol Genet Genom 268:122–129

Santos CAF, Simon PW (2006) Heritabilities and minimum gene number estimates of carrot carotenoids. Euphytica 151:79–86

Simon PW (1996) Inheritance and expression of purple and yellow storage root color in carrot. J Hered 87:63–66

Simon PW (2000) Domestication, historical development, and modern breeding of carrot. Plant Breeding Rev 19:157–190

Simon PW, Wolff XY (1987) Carotenes in typical and dark orange carrots. J Agric Food Chem 35:1017–1022

Simon PW, Peterson CE, Gabelman WH (1990) B493 and B9304, carrot inbreds or use in breeding, genetics and tissue culture. HortScience 25:815

Simon PW, Pollak LM, Clevidence BA, Holden JM, Haytowitz DB (2009) Plant breeding for human nutrition. Plant Breeding Rev 31:325–392

Surles RL, Weng Ning, Simon PW, Tanumihardjo SA (2004) Carotenoid profiles and consumer sensory evaluation of specialty carrots (*Daucus carota.* L.) of various colors. J Agric Food Chem 52:3417–3421

Thorup TA, Tanyolac B, Livingstone KD, Popovsky S, Paran I, Jahn M (2000) Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc Natl Acad Sci (USA) 97:11192–11197

Vivek BS, Simon PW (1999) Linkage relationships among molecular markers and storage root traits of carrot (*Daucus carota* L. ssp. *sativus*). Theor App Genet 99:58–64

Wright S (1934) An analysis of variability in number of digits in an inbred strain of guinea pigs. Genetics 19:506–536

Xiong L, Liu K, Dai X, Xu C, Zhang Q (1999) Identification of genetic factors controlling domestication related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. Theor Appl Genet 98:243–251

Xu S, Atchley WR (1996) Mapping quantitative trait loci for complex binary diseases using line crosses. Genetics 143:1417–1424

Yau YY, Simon PW (2003) A 2.5-kb insert eliminates acid soluble invertase isozyme II transcript in carrot (*Daucus carota* L.) roots, causing high sucrose accumulation. Plant Mol Biol 53:151–162

Yi N, Xu S (1999) A random model approach to mapping quantitative trait loci for complex binary traits in outbred populations. Genetics 153:1029–1040

Zou F, Yandell BS, Fine JP (2003) Rank-based statistical methodologies for quantitative trait locus mapping. Genetics 165:1599–1605