Mol Genet Genomics (2002) 268: 122–129
DOI 10.1007/s00438-002-0735-9

ORIGINAL PAPER

C.A.F. Santos · P.W. Simon

QTL analyses reveal clustered loci for accumulation of major provitamin A carotenes and lycopene in carrot roots

Received: 28 April 2002 / Accepted: 15 July 2002 / Published online: 9 August 2002

Abstract QTLs associated with products of the carotenoid pathway, including lycopene and the provitamin A carotenes α- and β-carotene, were investigated in two unrelated F_2 carrot populations, derived from crosses between orange cultivated B493 and white wild QAL (Population 1), and orange cultivated Brasilia and dark-orange cultivated HCM (Population 2). The mapping populations of 160 and 180 individuals, respectively, were analyzed with single-marker and interval-mapping statistical approaches, using coupling linkage maps for each parent. Single markers were selected for further analysis based on the Wilcoxon sum-rank non-parametric test. Interval mapping performed with Population 2 detected four, eight, three, one and five putative QTLs associated with accumulation of ξ-carotene, α-carotene, β-carotene, lycopene and phytoene, respectively. Among these, the major QTLs explained 13.0%, 10.2%, 13.0%, 7.2% and 10.2% of total phenotypic variation. In Population 1 single-marker analysis identified loci explaining up to 13.8%, 6.8%, 19.3%, 5.7%, and 17.5%, respectively, of total phenotypic variation for these same carotenoids. Overall analysis demonstrated clustering of these QTLs associated with the carotenoid pathway: the AFLP loci AACCAT178-Q and AAGCAG233-Q, on linkage group 5, explained 17.8%, 22.8% and 23.5% of total phenotypic variation for ζ-carotene, phytoene and β-carotene in Population 1. Two major clusters of QTLs, with LOD scores greater than 1.8, mapped to intervals no larger than 2 cM for ζ-carotene, β-carotene, α-carotene and lycopene on linkage group 3, and for ζ-carotene and phytoene on linkage group 9, and these explained 3.7% to 13.0% of variation for each carotenoid product. Thus, these results suggest that clustering of related pathway loci is favored during evolution, since closely linked "pathway mates" are not easily separated by recombination.

Keywords Biochemical pathway · *Daucus carota* L. · Linked pathway genes · Wilcoxon sum-rank test

Communicated by R. Hagemann

C.A.F. Santos
Plant Breeding and Plant Genetics Program,
University of Wisconsin, 1575 Linden Drive,
Madison, WI 53706, USA

P.W. Simon (✉)
USDA-Agricultural Research Service,
Vegetable Crops Research Unit Department of Horticulture,
1575 Linden Drive, University of Wisconsin, Madison, WI 53706
E-mail: psimon@facstaff.wisc.edu
Fax: +1-608-2624743

Present address: C. A. F. Santos
Embrapa (Brazilian Agricultural Research
Corporation) Semi-Arido, Caixa Postal 23.56300-970,
Petrolina PE, Brazil

Introduction

The biosynthesis of carotenoids is a well established and extensively studied biochemical pathway in many plants (Cunningham and Gantt 1998; Sandmann 1998), fungi and microorganisms (Armstrong 1994; Sandmann 1998). The formation of the colorless carotene phytoene from two molecules of geranylgeranyl diphosphate or pyrophosphate (GGPP) is the first step in the pathway of carotenoid biosynthesis. Phytoene undergoes a series of four desaturation reactions that result first in the formation of phytofluene, and then, in turn, ξ-carotene, neurosporene and the red-colored lycopene. The four sequential desaturations undergone by phytoene are catalyzed by two related enzymes in plants: phytoene desaturase (PDS) and ξ-carotene desaturase (ZDS). In contrast, bacteria and fungi achieve the same result with a single gene product (Cunningham and Gantt 1998). A bacterial carotene desaturase, which is capable of introducing all four double bonds required, has been used to engineer β-carotene accumulation in the endosperm of "golden rice" (Ye et al. 2000).

A single gene product, lycopene β-cyclase (LCYB), catalyzes the formation of the bicyclic β-carotene (with two β rings) from the linear, symmetrical lycopene in

plants and cyanobacteria (Cunningham and Gantt 1998), as demonstrated in studies with *Erwinia herbicola* and tomatoes (Sandmann 1998). In the case of α-carotene, with one β and one ε ring, two different enzymes, LCYB and lycopene ε-cyclase (LCYE), are involved (Sandmann 1998). Whether the route to α-carotene biosynthesis proceeds only via the ε ring first or can also proceed via the β ring first has not been determined, and carotenoids with two ε rings are not commonly found in plants (Cunningham and Gantt 1998). Hydroxylation at carbon three of each ring of the hydrocarbons β-carotene and α-carotene will produce the xanthophyll pigments zeaxanthin, with two β rings, and lutein, with one β ring and one ε ring (Cunningham and Gantt 1998). Only algae and higher plants have the potential to form α-carotene or lutein, while zeaxanthin is synthesized by some heterotrophic bacteria and all organisms that carry out oxygenic photosynthesis (Sandmann 1998).

Vertebrates do not synthesize carotenoids and depend on dietary carotenoids for synthesis of their retinoids, such as retinal (the main visual pigment), retinol (vitamin A) and retinoic acid (a substance controlling morphogenesis) (Giuliano et al. 2000). Consumption of horticultural crops provides more than 70% of vitamin A worldwide (Simpson 1983), with carrots accounting for 30% of the total intake of vitamin A precursor carotenoids in countries like the United States (Simon 1992). Lycopene is another carotene which accumulates in carrots and it too has health benefits, as it is implicated in the prevention of prostate (Giovannucci et al. 1995) and other cancers.

Efforts to improve carrots have included a focus on breeding for increased carotenoid content (Simon 2000), but none of these efforts have attempted to map the major genes for total root carotenoid content, or those genes that condition the accumulation of individual carotenoids. Marker-based analysis of QTLs (quantitative trait loci) for carotenoid accumulation provides a powerful tool for comprehensively mapping the key steps in this biochemical pathway.

Marker-based QTL analysis has proven to be useful in estimating the number of major loci that affect a given character, in elucidating the evolution of phenotypic traits and in providing insight into genome organization and facilitating pathway analysis. Specific examples of pathway dissection using QTL studies have included investigations of (1) flavonoid levels and the loci that encode enzymes of the general phenylpropanoid pathway that affect concentrations of silk maysin, a corn earworm resistance factor (Byrne et al. 1996, 1998; McMullen et al. 1998); (2) activities and concentrations of key enzymes of carbohydrate metabolism in maize (Prioul et al. 1999); (3) erucic acid levels in *Brassica napus* (Fourmann et al. 1998); (4) identification of QTLs associated with acylsugar accumulation in wild tomato (Blauth et al. 1998); and (5) mapping of cloned structural genes of the carotenoid biosynthetic pathway in pepper (Thorup et al. 2000). QTL studies have also been applied to study the distribution of related developmental loci in the maize genome (Khavkin and Coe 1997). The contribution of individual loci to the total variation of traits evaluated by QTL analysis typically ranges from less that 1% up to 15%, and the collective contribution of all loci mapped usually accounts for 5% to 40% of the total trait variation.

The goals of this study were: (1) to estimate the effects of QTLs associated with the five key products of the carotenoid biosynthetic pathway in carrots, and (2) to analyze the distribution of putative carotenoid QTLs on the linkage groups, using two different carrot crosses: B493 (orange) × QAL (white carrot) and Brasilia (medium orange) × HCM (dark orange). Individual pigments were quantified and QTL analyses were conducted for phytoene, ξ-carotene, β-carotene, α-carotene and lycopene.

Materials and methods

Plant material and phenotypic data

The F_2 mapping populations were derived from single F_1 plants resulting from the crosses B493 × QAL (Population 1) and Brasilia × HCM (Population 2). B493 is a dark-orange inbred carrot with a carotene content ranging from 180 to 210 ppm (Simon et al. 1990). QAL is a white wild carrot (*D. carota* var. *carota*) well distributed in temperate regions of Eastern North and South America and from the Atlantic coast of Western Europe to Western China (Rubatzky et al. 1999), and the parent plant used in this study was from Madison, Wis. HCM is a very dark orange population with an average carotene content ranging from 460 to 499 ppm, which was developed from a cross between Asian and European germplasm (Simon et al. 1989). Brasilia is a typical orange carrot developed in Brazil for production in warmer production areas (Hamerschmidt 1993). The population sizes were 180 and 160 plants, respectively, for Population 1 and Population 2. Major carotenes were quantified by high-performance liquid chromatography (HPLC), as described by Simon and Wolff (1987).

DNA extraction and linkage analysis

Total genomic DNA was isolated from freeze-dried leaves following the protocol of Doyle and Doyle (1990), with minor modifications tested by Boiteux et al. (1999) for carrots. The AFLP method was performed essentially as described in Vos et al. (1995), with minor adaptations for carrot DNA introduced by Vivek and Simon (1999). DNA restriction digestion, ligation, and selective amplification were performed with 1/4 and the pre-selective amplification reaction was performed with 1/10 of the recommended volumes, respectively. Linkage analyses were performed with the software JoinMap version 2.0 (Stam and van Ooijen 1995). AFLP fragments, scored as 0 or 1 for absence or presence, were codified to symbols required by JoinMap. Where the parental allelic phase was unknown, markers were double-scored and the phase was deduced by association with phase-known linkages (Vivek and Simon 1999) and based on the warning message issued by the module JMREC, in the JoinMap software, when the recombination fraction exceeded 0.6 (Stam and van Ooijen 1995). Segregation distortion tests were performed with JoinMap and Excel spreadsheets. Linkage groups were assigned with the JoinMap software using LOD scores ranging from 3.0 to 6.5 and a maximum recombination fraction of 0.49 as recommended (Stam and van Ooijen 1995). The options used in the JoinMap module JMMAP were: LOD score of 2.0, recombination threshold of 0.49, jump Chi-square threshold of 6.0 and triplet jump Chi-square threshold of 6.0. The recombination fraction frequencies were converted to Kosambi centiMorgans

(cM) (Kosambi 1944). Individual coupling linkage maps were developed for each population and drawn with DrawMap software (van Ooijen 1994) kindly provided by the developer.

QTL identification: single-marker and interval analysis

The QTLs were determined to be located near a marker if phenotypic values for the trait were significantly different among the marker genotypes. Tests for normality for all carotenoid products in the F_2 populations were obtained with kurtosis and the Wilk-Shapiro test in the procedure Univariate (SAS 1989). The Wilcoxon rank-sum test, implemented using the SAS NPAR1WAY procedure (SAS 1989), was used to detect significant differences between the two phenotypic marker classes. t-Tests, assuming both equal variance and unequal variance, were also implemented with the SAS t-test procedure (SAS 1989). To compare the Wilcoxon rank-sum and t-test values, simple correlations were calculated with the correlations procedure (SAS 1989). According to Hamilton (1990) the Wilcoxon sum-rank test is a more robust test than parametric tests; it tests the equality of the medians rather than the means and it makes use of ranks, rather than raw data.

Loci with significant values by the two-sided Wilcoxon Z-test, at a p-value of 0.05, were selected and regression analyses were performed using the default stepwise selection option in the SAS (1989) regression procedure to find the best small set of AFLP loci associated with a given character. A final analysis was conducted with the chosen group of markers to determine the coefficient of determination using the regression procedure or using the GLM procedure when some interaction factor was included in the model (SAS 1989).

Interval analyses were performed with PlabQTL version 1.1 [H. F. Utz and A. E. Melchinger (1996) PLABQTL: a program for composite interval mapping of QTL. J Agric Genomics 2(1); available online at http://www.ncgr.org/ag/papers96/paper196/indexp196.html) using the coupling linkage map for each parental line. A threshold LOD score of 2.0 was chosen to identify QTL regions with the single interval mapping. Within PlabQTL the QIN program was chosen to analyze only the additive confounded dominant model, since only two phenotypic classes were available to establish contrasts (Liu 1997).

Results and discussion

Linkage analyses

A total of 250 and 287 molecular markers were used to construct linkage groups in F_2 Population 1 (B493 × QAL) and Population 2 (Brasilia × HCM). Grouping only markers which were linked in coupling phase resulted in the same nine linkage groups with LOD scores ranging from 3.0 to 7.5, 3.0 to 5.5, 3.5 to 5.0 and 2.5 to 8.5 in the Brasilia-, HCM-, B493- and QAL-coupling F_2 populations, respectively. According to Knapp et al. (1995), estimates of recombination fraction are seriously biased with dominant markers in repulsion phase in F_2, due to the low frequency of the double recessive class. However, because a locus can be assigned to a group on the basis of $(k^2-k)/2$ tests the same authors point out that grouping of loci seldom poses a problem. Furthermore, groupings obtained with LOD thresholds in the range 4.0–7.0 are safe to work with in general (Stam and van Ooijen 1995). In contrast, ordering of mixed-phase dominant markers poses much more of a problem than grouping (Knapp et al. 1995) because the recombination frequency estimates are frequently biased. Using the inverse of the variance of recombination fraction obtained with the method of maximum likelihood, Mather (1936) demonstrated that repulsion-linked dominant markers in F_2 populations are most efficient in the absence of linkage. However, valid maps of dominant markers can be built by using sub-sets of markers linked in coupling, thereby creating two parental coupling maps (Knapp et al. 1995), which will be close to those for backcross mating maps when the recombination fraction is around 10 cM (Mather 1936). In our study, the average marker spacing for the two F_2 carrot populations were 4.78, 4.80, 5.54 and 5.13 cM in the Brasilia, HCM, B493 and QAL coupling-phase maps, respectively. Gaps (greater than 10 cM) were observed more frequently in the coupling linkage maps of B493 and QAL than in the Brasilia- and HCM-coupling maps (Santos 2001).

Wilcoxon sum-rank as a preliminary test to identify candidate loci involved in the carotenoid pathway

Correlation values between the non-parametric Wilcoxon and the t-test p-values were greater than 0.93 for β-carotene and α-carotene in Population 2, which had a well-fitted normal probability distribution. All the other traits had correlations around 0.5 in population 1 and 0.8 in Population 2, and they deviated from the normal distribution.

Since the Wilcoxon rank-sum test is a non-parametric statistical approach to test the difference between two medians without assuming normal distribution of data or equality of variances between two groups (SAS 1989), it has been applied to the genetic analysis of metabolic pathways of maysin in corn (Byrne et al. 1996), carbohydrate metabolism in maize (Prioul et al. 1999) and intestinal adenoma size in mice (Gould et al. 1996). In this study, single markers were selected for further analysis based on the Wilcoxon sum-rank test because (1) the correlation with t-test values was greater than 0.9 when the normal distribution fits well, (2) it is a more appropriate test in the absence of normal distribution, and (3) 3:1 segregation of dominant markers can lead to unequal variance and skewed distribution between the two classes of a particular molecular marker.

ξ-Carotene, β-carotene, α-carotene, lycopene and phytoene QTLs: single-marker analysis

Using a stepwise option as an initial step to select AFLPs, multiple regression models explained 43.4% and 44.2% of total phenotypic variation for phytoene in the Brasilia × HCM and B493 × QAL populations, respectively, for example (Table 1). For β-carotene, up to eight AFLP loci explained 26.4% and 40.4% of the total phenotypic variation in Populations 1 and 2, respectively (Table 1). The major QTLs explained 4.7–8.0% of the

total phenotypic variation for ξ-carotene, β-carotene, α-carotene, lycopene, phytoene in the Brasilia × HCM F_2 population (Table 1). The individual contribution of AFLP loci was quite large in the B493 × QAL F_2 population, ranging from 5.7% for lycopene to 19.3% for β-carotene (Table 1). A previous study of the carotenoid pathway enzymes in Solanaceae reported QTLs explaining much less of the total variation than reported in this study: 0.03% for ξ-carotene desaturase to 12% for phytoene synthase (Thorup et al. 2000).

Interval mapping analysis

No interval mapping analysis for QAL- and B493-coupling is presented because of strong deviation from the normal distribution and kurtosis ranging from 12.7 to 78.3. Normal distribution is an assumption required to test the presence of a QTL by a simple parametric test or interval mapping (Kruglyak and Lander 1995).

Four putative QTLs were associated with ζ-carotene accumulation in three different linkage groups of Population 2: two coupled with Brasilia and two coupled with HCM, with LOD scores ranging from 1.8 to 3.4 (Table 2). LOD scores of 1.8 and 3.4 correspond to *p*-values of 0.016 and 0.004, respectively (Ott 1999). The multiple-AFLP model for ζ-carotene accounted for 16.3% of the phenotypic variation. Since the R^2 value was estimated by fitting the nearest AFLP marker locus rather than by estimates on the QTL region per se, this value should be considered a minimum estimate, as suggested by Byrne et al. (1998). With the exception of one putative QTL from HCM, all others had positive effects upon the total phenotypic variation of ζ-carotene, and only the AGGCAT462-H locus was identified in the previous single marker analysis.

Eight putative QTLs were detected for α-carotene, mapping to linkage groups 3 (2), 8 (2), 6 (2), 4 and 1, with LOD scores ranging from 1.7 to 4.7 (Table 2). The multiple R^2 for the eight AFLPs mapping to the left of the eight putative QTLs accounted for 40% of the total variation in α-carotene, which was similar to the model fitted in the single-marker analysis. Single locus associations accounted for up to 13% of the overall phenotypic variation in this analysis, while the maximum was 8% in the previous single marker analysis. A large negative contribution was observed with three out of four putative QTL regions from Brasilia, while a marked

Table 1. Numbers of AFLP loci selected by the stepwise regression procedure and their coefficients of determination for the general fitted models and for the major QTLs contributing to phenotypic variation of the indicated carotenes in the Brasilia × HCM and B493 × QAL F_2 populations

Character	Number of AFLP loci		Coefficient of determination (R^2)			
			Model		Major QTL	
	Bsb × HCM	B493 × QAL	Bsb × HCM	B493 × QAL	Bsb × HCM	B493 × QAL
ξ-Carotene	7	8	28.3	21.7	8.4	13.8
β-Carotene	8	4	40.4	26.4	4.7	19.3
α-Carotene	9	13	36.6	37.7	8.0	6.8
Lycopene	9	7	35.1	15.8	5.2	5.7
Phytoene	11	12	43.4	44.2	7.4	17.5

Table 2. Summary of QTLs associated with carotenoid pathway enzymes in the Brasilia × HCM F_2 population. QTLs were estimated in the Brasilia and HCM coupling maps using the PlabQTL software

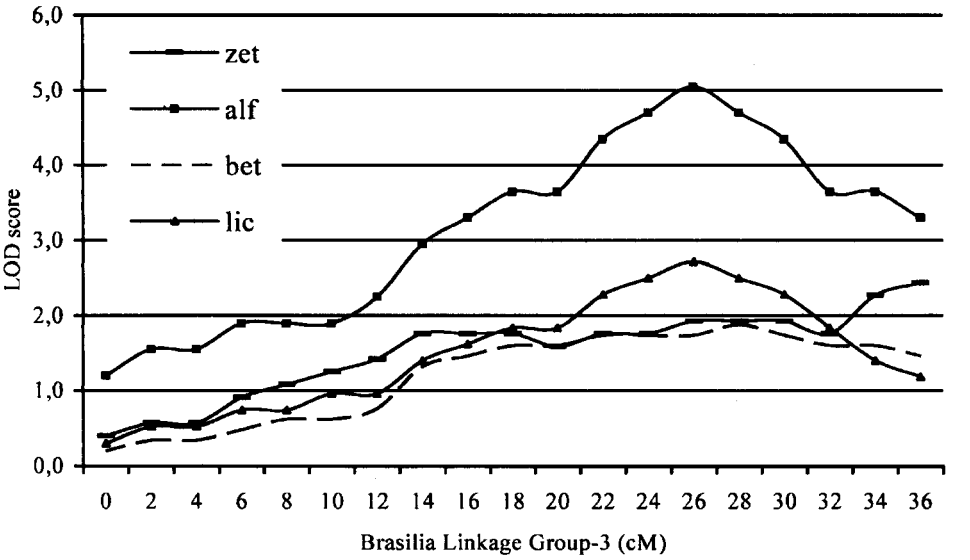
Linkage group (parent)	Position (cM)	Character	Left AFLP locus	LOD score	Coefficient of determination (%)	Additive effect
1 (BSB)	2	Phytoene	AACCTA079-B	3.4	10.2	−110.5
1 (HCM)	60	α-Carotene	GAGCAC171-H	1.7	6.0	127.0
1 (HCM)	76	Phytoene	GGGCAT139/260	2.7	7.6	−89.6
2 (BSB)	24	ζ-Carotene	GATCAT092-B	2.4	7.4	52.3
3 (BSB)	24	α-Carotene	AGCCTT057-B	4.7	13.0	−133.2
3 (BSB)	34	ζ-Carotene	AACCAT121-B	2.2	3.7	35.5
3 (BSB)	26	β-Carotene	AGCCTT057-B	1.7	4.5	73.5
3 (BSB)	24	Lycopene	AGCCTT057-B	2.5	7.2	−9.2
3 (HCM)	8	ζ-Carotene	AGGCAT462-H	1.8	4.3	−42.2
3 (HCM)	20	α-Carotene	AACCAA358-H	3.9	10.9	−120.6
3 (HCM)	56	β-Carotene	GGGCTG199-H	2.5	7.2	−96.4
4 (BSB)	40	α-Carotene	AAGCTC163-B	2.9	8.3	106.7
6 (HCM)	28	α-Carotene	GGACAG109-H	3.3	9.2	114.4
6 (HCM)	44	α-Carotene	GGTCTT350-H	3.4	9.7	123.6
8 (BSB)	18	α-Carotene	GGGCTG109-B	3.2	9.1	−106.8
8 (BSB)	48	α-Carotene	AAGCTC322-B	3.5	10.3	−121.1
9 (HCM)	92	ζ-Carotene	GGGCTG358-H	3.4	9.6	42.07
9 (HCM)	28	β-Carotene	GATCTC167-H	4.8	13.2	138.5
9 (HCM)	34	Phytoene	GATCTC167-H	3.5	9.7	136.3
9 (HCM)	76	Phytoene	ACCCAC425-H	3.6	10.2	135.2
9 (HCM)	92	Phytoene	GGGCTG358-H	3.2	9.0	107.8

positive contribution to the α-carotene variation was observed from the HCM parent.

Three putative QTLs were detected for β carotene with the interval mapping analysis: two in the HCM and one in the Brasilia coupling map (Table 2), and together they accounted for 21% of total variation. Two of these putative loci mapped to linkage group 3 of the Brasilia and HCM coupling maps and one mapped to linkage group 9 of HCM. None of the AFLP loci fitted in the single-marker analysis was among the AFLP loci fitted in the interval mapping analysis. One QTL region from HCM had a negative effect upon overall β-carotene content, and the other two putative QTLs made positive contributions, with linkage group 9

Table 3. Coefficients of determination of two leading QTLs which simultaneously affect the phenotypic variation of ξ-carotene, β-carotene, α-carotene, lycopene and phytoene in the B493 × QAL F_2 population

Character	Coefficient of determination (R^2)		
	AACCAT178-Q	AAGCAG233-Q	AACCAT178-Q and AAGCAG233-Q
ξ-Carotene	13.8	13.7	17.8
β-Carotene	18.1	19.3	23.5
α-Carotene	–	–	–
Lycopene	5.7	6.0	8.2
Phytoene	17.5	2.0	22.8



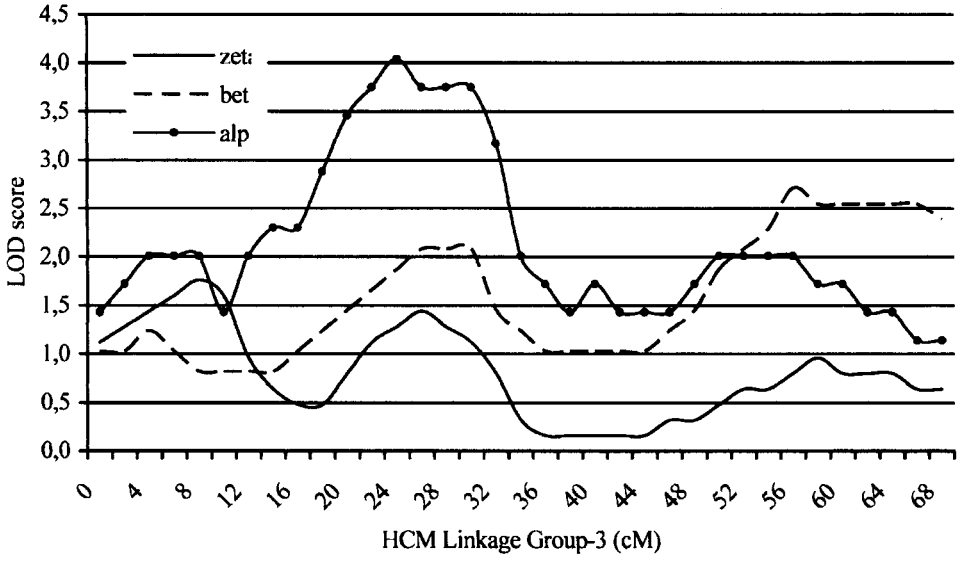


Fig. 1. QTL likelihood map showing LOD scores for ζ-carotene (zet), β-carotene (bet), α-carotene (alf) and lycopene (lic) content detected on Brasilia linkage group 3 and HCM linkage group 3

from the HCM parent accounting for 13.2% of total variation.

Only one putative QTL for lycopene was detected, on the Brasilia map. This QTL contributed 7.2% of total variation in lycopene content (Table 2). Five putative QTLs were associated with phytoene content: three mapped on HCM linkage group 9 and one each on Brasilia linkage group 1 and HCM linkage group 1 (Table 2). Together these explained 28% of the total phenotypic variation for phytoene and none of the AFLP loci was identified in the previous single-marker analysis. The single-marker analysis for phytoene explained almost twice as much of the total variation in phytoene levels as was explained by the interval mapping analysis. However, the interval mapping analysis fitted only five markers to explain slightly more than half the variation explained by the twelve AFLP loci fitted in the single analysis (Table 1). The highest R^2 value for any locus was 10.2%, while the highest effect for any locus in the single-marker analysis was 7.4%.

Clustered loci in a few regions account for most of the phenotypic variation in carotenoid pathway products

Two AFLP loci, AACCAT178-Q and AAGCAG233-Q, were associated with phenotypic variation for phytoene, ζ-carotene, lycopene and β-carotene (Table 3). They were separated by 9.5 cM and LOD score of 23 on linkage group 5 of the B493 × QAL F_2 population. Other AFLPs associated with more than one carotenoid were GGACAG480-Q, which was included in the phytoene, ζ-carotene and lycopene models, and AGCCTT349-Q, which was included in the phytoene and

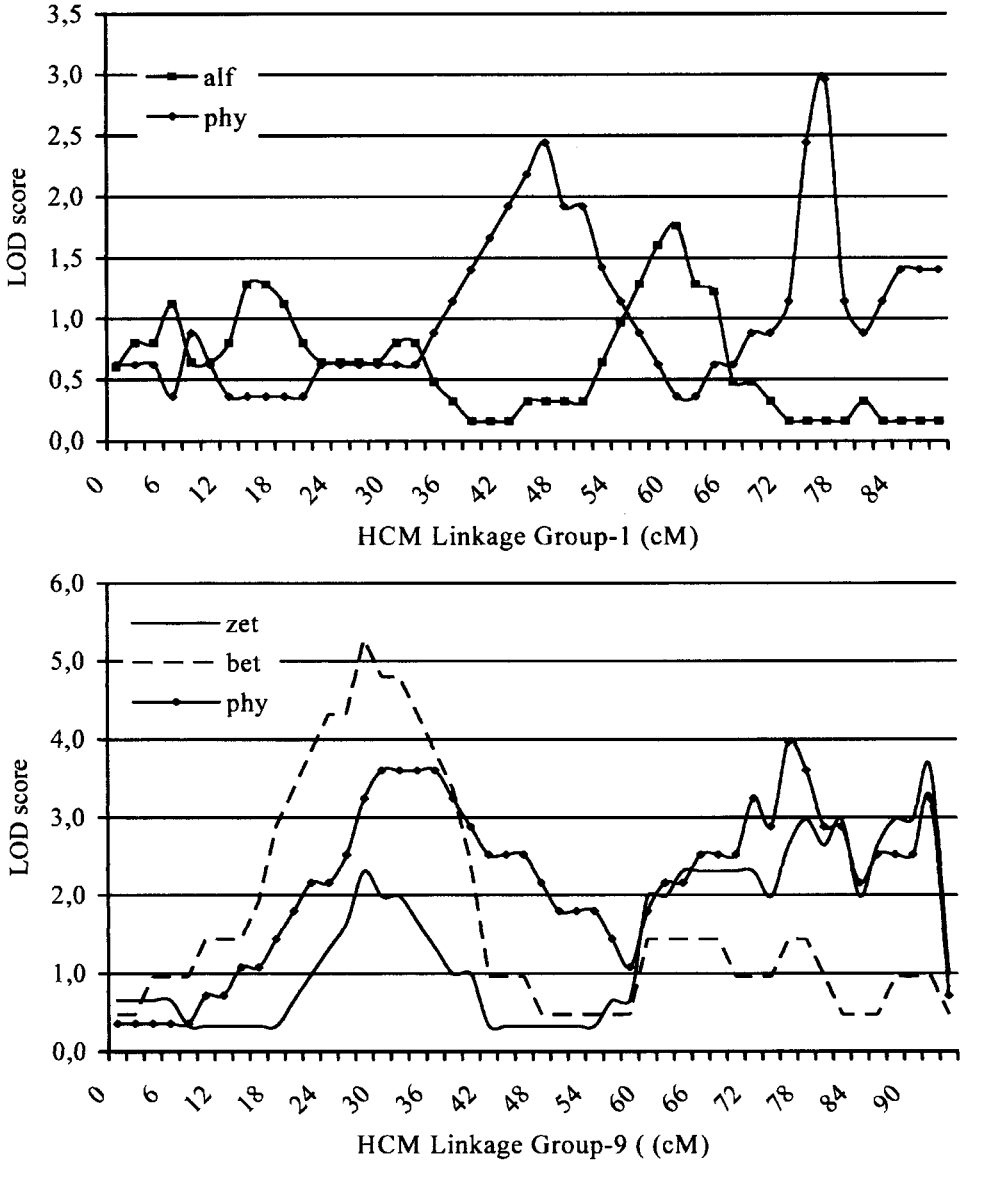


Fig. 2. QTL likelihood map showing LOD scores for ζ-carotene (zet), β-carotene (bet), α-carotene (alf) and phytoene (phy) content detected on the HCM linkage groups 1 and 9

α-carotene models. All these loci mapped to the linkage group 5 of Population 1. These four AFLP loci common to different products of the carotenoid pathway encompassed a total region of 45 cM, with AACCAT178-Q and AAGCAG233-Q being most closely linked. These two loci, plus their interaction term, explained 17.8%, 22.8% and 23.5% of total phenotypic variation for ζ-carotene, phytoene and β-carotene content, respectively. Thorup et al. (2000) reported no R^2 greater than 12% for phytoene synthase among ten structural genes of the *Capsicum* carotenoid biosynthetic pathway in an F_2 population phenotypied with AFLP, RAPD and RFLP markers, and anchored to the *Lycopersicum* map. In this carrot study we report a single putative QTL which explains approximately 20% of total variation for phytoene, ζ-carotene and β-carotene.

Interval analysis identified a cluster of putative QTLs near 26 cM which affected ζ-carotene, β-carotene, α-carotene and lycopene content, mapping to linkage group 3 of Brasilia (Table 2 and Fig. 1), with LOD scores of 2.2, 1.7, 4.7 and 2.5, respectively. Phytoene, a key early compound in the carotenoid pathway, was associated with a putative QTL at position 0 of the Brasilia linkage group 3 but with a LOD of only 1.4. Putative QTLs associated with carotenoids and LOD scores larger than 1.8 were also observed on linkage group 3 of the HCM parent and associated with ζ-carotene, β-carotene and phytoene accumulation (Table 2 and Fig. 1). Other putative QTLs with LOD scores higher than 2.2 were linked and associated with phytoene, ζ-carotene and β-carotene accumulation in the HCM-coupling linkage group 9, where three other QTLs were detected for phytoene (Table 2 and Fig. 2). Putative QTLs for α-carotene and phytoene content mapping to linked regions were also observed in linkage group 1 of the HCM parent (Fig. 2). QTL regions contributing to maysin concentration and larval weight of corn earworm have been reported by Byrne et al. (1998) as mapping to the same putative QTL intervals. Li et al. (1999) reported that defense response (DR) loci were not randomly distributed throughout the wheat genome, but rather were located in clusters and/or in distal gene-rich regions of the chromosomes. Khavkin and Coe (1997) found that genes for growth, development and stress response were located in functional clusters, 10–30 cM long, and distributed non-randomly along all ten chromosomes of corn.

Possible origins and practical implications of clustering of pathway loci

A simulation performed by Williams and Hurst (2000) suggested the attractive idea that proteins encoded by linked genes evolve at similar rates, and it is reasonable to assume that there are evolutionary mechanisms that act to keep the genes for enzymes of the same biosynthetic pathway together, as reported in this study with two different mapping populations, and also by Byrne et al. (1998) in their studies of the maysin pathway. Khavkin and Coe (1997) speculated that clusters of developmental genes in corn were in fact the major QTLs and that many plant reactions to abiotic stress, when mediated by the growth machinery, rely on gene clusters. Clustering of genes is advantageous since the products are maximized and the "pathway mates" are not easily lost due to recombination events. It is well known that carotenoids protect chlorophylls from photo-oxidation, and they act as light-harvesting pigments and photoreceptors (e.g. Simon 1997). Yet it is challenging to explain how, while all plants have the same protective carotenoids in the leaves, domesticated carrot roots developed the unusual orange color when many other plant species did not.

Focused analysis of clustered, functionally related loci could simplify genome sequencing. Evaluation of BAC libraries and cloning by targeting the analyses to these information-rich regions of the genome may often be preferable to scattered prospecting of entire genomes. According to Khavkin and Coe (1997), the association of functionally related genes helps the cooperative recruitment of transcription factors into multicomponent regulatory modules of high affinity. Efforts to identify such hypothesized related transcription factors could be focused primarily on the clustered regions affecting the phenotypic variation of major products of the carotenoid pathway in carrots.

Acknowledgements This work was carried out in compliance with U.S. laws governing genetic experimentation.

References

Armstrong GA (1994) Eubacteria show their true colors: genetics of carotenoid pigment biosynthesis from microbes to plants. J Bacteriol 176:4795–4802

Blauth SL, Churchill GA, Mutschler MA (1998) Identification of quantitative trait loci associated with acylsugar accumulation using intraspecific populations of the wild tomato, *Lycopersicon pennellii*. Theor Appl Genet 96:458–467

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 Am Soc Hort Sci 124:32–38

Byrne PF, McMullen MD, Snook ME, Musket TA, Theuri JM, Widstrom NW, Wiseman BR, Coe EH (1996) Quantitative trait loci and metabolic pathways: genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. Proc Natl Acad Sci USA 93:8820–8825

Byrne PF, McMullen MD, Wiseman BR, Snook ME, Musket TA, Theuri JM, Widstrom NW, Coe EH (1998) Maize silk maysin concentration and corn earworm antibiosis: QTLs and genetic mechanisms. Crop Sci 38:461–471

Cunningham FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol 49:557–583

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

Fourmann M, Barret P, Renard M, Pelletier G, Delourme R, Brunel D (1998) The two genes homologous to Arabidopsis *FAE1* co-segregate with the two loci governing erucic acid content in *Brassica napus*. Theor Appl Genet 96:852–858

Giovannucci E, Ascherio A, Rimm EB, Stampfer MJ, Colditz GA, Willett WC (1995) Intake of carotenoids and retinol in

relation to risk of prostate cancer. J. Natl Cancer Inst 87:1767–1776
Giuliano G, Aquilani R, Dharmapuri S (2000) Metabolic engineering of plant carotenoids. Trends Plant Sci 5:406–409
Gould KA, Dietrich WF, Borenstein N, Lander ES, Dove WF (1996) *Mom1* is a semi-dominant modifier of intestinal adenoma size and multiplicity in *Min*/+ mice. Genetics 144:1769–1776
Hamerschmidt I (1993) Produção de hortaliças e assistência técnica no Brasil. Hort Bras 11:156–157
Hamilton LC (1990) A First course in applied statistics. Brooks/Cole Publishing Company, Pacific Grove, Calif.
Khavkin E, Coe E (1997) Mapped genomic locations for developmental functions and QTLs reflect concerted groups in maize (*Zea mays* L.). Theor Appl Genet 95:343–352
Knapp SJ, Hooloway JL, Bridges WC, Liu BH (1995) Mapping dominant markers using F_2 matings. Theor Appl Genet 91:74–81
Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175
Kruglyak L, Lander ES (1995) A nonparametric approach for mapping quantitative trait loci. Genetics 139:1421–1428
Li WL, Faris JD, Chittoor JM, Leach JE, Hulbert SH, Liu DJ, Chen PD, Gill BS (1999) Genomic mapping of defense response genes in wheat. Theor Appl Genet 98:226–233
Liu BH (1997) Statistical genomics: linkage, mapping, and QTL analysis. CRC Press, Boca Raton, Fla.
Mather K (1936) Types of linkage data and their value. Ann Eugen 6:399–410
McMullen MD, Byrne PF, Snook ME, Wiseman BR, Lee EA, Widstrom NW, Coe EH (1998) Quantitative trait loci and metabolic pathways. Proc Natl Acad Sci USA 95:1996–2000
Ott J (1999) Analysis of human genetic linkage (3rd edn). Johns Hopkins University Press, Baltimore, Md.
Prioul JL, Pelleschi S, Sene M, Thevenot C, Causse M, de Vienne D, Leonardi A (1999) From QTLs for enzyme activity to candidate genes in maize. J Exp Bot 50:1281–1288
Rubatzky VE; Quiros CF; Simon PW (1999) Carrots and related vegetable Umbelliferae (Crop production science in horticulture, vol 10). CAB International, Wallingford, UK
Sandmann G (1998) Carotenoid biosynthesis in microorganisms and plants. Eur J Biochem 223:7–24
Santos CAF (2001) Biometrical studies and quantitative trait loci associated with major products of the carotenoid pathway of carrot (*Daucus carota* L.). PhD Thesis, University of Wisconsin-Madison
SAS (1989) SAS/STAT User's guide, version 6 (4th edn), vol 2. SAS Institute, Cary, N.C.
Simon PW (1992) Genetic improvement of vegetable carotene content. In: Bills DD, Kung S (eds) Biotechnology and nutrition. Butterworth-Heinemann, Boston, Mass., pp 291–300
Simon PW (1997) Plant pigments for color and nutrition. Hort Sci 32:12
Simon PW (2000) Domestication, historical development, and modern breeding of carrot. Plant Breed Rev 19:157–190
Simon PW, Wolff XY (1987) Carotenes in typical and dark orange carrots. J Agricul Food Chem 35:1017–1022
Simon PW, Wolff XY, Peterson CE, Kammerlohr DS, Rubatzky VE, Strandberg JO, Bassett MJ, White JM (1989) High carotene mass carrot population. Hort Sci 24:174
Simon PW, Peterson CE, Gabelman WH (1990) B493 and B9304, carrot inbreds for use in breeding, genetics, and tissue culture. Hort Sci 25:815
Simpson KL (1983) Relative value of carotenoids as precursors of vitamin A. Proc Nutrition Soc 42:7–17
Stam P, van Ooijen JW (1995) JOINMAP™ version 2.0: software for the calculation of genetic maps. CPRO-DLO, Wageningen, The Netherlands
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
Van Ooijen JW (1994) A computer program for drawing genetic linkage maps. CPRO-DLO, Wageningen, The Netherlands
Vivek BS, Simon PW (1999) Linkage relationships among molecular markers and storage root traits of carrot (*Daucus carota* L. spp. *sativus*. Theor Appl Genet 99:58–64
Vos P, Hogers R, Bleeker M, Reijans M, van der Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
Williams EJB, Hurst L (2000) The proteins of linked genes evolve at similar rates. Nature 407:900–903
Ye X, Salim AB, Kloti A, Jing Z, Lucca P, Beyer P, Potrykus I (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287:303–305