Theor Appl Genet (2014) 127:2051–2064
DOI 10.1007/s00122-014-2360-5

ORIGINAL PAPER

High-density single nucleotide polymorphism (SNP) array mapping in *Brassica oleracea*: identification of QTL associated with carotenoid variation in broccoli florets

Allan F. Brown · Gad G. Yousef · Kranthi K. Chebrolu · Robert W. Byrd · Koyt W. Everhart · Aswathy Thomas · Robert W. Reid · Isobel A. P. Parkin · Andrew G. Sharpe · Rebekah Oliver · Ivette Guzman · Eric W. Jackson

Received: 19 February 2014 / Accepted: 13 July 2014 / Published online: 14 August 2014

Abstract

Key message **A high-resolution genetic linkage map of *B. oleracea* was developed from a *B. napus* SNP array. The work will facilitate genetic and evolutionary studies in *Brassicaceae*.**

Abstract A broccoli population, VI-158 × BNC, consisting of 150 $F_{2:3}$ families was used to create a saturated *Brassica oleracea* (diploid: CC) linkage map using a recently developed rapeseed (*Brassica napus*) (tetraploid: AACC) Illumina Infinium single nucleotide polymorphism (SNP) array. The map consisted of 547 non-redundant SNP markers spanning 948.1 cM across nine chromosomes with an average interval size of 1.7 cM. As the SNPs are anchored to the genomic reference sequence of the rapid cycling *B. oleracea* TO1000, we were able to estimate that the map provides 96 % coverage of the diploid genome. Carotenoid analysis of 2 years data identified 3 QTLs on two chromosomes that are associated with up to half of the phenotypic variation associated with the accumulation of total or individual compounds. By searching the genome sequences of the two related diploid species (*B. oleracea* and *B. rapa*), we further identified putative carotenoid candidate genes in the region of these QTLs. This is the first description of the use of a *B. napus* SNP array to rapidly construct high-density genetic linkage maps of one of the constituent diploid species. The unambiguous nature of these markers with regard to genomic sequences provides evidence to the nature of genes underlying the QTL, and demonstrates the value and impact this resource will have on *Brassica* research.

Communicated by Maria Laura Federico.

A. F. Brown (✉) · G. G. Yousef · K. K. Chebrolu · R. W. Byrd · K. W. Everhart · A. Thomas · I. Guzman
Department of Horticultural Science, Plants for Human Health Institute, North Carolina State University, Kannapolis, NC 28081, USA
e-mail: afbrown2@ncsu.edu

R. W. Reid
UNC Charlotte Bioinformatics Services Division, University of North Carolina-Charlotte, 150 Research Campus Dr., Kannapolis, NC 28081, USA

I. A. P. Parkin
Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, SK S7N0X2, Canada

A. G. Sharpe
National Research Council Canada, 110 Gymnasium Place, Saskatoon, SK S7N0W9, Canada

R. Oliver · E. W. Jackson
General Mills Crop Biosciences, Kannapolis, NC 28081, USA

Introduction

One of the most significant challenges to human health and economic security is from chronic health disorders such as type 2 diabetes, obesity, heart disease, and multiple forms of cancer (Thompson and Thompson 2009). Fortifying fruits and vegetables through enhancement or modification of health promoting secondary metabolites is a straightforward preventative approach with significant potential impacts on health. Broccoli (*Brassica oleracea* L. *italica*) is an ideal model system for studying genetic factors that influence the accumulation of these compounds due to the unique profiles of an array of secondary metabolites that include glucosinolates, carotenoids, tocopherols, ascorbate, phyloquinone, sterols, phenolic acids, and flavonoids. In addition, the conserved sequence homology and synteny of the species to both the related model plant *Arabidopsis*

thaliana and to economically important amphidiploid oilseed crops of the same genus (rapeseed) provides opportunities to leverage a wealth of available genomic resources (Ayele et al. 2005; Babula et al. 2003; Li and Quiros 2003; Qiu et al. 2009; Schranz et al. 2007; Town et al. 2006).

Broccoli florets accumulate significant levels of five carotenoids: lutein, β-carotene, neoxanthin, violaxanthin and lutein epoxide, with trace levels of additional compounds occasionally observed (de Sá and Rodriguez-Amaya 2004; Farnham and Kopsell 2009; Guzman et al. 2012; Heinonen et al. 1989; Holden et al. 1999; Kurilich et al. 1999; Müller 1997; Murkovic et al. 2000; Zhang and Hamauzu 2004). While carotenoids as a class have generated interest due to anti-oxidant and pro-vitamin A (retinol) properties, lutein (3R,3′R,-6′β, ε-carotene-3,3′diol) provides additional benefits. A di-hydroxy xanthophyll carotenoid found in dark green vegetables, lutein is one of only two carotenoids (along with the stereoisomer zeaxanthin) found in the lens and macular region of the human retina and is associated with decreased risks of retinal degenerative diseases. Increased lutein consumption leads to a significant increase in both blood serum concentration and macular pigment optical density, measurements that are highly correlated to reduced onset and progression of age-related macular degeneration and cataracts (Sujak et al. 1999). In the plant, lutein provides important functions in photosystem structure and in adaptation to high light stress environments (DellaPenna and Pogson 2006; Hirschberg 2001). Significant genetic variation of lutein and other carotenoids among broccoli inbreds and hybrids has been previously observed (Farnham and Kopsell 2009; Guzman et al. 2012).

In-depth reviews of carotenoid metabolism and regulation in *A. thaliana* are available (Cazzonelli and Pogson 2010; Ruiz-Sola and Rodríguez-Concepción 2012) and provide a convenient genetic model for what is likely a homologous system in *B. oleracea*. Briefly, the methylerythritol phosphate pathway produces geranylgeranyl diphosphate (GGPP) as a precursor for the synthesis of isoprenoids such as carotenoids, tocopherols, phylloquinones, and the phytol tail of chlorophyll. The first committed step of the carotenoid pathway occurs when phytoene synthase (PSY)

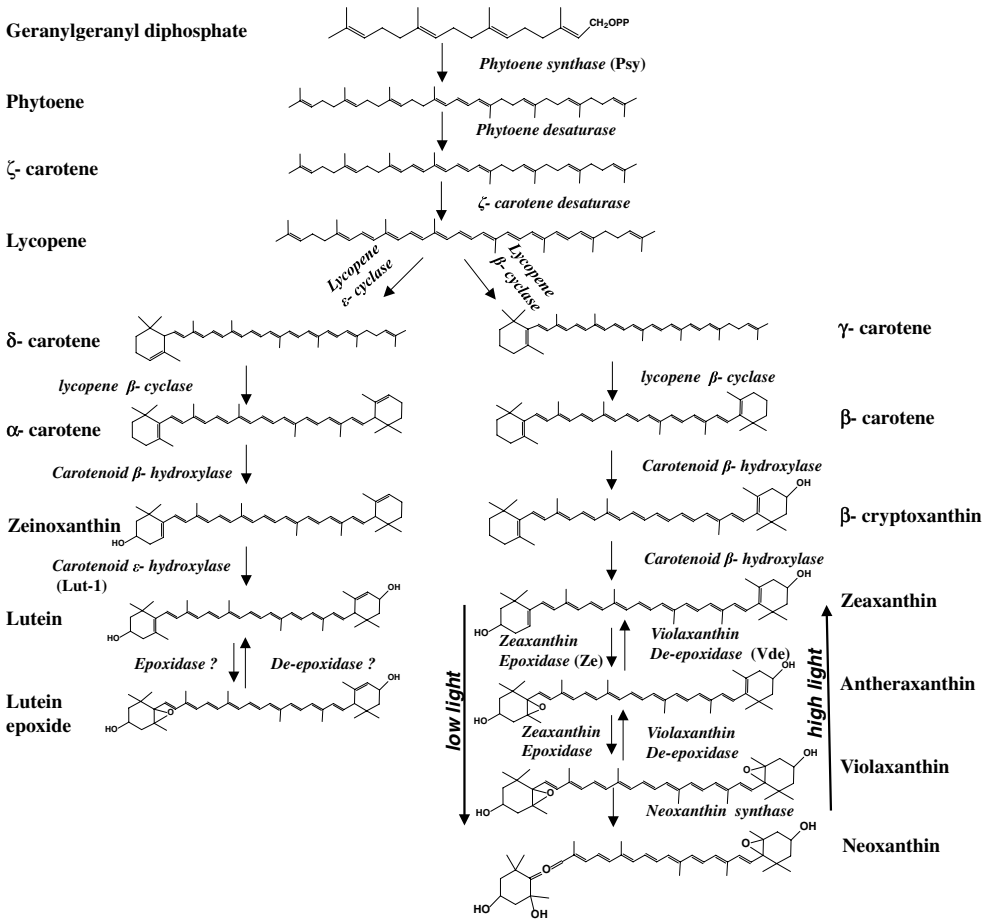

Fig. 1 A simplified scheme for the biosynthesis pathway for carotenoids (García-Plazaola et al. 2007; Ruuiz-Sola and Rodriguez-Concepcion 2012) isolated in broccoli florets

condenses two molecules of GGDP to form phytoene (Fig. 1). Phytoene isomerase and ζ-carotene desaturase catalyze two dehydrogenation reactions to convert phytoene to lycopene with the aid of an electron carrier plastoquinone (Ruiz-Sola and Rodríguez-Concepción 2012). Following the synthesis of lycopene, the first branching point in the pathway is catalyzed by closely related enzymes: a β-cyclase which mediates the formation of the β-rings of β-carotene, and a ε-cyclase which performs a similar function with ε-rings to produce α-carotene. In Arabidopsis, xanthophylls are generated from carotenoids through two non-heme iron-oxygenases and two cytochrome P450 enzymes that regulate the synthesis of zeaxanthin and lutein, respectively. Violaxanthin and lutein epoxide are products of two distinct reversible xanthophyll cycles that are thought to be catalyzed by the same enzymes: violaxanthin de-epoxidase and zeaxanthin epoxidase (García-Plazaola et al. 2007). The importance of light, temperature and developmental cues on the regulation of carotenoids is recognized, but the specific mechanisms involved are not well understood (Ruiz-Sola and Rodríguez-Concepción 2012).

A considerable number of molecular markers (and mapping populations) have been developed for genetic studies in *B. oleracea* (Babula et al. 2003; Bohuon et al. 1996; Brown et al. 2007; Camargo et al. 1997; Gao et al. 2007; Hu et al. 1998; Iniguez-Luy et al. 2009; Kianian and Quiros 1992; Lan and Paterson 2000; Landry et al. 1992; Li and Quiros 2003; Ramsay et al. 1996; Sebastian et al. 2000; Slocum et al. 1990). Marker systems utilized in these studies have included restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), sequence-related amplified polymorphisms (SRAP), and simple sequence repeats (SSR). For the most part, these maps and mapping populations have been generated using parents of different subspecies (or varietal groups) such as broccoli and cabbage, in part to exploit the greater degree of genetic variation that exists between the subspecies (Brown et al. 2007). While these resources have been useful in mapping QTL associated with shared traits common among subspecies such as disease resistance or time to flowering (Nagaoka et al. 2010) they are of limited value in subspecies traits such as the tissue-specific accumulation of secondary metabolites in the specialized florets of broccoli.

The recent development of an Illumina high-density (60K iSelect format) single nucleotide polymorphism (SNP) array for rapeseed *(Brassica napus* AACC*)* should provide rapid marker genotyping on a scale not previously observed and should provide ample markers to generate saturated maps of the constituent diploid species *B. oleracea* (CC) and *B. rapa* (AA). The development of the array addresses an additional concern of *Brassica* researchers regarding the difficulty of aligning maps constructed with different sets of molecular markers. As the SNP markers are anchored to the reference genome of the *B. oleracea* rapid cycling 'TO1000' (Parkin et al. submitted), the markers represent unambiguous physical loci and have the additional benefit of providing ready access to the identification of putative candidate genes that co-localize with QTL. The objectives of this study were: (1) to utilize the recently developed SNP array to generate a saturated genetic linkage map of an $F_{2:3}$ broccoli population that is segregating for carotenoid concentration in florets and (2) utilize the resultant map to identify QTL (and potential candidate genes) associated with the accumulation of individual or total carotenoids in broccoli florets over 2 years.

Materials and methods

Plant material

The development of the mapping population 'VI-158' × 'Broccolette Neri *E. Cespuglio*' (BNC) has been previously described (Brown et al. 2007). 'VI-158' is a doubled haploid Calabrese type derived from the F_1 hybrid 'Viking'. 'BNC' is a highly branched landrace accession with dark green leaves and white flowers that produces a small centralized head, the size of a single standard floret. The divergent phytochemical profiles of these lines have been widely reported (Brown et al. 2002; Eberhardt et al. 2005; Guzman et al. 2012; Kushad et al. 1999). Two replicates of 150 $F_{2:3}$ families were transplanted to adjacent fields at the Piedmont Research Station, Salisbury, NC, on Sept. 11 2009 and 2010. Plants were approximately 5 weeks old at the time of transplant. Ten plants per replicate were spaced 30 cM apart on raised beds with black plastic mulch and drip irrigation. Standard practices were applied for pest control and fertilization.

Harvest of florets began on Nov. 13 and Nov. 18 in 2009 and 2010, respectively. Harvesting was completed on Dec. 18 and Dec. 1 in 2009 and 2010, respectively. As this population segregates for head size and harvest date, multiple harvests were conducted at a uniform stage of compactness. This approach was used previously with the same population to identify year consistent QTL associated with harvest date and head size (Brown et al. 2007). No fewer than five heads were harvested from each replicate (a minimum of ten progeny total per family). To standardize the floret tissue samples, heads were cut to similar sized florets with equal proportions of stalk tissue and immediately flash frozen in liquid nitrogen. Samples were then stored at −80 °C until lyophilized. Florets were lyophilized using a freeze dryer (VirTis 24Dx48; SP Scientific, Stone Ridge, NY) with a temperature-controlled chamber for

samples. Freeze-dried florets were then stored at −80 °C until extraction as described below. Lyophilized tissue was ground to a fine powder using a coffee bean grinder. Samples were stored at −20 °C in the dark until analysis. As multiple harvest dates occurred for each replication, proportional ground tissue from each date (weighted by the number of heads harvested on each date) was combined into a bulked sample for analysis from each replication.

Metabolite extraction

Floret tissue was analyzed using methodology described by Guzman et al. (2012). Briefly, under yellow lights, 0.2 g of freeze dried finely ground tissue was extracted with 5 mL pure ethanol and shaken for 20 min on ice at 180 rpm. Three extractions with ethanol produced a color-free pellet. After each extraction, the extracts were centrifuged at 2,900g_n for 15 min at 10 °C to collect the supernatant. The supernatants were pooled, completely dried under nitrogen gas, and re-suspended in 1.5 mL ethanol. The ethanolic extract was filtered through a 0.2 μm polytetrafluoroethylene filter (Nalgene; Rochester, NY, USA) into an amber HPLC vial and stored under nitrogen gas at −80 °C until chromatography analysis. Triplicate extractions were performed for each sample.

Chromatography

For the characterization of carotenoids, a Waters ACQUITY Ultra Performance Liquid Chromatography (UPLC) system with a photodiode array (PDA) detector was used. The column used was a 2.1 × 100 mm i.d., 1.8 μm (Waters ACQUITY HSS T3, Waters; Milford, MA, USA) set at 35 °C. The solvent system consisted of 0.05 % ammonium acetate (A) and 74:19:7 (v/v/v) acetonitrile:methanol:chloroform (B). The solvent elution included three linear phases (0–18 min. at 75 % B; 18–19 min. from 75 % B to 100 % B; and 19 to 30 min. at 100 % B to 98 % B). Flow rate was set at 0.4 mL/min and the injection volume was 5 μL. Spectra were observed for the wavelengths between 240 and 670 nm. Carotenoids were detected at 450 nm and identified according to retention time and spectrum comparison with published data (Guzman et al. 2012; Kurilich et al. 1999; Murkovic et al. 2000).

DNA extraction and SNP analyses

DNA was extracted from freeze-dried leaf tissue of bulked F3 plants (minimum of 12 plants/family) using the modified CTAB procedure described by Kidwell and Osborn (1992) and hybridized to the Illumina Infinium iSelect (Brassica60k_Cons_ParkinAAFC) array according to manufacturers' protocols. The arrays were scanned with Illumina iSCAN and alleles were defined with the Genotyping Module v.1.9.5 of the Illumina GenomeStudio software. Parental lines from the mapping population were used to identify polymorphic assays, and non-polymorphic SNPs were eliminated. Heterozygote controls, derived from combining DNA from both mapping parents, were used to facilitate and validate heterozygous genotype calls. Allele assignments were evaluated visually using both polar and Cartesian coordinates, and adjusted manually based on parental and heterozygous controls.

Data analysis

Statistical analyses were performed using SAS software (version 9.2 for Windows; SAS Institute, Cary, NC). Means, standard deviations, ranges and tests of normality were obtained with Proc Univariate Normal Plot and analysis of variance (ANOVA) was conducted for all traits with all factors (genotype, years, replication, genotype × year) considered random using Proc GLM procedures. The linear random model used *was* $y_{ijk} = u + G_i + E_j + R_k + G \times E_{ij} = e_{ijk}$; where y_{ijk} is the trait measurement associated with individual *ijk*, u is the mean or common effect of the population, G_i is genetic effect associated with family i, E_j is the year effect, R_k is the block (replication) effect, $G \times E$ is the genotype × year interaction. Pearson correlation coefficients (r) were generated among all compound combinations, average harvest date and average head size, using Proc Corr. statement.

Map construction was accomplished using JoinMap® 4.0 (Ooijin 2009) with Kosambi mapping function. Logarithm of odds (LOD) scores was generated from 10.0 to 20.0 with a maximum recombination fraction of 0.35. Polymorphic markers were assigned to 1 Mb sized bins according to expected locations on the genomic reference sequence. A preliminary map was constructed by randomly selecting two markers from each bin. All markers were then added sequentially from bins in successive rounds of analysis. Markers that were redundant, mapped out of sequence or grossly altered the order of previously selected markers were removed. After a single pass through each chromosome, the process was repeated, from the opposite end of the chromosome. The final map had to meet the strict criteria that no markers could be rejected by the default goodness of fit criteria (sharp reduction in scores or estimated negative distances) that would require a second or third round of analysis. Analysis was conducted using both regression and maximum likelihood approaches and produced similar map orders. A LOD score of 2.0 was used to determine the optimal marker order by applying a ripple value of 1.0 and a jump threshold of 5.0.

MapQTL® 6 (Ooijin 2009) was used to identify QTL associated with individual and total carotenoids. Analysis

was conducted by year and also using the average family value for both years. Preliminary QTL analysis of the individual compounds produced nearly identical results for neoxanthin and violaxanthin and these compounds were reported as a single class of β-xanthophylls for analysis. Single factor analysis was performed using the MapQTL Kruskal–Wallis non-parametric test and the results were used to select markers for inclusion in the programs for automated co-factor selection. Several iterations were conducted to produce an optimal set of co-factors for each analysis. Non-restricted multiple-QTL mapping (MQM) was used with the default settings adjusted to a scan distance of 0.2 cM. Genome-wide threshold values ($P < 0.05$) for declaring the presence of QTLs were estimated from 1,000 permutations of each phenotypic trait using the programs provided bootstrapping algorithm. Confidence intervals were established using a 1 LOD drop off on either side of the maximum score. QTL from each analysis was considered the same if the combined confidence intervals overlapped and the magnitude and direction of the QTL effect were common between years. Geneious® version 6.1.5 (Biomatters: http://www.geneious.com) was used to conduct protein to nucleotide Blast searches in all 6 possible reading frames to the TO1000 *B. oleracea* reference genome draft (Agriculture and Agri-Food Canada, Saskatoon, Canada) for putative candidate genes occurring within or adjacent to significant intervals. A list of the *Arabidopsis* protein sequences used as queries for these searches is found in Table 1. The closest flanking A genome markers were also used to conduct positional searches of the publically available *B. rapa* genome sequence (Chiifu-401, version 1.5) through the Brassica database (BRAD) (http://www.plantgdb.org/BrGDB) using the same *Arabidopsis* sequence queries. The alignment program Exonerate (Slater and Birney 2005) was used to determine an amino acid alignment percentage for the putative candidate genes and a 75 % identity match was required to declare a putative candidate from either sequence. While a 75 % match is not stringent enough to discriminate between paralogues of the same gene, the intention was to identify a limited set of potential candidates from the reference sequence for future investigation.

Table 1 *Arabidopsis* protein sequences used for Blast searches of *B. oleracea* and *B. rapa* genome sequences

Protein	Entry	Enzyme
Phytoene synthase	AT5G17230	EC:2.5.1.32
Phytoene desaturase 3	AT4G14210	EC:1.3.5.5
15-cis-zeta-carotene isomerase	AT1G10830	EC:5.2.1.12
Zeta-carotene desaturase	AT3G04870	EC:1.3.5.6
Carotenoid isomerase	AT1G06820	EC:5.2.1.13
Lycopene epsilon cyclase	AT5G57030	EC:5.5.1.18
Lycopene beta cyclase	AT3G10230	EC:5.5.1.19
Beta-carotene 3-hydroxylase 1	AT4G25700	EC:1.14.13.129
Carotene epsilon-monooxygenase	AT3G53130	EC:1.14.99.45
Zeaxanthin epoxidase	AT5G67030	EC:1.14.13.90
Violaxanthin de-epoxidase	AT1G08550	EC:1.10.99.3
Lutein deficient	AT1G31800	EC:1.14.-.-]

Results

Mapping

The Illumina *Brassica* Infinium array consists of 52,157 SNP markers from the C (*B. oleracea,* $2n = 18$) and A (*B. rapa,* $2n = 20$) genomes of *B. napus* (AACC, $2n = 38$), 30,420 of which are physically positioned on the C genome. A total of 4,990 useful polymorphisms were revealed in this population. These consisted of 4,645 SNPs assigned to the C genome, 238 SNPs currently unassigned, and 107 SNPs assigned to the A genome. SNP marker nomenclature includes a genome designation (A or C) followed by the position of the SNP as referenced by the 'Chiifu-401' or 'TO1000' genome sequence, respectively. Unassigned SNPs are designated by their respective sequenced scaffold number and relevant base pair position. Markers were added sequentially to map files in bins representing 1 Mbp of the reference genomic sequence. Efforts were made to retain the predicted order of markers. As the objective was to produce a saturated linkage map for QTL analysis, redundant markers were removed from the map. Non-sequential markers that deviated from expected positions were included only when the following conditions were met: three or more closely linked (or redundant) markers identified the same change in order, markers were needed to fill gaps larger than 2 cM, and the addition of non-sequential markers did not grossly alter the order of sequential markers previously added. After linkage groups were formed, "A" genome-specific SNP markers were added to allow for positional cross-taxonomic comparisons to *B. rapa*. Unassigned SNPs were added if they proved informative (non-redundant) and if multiple SNPs from the same sequencing scaffold mapped to the same relative position on the chromosome.

The resultant genetic linkage map consisted of 547 non-redundant markers spanning 948.1 cM across nine chromosomes (Fig. 2). Average interval length on map was 1.7 cM with no interval exceeding 7.5 cM, and only 14 intervals (~3 %) exceeding 5 cM. The map covered 429,265,051 base pairs of the current reference assembly of 446,905,700 (96 %). Of the markers that were removed due to the previously described criteria, only 67 (~0.014 %) were predicted

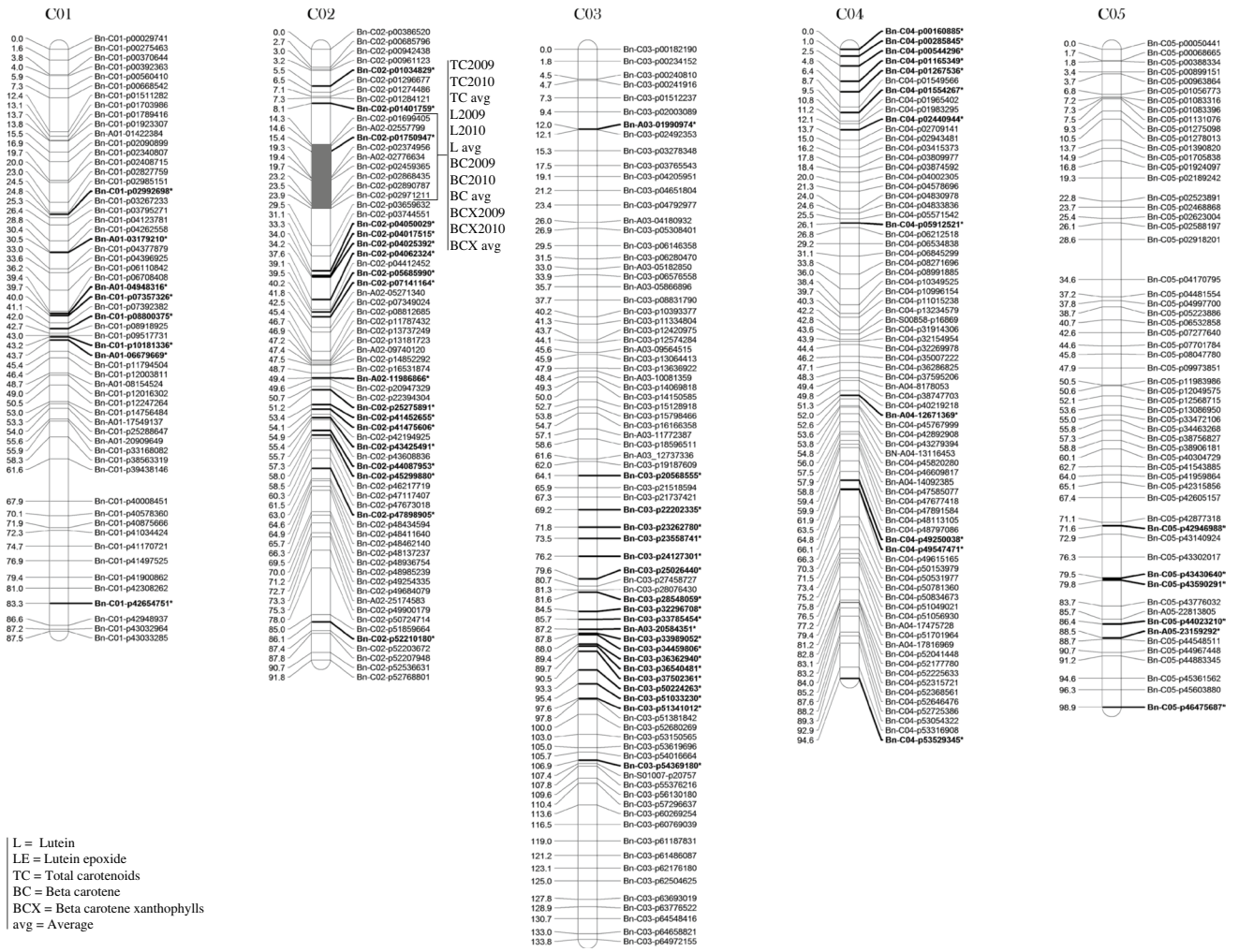

Fig. 2 Genetic linkage map for the $F_{2:3}$ broccoli population (VI158 × BNC) constructed using single nucleotide polymorphism (SNP) markers. Markers in *bold* and with a *asterisk* are regions of segregation distortion ($p \leq 0.05$), with a Chi-square test of $p \leq 0.05$. *C0* Chromosome

(based on the reference genome) to occur beyond the outermost markers of our map. For the most part, the synteny of the mapped SNPs was in consensus with the reference genomic sequence with only nine minor changes in order across all chromosomes. In all cases, two adjacent markers were inverted and efforts to resolve these discrepancies were not affected by the inclusion of redundant or adjacent markers.

Some degree of segregation distortion was apparent on all chromosomes with 136 of 547 (25 %) markers deviating from expected Mendelian ratios (indicted in bold type in Fig. 2). Distortion was confined to specific chromosomal regions and was characterized by similar patterns. For example, the first 13 markers on chromosome 9 (C09) and the first 25 markers of C07 were underrepresented in the homozygous state by the B (BNC) allele. Conversely, a smaller region near the distal end of C05 was underrepresented by the A (VI-158) allele. Attempts to substitute redundant or adjacent markers in these regions did not affect the observed distortion.

Data analysis

Frequency distributions for all carotenoids approximated normality (Fig. 3). While variation due to year had a major impact, significant genetic effects were observed for all carotenoids evaluated in this study (Table 2). No significant genotype by year or replication effects was evident (with the exception of neoxanthin). Standard deviations and range of individual and total carotenoid accumulation were relatively consistent between years (Table 3), but mean concentrations were higher in 2010 for all compounds except lutein. A threefold variation in lutein among families was observed in both years with a mean family concentration of 105.9 µg/g dried weight (DW) in 2009

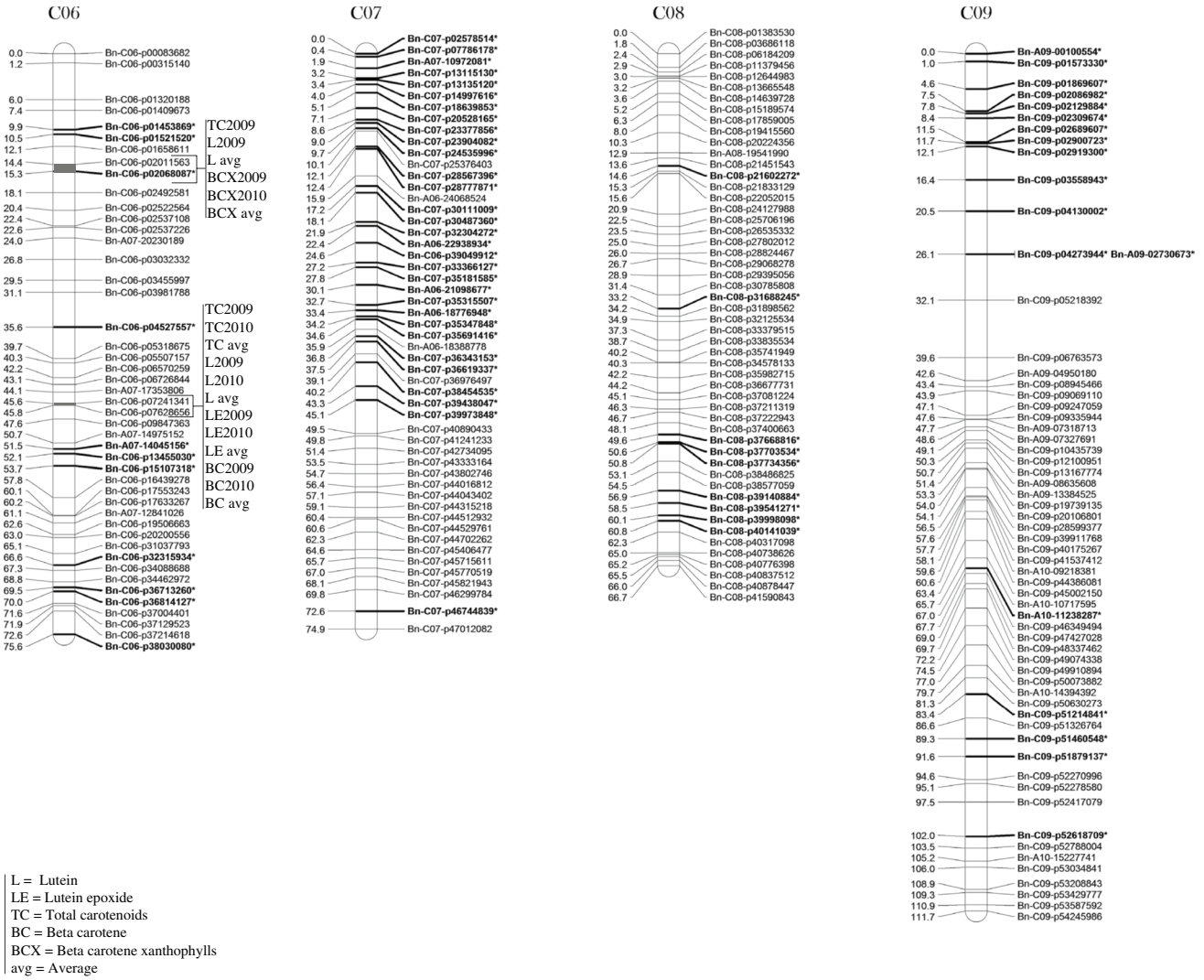

Fig. 2 continued

and 105.8 µg/g in 2010. The same family (VB042) accumulated the highest concentration of lutein in both years of the study and exceeded the next highest family value by 15 % (data not shown). A twofold to threefold variation was also observed in regards to concentrations of additional individual and total carotenoids with higher levels observed in 2010. With the exception of lutein epoxide, all carotenoid concentrations were highly correlated among families in both years of the study and ranged between $r = 0.82$ and 0.97 (Table 4). Lutein epoxide was significantly correlated with other carotenoids but the relationship was considerably weaker ($r = 0.40$–0.57). Head weight was not significantly correlated with carotenoid concentration (with the exception of lutein epoxide in 2010). Harvest date was inversely correlated with all carotenoids (with lutein epoxide again as the sole exception) with r ranging from -0.42 to -0.73.

QTL analysis

Three QTLs (designated *bocarot1*, *bocarot2*, and *bocarot3*) were significantly associated with total or individual carotenoids in both years of analysis and met or exceeded the 95 % genome-wide significance level (Table 5). Near the proximal end of chromosome 2, *bocarot1*, was associated with total carotenoids, lutein, β-carotene, and β-xanthophylls in both years. Estimates of the most likely position of this QTL ranged from 13.1 to 21.7 cM in 12 independent analyses of individual compounds (by year and the average of the 2 years). This region represents approximately 900 kb of the reference genome (1.7–2.6 Mbp). The amount of phenotypic variation explained by *bocarot1* ranged from a low of 6.9 % (β-carotene 2009) to a high of 24 % (total carotenoids 2010). This QTL was not associated with lutein epoxide in either year of study. The

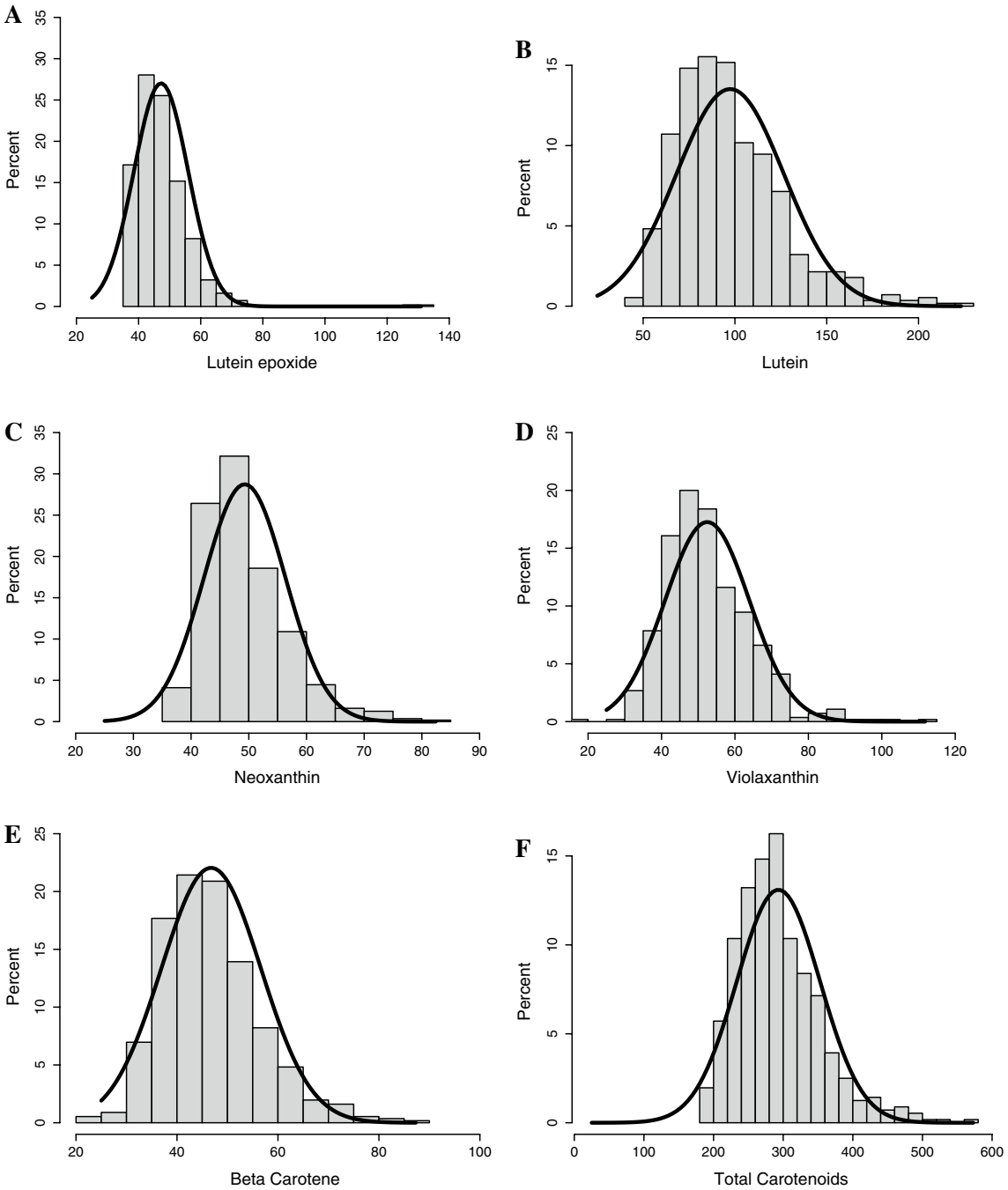

Fig. 3 Phenotypic distribution for individual carotenoids in the floret of $F_{2:3}$ broccoli population (VI-158 × BNC) across 2 years (2009 and 2010) at Salisbury, NC. (**a** Lutein epoxide, **b** lutein, **c** Neoxanthin, **d** violaxanthin, **e** β-carotene, **f** total carotenoids)

allele responsible for enhanced carotenoids was contributed by the parent VI-158 and was additive in nature.

Two linked QTLs were detected on chromosome 6 (*bocarot2* and *bocarot3*) approximately 28 cM apart and separated by 13 intervals. The first (*bocarot2)* was significantly associated with the β-xanthophylls in both years (and the average of both years) explaining up to 22.4 % of the phenotypic variation. This QTL was also significantly associated with lutein in 2009 (and lutein average of both years) and for total carotenoids in 2009 but had a lesser effects on these measurements as evidenced by the percentage of phenotypic variation explained (6.3–8 %). The most likely position of this QTL was estimated to be between 12.7 and 16.3 cM and represents a region of about 400 kb of the reference genome. It should be noted that a non-significant (LOD ~3.0) peak could be detected for lutein in

Table 2 Source of variation (ANOVA) for carotenoids in the $F_{2:3}$ broccoli population evaluated in 2009 and 2010, Salisbury, NC

Source	Lutein epoxide	Lutein	Neoxanthin	Violaxanthin	β-Carotene	Total carotenoids
Genotype (G)	130**[a]	1,824**	95**	269**	217**	7,552**
Year (Y)	344*	40,757**	3,531**	4,400**	465**	13,5243**
Replication	42	5	88*	20	1	128
G × Y	58	298	19	25	44	1,417
Error	53	428	24	69	52	1,842
R^{2}[b]	0.65	0.75	0.76	0.74	0.73	0.75
CV[c]	15.5	21.2	10.1	15.8	15.4	14.6

[a] Mean square values followed by * or ** are significantly different at $P < 0.05$ or $P < 0.01$, respectively

[b] R^2, variation explained by model

[c] Coefficient of variation

Table 3 Means, standard deviations, and ranges of carotenoid concentrations in floret of parental lines and derived $F_{2:3}$ broccoli families

Compound/year	Parental lines		$F_{2:3}$ families		
	VI-158	BNC	Mean[a]	SD[b]	Range
Lutein					
2009	64.15	108.4	105.9	27.6	63.0–197.8
2010	–	–	105.8	26.0	57.8–210.0
β-Carotene					
2009	40.4	59.8	47.6	9.2	28.9–76.3
2010	–	–	55.6	10.4	38.2–93.5
Neoxanthin					
2009	41.0	53.2	51.8	6.6	40.8–71.2
2010	–	–	56.2	5.6	46.3–75.7
Violaxanthin					
2009	36.0	56.4	55.3	11.1	33.9–91.3
2010	–	–	59.9	10.4	37.1–90.5
Lutein epoxide					
2009	44.3	46.6	49.3	17.4	38.1–218.6
2010	–	–	56.0	7.9	43.0–100.5
Total carotenoids					
2009	260.0	360.2	347.5	58.9	249.4–514.9
2010	–	–	379.1	53.5	275.5–572.1

[a] μg/g freeze-dried tissue

[b] Standard deviation

2010 at this location but fell below the genome-wide significance level.

The second QTL on chromosome 6 (*bocarot3*) was significantly associated with total carotenoids, lutein and β-carotene in both years (and the average) but was not associated with β-xanthophylls (non-significant peaks were not detected at this region). The most likely position of this QTL was estimated to occur between 45 and 46 cM (420 kb) and explained up to 18.9 % of the phenotypic variation associated with these compounds. *Bocarot3* was significant for the accumulation of lutein epoxide in 2009 and 2010 explaining 12.3 and 11.1 % of the phenotypic variation, respectively. The alleles for enhancing carotenoids at both *bocarot2* and *bocarot3* were contributed by the parent 'BNC' and were primarily additive.

Twenty-two putative hits to Arabidopsis carotenoid protein sequences were identified from protein to nucleotide BLAST searches of the *oleracea* (11) and *rapa* (11) genome sequences (Table 6). Two A genome SNPs (BN_A02_02557799 and BN_A02_02776634) occurring within the *bocarot1* confidence interval on C02 were used to conduct positional searches of the *B. rapa* sequence through the *Brassica* database (BRAD). An annotated copy of PSY that spans the region from 3,121,495 to 3,119,201 bp was identified on A02. Similar BLAST searches of the provided *B. oleracea* 'TO1000' draft sequence, also identifies an annotated copy of PSY that is predicted to span the region on C02 from 4,211,748 to 4,209,741 bp. No other known carotenoid candidate gene could be identified in the proximity of *bocarot1*. In a similar manner, two A genome SNPs that flanked *bocarot3* on C06 (BN_A07_17353806 and BN_A07_14975152) were used to search the *B. rapa* database and identified a significant match to Carotene epsilon-monooxygenase (Lut1) on A07 that spanned the region of 14698960 to 14701659 bp. The same gene was identified on C06 spanning the region of 21,285,557 to 21,282,830 Mbp. Unfortunately, no A genome markers occurred in the region of *bocarot2*, but it should be noted that a copy of zeaxanthin epoxide was identified approximately 2.5 Mbp away from Lut1 in *B. rapa*. Searches of the *B. oleracea* sequence failed to identify a potential candidate for this QTL on C06.

Discussion

Map construction

This is the first report of the use of a high-density *B. napus* SNP array to create a saturated linkage map of *B.*

Table 4 Correlation coefficients (*r*) among levels of individual carotenoids in broccoli floret in the $F_{2:3}$ population VI-158 × BNC in 2009 (A) and 2010 (B)

	Neoxanthin	Violaxanthin	Lutein epoxide	Lutein	*β*-Carotene	DTH[a]
(A) 2009						
Violaxanthin	0.91945 (<0.0001)[b]					
Lutein epoxide	0.4523 (<0.0001)	0.43148 (<0.0001)				
Lutein	0.97043 (<0.0001)	0.89485 (<0.0001)	0.44584 (<0.0001)			
β-Carotene	0.88282 (<0.0001)	0.85891 (<0.0001)	0.40453 (<0.0001)	0.88146 (<0.0001)		
DTH	–0.71508 (<0.0001)	–0.61367 (<0.0001)	–0.25017 (0.0036)	–0.73484 (<0.0001)	–0.59464 (<0.0001)	
Head weight	–0.11962 (0.1686)	–0.10635 (0.2213)	–0.16651 (0.0545)	–0.1139 (0.1901)	–0.08986 (0.3018)	–0.23953 (0.0053)
(B) 2010						
Violaxanthin	0.89575 (<0.0001)					
Lutein epoxide	0.44610 (<0.0001)	0.41308 (<0.0001)				
Lutein	0.96543 (<0.0001)	0.84104 (<0.0001)	0.47296 (<0.0001)			
β-Carotene	0.90271 (<0.0001)	0.82255 (<0.0001)	0.57010 (<0.0001)	0.90663 (<0.0001)		
DTH	–0.59674 (<0.0001)	–0.46781 (<0.0001)	0.05032 (0.5450)	–0.59987 (<0.0001)	–0.42331 (<0.0001)	
Head weight	–0.10116 (0.2228)	–0.00305 (0.9707)	–0.29038 (0.0004)	–0.12221 (0.1403)	–0.17334 (0.0358)	–0.22147 (0.0070)

[a] Days to harvest

[b] Values within parenthesis refer to the significance of correlation among carotenoids (i.e., *P* values)

oleracea. The map generated compares favorably with previous RFLP and PCR-based linkage maps of *B. oleracea* in density and coverage (Gao et al. 2007; Iniguez-Luy et al. 2009). The map provided near-complete uniform coverage of the genome with an estimated coverage of ~96 % of the current reference assembly of the rapid cycling *B. oleracea* accession 'TO1000'. Given the interval size and coverage, it is unlikely that any QTL segregating in this population will occur more than 3 cM from the closest marker.

Providing unambiguous genomic locations of markers should be extremely valuable in future genetic and evolutionary studies. Although debated, most estimates of the divergence of the diploid *Brassica* species (*B. oleracea* and *B. rapa*) suggest this event occurred less than 4 million years ago and only recently converged to produce the amphidiploid species *B. napus* through repeated natural hybridizations and genome doubling (Arias et al. 2014). We observed that approximately 17 % of the *B. napus* potential C genome markers produced useful polymorphisms in our population and produced a map with only nine minor changes in order compared to the genome reference. Parkin et al. (1994) and others have previously suggested that the C genomes of *B. napus* and *B. oleracea* are essentially identical and our work provides additional support for this observation. In contrast to the C genome markers, less than 1 % of the *B. napus* A genome SNPs were informative, but these markers did display significant co-linearity on several chromosomes. Twelve A09 and A10 SNP markers, for example, mapped to C09 in a continuous manner and suggested that divergence in base chromosomal number had not disrupted homology between A09/A10 and C09. The relationship between C06 and C07 and the A genome conforms to the observed homoeologous relationships between the A and C genomes in *B. napus* (Parkin et al. 2005). All but one of the polymorphic A07 SNPs mapped to C06 and all polymorphic A06 SNPs mapped to C07. Our searches for putative carotenoid candidate sequences identified homologous sequences on C06 and A07, as well as C07 and A07.

In this population, 1 cM corresponds to ~500 kb of genome sequence on average, but these estimates vary dramatically across the genome. One of the largest intervals on the map (6.3 cM) occurs between markers Bn-C1-39438146 and Bn-C1-40008451, which represents an estimated physical distance of 570 kb. Conversely, an interval on C05 flanked by the markers BN-C5-p13086950 and BN-C5-p33472106 has an estimated recombination distance of 1.4 cM but represents more than 20 Mbp of the reference genome. None of the estimated 136 SNPs on the array predicted to occur in this interval were informative; however, the centromere for C05 has been localized within this region (Parkin et al. submitted), which likely explains the paucity of markers and the observed reduction in recombination. Such regions showing extensive disparity between the physical and genetic distances could potentially indicate areas of increased/decreased recombination or they could represent areas of the draft sequence that might require further refinement.

We observed that 25 % of the 547 SNP markers placed on the map displayed segregation distortion in this F_2-derived population. This represents the same estimate of segregation distortion we had previously made with this population using SSR and SRAP markers (Brown

Table 5 QTL associated with major carotenoids isolated in the $F_{2:3}$ broccoli population VI-158 × BNC evaluated in 2009 and 2010, Salisbury, NC

	(*bocarot1*)[a]/Chromosome 2 (13.1–21.7 cM)						(*bocarot2*)/Chromosome 6 (12.7–16.3 cM)						(*bocarot3*)/Chromosome 6 (45 to 46 cM)					
Compound	Pos[b]	LOD[c]	CI[d]	% σ[e]	A[f]	D[g]	Pos.	LOD	CI	% σp	A	D	Pos.	LOD	CI	% σp	A	D
Total Carotenoids																		
2009	19.4	6.6	19–20	10.4	28.4	3.3	16.3	4.8	14–16.3	7.3	−23.6	−22.3	45.8	5.4	45–46	8.4	−31.6	−4.1
2010	18.4	10.5	15–20	24.1	36.4	13.7	–	–	–	–	–	–	45.6	7.6	45–46	16.7	−30.2	−10.9
Average	17.4	9.2	14–20	17.1	31.7	9.5	–	–	–	–	–	–	45.6	4.9	45–46	8.5	−24.9	−6.7
Lutein																		
2009	15.4	8.0	14–19	13.3	18.2	-2.1	12.7	5.1	12–15	8.0	-13.1	-9.5	45.8	4.0	45–46	6.2	−13.5	0.4
2010	19.3	9.2	18–20	17.6	13.5	11.9	–	–		–	–	–	45.6	8.4	45–46	18.1	−16.7	−4.8
Average	13.1	9.2	13–15	16.7	16.8	1.9	14.4	3.8	14–16	6.3	-9.6	-8.2	45.6	5.1	45–46	8.7	−13.1	−2.4
β-Carotene																		
2009	15.4	3.8	15–19	6.9	4.2	-0.1	–	–	–	–	–	–	45.6	5.8	45–46	11.2	−6.3	0.1
2010	14.6	7.1	14–19	15.4	6.0	1.9	–	–	–	–	–	–	45.8	8.4	45–46	18.9	−6.6	−3.4
Average	14.3	6.3	14–15	12.6	5.1	1.3	–	–	–	–	–	–	45.6	6.1	45–46	12.2	−5.3	−3.1
β-Xanthophylls[h]													–	–	–	–	–	–
2009	15.4	4.2	15–19	9.7	8.8	-1.3	14.1	8.1	11–15	20.2	−11.1	-8.0	–	–	–	–	–	–
2010	21.7	6.9	15–22	17.1	8.7	5.3	14.4	5.8	12–15	12.1	−7.9	−4.8	–	–	–	–	–	–
Average	17.4	5.6	15–20	11.5	8.0	2.1	14.4	10.0	12–15	22.4	−10.3	−6.7	–	–	–	–	–	–
Lutein epoxide																		
2009	–	–	–	–	–		–	–	–	–	–	–	45.6	5.5	45–46	12.3	−4.5	−0.3
2010	–	–	–	–	–		–	–	–	–	–	–	45.6	4.7	45–46	11.1	−3.6	−2.5
Average	–	–	–	–	–		–	–	–	–	–	–	45.1	4.5	45–46	10.1	−3.5	−2.3

[a] QTL associated with carotenoids is shown in parenthesis proceeding chromosome number

[b] Most likely QTL position

[c] Maxima of logarithm of odds score

[d] Confidence interval (cM)

[e] Estimated percentage of phenotypic variation explained by QTL

[f] Additive gene effect estimate

[g] Dominance gene effect estimate

[h] Xanthophylls include neoxanthin and violaxanthin

et al. 2007) but is considerably lower than what has been observed (up to 65 %) in some *B. oleracea* doubled haploid populations (Pink et al. 2008).

Carotenoid QTL

Three QTLs associated with carotenoid variability were observed consistently across years and compounds. The first (*bocarot1)* was associated with TC and all individual carotenoids except lutein epoxide. The results suggest this likely represents a major gene contributing up to 25 % of the total phenotypic variation. As this QTL impacts all individual carotenoids (except lutein epoxide) it possibly represents either an unknown regulatory gene or a common enzymatic step in the carotenoid pathway such as phytoene synthase. A previous *B. oleracea* study identified 3 copies of PSY on C02, C03, and C09; and provided some evidence that these genes were expressed differently during flower development (Cárdenas et al. 2012). Our protein to nucleotide searches also identified three copies of PSY on C02, C03, and C09 in *B. oleracea* and three homologous copies on A02, A03, and A10 in *B. rapa*.

The positioning of the C02 PSY in the current draft genome sequence of the rapid cycling 'TO1000' spans the region from 4,211,748 to 4,209,741 bp which is adjacent (but outside) the confidence interval for *bocarot1* (Tables 5, 6). The A genome markers that occur within the confidence interval (BN_A02_2557799 and BN_A02_2776634), however, are less than 500 Kbp from the estimated position of A02 PSY in the *B. rapa* genome. While further work with gene specific markers will be required to verify that PSY is indeed *bocarot1*, the nature of the QTL (impacting all downstream carotenoids) and the close proximity

Table 6 Alignments of select carotenoid pathway proteins to *B. oleracea* and *B. rapa* genomes[a]

Gene name	NCBI accession #	Chromosome #	Location start	Location stop
Brassica oleracea				
Phytoene synthase	NP_197225.1	C02	4211748	4209741
–	–	C03	4171933	4169843
–	–	C09	47556500	47558540
Lycopene beta cyclase	AAB53336.1	C03	6689338	6686683
Lycopene epsilon cyclase	NP_001078131.1	C05	42958391	42959498
Beta-carotene 3-hydroxylase 1	NP_194300.1	C01	12609382	12607478
–	–	C07	43813636	43811674
Carotene epsilon-monooxygenase	NP_190881.2	C06	21285557	21282830
Protein LUTEIN DEFICIENT 5	AEE31394.1	C08	10242093	10239111
Violaxanthin de-epoxidase	VDE_ARATH	C05	3252578	3254680
Zeaxanthin epoxides, chloroplastic	Q9FGC7	C07	23997316	24000569
Brassica rapa				
Phytoene synthase	NP_197225.1	A02	3121495	3119201
–	–	A03	3312163	3310193
–	–	A10	11670631	11672730
Lycopene beta cyclase	AAB53336.1	A10	7779942	7782294
–	–	A03	5284827	5282239
Lycopene epsilon cyclase	NP_001078131.1	A05	22867029	22868121
Beta-carotene 3-hydroxylase 1	NP_194300.1	A01	8422205	8420294
Carotene epsilon-monooxygenase	NP_190881.2	A07	14698960	14701659
Protein LUTEIN DEFICIENT 5	AEE31394.1	A08	8796128	8793136
Violaxanthin de-epoxidase	VDE_ARATH	A06	2964331	2966421
Zeaxanthin epoxides, chloroplastic	Q9FGC7	A07	12102551	12105803

a Select carotenoid pathway protein sequence identified in *Arabidopsis thaliana* was aligned to both *B. oleracea* 'TO1000' and *B. rapa* (Chiifu-401, version 1.5) genomes using Exonerate (Slater and Birney 2005). Parameters were as follows: model = protein 2 genome, query type = protein, target type = DNA, percent cut off = 75 %

of the gene in both diploid genomes suggest it is a strong candidate.

Two additional QTLs (*bocarot2* and *bocarot3)* were identified on C06 approximately 28 cM apart and separated by 13 intervals. In contrast to *bocarot1,* these linked QTLs were associated with the accumulation of individual carotenoids in broccoli florets. These QTLs could represent genes specific to the α- and β-carotene pathways. Multiple "A" genome SNPs (A07_14045156, A07_14975152, and A07_17353806) mapped to the region of *bocarot3* and would suggest that carotene epsilon-monooxygenase *(*Lut-1) is a putative candidate identified between 14,698,960 and 14,701,659 bp in the A genome due to its proximity, its association with lutein, and its lack of association with β-xanthophylls. This hypothesis does not, however, explain how the expression of Lut1 might impact the upstream compound β-carotene. Searches of the *B. oleracea* 'TO1000' genomic draft places LUT1 on C06, but again predicts the location to be a considerable distance from our estimated location. The closest A genome marker to *bocarot2* occurs roughly 9 cM away on C06 which makes it difficult to identify potential candidates, but it is interesting to note that our searches of the *B. rapa* genome suggest that Lut1 occurs approximately 2.5 Mbp away from zeaxanthin epoxidase on the same chromosome. Differences in the specificity of zeaxanthin epoxidase could explain why it has a major impact on β-xanthophylls but a reduced impact on lutein (only significant in 1 year).

The combination of the three QTLs resulted in higher levels of total carotenoids and lutein than has been previously described in broccoli and significantly higher than either parent. The QTL designated *bocarot1* was contributed by VI-158 which is a sister line of USVL066, previously reported to contain among the highest concentrations of total carotenoids in broccoli inbred material across multiple environments (Farnham and Kopsell 2009). The QTL designed as *bocarot2* and *bocarot3* was contributed from the wild-type parent BNC. Statistical analyses using the closest flanking markers to these QTLs did not reveal significant interaction between the QTL (data not shown) and the additive nature of the alleles suggests that to improve carotenoid concentration it will require introgression of these QTLs into both inbreds to produce a hybrid with enhanced levels of lutein and other carotenoids. The highly correlated nature of the individual carotenoids suggests that increasing lutein can be accomplished without adversely affecting levels of either β-carotene (which provides pro-vitamin A) or the β-xanthophylls which play

a role in plant stress and are precursors of ABA. No significant relationship was observed between head size and carotenoid concentration but significant negative correlations were observed between harvest date and carotenoid concentration. As harvesting for both years began in early to mid November and ended in mid to late December, both day length and temperature conditions differed between early and late maturing families. Lefsrud et al. (2006) noted increased photoperiod had significant impacts on carotenoid accumulation in kale which may be the same phenomena we are observing in broccoli.

Author contributions Dr. Allan Brown constructed genetic map, conducted QTL analysis and was primary author. Dr. Gad Yousef conducted all statistical analysis of phenotypic data. Drs. Ivette Guzman, Kranthi Chebrolu and Mr. Robert Byrd conducted carotenoid extraction and analysis. Dr. Rebekah Oliver conducted SNP calling analysis. Dr. Rob Reid, Mr. Koyt Everhart, and Ms. Aswathy Thomas conducted bioinformatics analysis and identification of candidate sequences from genomic drafts of B. *rapa* and B. *oleracea*. Drs. Isobel Parkin and Andrew Sharpe contributed to the writing, provided key information regard genomic locations of SNPs and provided access to unpublished B. *oleracea* draft sequence. Dr Eric Jackson (and General Mills) provided funding for study and contributed to analysis of SNP markers and map construction.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The experiment conducted complies with the laws of the United States and Canada.

References

Arias T, Beilstein MA, Tang M, McKain MR, Pires JC (2014) Diversification times among Brassica (Brassicaceae) crops suggest hybrid formation after 20 million years of divergence. Am J Bot 101:86–91

Ayele M, Haas BJ, Kumar N, Wu H, Xiao Y, Van Aken S, Utterback TR, Wortman JR, White OR, Town CD (2005) Whole genome shotgun sequencing of *Brassica oleracea* and its application to gene discovery and annotation in *Arabidopsis*. Genome Res 15:487–495

Babula D, Kaczmarek M, Barakat A, Delseny M, Quiros CF, Sadowski J (2003) Chromosomal mapping of *Brassica oleracea* based on ESTs from *Arabidopsis thaliana*: complexity of the comparative map. Mol Genet Genomics 268:656–665

Bohuon EJR, Keith DJ, Parkin IAP, Sharpe AG, Lydiate DJ (1996) Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*. Theor Appl Genet 93:833–839

Brown AF, Yousef GG, Jeffery EH, Klein PB, Wallig MA, Kushad MM, Juvik JA (2002) Glucosinolate profiles in broccoli: variation in levels and implications in breeding for cancer chemoprotection. J Am Soc Hortic Sci 127:807

Brown AF, Jeffery EH, Juvik JA (2007) A polymerase chain reaction-based linkage map of broccoli and identification of quantitative trait loci associated with harvest date and head weight. J Am Soc Hortic Sci 132:507–513

Camargo LEA, Savides L, Jung G, Nienhuis J, Osborn TC (1997) Location of the self-incompatibility locus in an RFLP and RAPD map of *Brassica oleracea*. J Hered 88:57

Cárdenas P, Gajardo H, Huebert T, Parkin I, Iniguez-Luy F, Federico M (2012) Retention of triplicated phytoene synthase (PSY) genes in *Brassica napus* L. and its diploid progenitors during the evolution of the *Brassiceae*. Theor Appl Genet 124:1215–1228

Cazzonelli IC, Pogson BJ (2010) Source to sink: regulation of carotenoid biosynthesis in plants. Trends Plant Sci 15:266–274

de Sá MC, Rodriguez-Amaya DB (2004) Optimization of HPLC quantification of carotenoids in cooked green vegetables—comparison of analytical and calculated data. J Food Compos Anal 17:37–51

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

Eberhardt VM, Kobira K, Keck AS, Juvik AJ, Jeffery EH (2005) Correlation analyses of phytochemical composition, chemical, and cellular measures of antioxidant activity of broccoli (*Brassica oleracea* L. Var. *italica*). J Agric Food Chem 53:7421–7431

Farnham MW, Kopsell DA (2009) Importance of genotype on carotenoid and chlorophyll levels in broccoli heads. HortScience 44:1248–1253

Gao M, Li G, Yang B, Qui D, Farnham M, Quiros C (2007) High-density *Brassica oleracea* linkage map: identification of useful new linkages. Theor Appl Genet 115:277–287

García-Plazaola JI, Matsubara S, Osmond CB (2007) The lutein epoxide cycle in higher plants: its relationships to other xanthophyll cycles and possible functions. Funct Plant Biol 34:759–773

Guzman I, Yousef GG, Brown AF (2012) Simultaneous extraction and quantitation of carotenoids, chlorophylls, and tocopherols in *Brassica* vegetables. J Agric Food Chem 60:7238–7244

Heinonen MI, Ollilainen V, Linkola EK, Varo PT, Koivistoinen PE (1989) Carotenoids in Finnish foods: vegetables, fruits, and berries. J Agric Food Chem 10:655–659

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

Holden JM, Eldridge AL, Beecher GR, Buzzard IM, Bhagwat S, Davis CS, Douglass LW, Gebhardt S, Haytowitz D, Schakel S (1999) Carotenoid content of U.S. foods: an update of the database. J Food Compos Anal 12:169–196

Hu J, Sadowski J, Osborn TC, Landry BS, Quiros CF (1998) Linkage group alignment from four independent *Brassica oleracea* RFLP maps. Genome 41:226–235

Iniguez-Luy F, Lukens L, Farnham M, Amasino R, Osborn T (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa* and *B. oleracea*. Theor Appl Genet 120:31–43

Kianian SF, Quiros CF (1992) Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. Theor Appl Genet 84:544–554

Kidwell KK, Osborn TC (1992) Simple plant DNA isolation procedures. In: Beckman JS, Osborn TC (eds) Plant genomes: methods for genetic and physical mapping, pp 1–13 Kluwer Academic, Netherlands

Kurilich AC, Tsau GJ, Brown AF, Howard L, Klein PB, Jeffery EH, Kushad M, Wallig MA, Juvik JA (1999) Carotene, tocopherol, and ascorbate contents in subspecies of *Brassica oleracea*. J Agric Food Chem 47:1576–1581

Kushad M, Brown AF, Kurilich A, Juvik JA, Klein PP, Wallig M, Jeffery E (1999) Variation of glucosinolates in vegetable crops of *Brassica oleracea*. J Agric Food Chem 47:1541–1548

Lan TH, Paterson AH (2000) Comparative mapping of quantitative trait loci sculpting the curd of *Brassica oleracea*. Genetics 155:1927

Landry BS, Hubert N, Crete R, Chang MS, Lincoln SE, Etoh T (1992) A genetic map for *Brassica oleracea* based on RFLP markers detected with expressed DNA sequences and mapping of resistance genes to race 2 of *Plasmodiophora brassicae* (Woronin). Genome 35:409–420

Lefsrud MG, Kopsell DA, Kopsell DE, Randle WM (2006) Kale carotenoids Are unaffected by, whereas biomass production, elemental concentrations, and selenium accumulation respond to, changes in selenium fertility. J Agric Food Chem 54:1764–1771

Li G, Quiros CF (2003) In planta side-chain glucosinolate modification in *Arabidopsis* by introduction of dioxygenase *Brassica* homolog BoGSL-ALK. Theor Appl Genet 106:1116–1121

Müller H (1997) Determination of the carotenoid content in selected vegetables and fruit by HPLC and photodiode array detection. Z Lebensm Unters F A 204:88–94

Murkovic M, Gams K, Draxl S, Pfannhauser W (2000) Development of an Austrian Carotenoid Database. J Food Compos Anal 13:435–440

Nagaoka T, Doullah MAU, Matsumoto S, Kawasaki S, Ishikawa T, Hori H, Okazaki K (2010) Identification of QTLs that control clubroot resistance in *Brassica oleracea* and comparative analysis of clubroot resistance genes between *B. rapa* and *B. oleracea*. Theor Appl Genet 120:1335–1346

Ooijin VJW (2009) MapQTL® 6. Kyazma, Netherlands

Parkin I, Magrath R, Keith D, Sharpe A, Mithen R, Lydiate D (1994) Genetics of aliphatic glucosinolates. II. Hydroxylation of alkenyl glucosinolates in *Brassica napus*. Heredity 72:594–598

Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171:765–781

Pink D, Bailey L, McClement S, Hand P, Mathas E, Buchanan WV, Astley D, King G, Teakle G (2008) Double haploids, markers and QTL analysis in vegetable brassicas. Euphytica 164:509–514

Qiu D, Gao M, Li G, Quiros C (2009) Comparative sequence analysis for *Brassica oleracea* with similar sequences in *B. rapa* and *Arabidopsis thaliana*. Plant Cell Rep 28:649–661

Ramsay LD, Jennings DE, Kearsey MJ, Marshall DF, Bohuon EJ, Arthur AE, Lydiate DJ (1996) The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci. Genome 39:558–567

Ruiz-Sola MÁ, Rodríguez-Concepción M (2012) Carotenoid biosynthesis in arabidopsis: a colorful pathway. The Arabidopsis Book 10:e0158

Schranz ME, Song BH, Windsor AJ, Mitchell-Olds T (2007) Comparative genomics in the *Brassicaceae*: a family-wide perspective. Curr Opin Plant Biol 10:168–175

Sebastian RL, Howell EC, King GJ, Marshall DF, Kearsey MJ (2000) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. Theor Appl Genet 100:75–81

Slater G, Birney E (2005) Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31

Slocum MK, Figdore SS, Kennard WC, Suzuki JY, Osborn TC (1990) Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theor Appl Genet 80:57–64

Sujak A, Gabrielska J, Grudziński W, Borc R, Mazurek P, Gruszecki WI (1999) Lutein and zeaxanthin as protectors of lipid membranes against oxidative damage: the structural aspects. Arch Biochem Biophys 371:301–307

Thompson MJ, Thompson HJ (2009) Biomedical agriculture: a systematic approach to food crop improvement for chronic disease prevention. Adv Agron 102:1–54

Town CD, Cheung F, Maiti R, Crabtree J, Haas BJ, Wortman JR, Hine EE, Althoff R, Arbogast TS, Tallon LJ, Vigouroux M, Trick M, Bancroft I (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell 18:1348–1359

Zhang D, Hamauzu Y (2004) Phenolics, ascorbic acid, carotenoids and antioxidant activity of broccoli and their changes during conventional and microwave cooking. Food Chem 88:503–509