Theor Appl Genet (2011) 122:1075–1090
DOI 10.1007/s00122-010-1512-5

ORIGINAL PAPER

The identification and mapping of candidate genes and QTL involved in the fatty acid desaturation pathway in *Brassica napus*

A. M. Smooker · R. Wells · C. Morgan · F. Beaudoin · K. Cho · F. Fraser · I. Bancroft

Received: 1 July 2010 / Accepted: 4 December 2010 / Published online: 24 December 2010

Abstract We constructed a linkage map for the population QDH, which was derived from a cross between an oilseed rape cultivar and a resynthesised *Brassica napus*. The linkage map included ten markers linked to loci orthologous to those encoding fatty acid biosynthesis genes in *Arabidopsis thaliana*. The QDH population contains a high level of allelic variation, particularly in the C genome. We conducted quantitative trait locus (QTL) analyses, using field data obtained over 3 years, for the fatty acid composition of seed oil. The population segregates for the two major loci controlling erucic acid content, on linkage groups A8 and C3, which quantitatively affect the content of other fatty acids and is a problem generally encountered when crossing "wild" germplasm with cultivated "double low" oilseed rape cultivars. We assessed three methods for QTL analysis, interval mapping, multiple QTL mapping and single marker regression analysis of the subset of lines with low erucic acid. We found the third of these methods to be most appropriate for our main purpose, which was the study of the genetic control of the desaturation of 18-carbon fatty acids. This method enabled us to decouple the effect of the segregation of the erucic acid-controlling loci and identify 34 QTL for fatty acid content of seed oil, 14 in the A genome and 20 in the C genome. The QTL indicate the presence of 13 loci with novel alleles inherited from the progenitors of the resynthesised *B. napus* that might be useful for modulating the content or extent of desaturation of polyunsaturated fatty acids, only one of which coincides with the anticipated position of a candidate gene, an orthologue of *FAD2*.

Communicated by I. Rajcan.

A. M. Smooker and R. Wells contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-010-1512-5) contains supplementary material, which is available to authorized users.

A. M. Smooker · R. Wells · C. Morgan · F. Fraser · I. Bancroft (✉)
Department of Crop Genetics, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK
e-mail: ian.bancroft@bbsrc.ac.uk

F. Beaudoin
Biological Chemistry Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

K. Cho
Highland Agriculture Research Centre, Rural Development Administration, Pyeongchang 232-955, Korea

Introduction

Oilseed rape (*Brassica napus*), some types of which are referred to as canola, is grown in many parts of the world. Its principal product is high-value oil, which is extracted from its seeds. Rapeseed oil usage falls within two broad categories, edible and industrial. In the latter, because of the increasing threat of the loss of fossil oil sources, there is a growing interest in the use of rapeseed oil as a substitute for mineral oil. Specialist uses require modification of fatty acid (FA) composition to adapt the physical properties of rapeseed oil to those required by the potential end-uses. A detailed understanding of the genetic control of the fatty acid composition of rapeseed oil is central to knowledge-based breeding of a range of cultivars producing oil of specific composition.

Three of the major FAs of rapeseed oil are oleic acid (C18:1, where 18 refers to the number of carbon atoms in the hydrocarbon chain and 1 refers to the number of double

bonds between carbon atoms in that chain), linoleic acid (C18:2) and linolenic acid (C18:3). These unsaturated fatty acids are synthesised as a result of the fatty acid desaturation pathway, in which double bonds are introduced, sequentially, starting from stearic acid (C18:0). In seed oil, these conversions are controlled by enzymes encoded by the genes *FAB2* (C18:0–C18:1) (Kachroo et al. 2007), *FAD2* (C18:1–C18:2) (Miquel and Browse 1992) and *FAD3* (C18:2–C18:3) (Browse et al. 1993). Polyunsaturated fatty acids (PUFAs), such as linoleic and linolenic acid, are essential nutrients in the human diet and have been linked to specific health benefits including the prevention of coronary heart disease (Simopoulos 2008). For these reasons there has been much interest in increasing the proportion of these PUFAs within rapeseed oil for improved human nutrition.

Acting in parallel with the desaturation pathway, the product of the *FAE1* gene encodes a fatty acid elongase that catalyses the extension of the hydrocarbon backbone, so increasing fatty acid chain length. This produces very long chain fatty acids (VLCFA) including erucic acid (C22:1) (James et al. 1995), an important constituent of non-food rapeseed oil (Lühs and Friedt 1993). The elongation and desaturation pathways compete for the same precursor, C18:1. It is therefore important to consider the effects of allelic variation at *FAE1* loci when studying genetic variation for the quantitative control of PUFA content of seed oil.

Few studies taking a quantitative genetic approach, i.e. aiming to identify quantitative trait loci (QTL) in *Brassica* species have investigated the fatty acid desaturation pathway in its entirety, from C18:0 to C18:3. Zhao et al. (2008) reported the mapping of QTL controlling fatty acid composition in a doubled haploid (DH) *B. napus* population that segregated for total seed oil content. However, effects of the two important *FAE1* loci on linkage groups A8 and C3, as previously described by Qiu et al. (2006), were not clearly observed segregating in this population because both parents had a high C22:1. For the development of cultivars with either very high PUFA content (for dietary supplements) or very low PUFA content (for most other applications), low C22:1 oil is required. Studies in high C22:1 populations therefore have limited utility.

Much understanding of lipid bioresynthesised pathways and the genetic control of FA composition of seed oil has been obtained through the use of the model plant species *Arabidopsis thaliana*. The *Brassica* species and *Arabidopsis* diverged from a common hexaploid ancestor approximately 20 million years ago (Koch et al. 2000; Lysak et al. 2005). Comparative mapping between the diploid *Brassica* species (e.g. *B. oleracea*) and *Arabidopsis* suggest that the *Brassica* genomes are composed of three rearranged variants of an ancestral genome, with extensive conservation of gene order compared with that of *A. thaliana* (O'Neill and Bancroft 2000; Town et al. 2006).

B. napus is an amphidiploid species believed to have arisen less than 10,000 years ago following hybridisation events between *B. oleracea* (from which the species inherited a version of what is termed the *Brassica* C genome) and *B. rapa* (from which species inherited a version of what is termed the *Brassica* A genome) (Rana et al. 2004). Although *B. napus* contains the genomes of two species, genetically it behaves as a diploid. Because of the extensive triplication observed within the genomes of each of its progenitor species, any given *A. thaliana* gene could have up to six orthologues in *B. napus*. The close relationship between *Arabidopsis* and the *Brassica* species allows the use of comparative genetics to predict the positions of orthologous genes within the *Brassica* genome (Parkin et al. 2005). The deletion from *Brassica* genomes of one or more copies of predicted orthologues of *Arabidopsis* genes is common (O'Neill and Bancroft 2000; Town et al. 2006), so there will frequently be less than the "full" complement of six.

Using *A. thaliana*-based candidate gene probes, Scheffler et al. (1997) reported the number and positions on a *B. napus* linkage map of the desaturase-encoding genes *FAD2* and *FAD3*. Four loci were reported for *B. napus* orthologues of *FAD2* (*BnaFAD2* loci) and six loci were reported for *FAD3* (*BnaFAD3* loci) (Scheffler et al. 1997). The four *BnaFAD2* loci were mapped to linkage groups A1, A5, C1 and C5 (Scheffler et al. 1997). Schierholt et al. 2000 mapped a *BnaFAD2* locus to A5. Hu et al. (2006) identified two QTL for C18:2 content in *B. napus*, with one having a major effect and one having a minor effect. The major QTL mapped to A5 and the minor QTL mapped to A1. It is unclear whether functional variants of *FAD2* orthologues are present in the *Brassica* C genome (Schierholt et al. 2000; Tanhuanpää et al. 1998; Zhao et al. 2008). Five *BnaFAD3* loci were mapped, to linkage groups A3, A4, A5 and at two loci on C4 (Scheffler et al. 1997). A further, monomorphic, RFLP band was observed, indicating that there are a total of six *BnaFAD3* loci. Hu et al. 2006 mapped two major QTL for C18:3 content in *B. napus*: one on linkage group A4 and one on C4. Both of these loci were suggested to correspond to functional *FAD3* orthologues. It is not known how many other *BnaFAD3* loci encode functional genes.

Fourmann et al. (1998) used consensus primers designed from known Arabidopsis and *B. napus FAE1* sequence to amplify, by PCR, two *FAE1* orthologues from *B. napus*: one from the A genome and one from the

C genome. On mapping these, they were found to co-segregate with two loci that were responsible for C22:1 content. In their study, BN-FAE1.1 mapped to the *B. napus* linkage group DY9 and BN-FAE1.2 mapped to linkage group SD7. In the current nomenclature of linkage groups, these correspond to A8 and C3, respectively. Major QTL for C22:1 content have been identified on linkage groups A8 and C3 of *B. napus* in further studies (Qiu et al. 2006).

Extensive gene duplication occurred in the common ancestor of *Arabidopsis* and *Brassica* species. This took the form of both whole genome duplication and tandem duplication of individual genes. Kachroo et al. (2007) describe seven homologues in the *A. thaliana* genome that form the stearoyl-ACP desaturase (S-ACP-DES)-encoding gene family. The homologues are present on four of the five *A. thaliana* chromosomes with single copies on chromosome 1 and chromosome 2, a tandem pair on chromosome 5, and a tandem triplet on chromosome 3 (Kachroo et al. 2007). The single copy on *A. thaliana* chromosome 2 (At2g43710) is called the *SSI2* or *FAB2* gene (Kachroo et al. 2001) and is involved in the first step of the desaturation pathway to produce oleic acid. The mutation *fab2* reduces S-ACP-DES activity almost completely, though the fact that C18:1 is not eliminated suggests that at least one of the other homologues encodes an active S-ACP-DES (Kachroo et al. 2001). The number and positions of S-ACP-DES-encoding genes has not yet been determined in *B. napus*, although the *A. thaliana* S-ACP-DES triplet on chromosome 3 (At3g02610, At3g02620, At3g02630) is represented in *B. napus* by only two loci, one on linkage group A3 and the other on linkage group C3, each with a single gene (Cho et al. 2010).

Allelic diversity amongst the germplasm exploited by oilseed rape breeders is limited because of its restricted origins and history of selective breeding. Because of this, we sought to identify variation for PUFA synthesis, for characterization by quantitative genetics approaches, from beyond the conventional germplasm. We selected a population that had already been developed from a resynthesised *B. napus* (i.e., derived from a cross between *B. rapa* and *B. oleracea* parents, followed by chromosome doubling) and so is expected to maximise allelic variation relative to the typical European winter oilseed rape variety, Tapidor, with which it was crossed to produce a DH population. We aimed to identify QTL for seed oil composition and to position at least some of the *B. napus* orthologues of the fatty acid biosynthesis genes *FAB2*, *FAD2*, *FAD3* and *FAE1* relative to these QTL. In particular, we were seeking to identify QTL indicative of novel genetic variation controlling the quantitative abundance in seed oil of 18-carbon PUFAs.

Methods and materials

Experimental plant material

The *B. napus* Q population is a doubled haploid (DH) population derived from a cross between a resynthesised *B. napus* line, developed by hybridisation of the spring turnip rape *B. rapa oleifera* '29' var. Maleksberger (hereafter *B. rapa* '29') and the Mediterranean cliff-dwelling wild species *B. oleracea atlantica*, with the European *B. napus* cultivar, Tapidor (Mithen and Magrath 1992). The resulting F_1 plants were used to develop a population of 92 DH lines, named QDH, which was used for linkage map construction. Due to the limited fertility and, consequently, seed availability of some lines, a subset of 86 lines was used for field trials and fatty acid profiling.

Field trials

Eighty-six lines of the QDH population, Tapidor and *B. rapa* '29' were grown in field trials in 2006, 2007 and 2008 at Norwich, UK. *B. oleracea atlantica* was unsuitable for field conditions within the UK and so could not be included in the trials. However, seeds for fatty acid profiling were obtained from glasshouse-grown material.

Because the Q population segregates for vernalisation response and winter hardiness, all lines were grown in pots under glasshouse conditions (18°C, 16 h photoperiod) until they reached the 4-leaf stage, were then vernalised for 6 weeks at 4°C and placed in cold frames to harden off for 1 week. They were then transplanted into the field in three randomised blocks consisting of paired rows spaced 1.5 m apart. Within each pair of rows plants were staggered so that they were 0.5 m apart. The second primary raceme of each plant was bagged to obtain self-fertilised seed. Plants were harvested at maturity, dried in a glasshouse and the bagged racemes threshed. Harvested seed was cleaned using a winnowing air separator.

Fatty acid quantification

Fatty acid methyl esters (FAMEs) were prepared and extracted from samples of five selfed-seeds per plant by slicing them in half with a razor, adding 1 ml of methanol:toluene:2,2-dimethoxypropane:sulphuric acid (33:14:2:1) and 0.5 ml heptane and then incubating at 80°C for 1 h. FAMEs were analysed by gas chromatography (GC). Quantification data for the five main fatty acids, expressed as percentage of total identified fatty acids, were used in this study: C18:0, C18:1, C18:2, C18:3 and C22:1. Within-plant variation was determined by extracting fatty acids from three replicate seed samples per plant. No significant

differences were observed amongst these replicates; therefore, the mean values were used for further analysis.

Measurement of flowering time

Flowering time was recorded in each year as the number of days after 1st April until the emergence of the first flower on each plant.

Measurement of seed oil content

Samples of selfed seed weighing 200 mg were analysed by NMR spectroscopy using a benchtop Oxford MQC NMR machine (Oxford Instruments, Oxford). The oil content of the seeds was expressed as % seed dry weight.

Trait data statistical analysis

Of the 86 lines grown in the field trials analysis was carried out using 80 lines that had a minimum of one replicate occurring in at least 2 years, except for oil content where only 75 lines produced sufficient seed for analysis. Statistical analyses were carried out using Genstat (Payne et al. 2009).

Marker screening and genotyping

DNA was prepared from all 92 lines of the Q population plus Tapidor, *B. rapa* '29' and *B. oleracea atlantica* using a CTAB DNA extraction method based on that of Doyle and Doyle, 1987. DNA was quantified using a Nanodrop ND1000 spectrophotometer and for each sample a working dilution of 20 ng/μl was prepared.

Three hundred and ninety-five BBSRC simple sequence repeat (SSR) markers (Lowe et al. 2004), 343 BRMS SSR markers from the National Institute of Vegetable and Tea Science, Japan, (Suwabe et al. 2002; Suwabe et al. 2003) and 79 public SSR markers from the AAFC Saskatoon Research Centre (http://brassica.agr.gc.ca/index_e.shtml) were screened on a panel of five randomly selected lines of the population, *B. napus* 'Tapidor', *B. oleracea atlantica* and *B. rapa* '29'. The protocol adopted for the analysis of SSR markers was as described by Lowe et al. (2004). PCR was carried out on MWG Ag Biotech Primus96 plus thermo-cycler.

A further 111 SSR markers were identified from fully sequenced *B. rapa* BACs from the Multinational *Brassica* Genome Project (MBGP) (Trick et al. 2007) by passing the finished BAC sequences through an annotation pipeline to detect SSR repeats, available at http://brassica.bbsrc.ac.uk/.

Fifty-three allele-specific and InDel markers previously developed in an EU-funded project Integrated Marker System for Oilseed Rape Breeding (IMSORB; Qiu et al. 2006) were screened for polymorphism using two-stage PCR as described at http://brassica.bbsrc.ac.uk/IMSORB/. Large InDels were visualised in the original ~500 bp PCR product on 1.5% agarose. Small InDels and SNPs were resolved on 3% low melting point metaphor agarose following amplification of 100–200 bp PCR products using the original ~500 bp amplification products as templates.

Homologue-linked marker development

Probes for the genes *FAB2* (At2g43710) (forward primer 5′-CAGGGAAGTGCATGTTCAAG-3′, reverse primer 5′-GTTGACATGAGGCAGAT CGA-3′), *FAD2* (At3g12120) (GGGTGCAGGTGGAAGAATG, TTGT TGTACCAGTACACACC), *FAD3* (At2g29980) (TTACC AGAAAGGGTGTACAAG, TGATTTT AGATTTGTC AGAAGC) and *FAE1* (At4g34520) (ACGTCCGTTA ACGTTAAGC, CGGATCATCACATGCCACG) were produced by PCR from *Arabidopsis* BACs predicted to contain the genes of interest, as identified at http://brassica.bbsrc.ac.uk/IGF/. Colony hybridisation using these probes was performed as described by O'Neill and Bancroft (2000) with the JBnY *B. napus* genomic clone library (Cho et al. 2010). To confirm the presence of the gene of interest, Southern hybridisation was conducted on positively hybridising clones as described by Rana et al. (2004). In order to increase specificity to *FAB2*, so excluding other members of the S-ACP-DES-encoding gene family, a more specific probe produced from exon 2 of At2g43710 (CAGGGAAGTGCATGTTCAAG, GAGA GAGGGCTAGAGAGCTCC) was used for Southern hybridisation.

Genomic clones identified by Southern hybridisation were grouped into bins representing different homologues on the basis of the molecular weights of the hybridising bands on the Sothern blots, as described by O'Neill and Bancroft (2000). Clones were end-sequenced to confirm alignment to the expected region of the *A. thaliana* genome and to scrutinise sequences for marker development. One clone per bin was chosen for marker development.

To increase the availability of sequence data from genomic clones for marker development, a BAC sub-cloning strategy was employed. Double blunt-ending restriction digests were performed on genomic clone DNA using pairs made up from the restriction enzymes DraI, EcoRV, HpaI and StuI. Fragments were cloned using Lucigen's CloneSmart® cloning system (Lucigen 2008). Sub-clone inserts were screened by colony PCR for size, with inserts $\gtrsim$400 bp subsequently being sequenced. The *Arabidopsis thaliana* Integrated Database (http://ATIDB.org/) was also used to obtain genomic clones and linked sequences from the *B. rapa* genome project that aligned to the regions of *Arabidopsis* containing the genes of interest.

One BAC clone, KBrH090I02, was identified in this manner.

End-sequences and/or sub-clone insert sequences were used to design locus-specific primers. Sequencing of the PCR product from Tapidor, *B. rapa* '29' and *B. oleracea atlantica* was performed to identify polymorphisms between the parents that could be exploited for marker use. Genotyping of markers was performed on the Q population using the most appropriate technique, determined by the specific type and size of polymorphism. These included agarose, polyacrylamide and flurolabelled capillary electrophoresis.

Map construction

The genetic map was constructed using Joinmap (Stam 1993; Van Ooijen and Voorips 2001) using Kosambi's formula (Kosambi 1944) with the default recombination (REC) threshold (0.4) and logarithm of the odds of the differences (LOD) value (1.0) (Van Ooijen and Voorips 2001). Following the initial round of mapping any markers showing a high Chi-square contribution, the markers average contribution to the goodness-of-fit, which were less likely to be mapped correctly, were removed and the map re-calculated. The locus genotype frequencies, ratios of parent allele A to B, were checked to study any segregation distortion, which where present should be more or less the same for loci in the same region on the map.

Linkage groups believed to map to the same chromosome as predicted from anchored Saskatoon and BBSRC markers but failing to map together under default conditions, were remapped with thresholds reduced to less stringent values (REC = 0.5, LOD = 0).

QTL analysis

QTL were initially identified by interval mapping using MapQTL (Van Ooijen et al. 2004). The statistical significance of the QTL were determined as the logarithm of odds (LOD) and the percentage of the total trait variation explained calculated for each QTL. Permutation analysis (1,000 replications) was used to determine the significance level for LOD. Additivity was also determined as half the difference between the means of the two parental alleles (resynthesised − Tapidor thus a positive value indicates that the increasing allele originates from the resynthesised parent) at each locus. Interval mapping may be affected by departures from normality. Where it was possible to transform the data no significant effects were observed on QTL position or effect. However, as it is not possible to transform bimodally distributed data, as observed for C18:1 and C22:1, these were also analysed using the non-parametric Kruskal–Wallis test.

Because the QDH population segregates for C22:1 content of seed oil, we anticipated that two loci, containing orthologues of the *A. thaliana FAE1* locus, would have major effects on the content of many fatty acids. The effects of these major QTL were investigated in two ways. Firstly, MQM mapping was carried out on the same data set after selecting these loci as cofactors. The corresponding markers for these loci were JICB0633 (on linkage group A8) and JICB0676 (on linkage group C3). Secondly, because these major QTL were likely to mask the apparent expression of other QTL with smaller effects, a subset of 25 phenotypically zero (or near zero) C22:1 lines were analysed, i.e. lines that inherited non-functional *BnaFAE1* alleles at both loci (originating from the Tapidor parent, which is a low C22:1 oilseed rape cultivar). Because of the relatively small number of such lines available, regression analysis was carried out to identify single marker QTL using Genstat (Payne et al. 2009). It should be noted that the small population size reduces the level of recombination amongst the markers and as a consequence, whilst enabling successful QTL detection, this approach has reduced capability for localization of QTL.

QTL were scored where criteria of both significance and reproducibility were satisfied: significance threshold of $P < 0.05$ and occurrence in at least 2 of the 3 years. Many QTL were broad and had peaks at different positions in each year. In those cases, the instance with the lowest P value has been used to describe the QTL effect.

Results

Identification and scoring of anonymous markers

A total of 928 SSR markers were screened for polymorphism on the parents of the Q population, identifying 378 polymorphic primer pairs, with an average polymorphism rate of 40% across the different primer sources. Twenty-six of the 111 (23%) of the putative SSR markers identified by BAC annotation were polymorphic in the population. Of the IMSORB SNP and InDel markers, 17 of the 53 (32%) were polymorphic. Many of the SSR primers revealed multiple polymorphic loci, resulting in a total of 503 loci being available for linkage mapping.

Development and scoring of markers linked to candidate genes

Colony and Southern hybridisation of gene-specific probes to the *B. napus* JBnY BAC library (that had been derived from genomic DNA of oilseed rape cultivar Tapidor) enabled the identification of 13 BAC clones containing homologues of *A. thaliana FAB2*, six clones containing

homologues of *A. thaliana FAD2*, 13 clones containing homologues of *A. thaliana FAD3* and seven clones containing homologues of *A. thaliana FAE1*. BAC end sequence from these clones or sub-clone sequences were used to develop markers. In addition, public sequence databases were searched for sequenced *Brassica* BAC clones containing candidate genes. PCR primers were designed and amplification products were generated from the parent *Brassica* lines and a small panel of QDH lines in order to identify polymorphisms segregating within the population. In total 17 markers were developed and assayed across the QDH population using appropriate assay methods for the polymorphism type. Details of primers and assay methods are shown in Supplemental Table 1.

Linkage map construction

The final constructed QDH linkage map consisted of 357 molecular markers over 19 linkage groups, totalling a genetic distance of 1,381 cM, as shown in Fig. 1 for the A genome and Fig. 2 for the C genome. A greater number of markers mapped to the C genome (210) than to the A genome (147). Of all the mapped markers, 55% showed significant segregation distortion ($P < 0.05$). This skewing of marker alleles affected almost all linkage groups, with least on C1, C2 and the top of C3. The majority of the distorted loci (59%) were biased in favour of the Tapidor parent. The identity and orientation of the linkage groups were determined using mapping data accompanying markers and by comparison with marker positions in the reference *B. napus* linkage map developed using the TNDH mapping population (http://www.jic.ac.uk/staff/ian-bancroft/research_page3.htm#linkage; Qiu et al. 2006).

The linkage map includes nine markers developed from BAC-derived sequences that produced low copy amplification products and are likely to be linked to the targeted candidate gene loci. Table 1 shows the mapping positions of these on the QDH linkage map, along with a further marker linked to *FAD2*, KBrB063G23.29, which was identified in the sequence of a published *B. rapa* BAC clone. All of the positions of these markers conform to linkage groups with anticipated collinearity to the appropriate segments of the genome of *A. thaliana* (Parkin et al. 2005). Markers JICB0677 and JICB0680 map within 1.2 cM of each and are most likely linked to the same *FAB2* homologue.

Phenotypic variation within the QDH population

Mean trait values for the analysed lines are summarised in Table 2, together with summary statistics for QDH. Histograms of the fatty acid distributions are given in Supplemental Figures 1 to 5. Tests of normality showed that the distribution was normal for C18:3, near normal for C18:0 and C18:2 while C18:1 and C22:1 showed bi-modal distributions, resulting from the presence or absence of the FAE1 loci on A8 and C3. Tapidor and *B. oleracea atlantica* have low (~0%) and high (47%) C22:1, respectively. *B. rapa* '29' has an intermediate C22:1 content (19.7%). The presence or absence of C22:1 had a proportional effect on the other fatty acids, thus Tapidor had higher levels of unsaturated C18 fatty acids than *B. oleracea atlantica*. This was most notable for C18:1, where in Tapidor it was over three times that found in *B. oleracea atlantica* (61.2 and 17.1%, respectively). The QDH population showed transgressive segregation for all traits measured, with mean values similar to those of *B. rapa* '29'. No significant difference was observed in oil content between Tapidor and *B. rapa* '29'. Date of flowering varied by 32 days, but further analyses (not shown) showed no correlation with fatty acid composition or oil content.

Because it was possible to grow only one plant of each line per block, replication and blocks are confounded for the statistical analysis of data. However, analysis of the data using blocks as the main effect with lines as replicates showed that there were no significant differences between blocks or interactions with lines. Consequently, blocks have been treated as replicates in our analysis. Analysis of variance (ANOVA) was carried out using all lines where there was at least one replicate plant measured in a minimum of 2 years with between lines and between years as main effects and testing for line × year interaction. The results are shown in Table 3. Significant differences ($P < 0.001$) were found for all traits (except flowering time line × year interaction) among lines, among years and for the line × year interaction. Expected mean squares (EMS) were calculated to estimate trait heritability and relative biological importance of year and interactions. Heritabilities were highly correlated with the mean value of the trait ($r = +0.917$) while the residual and interaction were equally, but negatively, correlated ($r = -0.912$ and -0.842, respectively). As an example C18:0, with a trait mean of 1.57%, had a residual accounting for 23% of the total variance and a heritability of 40.8% in contrast to C22:1, with a trait mean of 20.59% which had a residual of only 1% and a heritability of 95.5%. Heritability of fatty acid composition is high because it is, to a large extent, determined by the presence or absence of erucic acid which is known to be primarily controlled by two major genes which show little interaction with the environment. Heritabilities similar to those described here have previously been reported by Schierholt and Becker (2001) who found that fatty acids such as C18:1 which are predominantly controlled by a few major genes, have extremely high heritabilities (99%), while others, such as C18:3 which are under more complex genetic control, have lower heritabilities

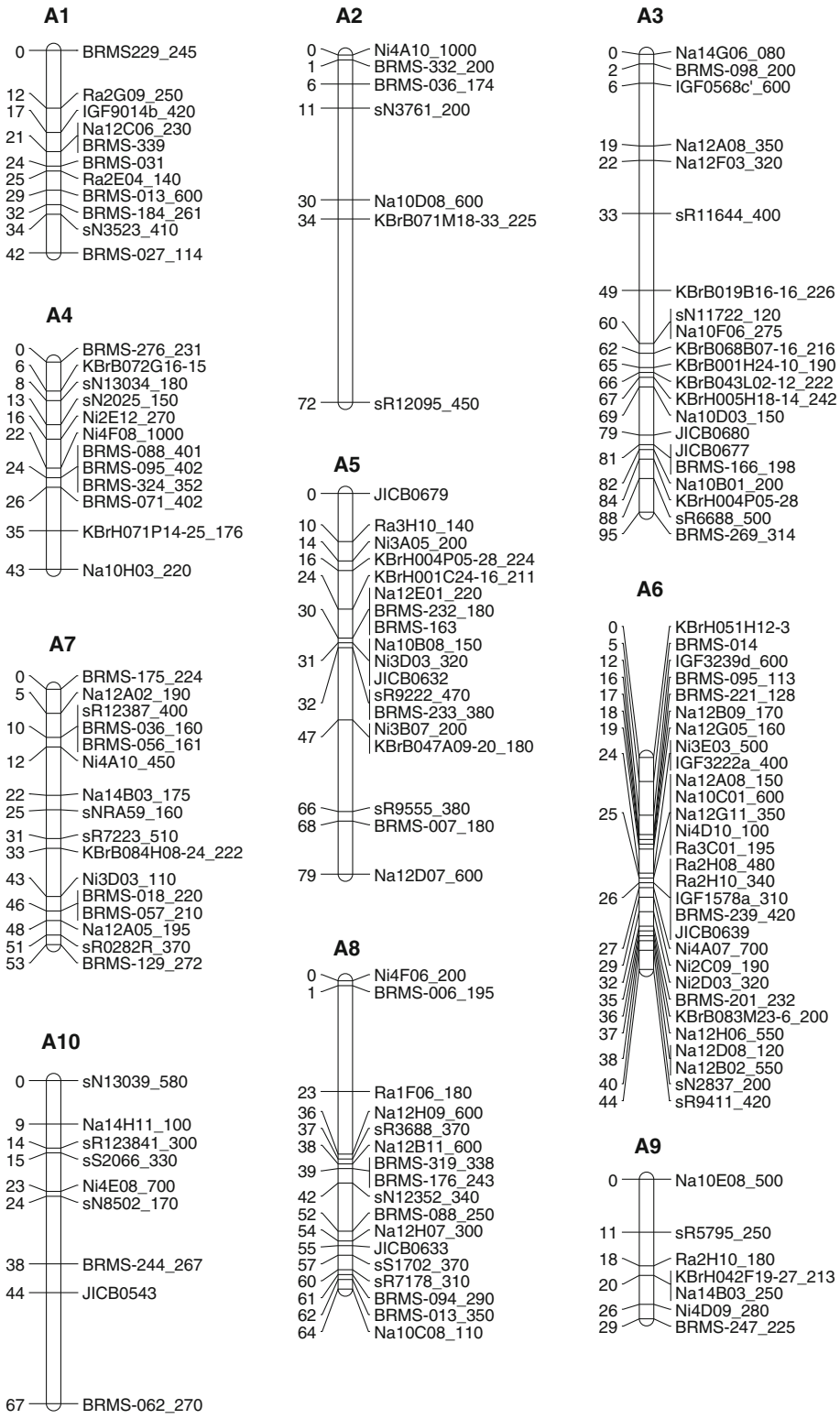

Fig. 1 Genetic linkage map of the QDH population: A genome. Mapped using the Kosambi mapping function. Marker distances are given in cM

(58%). Mean correlations among traits were determined by averaging the z-transformed correlations for each year and are summarised in Table 4. Associations amongst fatty acids are subject to autocorrelation because they are expressed as percentages of total identified fatty acids and also were greater where the combined percentage of the fatty acids were high thus, most notably, r was −0.968 between C18:1 and C22:1 acid. Of note were the correlations of C18:3 with C18:1 and C22:1, which were lower than expected and non-significant, suggesting independent control of these fatty

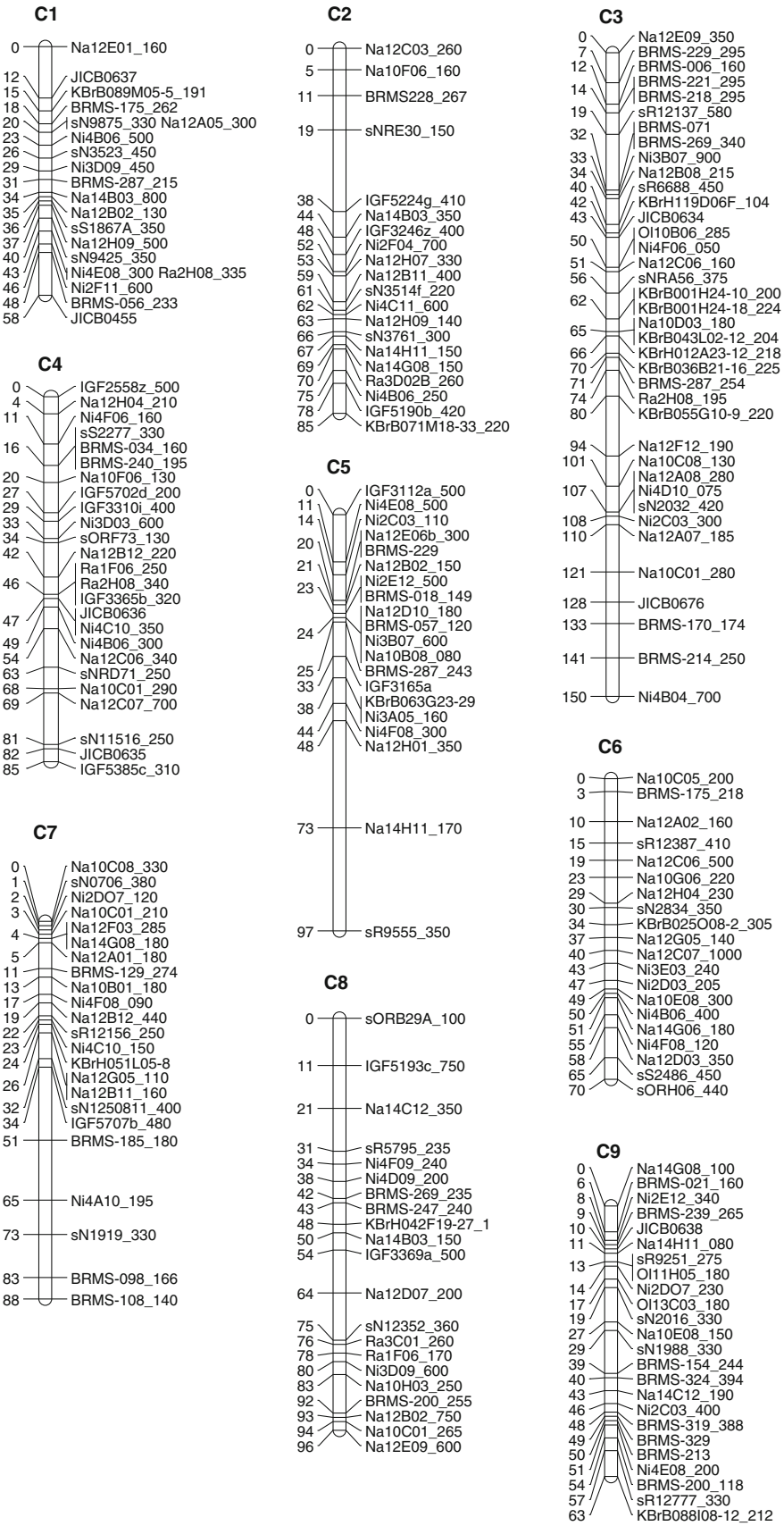

Fig. 2 Genetic linkage map of the QDH population: C genome. Mapped using the Kosambi mapping function. Marker distances are given in cM

Table 1 Mapping positions of markers linked to candidate genes

Candidate gene	Marker	Linkage group	Mapping position (cM)	Locus name
FAB2	JICB0677	A3	79.3	*BnaFAB2*_A3
FAB2	JICB0680	A3	80.5	*BnaFAB2*_A3
FAD2	JICB0637	C1	11.9	*BnaFAD2*_C1
FAD2	JICB0679	A5	0.0	*BnaFAD2*_A5
FAD2	KBrB063G23.29	C5	37.7	*BnaFAD2*_C5
FAD3	JICB0634	C3	43.5	*BnaFAD3*_C3
FAD3	JICB0636	C4a	46.7	*BnaFAD3*_C4a
FAD3	JICB0635	C4b	82.5	*BnaFAD3*_C4b
FAE1	JICB0633	A8	55.3	*BnaFAE1*_A8
FAE1	JICB0676	C3	128.0	*BnaFAE*1_C3

Table 2 Summary of seed fatty acid content, seed oil and time to flowering for QDH and its parental/ancestral lines

Trait	Tapidor	*B. rapa* '29'[a]	SEM[c]	*B. oleracea atlantica*[b]	QDH			
	Mean	Mean			Mean	Min	Max	SD
C18:0[d]	1.02	1.04	0.058	1.16	1.57	0.22	4.67	0.539
C18:1[d]	61.21	30.41	2.036	17.11	34.60	7.77	71.52	18.88
C18:2[d]	19.12	20.11	0.347	13.55	17.57	8.87	30.28	3.214
C18:3[d]	10.88	8.26	0.190	8.54	8.01	4.03	14.05	1.763
C22:1[d]	0.00	19.74	1.711	47.03	20.59	0.00	50.77	15.87
C18:2/C18:1[e]	0.31	0.66	0.032	0.79	0.65	0.19	2.45	0.300
C18:3/C18:2[e]	0.57	0.41	0.012	0.63	0.465	0.21	0.75	0.112
% Oil[f]	46.81	41.18	0.526	–	40.22	18.02	50.41	4.878
Flowering[g]	43.56	32[h]	0.767	–	44.45	21.00	73.00	7.112

Data taken from the analysis of 78 lines grown in 3 years trials (see text)

[a] Fatty acid data for *B. rapa* from one plant only grown in 2007

[b] Not grown in trial, data taken from glasshouse material

[c] Standard error of the mean calculated from ANOVA of 3 years trials and is appropriate for all lines where there was full replication

[d] Fatty acid expressed as percentage of total fatty acid content

[e] Ratio of fatty acids

[f] Seed oil content expressed as a percentage of seed dry weight

[g] Expressed as the number of days from 1st April to the emergence of the first flower

[h] Only two plants flowered in 2006 and 2007 while no plants flowered in 2008 and were therefore not included in the ANOVA. The SEM calculated separately for *B. rapa* = 8.3

acids. Correlations of the content of other fatty acids with the minor constituent C18:0 were all very low. Oil content showed no correlation with C22:1 or C18:1, though there was a significant but low negative association with C18:2 ($r = -0.333$) and a small positive association ($r = +0.456$) with the ratio C18:3:C18:2. No correlations with flowering time were observed.

QTL analysis by interval mapping

Our initial approach to QTL analysis used interval mapping (IM). IM is most reliable where the data is normally distributed; however, for those fatty acids where transformation resulted in a more normal distribution, no significant effects on the position and probability levels of the detected QTL were observed. QTL analysis using Kruskal–Wallis analysis provided further confirmation of the marker–trait association, and indicated that the results of interval mapping were not influenced by segregation distortion or non-normal distribution of certain traits (data not shown). All the identified QTL are given in Supplemental Table 2 while a summary is given in Table 5. As anticipated for a cross between high and low C22:1 parent lines, two major QTL were identified for C22:1, on linkage group

Table 3 Analysis of variance and expected mean squares of seed oil content and fatty acid traits, and flowering time, in QDH population grown over 3 years

Trait	Variance ratio			Mean square	Expected means squares			
	Line	Year	Line × year	Residual	Line[a]	Year	Line × year	Residual
C18:0	12.4	39.3	3.9	0.1182	40.8	5.1	31.4	22.7
C18:1	543.8	161.3	3.6	5.986	96.4	1.1	1.4	1.1
C18:2	42.8	172.0	3.0	1.675	69.1	10.6	9.7	10.6
C18:3	24.3	164.8	2.5	0.7426	57.9	15.3	11.0	15.9
C22:1	724.7	262.8	7.0	3.115	95.5	1.3	2.4	0.8
C18:2/C18:1	48.5	62.7	2.2	0.01382	79.5	3.9	6.0	10.7
C18:3/C18:2	26.0	262.6	2.2	0.002725	55.9	21.9	7.9	14.3
Oil	22.0	68.0	3.6	7.824	55.8	6.7	20.9	16.6
Flowering	12.9	98.7	1.9	18.7	48.5	14.9	10.7	26.0

All variance ratios are significant at $P < 0.001$. Line *df* for fatty acids and flowering time = 79, oil *df* = 74. Lines were included in the analysis where there was at least one replicate in a minimum of 2 years

[a] Expected mean square for Line = Heritability h^2. Traits as for Table 2

Table 4 Correlation matrix for fatty acid and oil traits and flowering (*df* = 78)

	C18_0	C18_1	C18_2	C18_3	C22_1	Oil
C18_1	0.233					
C18_2	−0.043	0.599				
C18_3	−0.425	0.088	0.340			
C22_1	−0.229	−0.968	−0.668	−0.206		
Oil	−0.192	−0.088	−0.333	0.191	0.095	
Flowering	0.010	0.009	−0.113	−0.200	0.026	−0.082

A8 (peak LOD at 57.3 cM) and C3 (peak LOD at 128 cM). These QTL accounted for 36.8 and 45.8% of the variation, respectively, and were coincident in position with markers for the *BnaFAE1* loci. The increasing alleles were inherited from *B. rapa* '29' and *B. oleracea atlantica*, respectively. Seven further QTL were identified for C22:1, accounting for between 8.7 and 15.5% of the variation identified. In three cases, i.e. two QTL on C3 (peak LOD at 7.3 cM; peak LOD at 101.1 cM) and one on C4 (peak LOD at 20.4 cM), the increasing alleles were inherited from *B. oleracea atlantica*. In four cases, i.e. QTL on C6 (peak LOD at 65.2 cM), C7 (peak LOD at 19.3 cM), C8 (peak LOD at 42.9 cM) and C9 (peak LOD at 7.7 cM), the increasing alleles were inherited from Tapidor.

The presence of the biochemical capacity to synthesise C22:1 has a major impact on the oil content of C18:1 and C18:2. This is confirmed by the identification of coincident QTL for C18:1 and C18:2 content with the major C22:1 QTL on linkage groups A8 and C3, which have the opposite source of increasing alleles, i.e. Tapidor. Similarly, most of the minor C22:1 QTL show reciprocal relationships with C18:1, i.e. increasing C22:1 alleles from *B. rapa* '29' or *B. oleracea atlantica* coincide with increasing C18:1 alleles from Tapidor (linkage group C3 peak LOD at 7.3 cM, linkage group C3 peak LOD at 101.1 cM, linkage group C4 peak LOD at 20.4 cM) and increasing C22:1 alleles from Tapidor coincide with increasing C18:1 alleles from *B. rapa* '29' or *B. oleracea atlantica* (linkage group C6 peak LOD at 65.2 cM, linkage group C8 peak LOD at 42.9 cM, linkage group C9 peak LOD at 7.7 cM). Consequently, in order to conduct QTL analyses that are informative for understanding the regulation of the desaturation of 18-carbon fatty acids, we need to first take account of the segregation of the *BnaFAE1* loci.

QTL analysis by multiple QTL mapping

The first approach that we took in order to take account of the segregation of the *BnaFAE1* loci was multiple QTL mapping (MQM). The results, using the major C22:1 QTL on linkage groups A8 and C3 as cofactors, are summarised in Table 6 while details of all the identified QTL are given in Supplemental Table 3. This analysis reduced the number of QTL identified to 23. The major C22:1 QTL on linkage groups A8 and C3 were identified, with increasing alleles from *B. rapa* '29' and *B. oleracea atlantica*, respectively, as were coincident QTL for C18:1 and C18:2 that showed increasing alleles from Tapidor. MQM analysis, however, decreased the proportion of the phenotype explained by these two loci from 82.6 to 73.9 and 71.8 to 64.4 on A8 and C3, respectively. MQM reduces type 1 errors and we see fewer identified QTL when compared to IM, thus only 19 QTL identified by IM were also identified by MQM, albeit

Table 5 Means of QTL over 3 years identified by interval mapping

Linkage group	Fatty acid[a]	QTL Peak (cM)	QTL significance	QTL span (cM)	% explained ± SD[c]	Additivity ± SD	Source of increasing allele
A1	C18:3	26.0	**	23.6–28.9	11.6 ± 2.3	0.35 ± 0.25	Tapidor
A6	C18:2[b]	32.4	***	23.8–43.6	11.9 ± 4.0	1.01 ± 0.10	*B. rapa* '29'
A8	C18:1	57.3	***	36.3–64.5	35.1 ± 1.2	11.41 ± 0.33	Tapidor
	C18:2	57.3	***	42.0–64.5	32.7 ± 6.0	1.61 ± 0.20	Tapidor
	C18:3	60.6	***	59.5–64.5	17.3 ± 8.6	0.61 ± 0.12	Tapidor
	C22:1	57.3	***	37.3–64.5	36.8 ± 2.3	9.75 ± 0.55	*B. rapa* '29'
A9	C18:0[b]	16.0	*	11–20.3	6.7 ± 3.6	0.13 ± 0.04	*B. rapa* '29'
C1	C18:1[b]	23.8	*	19.8–25.8	9.2 ± 1.2	5.76 ± 0.52	Tapidor
C2	C18:3[b]	74.9	***	47.6–74.9	13.0 ± 4.6	0.55 ± 0.09	*B. oleracea atlantica*
C3	C18:1	13.8	**	7.3–13.8	11.0 ± 1.3	6.33 ± 0.16	Tapidor
	C18:1	101.1	*	101.1	9.9 ± 0.7	6.38 ± 0.47	Tapidor
	C18:1	128	***	120.5–150	36.7 ± 0.6	11.67 ± 0.47	Tapidor
	C18:2[b]	7.3	*	7.3	9.2 ± 0.9	0.90 ± 0.05	Tapidor
	C18:2	128	***	93.6–150	26.3 ± 2.4	1.54 ± 0.25	Tapidor
	C18:3	113.6	**	10.0.5–120.5	9.2 ± 3.9	0.49 ± 0.08	*B. oleracea atlantica*
	C22:1[b]	7.3	*	7.3–13.8	8.9 ± 1.5	4.69 ± 0.26	*B. oleracea atlantica*
	C22:1	101.1	**	93.6–101	12.2 ± 1.0	5.90 ± 0.64	*B. oleracea atlantica*
	C22:1	128.7	***	120.5–150	45.8 ± 1.0	10.87 ± 0.80	*B. oleracea atlantica*
C4	C18:1	20.4	**	11.2–20.4	12.3 ± 1.1	7.03 ± 0.45	Tapidor
	C22:1	20.4	**	11.2–20.4	15.5 ± 1.2	6.57 ± 0.44	*B. oleracea atlantica*
C5	C18.3	12.8	***	0–33.1	16.5 ± 5.0	0.65 ± 0.09	Tapidor
	C18.3	33.1	***	33.1	17.1 ± 3.9	0.64 ± 0.09	Tapidor
C6	C18:1	65.2	**	65.2	13.2 ± 1.3	7.21 ± 0.38	*B. oleracea atlantica*
	C18:2	30.5	*	30.8	9.8 ± 2.4	1.01 ± 0.18	*B. oleracea atlantica*
	C18:2	62.7	**	57.8–65.2	14.1 ± 3.9	1.18 ± 0.27	*B. oleracea atlantica*
	C22:1	65.2	***	57.8–65.2	15.0 ± 0.1	6.41 ± 0.53	Tapidor
C7	C18:2	2.3	**	2.3	9.8 ± 1.6	1.22 ± 0.15	Tapidor
	C22:1	19.3	*	19.3	8.7 ± 0.9	5.21 ± 0.32	Tapidor
C8	C18:1	42.9	**	0–63.6	12.7 ± 1.7	6.73 ± 0.23	*B. oleracea atlantica*
	C22:1	42.9	***	20.7–63.6	13.3 ± 0.7	5.76 ± 0.32	Tapidor
C9	C18:1	7.7	**	0–12.9	12.4 ± 1.1	7.41 ± 0.65	*B. oleracea atlantica*
	C22:1	7.7	*	7.7	9.1 ± 1.1	5.35 ± 0.74	Tapidor

Peak position, % explained and additivity are means of the QTL over all years regardless of significance in any one year while the spread is the maximum range where $P < 0.05$ for the three QTL

QTL probabilities: *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$

[a] Fatty acids expressed as percentage of total fatty acid content

[b] QTL identified at a minimum of $P < 0.05$ in 2 years only

[c] % of the phenotypic variance explained by QTL

with slight changes in position for some, which accounted for less of the phenotypic variation. Thirteen of the QTL that were found by IM were not identified by MQM (C18:0 content on A9 peak LOD at 19.8 cM; C18:1 content on C1 peak LOD at 25.8 cM, C3 peak LOD at 13.8 cM, C3 peak LOD at 101.1 cM, C4 peak LOD at 20.4 cM, C6 peak LOD at 65.2 cM; C18:2 content on C3 peak LOD at 7.3 cM, C6 peak LOD at 30.5 cM, C6 peak LOD at 65.2 cM; C22:1 content on C3 peak LOD at 7.3 cM, C3 peak LOD at 101.1 cM, C6 peak LOD at 65.2 cM, C7 peak LOD at 19.3 cM). However, MQM, which also reduces type 2 errors, identified four new QTL (C18:2 content on A3 peak LOD at 64.9 cM, A10 peak LOD at 20.7 cM, C9 peak LOD at 27.4 cM; C18:3 content on C9 peak LOD 38.6 cM) and one which resolved into two QTL (C18:2 content on A8 peak LODs at 54.2 and 61.5 cM) for the content of 18-carbon PUFAs that were not found by IM, which are highlighted in Table 6 in bold typeface.

Table 6 Means of QTL over 3 years identified by MQM analysis using *FAE1* loci on A8 and C3 as cofactors

Linkage group	Fatty Acid[a]	QTL Peak (cM)	QTL significance	QTL span (cM)	% explained ± SD[c]	Additivity ± SD	Source of increasing allele
A1	C18.3	26.0	**	23.6–31.6	10.7 ± 2.4	0.52 ± 0.06	Tapidor
A3	**C18:2[b]**	**64.9**	**	**64.9–68.6**	**7.7** ± 3.6	**0.82** ± 0.13	**Tapidor**
A6	C18:2[b]	23.8	**	16.5–34.8	4.9 ± 1.9	0.70 ± 0.11	*B. rapa* '29'
A8	C18:1	54.2	***	54.2–57.3	30.6 ± 1.0	10.76 ± 0.65	Tapidor
	C18:2	**54.2**	***	**54.2–64.5**	**17.0 ± 5.0**	**1.26 ± 0.16**	**Tapidor**
	C18:2	**61.5**	***	**59.6–62.4**	**16.0 ± 1.2**	**2.07 ± 0.25**	**Tapidor**
	C18:3	60.6	***	59.5–64.5	15.0 ± 6.47	0.99 ± 0.17	Tapidor
	C22:1	54.2	***	54.1–57.3	30.8 ± 1.0	5.56 ± 2.28	*B. rapa* '29'
A10	**C18:2[b]**	**20.7**	***	**13.6–38.0**	**7.4** ± 3.3	**0.87** ± 0.19	***B. rapa* '29'**
C2	C18:3[b]	65.7	***	47.6–74.9	10.9 ± 4.7	0.52 ± 0.09	***B. oleracea atlantica***
C3	C18:1	128	***	120.5–128	33.8 ± 0.7	11.19 ± 0.42	Tapidor
	C18:2	128	***	120.6–128	23.7 ± 3.4	1.48 ± 0.25	Tapidor
	C22:1	128	***	120.5–128	43.1 ± 1.4	10.52 ± 0.8	*B. oleracea atlantica*
C4	C22:1	23.1	**	4.1–28.6	3.5 ± 0.6	2.91 ± 0.25	*B. oleracea atlantica*
C5	C18:3	12.8	***	0–33.1	13.7 ± 3.8	0.60 ± 0.09	Tapidor
	C18:3	33.1	***	33.1	15.5 ± 3.1	0.62 ± 0.07	Tapidor
C7	C18:2	15.5	***	2.3–21.6	9.5 ± 1.5	1.18 ± 0.05	Tapidor
C8	C18:1	44.1	***	10.8–82.6	7.0 ± 1.7	4.84 ± 0.21	*B. oleracea atlantica*
	C22:1	37.2	***	20.7–92.5	5.3 ± 0.5	3.52 ± 0.15	Tapidor
C9	C18:1	7.7	**	0–18.6	7.6 ± 1.3	5.77 ± 0.58	*B. oleracea atlantica*
	C18:2	**27.4**	**	**27.4**	**6.5** ± 0.6	**0.88** ± 0.04	**Tapidor**
	C18:3[b]	**38.6**	**	**38.6–39.9**	**8.8** ± 2.9	**0.52** ± 0.10	**Tapidor**
	C22:1	7.7	***	0–16.8	4.7 ± 0.9	3.78 ± 0.62	Tapidor

LOD scores and peak position are given for the most significant QTL found while the spread is the maximum range where $P < 0.05$ for the three QTL; % explained and additivities are the means of 3 years regardless of significance in any one year. QTL for content of PUFAs highlighted by bold typeface were identified by MQM, but not IM analysis

QTL probabilities: *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$

[a] Fatty acids expressed as percentage of total fatty acid content

[b] QTL identified at a minimum of $P < 0.05$ in 2 years only

[c] % of the phenotypic variance explained by QTL

QTL analysis of low erucic acid lines

The third approach that we took in order to take account of the segregation of the *BnaFAE1* loci was to focus our analysis upon only the lines with low C22:1 content (i.e. those with the Tapidor allele at both A8 and C3 major C22:1 QTL). The absence of the *BnaFAE1* loci in this subpopulation resulted in near normal distributions for all the fatty acids in contrast to the entire population (Supplemental Figures 1–5). The number of lines involved, 25, was too low to use for IM, so we conducted the QTL analysis as a single marker (SM) regression analysis. A summary of the QTL is given in Table 7 while details of all the identified QTL are given in Supplemental Table 4. As expected, the major C22:1 QTL were eliminated. In fact, only one C22:1 QTL was identified (C7 peak LOD at 19.3 cM), suggesting that the occurrence of the minor C22:1 QTL identified in the whole population using IM was largely dependent upon there being at least one functional *BnaFAE1* locus present. This QTL had been identified previously by IM, but had not been identified by MQM. A further 33 QTL were identified, all relating to the content of 18-carbon fatty acids. Twelve of these had been identified previously by IM and/or MQM (C18:1 content on C1 peak LOD 17.7 cM, C8 peak LOD 47.9 cM; C18:2 content on A3 peak LOD 61.9 cM, A6 peak LOD 18.5 cM, A8 peak LOD 59.6–64.5 cM, C3 peak LOD 101–110.1 cM, C6 peak LOD 57.8 cM, C9 peak LOD 38.6 cM; C18:3 content on A1 peak LOD 16.9–24.6 cM, A8 peak LOD 59.6–64.5 cM, C3 peak LOD 101–110.1 cM, C9 peak LOD 38.6 cM). Two coincided with positions of QTL identified by IM or MQM, but with the opposite parent donating the increasing alleles. These are both QTL for C18:1 content, with the increasing allele coming from

Table 7 Means of QTL for the near-zero erucic acid sub-population over 3 years identified by single marker regression analysis

Linkage group	Fatty acid[a]	QTL peak (cM)	QTL significance	QTL span (cM)	% explained ± SD[d]	Additivity ± SD	Source of increasing allele
A1	**C18:1[c]**	**0**	**	**0**	**26.4**	**1.94**	***B. rapa* '29'**
	C18:1[b]	**16.8–24.6**	*	**16.9–34.1**	**17.7** ± 5.9	**2.10** ± 0.3	***B. rapa* '29'**
	C18:3	16.9–24.6	**	16.9–34.1	21.0 ± 5.7	0.67 ± 0.10	Tapidor
A3	**C18:1**	**62.9**	**	**49.1–68.6**	**31.5** ± 1.9	**0.08** ± 0.07	***B. rapa* '29'**
	C18:2	61.9	**	59.7–68.6	27.0 ± 4.8	1.58 ± 0.03	Tapidor
A4	**C18:1**	**16.7**	**	**13.3–26.1**	**19.7** ± 4.9	**2.15** ± 0.37	**Tapidor**
	C18:2	**16.7**	**	**13.3–26.1**	**29.2** ± 9.3	**1.66** ± 0.17	***B. rapa* '29'**
	C18:3	**13.3**	**	**13.3**	**21.7** ± 4.5	**0.69** ± 0.09	***B. rapa* '29'**
A5	**C18:0**	**75.4**	*	**68.3–79**	**15.3** ± 2.6	**0.20** ± 0.20	***B. rapa* '29'**
A6	**C18:1[b]**	**18.5**	**	**16.4–28.8**	**21.1** ± 12.2	**1.97** ± 0.59	**Tapidor**
	C18:2	18.5	**	16.4–28.8	27.7 ± 5.4	1.54 ± 0.20	*B. rapa* '29'
A8	C18:1	59.6–64.5	**	59.6–64.5	19.9 ± 6.8	3.37 ± 0.15	*B. rapa* '29'
	C18:2	59.6–64.5	***	59.6–64.5	29.1 ± 13.3	2.75 ± 0.62	Tapidor
	C18:3	59.6–64.5	*	59.6–64.5	19.1 ± 1.3	1.09 ± 0.10	Tapidor
C1	C18:1[b]	18.4	**	0–36.6	20.5 ± 11.3	2.03 ± 0.32	Tapidor
	C18:1	**42.7–43.2**	***	**42.7–58**	**28.2** ± 9.7	**2.21** ± 0.33	**Tapidor**
	C18:2	**18.4**	**	**0–36.6**	**28.1** ± 5.3	**1.63** ± 0.06	***B. oleracea atlantica***
	C18:2	**45.9**	**	**42.7–58**	**23.2** ± 5.8	**1.38** ± 0.20	***B. oleracea atlantica***
C2	**C18:3**	**47.6**	*	**0–50.4**	**16.7** ± 3.01	**0.56** ± 0.06	***B. oleracea atlantica***
C3	C18:1	101–110.1	**	93.6–110.1	18.3 ± 4.1	3.38 ± 0.14	*B. oleracea atlantica*
	C18:2	101–110.1	**	93.6–110.1	27.1 ± 9.9	2.75 ± 0.62	Tapidor
	C18:3	101–110.1	*	101–110.1	19.1 ± 1.3	1.08 ± 0.11	Tapidor
C5	**C18:1[b]**	**0**	**	**0–48.1**	**18.9** ± 9.1	**2.19** ± 0.61	***B. oleracea atlantica***
C6	C18:2[b]	57.8	*	57.8	12.3 ± 5.5	1.08 ± 0.22	*B. oleracea atlantica*
C7	**C18:1**	**16.9–22.7**	**	**0–23.8**	**22.7** ± 11.2	**3.33** ± 0.97	***B. oleracea atlantica***
	C18:2	**64.7**	**	**50.8–64.7**	**23.4** ± 7.0	**1.40** ± 0.05	**Tapidor**
	C22:1[b]	19.3	**	0–25.9	20.7 ± 12.2	0.42 ± 0.34	Tapidor
C8	C18:1[b]	61.6	*	47.9–82.6	14.6 ± 6.8	1.69 ± 0.49	*B. oleracea atlantica*
	C18:2[b]	**59.5**	*	**47.9–82.6**	**12.6** ± 7.8	**1.08** ± 0.19	**Tapidor**
	C18:3	**45.4**	**	**10.8–82.9**	**25.7** ± 7.1	**0.67** ± 0.06	**Tapidor**
C9	**C18:0**	**46.5**	***	**29.2–53.7**	**30.3** ± 24.7	**0.27** ± 0.19	***B. oleracea atlantica***
	C18:1[b]	**23.2**	***	**14.3–43.4**	**30.7** ± 22.7	**2.44** ± 0.94	***B. oleracea atlantica***
	C18:2	38.6	***	14.3–43.4	38.6 ± 11.4	1.97 ± 0.46	Tapidor
	C18:3	38.6	**	14.3–43.4	25.9 ± 9.5	0.70 ± 0.24	Tapidor

A range of peak positions are given where restricted recombination may have resulted in similar peak values for multiple adjacent markers. LOD scores and peak position are given for the most significant QTL found while the spread is the maximum range where $P < 0.05$ for the three QTL; % explained and additivities are the means of 3 years regardless of significance in any one year. QTL probabilities: *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$. QTL highlighted by bold typeface were identified only in the single marker regression analysis of the low C22:1 subset of lines

[a] Fatty acids expressed as percentage of total fatty acid content

[b] QTL identified at a minimum of $P < 0.05$ in 2 years only

[c] QTL in 1 year only

[d] % of the phenotypic variance explained by QTL

B. rapa '29' (A8 peak LOD 59.6–64.5 cM) or *B. oleracea atlantica* (C3 peak LOD 101–110.1 cM) and correspond in position to the major C22:1 QTL. The remaining 19 QTL relating to 18-carbon fatty acids were identified only in the analysis restricted to lines with low C22:1 content, and are highlighted in Table 7 in bold typeface.

Discussion

In order to maximise genome-wide allelic variation, we selected the DH mapping population QDH. This population was derived from a cross between a conventional European winter habit oilseed rape variety with low C22:1 content in seed oil (Tapidor) and a resynthesised *B. napus*, which had been produced by crossing *B. oleracea atlantica* (C genome donor; a Mediterranean cliff-dwelling wild species) and *B. rapa* '29' (A genome donor; a cultivated turnip rape). Doubled haploid mapping populations commonly display segregation distortion (Foisset et al. 1996). Many (55%) of the markers mapped in QDH showed significant segregation distortion ($P > 0.05$), a higher proportion than reported in previous studies using such populations (e.g. 26.4% by Lombard and Delourme 2001; 22% by Cheung et al. 1997). A greater proportion of polymorphic markers were identified and mapped to the C genome (59%) than mapped to the A genome (41%), which is indicative of the genome of *B. oleracea atlantica* being more diverged from the C genome of oilseed rape than the genome of *B. rapa* '29' is from the A genome of oilseed rape. A similar asymmetry between the genomes was observed for QTL content, with more being identified in the C genome than in the A genome. Thus, our attempt to maximise allelic diversity was more successful for the C genome linkage groups than for the A genome linkage groups, and resulted in a relatively high level of segregation distortion of markers.

The genetic control of the fatty acid composition of seed oil in *B. napus* has been studied previously in other populations: two low C22:1 × low C22:1 populations (Burns et al. 2003; Hu et al. 2006), a high C22:1 × high C22:1 population (Zhao et al. 2008) and a high C22:1 × low C22:1 population (Qiu et al. 2006). The C22:1 content of seed oil is controlled almost entirely by two *BnaFAE1* loci, which encode the elongase that competes for C18:1 with the product of *BnaFAD2*. We developed molecular markers for the QDH population for the positions in the genome for the two *BnaFAE1* loci, which mapped as expected to linkage groups A8 and C3, along with markers for the positions on the QDH linkage map of seven loci encoding *Brassica* orthologues of enzymes involved in the desaturation of 18-carbon fatty acids for accumulation in seed storage lipid. These include one for *BnaFAB2* (C18:0–C18:1) represented by a pair of tightly linked markers on linkage group A3, three for *BnaFAD2* (C18:1–C18:2) on linkage groups C1, A5 and C5, and three for *BnaFAD3* (C18:2–C18:3), one on linkage group C3 and two at different positions on linkage group C4. Although this set of markers probably represents all loci for *B. napus* orthologues of the Arabidopsis locus *FAE1*, it does not represent all of the loci inferred to be present for the *B. napus* orthologues of the Arabidopsis loci *FAD2* and *FAD3* (Scheffler et al. 1997). It is not known how many orthologues of *FAB2* are present in *B. napus*.

The *BnaFAE1* loci on linkage groups A8 and C3 were expected to control C22:1 content of seed oil, and the results of our QTL analysis by IM confirmed this. In total, IM revealed 32 QTL for fatty acid content, 7 in the A genome and 25 in the C genome. As the elongase encoded by *BnaFAE1* loci competes with the desaturase encoded by *BnaFAD2* for C18:1 substrate, segregation of *BnaFAE1* alleles also has a major influence on content of C18:1, C18:2 and C18:3. In an attempt to dissect out the influence of the *BnaFAE1* loci, we used MQM, which revealed an overlapping set of 23 QTL, 9 in the A genome and 14 in the C genome. Although the MQM analysis was unable to completely eliminate the effects of the major C22:1 loci, it successfully revealed four QTL and one that resolved into two adjacent QTL for the content of PUFAs that had not been identified by IM. As a third approach to the detection of QTL, we analysed fatty acid data from only the subset of 25 lines that showed a low C22:1 content of seed oil. This approach eliminated the effect of segregation of the major C22:1 loci. However, the relatively small number of lines necessitates the use of SM regression analysis, which has less power to determine the positions of QTL than, for example, IM. Nevertheless, this analysis revealed 34 QTL, 14 in the A genome and 20 in the C genome.

The QTL that we identified on linkage group C1, peak LOD 17.7 cM, coincides with the mapped position of an orthologue of *FAD2* (marker JICB0637). At this locus, the *B. oleracea atlantica* allele increases content of C18:2 and decreases content of C18:1, which is consistent with the basis of the QTL being functional variation of a *FAD2* orthologue. Hu et al. (2006) did not identify a QTL for C18:2 in this position. The variant that we have identified in *B. oleracea atlantica* is thus a novel allele of potential value for increasing the PUFA content of seed oil. None of the remaining QTL that we identified affecting desaturation of seed oil were consistent with the positions of orthologues of *FAB2*, *FAD2* or *FAD3*. However, using SM regression analysis, we were able to overcome the difficulties imposed by the segregation of major QTL controlling C22:1 content and identified in *B. rapa* '29' and *B. oleracea atlantica* novel alleles at numerous other loci that are of potential value for altering PUFA content or PUFA composition of seed oil. These include both alleles that might be useful for decreasing PUFA content/extent of desaturation for industrial applications such as bio-lubricants (A1 peak LOD 16.9–24.6 cM, A3 peak LOD 61.9 cM, C7 peak LOD 64.7 cM, C8 peak LOD 47.9 cM, C8 peak LOD 82.6 cM, C9 peak LOD 38.6 cM) and alleles that might be useful for increasing PUFA content/extent of desaturation for increased omega 3/6 essential fatty acids in

the diet (A4 peak LOD 13.3 cM, A6 peak LOD 18.5 cM, C1 peak LOD 17.7 cM, C1 peak LOD 42.7–45.9 cM, C2 peak LOD 47.6 cM, C6 peak LOD 57.8 cM). These may represent the positions of regulatory loci, as they do not coincide with loci encoding the enzymes known to be involved in PUFA biosynthesis (*i.e. BnaFAD2* and *BnaFAD3*).

A characteristic of IM is that potential QTL are treated as discrete entities and does not allow for epistasis among QTL and thus tends to overestimate the potential effect of the QTL. This can be seen clearly when the estimated proportion of the total phenotypic variation for all the QTL are combined. The proportion explained, with either Tapidor or the Synthetic parent as the increasing allele, is 46.1 and 119.2%, respectively giving a total of 165.3% of the observed variation; a significant overestimate of the effect of the individual QTL effects. MQM mapping, using the two major *BnaFAE1* loci, partially controls this overestimate and reduces the nine QTL identified in interval mapping to five and where, in consequence, the proportion of the variance explained by the QTL is reduced to 10.0 and 77.4% for Tapidor and the Synthetic parent, respectively which total together 87.4%. This change in the proportion of the variance explained is reflected in the combined additivities where, for C22:1, IM predicts that the QTL account for 119.1% compared to only 51.8% in MQM. MQM thus suggests that, when epistasis is taken into account, there are other potential QTL which remain unidentified in our analysis. When these calculations are also made for the other fatty acids it becomes clear that C18:1 behaves in a very similar manner to C22:1 as would be predicted by the close identity of the QTL positions. In contrast, MQM mapping for C18:2 and C18:3 detects fewer QTL than IM but makes less difference in the proportion of the variance detected 102.0 and 84.9% which is reduced to 92.7 and 74.6%, respectively. In contrast, however, the proportion of the variation estimated by the QTL increases from 39.6 and 33.0% to 43.3 and 37.6%, respectively. QTL analysis of the low-erucic sub-population detected the most QTL. This was reflected in the proportion of the variance explained which totalled 252.1, 265.6 and 149.2% for C18:1, C18:2 and C18:3, respectively with estimated additivities of 136.1, 165.8 and 105.5% of the population range for these fatty acids. These large overestimates are expected to be the results of several factors including epistatic interactions and the greater likelihood of type 1 errors resulting from the small population size.

By resynthesising *B. napus* from its progenitor species, we have accessed genetic variation from beyond the conventional *B. napus* germplasm. Using this, we have discovered novel QTL regulating the fatty acid composition of seed oil. These QTL can be used to help identify genes that have not previously been linked to the control of fatty acid profile in rapeseed and therefore may be used to develop tools for the fine-tuning of oil composition. Our study was restricted to the contribution of alleles from single donor accessions representing each of the *Brassica* A and C genomes. Further novel alleles are likely to exist amongst the broad range of genetic diversity available in collections of *B. rapa* and *B. oleracea*, which can be explored and exploited via resynthesised *B. napus*. Incorporating this unused allelic diversity into oilseed rape breeding programmes will further improve our ability to modulate fatty acid profiles for specific end-uses of rapeseed oil.

Acknowledgments This work was supported by a PhD studentship and competitive support grant to J.I.C. from the UK Biotechnology and Biological Sciences Research Council.

References

Browse J, McConn M, James D, Miquel M (1993) Mutants of *Arabidopsis* deficient in the synthesis of alpha-linolenate—biochemical and genetic-characterization of the endoplasmic-reticulum linoleoyl desaturase. J Biol Chem 268:16345–16351

Burns MJ, Barnes SR, Bowman JG, Clarke MHE, Werner CP, Kearsey MJ (2003) QTL analysis of an intervarietal set of substitution lines in *Brassica napus*: (i) seed oil content and fatty acid composition. Heredity 90:39–48

Cheung WY, Champagne G, Hubert N, Landry BS (1997) Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. Theor Appl Genet 94:569–582

Cho K, O'Neill CM, Kwon S-J, Yang T-J, Smooker A, Fraser F, Bancroft I (2010) Sequence-level comparative analysis of the *Brassica napus* genome around two stearoyl-ACP desaturase loci. Plant J 61:591–599

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Foisset N, Delourme R, Barret P, Hubert N, Landry BS, Renard M (1996) Molecular-mapping analysis in *Brassica napus* using isozyme, RAPD and RFLP markers on a doubled-haploid progeny. Theor Appl Genet 93:1017–1025

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

Hu XY, Sullivan-Gilbert M, Gupta M, Thompson SA (2006) Mapping of the loci controlling oleic and linolenic acid contents and development of FAD2 and FAD3 allele-specific markers in canola (*Brassica napus* L.). Theor Appl Genet 113:497–507

James DWJ, Lim E, Keller J, Plooy I, Ralston E, Dooner HK (1995) Directed tagging of the *Arabidopsis* fatty acid elongase-1 (FAE1) gene with the maize transposon activator. Plant Cell 7:309–319

Kachroo P, Shanklin J, Shah J, Whittle EJ, Klessig DF (2001) A fatty acid desaturase modulates the activation of defense signalling pathways in plants. Proc Natl Acad Sci USA 98:9448–9453

Kachroo A, Shanklin J, Whittle E, Lapchyk L, Hildebrand D, Kachroo P (2007) The Arabidopsis stearoyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. Plant Mol Biol 63:257–271

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). Mol Biol Evol 17:1483–1498

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. Theor Appl Genet 103:491–507

Lowe AJ, Moule C, Trick M, Edwards KJ (2004) Efficient large-scale development of microsatellites for marker and mapping applications in Brassica crop species. Theor Appl Genet 108:1103–1112

Lucigen (2008) CloneSmart® cloning kits

Lühs W, Friedt W (1993) Nonfood uses of vegetable oils and fatty acids. In: Murphy DJ (ed) Designer oil crops: breeding processing and biotechnology. VCH, Cambridge, pp 73–130

Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe Brassiceae. Genome Res 15:516–525

Miquel M, Browse J (1992) *Arabidopsis* mutants deficient in polyunsaturated fatty-acid synthesis—biochemical and genetic-characterization of a plant oleoyl-phosphatidylcholine desaturase. J Biol Chem 267:1502–1509

Mithen RF, Magrath R (1992) Glucosinolates and resistance to *Leptosphaeria maculans* in wild and cultivated *Brassica* species. Plant Breed 108:60–68

O'Neill CM, Bancroft I (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *albogalabra* that are homeologous to sequenced regions of chromosome 4 and 5 of *Arabidopsis thaliana*. Plant J 23:233–243

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

Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM (2009) GenStat for Windows introduction, 12th edn. VSN International, Hemel Hempstead

Qiu D, Morgan CL, Shi J, Long Y, Liu J, Li R, Zhuang X, Wang Y, Tan X, Dietrich E (2006) A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. Theor Appl Genet 114:67–80

Rana D, Boogaart T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lim YP, Bancroft I (2004) Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. Plant J 40:725–733

Scheffler JA, Sharpe AG, Schmidt H, Sperling P, Parkin IAP, Lühs W, Lydiate DJ, Heinz E (1997) Desaturase multigene families of *Brassica napus* arose through genome duplication. Theor Appl Genet 94:583–591

Schierholt A, Becker HC (2001) Environmental variability and heritability of high oleic acid content in winter oilseed rape. Plant Breed 120:63–66

Schierholt A, Becker HC, Ecke WG (2000) Mapping a high oleic acid mutation in winter oilseed rape (*Brassica napus* L.). Theor Appl Genet 101:897–901

Simopoulos AP (2008) The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. Exp Biol Med 233:674–688

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: Join Map. Plant J 3:739–744

Suwabe K, Iketani H, Nunome T, Kage T, Hirai M (2002) Isolation and characterization of microsatellites in *Brassica rapa* L. Theor Appl Genet 104:1092–1098

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Fujimura M, Nunome T, Fukuoka H, Matsumoto S, Hirai M (2003) Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa*. Theor Appl Genet 107:997–1002

Tanhuanpää P, Vilkki J, Vihinen M (1998) Mapping and cloning of FAD2 gene to develop allele-specific PCR for oleic acid in spring turnip rape (*Brassica rapa* ssp. *oleifera*). Mol Breed 4:543–550

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

Trick M, Bancroft I, Lim YP (2007) The *Brassica rapa* genome sequencing initiative. Genes Genomes Genomics 1:35–39

Van Ooijen JW (2004) MAPQTL5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma B. V., Wageningen, The Netherlands

Van Ooijen JW, Voorips RE (2001) Joinmap 3.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands

Zhao JY, Dimov Z, Becker HC, Ecke WG, Mollers C (2008) Mapping QTL controlling fatty acid composition in a doubled haploid rapeseed population segregating for oil content. Mol Breed 21:115–125