Theor Appl Genet (2006) 114:67–80
DOI 10.1007/s00122-006-0411-2

ORIGINAL PAPER

A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content

D. Qiu · C. Morgan · J. Shi · Y. Long · J. Liu · R. Li · X. Zhuang · Y. Wang · X. Tan · E. Dietrich · T. Weihmann · C. Everett · S. Vanstraelen · P. Beckett · F. Fraser · M. Trick · S. Barnes · J. Wilmer · R. Schmidt · J. Li · D. Li · J. Meng · I. Bancroft

Received: 9 June 2006 / Accepted: 8 September 2006 / Published online: 11 October 2006

Abstract We have developed a new DH mapping population for oilseed rape, named TNDH, using genetically and phenotypically diverse parental lines. We used the population in the construction of a high stringency genetic linkage map, consisting of 277 loci, for use in quantitative genetic analysis. A proportion of the markers had been used previously in the construction of linkage maps for *Brassica* species, thus permitting the alignment of maps. The map includes 68 newly developed Sequence Tagged Site (STS) markers targeted to the homologues of defined genes of *A. thaliana*. The use of these markers permits the alignment of our linkage map with the *A. thaliana* genome sequence. An additional 74 loci (31 newly developed STS markers and 43 loci defined by SSR and RFLP markers that had previously been used in published linkage maps) were added to the map. These markers increased the resolution of alignment of the newly constructed linkage map with existing *Brassica* linkage maps and the *A. thaliana* genome sequence. We conducted field trials with the TNDH population at two sites, and over 2 years, and identified reproducible QTL for seed oil content and erucic acid content. The results provide new insights into the genetic control of seed oil and erucic acid content in oilseed rape, and demonstrate the utility of the linkage map and population.

Introduction

Rapeseed oil is edible and its fatty acid composition has been extensively manipulated by breeding. Oil low in erucic acid contains a near optimal balance of fatty acids for human health and nutrition, and has found widespread application in the food industry.

Electronic supplementary material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00122-006-0411-2 and is accessible for authorized users.

Communicated by M. Kearsey.

D. Qiu and C. Morgan authors contributed equally to the work.

D. Qiu · J. Shi · Y. Long · J. Liu · R. Li · J. Meng (✉)
National Key Laboratory of Crop Genetic Improvement,
Huazhong Agricultural University, 1 Shizishan,
430070 Wuhan, China
e-mail: jmeng@mail.hzau.edu.cn

C. Morgan · P. Beckett · F. Fraser · M. Trick · I. Bancroft
John Innes Centre, Norwich Research Park, Colney,
Norwich, NR4 7UH, UK

X. Zhuang · Y. Wang · X. Tan · J. Li
Chinese Academy of Sciences,
Institute of Genetics, Building 917,
Datun Road, Chaoyang District,
100101 Beijing, China

E. Dietrich · T. Weihmann · R. Schmidt
Max Planck Institute of Molecular Plant Physiology,
14424 Potsdam, Germany

C. Everett · J. Wilmer
Biogemma UK, 200 Cambridge Science Park, Milton Road,
Cambridge, CB4 0GZ, UK

S. Vanstraelen · S. Barnes
SESVANDERHAVE N.V./S.A. Industriepark,
Soldatenplein Z2 nr 15, 3300 Tienen, Belgium

D. Li
Hybrid Rapeseed Research Centre of Shaanxi,
Dali 715105, China

Oil with other fatty acid composition characteristics, including oil high in erucic acid, has considerable potential for non-food applications (Lühs and Friedt 1993). Oilseed rape (*Brassica napus*) represents an excellent cash crop, particularly for developing countries where the byproduct of oils production, a high quality meal, is used as animal feed. Global demand is developing rapidly, but production is presently unable to meet that demand. An important obstacle to production is the availability of cultivars of oilseed rape, which are able to thrive under low input farming regimes. Understanding the genetic control of key agronomic traits in *B. napus*, such as seed oil and erucic acid content, and their environmental interactions, is a major objective, which will underpin the development of the crop for cultivation in a broader range of environments.

Many traits of agronomic importance are quantitative in nature and have complex genetic bases. The identification of Quantitative Trait Loci (QTL) represents a first step toward dissecting the molecular basis of such complex traits. A prerequisite for QTL mapping studies is the availability of genetic linkage maps, and many of these have been produced for *Brassica* species (e.g. Ferreira et al. 1994; Sharpe et al. 1995; Uzunova et al.1995; Foisset et al. 1996; Cheung et al. 1997; Lombard and Delourme 2001). QTL studies have been conducted in *Brassica* species for a variety of traits, such as flowering-time, glucosinolates and erucic acid contents in seeds, oil and oleic acid content in seeds and some other agronomic traits (Snowdon and Friedt 2004). However, few of the detected QTL have been successfully used for rapeseed breeding, perhaps because QTL for many traits are environmentally sensitive and QTL-environment interaction has been considered as a common phenomenon (Jansen et al. 1995; Wang et al. 1999; Piepho 2000; Yan et al. 1998; Cao et al. 2001; Xing et al. 2002).

The cultivated *Brassica* species are the group of crops most closely related to the widely used model plant, *Arabidopsis thaliana*. Although the *Brassica* and *Arabidopsis* lineages diverged only around 20 million years ago (Yang et al. 1999), comparative analyses of genome organization are not straightforward. This is primarily a consequence of the "diploid" *Brassica* species, including *B. rapa* and *B. oleracea*, being palaeohexaploids (Lagercrantz and Lydiate 1996; Lysak et al. 2005). In *B. napus*, the situation is even more complex, as the species arose by the hybridization of *B. rapa* and *B. oleracea*, although each of the progenitor genomes largely retains its original organization (Parkin et al. 2003). When studied at the level of linkage maps, the genomes of *Brassica* species show extensive segmental rearrangement compared with the genome of *A. thaliana* (Lagercrantz 1998). However, when the microstructure of genome segments are compared, extensive collinearity of conserved genes is revealed (O'Neill and Bancroft 2000; Rana et al. 2004). A genome-wide comparative analysis has been conducted using a linkage map of *B. napus* that consisted of over 1,000 linked RFLP loci that were mapped to homologous positions in the *A. thaliana* genome based on sequence similarity (Parkin et al. 2005). Twenty-one segments of the genome of *A. thaliana*, representing almost its entirety, could be duplicated and rearranged to generate the extant structure of the *B. napus* genome. The majority of the *A. thaliana* genome (11 segments) could be aligned to 6 segments of the *B. napus* genome, as expected given the triplicated nature of the genomes of both progenitor species. However, some segments of the *A. thaliana* genome aligned to 7 segments of the *B. napus* genome, indicative of additional segmental duplications in the *Brassica* lineage and others aligned to only 4 or 5 segments of the *B. napus* genome, indicative of segmental loss from the *Brassica* lineage. Although RFLP markers are too laborious for routine genetic analysis, this study highlights both the need for, and the potential utility of, linkage maps with comparative power.

Our aims were to develop a DH mapping population using genetically diverse oilseed rape lines and to develop a robust linkage map, based on this population, primarily using convenient PCR-based markers that enable integration with existing *Brassica* linkage maps and with the *A. thaliana* genome sequence. Our first application of these resources was to undertake quantitative genetic analyses of the genetic control of seed oil and erucic acid content, to relate our results to those of previous studies in oilseed rape, and to use comparative genomics to infer the likely presence of candidate genes for the control of erucic acid content in QTL-containing regions of the genome.

Methods and materials

Production of doubled haploid lines for the mapping population

Before development of the mapping population, doubled haploid (DH) lines had been developed from both of the parental genotypes. F_1 plants were made by

crossing Ningyou 7 onto Tapidor, several of which were then used to generate a doubled haploid (DH) population by microspore culture. A total of 202 DH lines were produced in the year 2002 and named as the TNDH population. A maximally informative subset of 188 of these (the number for which marker assays can be conducted in two full 96-well microtitre plates, with parental genotypes in each plate) was identified after preliminary genotyping with a range of molecular markers. These were used for map construction and the QTL analysis reported.

Field trials

Field trials were carried out over two seasons (2002–2003 and 2003–2004) and in two locations: Dali, a winter-type rapeseed growing area in Northwestern China, and Wuhan, a semi-winter-type rapeseed growing area in central China. On each of the four experimental occasions all 202 lines, together with their parents Tapidor and Ningyou 7 and their $F_{1,}$ were grown in three randomized blocks with plots consisting of 30 plants. Five replicate plants were sampled from each plot for analysis of percentage seed oil as a proportion of seed dry weight and percentage erucic acid content as a proportion of total seed fatty acids.

Seed quality trait measurement

The erucic acid content of seeds was measured for the first year by Gas Chromatography (GC) using standard methods (Kaushik and Agnihotri 2000) and in the second year by Near Infrared Spectroscopy (NIR) using standard methods (Mika et al. 2003). The seed oil content was measured in both years by Near Infrared Spectroscopy (NIR) using standard methods (Mika et al. 2003).

Molecular marker assays

RFLP marker probes were obtained from Professor T.C. Osborn, Wisconsin University. Marker assays followed the method described by Zhao et al. (2005), with polymorphisms detected using DNA samples digested with *Eco*RI or *Hin*dIII.

Primer sequences for SSR markers were obtained from various sources: John Innes Centre, UK (http://www.brassica.bbsrc.ac.uk/BrassicaDB/), Agriculture and Agri-Food Canada (http://www.brassica.agr.gc.ca/index_e.shtml), Plant Biotechnology Centre, La Trobe University, Australia (http://www.hornbill.cspp.latrobe.edu.au) and as described by Suwabe et al. (2002). The protocol adopted for the analysis of SSR markers was as described by Lowe et al. (2004).

The method for AFLP analysis was described by Vos et al. (1995) with the restriction enzyme combination changed from E(*Eco*RI)/M(*Mse*I) to S (*Sac*I)/M(*Mse*I). Meanwhile two kinds of methylation-sensitive AFLP markers, Ms-AFLP, were used in this study following Reyna-López et al. (1997) for detecting CG methlation status, and following Vuylsteke et al. (1999) for CNG methylation status.

Markers developed by sequencing PCR products from the parental lines were mainly assayed as single nucleotide polymorphisms (SNPs) or, where the amplified products from Tapidor and Ningyou 7 differed sufficiently in size, as insertion/deletion (InDel) markers. For SNP markers, either Pyrosequencing (http://www.pyrosequencing.com) or methods based on allele-specific amplification (ASA) were used. In ASA, sets of primers were designed that would amplify preferentially from the different alleles at a locus and the size or presence of the predominant amplification products used to determine the allele present in an individual line.

Markers developed from additional PCR amplicons that were not sequenced were assayed by either differential band amplification (DBA) or as cleaved amplified polymorphic sequences (CAPS). The DBA assay exploited differences in the presence or patterns of PCR bands amplified from the Tapidor and Ningyou 7 alleles. CAPS assays involved cleavage of PCR products with a cocktail of four restriction endonucleases with 4 bp specificity (*Alu*I, *Hae*III, *Mbo*I and *Mse*I), resolution of the fragments by polyacrylamide gel electrophoresis and visualisation by silver staining. Both methods are suitable for the analysis of amplicons comprising pairs of co-amplified loci.

Development of molecular markers for genome alignment

The TNDH population and associated linkage map have been developed as communal resources, with the aim of integrating existing oilseed rape linkage maps with the clone-based physical maps of *Brassica* genomes (http://www.brassica.bbsrc.ac.uk/IGF/) and the *A. thaliana* genome sequence (Arabidopsis Genome Initiative 2000). Some of the pre-existing *Brassica* RFLP markers have been aligned in silico to the *A. thaliana* genome via their DNA sequences (Parkin et al. 2005). These sequences were entered into the project database (http://www.brassica.bbsrc.ac.uk/IMSORB/) and are also available from BrassicaDB (http://www.brassica.bbsrc.ac.uk/BrassicaDB/). In addition,

we developed markers linked to a subset of *A. thaliana* gene-specific hybridization probes, the "IGF" probes (http://www.brassica.bbsrc.ac.uk/IGF/). Our approach was to hybridize the *A. thaliana* gene-specific probes to a bacterial artificial chromosome (BAC) library of oilseed rape cultivar Tapidor genomic DNA (Rana et al. 2004). Positive clones were then sorted, using Southern blots, into the different paralogues/homoeologues that they represent, as previously described (O'Neill and Bancroft 2000). Sequences were derived from the BACs by either directly end sequencing the BAC clones or by subcloning digested BAC inserts into plasmid vectors and sequencing these. These sequences were used to design locus-specific PCR primers and polymorphisms were identified between Tapidor and Ningyou 7 by sequencing of allelic PCR products. Markers were assayed across the population by the most appropriate method for the type of polymorphism detected, including Pyrosequencing, ASA, InDels, DBA and CAPS. Marker names were prefixed with "IGF", followed by the IGF probe number and suffixed by a letter denoting the locus.

The primers for amplifying marker STS02 were described by Pires et al. (2004) to identify *FLC* orthologous in *B. napus*. The primer pair of marker STS06 was designed from the *FRI* gene in *Arabidopsis* (AT4G00650, http://www.ncbi.nlm.nih.gov) to identify *FRI* orthologous of Arabidopsis with the method of TRAP described by Hu and Vick (2003) for mapping.

Results

Linkage map construction

The parents of the mapping population, oilseed rape cultivars Tapidor and Ningyou 7, were selected on the basis of their gene genetic dissimilarity and their contrasting trait characteristics and cultivation ranges. These two varieties had been found to be the most divergent lines analyzed using RFLP markers (Meng et al. 1996). Tapidor is a European winter-type cultivar (i.e. it has a strong vernalization requirement) with low seed erucic acid and glucosinolate content. Ningyou 7 is a Chinese semi-winter-type cultivar (i.e. it has little vernalization requirement) with high seed erucic acid and glucosinolate content. The breeding of Ningyou 7 included crosses with *B. rapa*, and the cultivar is less winter hardy than Tapidor. A Doubled Haploid mapping population was developed from a cross between DH lines of Tapidor and Ningyou 7, and named the TNDH population.

A variety of different types of molecular markers were used in the construction of the linkage map for the TNDH population. These included newly developed locus-specific sequence-tagged site (STS) markers in addition to non-locus-specific markers (AFLP, Ms-AFLP, RFLP and SSR) that had been mapped previously on other populations. These markers were used to construct a genetic linkage map using 188 DH lines. Mapping was carried out using JoinMap 3.0 (Stam 1993) with the Kosambi mapping function used to estimate genetic intervals in cM. The threshold for goodness of fit was set to $< = 5.0$ with LOD scores >1.0 and a recombination frequency <0.4. The scoring data were re-checked for all markers with poor goodness of fit scores and a proportion of assays were re-run in order to check apparent improbable double recombination events. Corrected data were used for remapping. The overall quality of the linkage map was increased by sequentially removing markers with mean chi-square values greater than 3.0. Additionally, where markers mapped to within 1 cM of each other, only one was selected for inclusion in the final map. This linkage map consisted of 277 loci, of which 115 were defined by AFLP markers, 23 were defined by RFLP markers, 71 were defined by SSR markers and 68 were defined by STS markers. We refer to the resulting linkage map as the high stringency linkage map. Linkage groups were named using the nomenclature convention internationally adopted for *Brassica* maps primarily via the set of SSR markers prefixed "sN", which mostly amplify single bands and were developed specifically to enable this. In addition, consensus alignments were made using further markers that have previously been assigned to linkage groups, but may result in multiple assayable bands and hence potential ambiguity. The resulting linkage map is illustrated in Fig. 1. It has an average mean chi-square value of 1.221 for the 19 linkage groups, a total length of 1,685 cM and an average distance between markers of 7.2 cM. A key to the coding and origin of the different classes of markers is given in Table 1. Details of the primer sequences of all the newly developed STS markers and the *Arabidopsis* gene models to which they relate are given in Supplementary Materials Table S1. Details are given in Supplementary Materials Table S2 of those SSR and RFLP markers that have been mapped in other *Brassica* populations or that can be anchored to the Arabidopsis genome via sequence homology. Primer sequence details of the newly developed AFLP and Ms-AFLP markers are given in Supplementary Materials Table S3.

In order to maximize the alignment of the TNDH linkage map with other *Brassica* linkage maps and also with the *A. thaliana* genome sequence we integrated a subset of informative markers that had been excluded

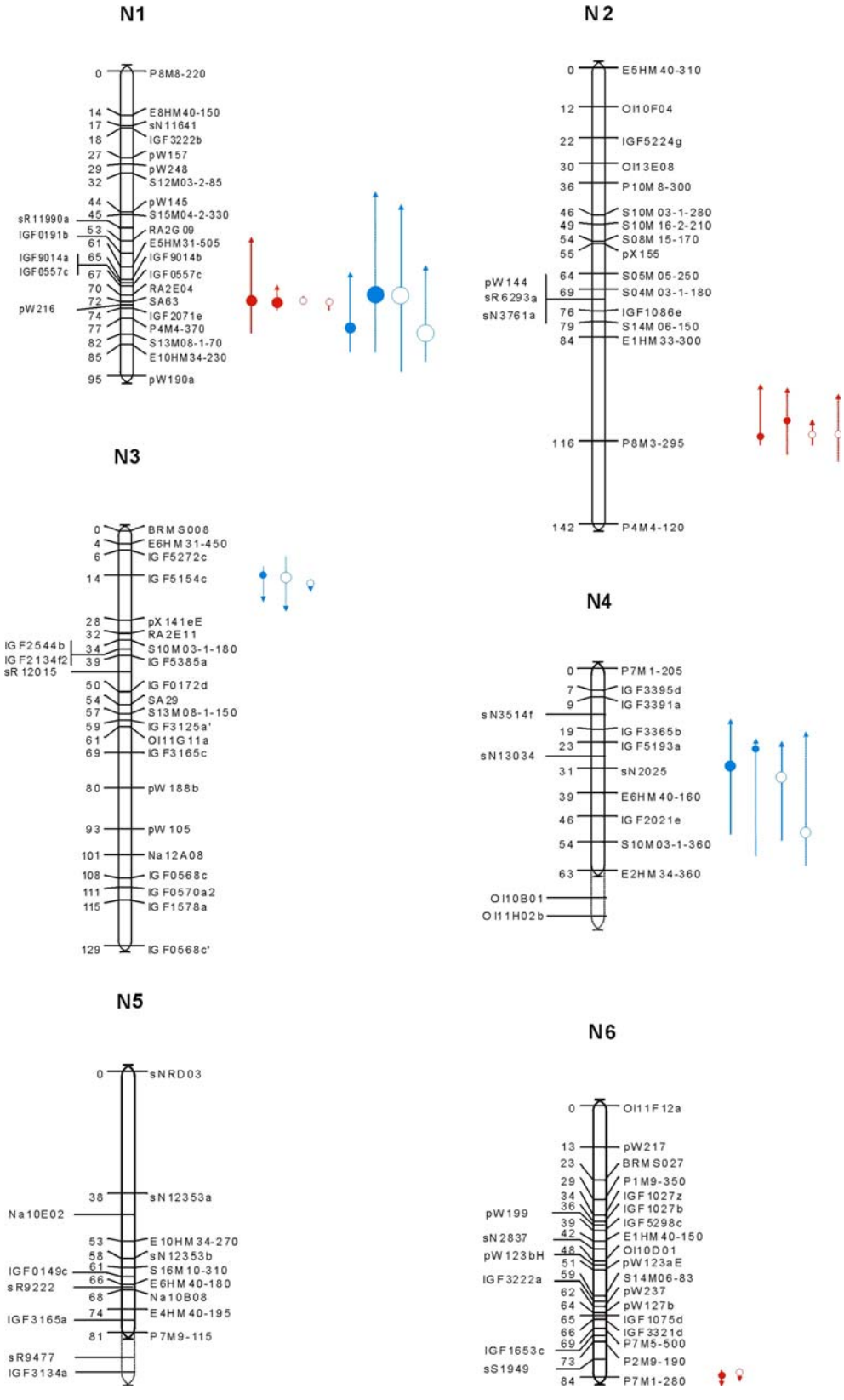
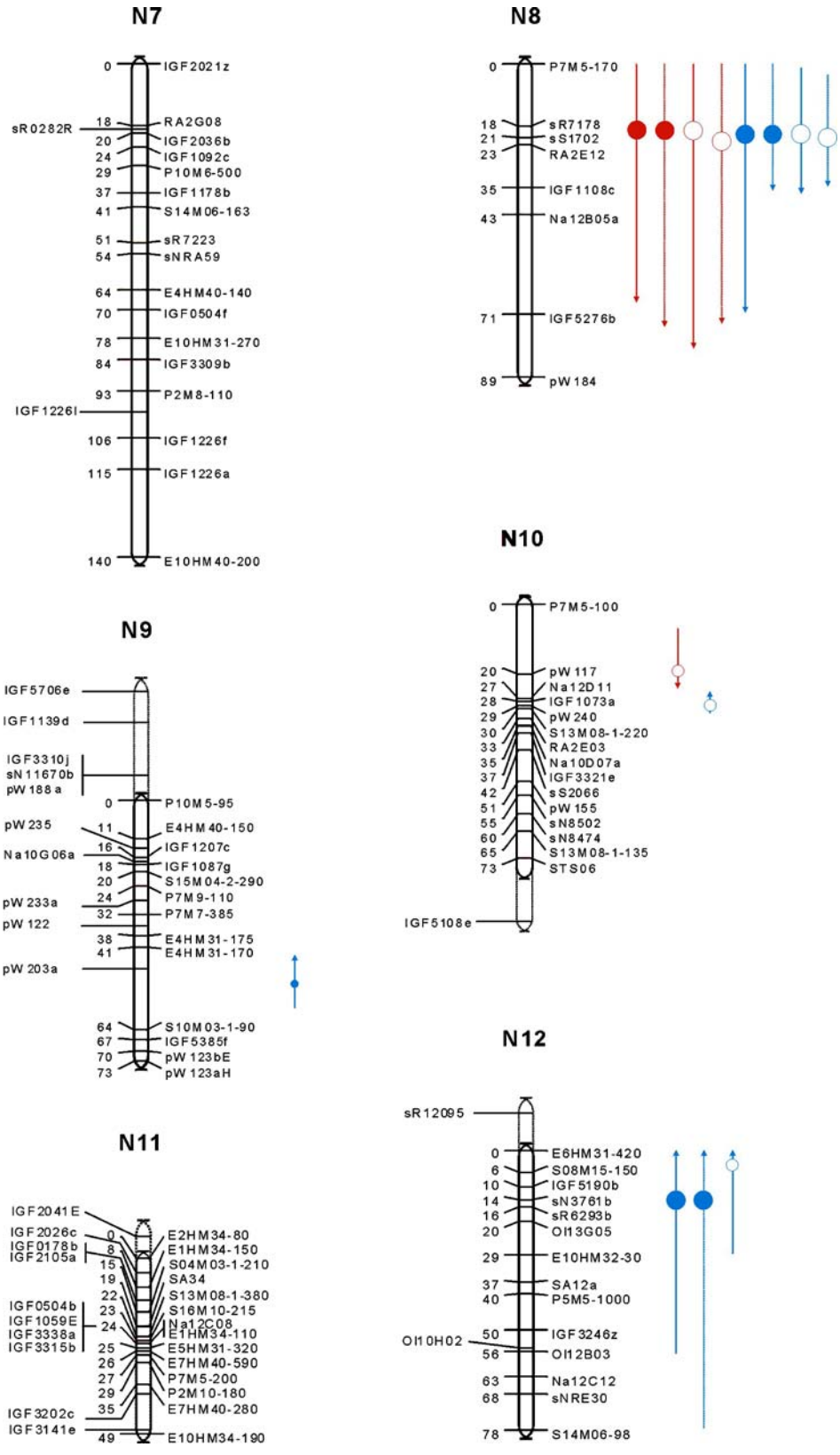


Fig. 1 Genetic Linkage map of TNDH population showing interval mapped QTL determined for erucic acid and oil content of seeds. *Red* Erucic acid content, *blue* seed oil content, *closed symbols* Wuhan, *open symbols* Dali, *solid lines* 2002–2003, *dashed lines* 2003–2004; *position of symbol* indicates peak of QTL. *Upward arrows* indicate Tapidor as the source of the increasing allele; *Downward arrows* indicate Tapidor as the source of the decreasing allele. *Line* indicates length of QTL significant at $P < 0.05$

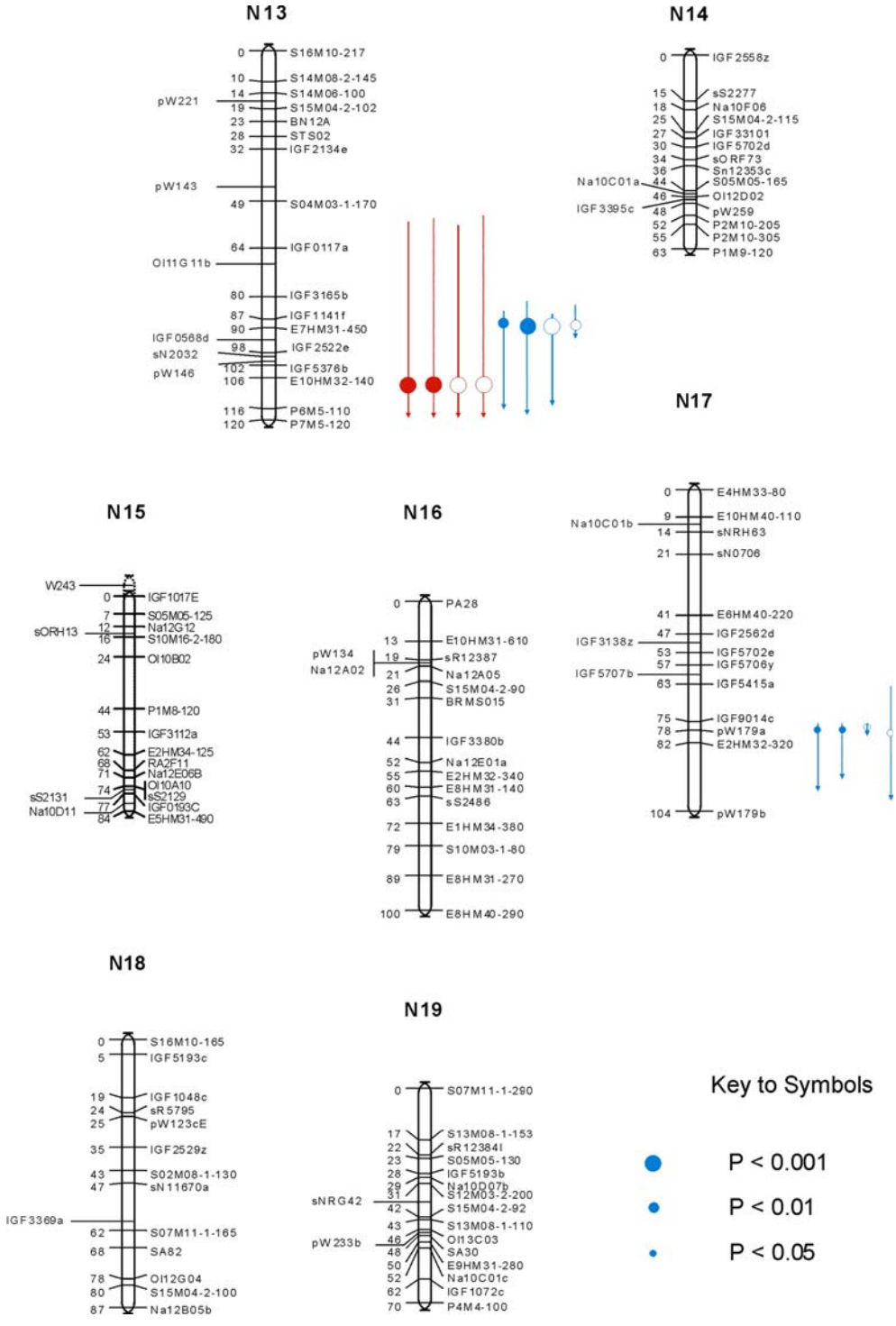
from the high stringency linkage map. This remapping was carried out at reduced stringency levels, for which the recombination level was set to 0.5 and the minimum required LOD value set to 0. These markers were integrated into the high stringency linkage map irrespective of their chi-square likelihood values, but

Fig. 1 continued



maintaining the map in fixed order format to avoid degrading its quality. The 74 loci corresponding to the integrated markers are shown separately in Fig. 1, on the left of the linkage maps, and were not included in subsequent QTL analyses by interval or MQM mapping. This remapping of markers also incorporated one

Fig. 1 continued



further AFLP into the linkage map, bringing the total number of markers to 352.

QTL mapping seed quality traits

The population and its parents, Tapidor and Ningyou 7, were grown in the field as four separate experiments. These were grown at two sites (Wuhan and Dali) in two years (2003 and 2004). There were no significant differences among these experiments for erucic acid or oil content of the harvested seed of the parents. The parents of the population differed in their erucic acid content, with mean values of 3.9 and 56.2% for Tapidor and Ningyou 7, respectively. In contrast, the mean values for seed oil content were similar at 43.6 and 42.0% for Tapidor and Ningyou 7, respectively. No

Table 1 Identification codes for all marker types used in linkage map construction

Marker types	Prefix	Example	Markers in high stringency map	All mapped markers
SNP/InDel	IGF....[a] STS	IGF5224g STS02	68	99
RFLP	pW....[b] pX....	pW248 pX141aE	23	35
SSR (Saskatoon)	sN....[c] sR.... SS....	sN13034 sR6293a sS1949	26	43
SSR (BBSRC)	Na.....[d] Ol.... RA...	Na10E02 Ol10F04 RA2E11	34	48
SSR (others)	BN...[e] BRMS...[f] PA SA....[g]	BN12A BRMS008 PA28 SA63	11	11
Ms-AFLP/AFLP	E....[h] P.... S....	E5HM40-310 P10M8-300 S10M03-1-120	115	116
Total			277	352

[a] IMSORB http://www.brassica.bbsrc.ac.uk/IMSORB/welcome.htm, [b] Sharpe et al. (1995), [c] Agriculture and Agri-Food/*Arabidopsis* Genomics Initiative (Lydiate and Sharpe 2003)

[d] Lowe et al. (2002), [e] Kresovich et al. (1995); Szewc-McFadden et al. (1996)

[f] Suwabe et al. (2002, 2003), [g] Plant Biotechnology Centre, La Trobe University, Australia http://www.hornbill.cspp.latrobe.edu.au/ssrresults/brassica_oleracea/

[h] E and P denote Ms-AFLP and S denotes AFLP

transgressive segregation was noted in the TNDH population for erucic acid, whereas transgressive segregation was observed for seed oil content, giving a range from 35 to 52%. The range and distribution of these traits is shown as histograms in Fig. 2.

To assess the utility of the population and linkage map for the analysis of traits of agronomic importance in oilseed rape, we used MapQTL (Van Ooijen 2004) to carry out QTL analyses of the oil content of seeds and the erucic acid content of seed oil using the 188 lines used for construction of the high stringency linkage map. Significance levels for the LOD scores were determined by permutation analysis (using 10,000 assortments) for each trait and linkage group separately. Single marker analysis was also carried out to assess individual marker trait associations and the presence of QTL on the unlinked markers not mapped in the high stringency map. The maximum extent of putative QTL ($P < 0.05$) for erucic acid and seed oil content are indicated in Fig. 1, on which is also indicated the position and strength of maximum likelihood of the QTL. Additional information including the position and LOD score for the peak of the QTL, the mean trait values for the Tapidor and Ningyou 7 alleles at each QTL and the percentage of the trait variance accounted by that QTL are given in Tables 2 and 3 for seed oil and erucic acid content, respectively. Two major QTL for erucic acid content were detected consistently on linkage groups N8 and N13, at the same map positions, for both experimental locations and over both years. Ningyou 7 was the donor of the increasing allele for both QTL. The average level of expression of lines carrying the allele from Ningyou 7 was 37.9% erucic acid on N8 compared to that of Tapidor, which gave 16.3% erucic acid (additivity is the difference between the mean of the two parents and either parent and in this case is therefore ±10.8%). Similarly on N13, the Ningyou 7 allele gave 33.6% erucic acid compared to the 16.2% for Tapidor allele. In contrast, the Tapidor allele gave higher levels of erucic acid on two additional but smaller QTL in each experiment. The QTL on N1 (mean LOD = 2.7; $P < 0.01$ for Wuhan and $P < 0.05$ for Dali) and N2 (mean LOD = 2.2; $P < 0.05$) had mean erucic acid levels for the Tapidor allele of 29.3 and 30.0% (Ningyou 7 = 21.1 and 22.4%), respectively. The proportion of the trait variation (averaged over location and year) accounted for by these four consistent QTL, i.e. those on N1, N2, N8 and N13, were 5.8, 6.6, 45.1 and 30.4% of the variation, respectively and collectively accounted for 87.9% of the total variation for erucic acid. In addition, there were two potential QTL ($P < 0.05$), on N10 and N6,

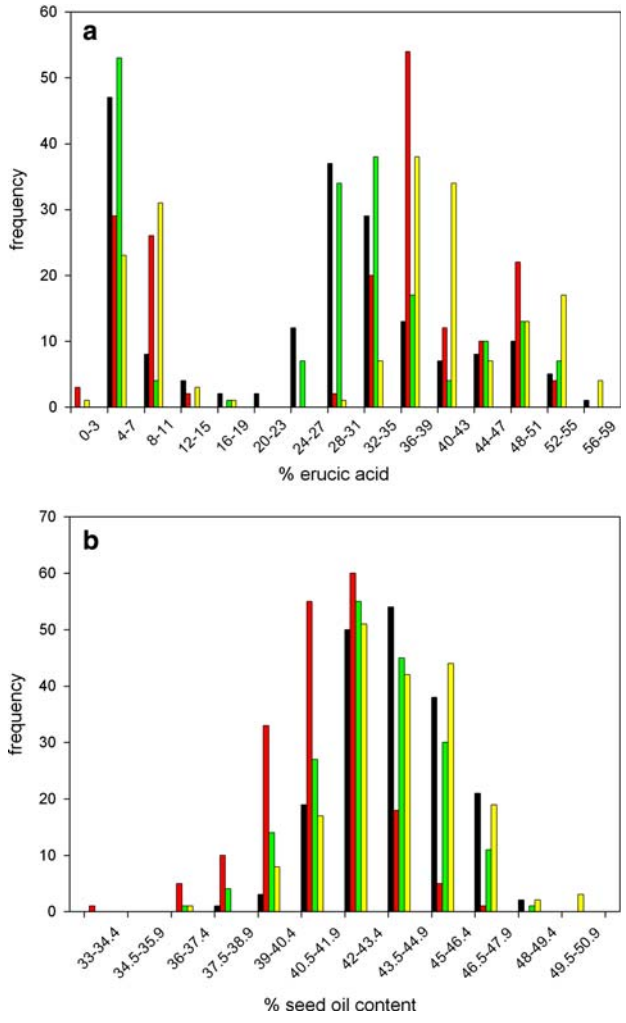


Fig. 2 Distribution frequency of **a** % seed erucic acid content and **b** % seed oil content. *Black bar* Wuhan 2003, *red bar* Wuhan 2004, *green bar* Dali 2003, *yellow bar* Dali 2004

which were detected in only a subset of the four experimental occasions.

Seven QTL for seed oil content were identified. Five of these, located on N1, N3, N4, N8 and N13, were identified on all four experimental occasions. The peak of the QTL on N8 (mean LOD = 4.1; $P < 0.001$) was coincident with that observed for erucic acid content, however, though the peak of the QTL on N13 (mean LOD = 3.3; $P < 0.001$ for Wuhan 2004 and Dali 2003; $P <= 0.01$ for Wuhan 2003 and Dali 2004) fell within QTL detected for erucic acid, its peak differed by 15 cM. The QTL identified on N1, N4 and N17 had mean LOD values of 4.0, 2.6 and 2.4, respectively. The remaining two QTL, which were not detected on all occasions, were located on N3 and N12 with mean LOD values of 2.2 and 3.9, respectively. All of these QTL were more environmentally sensitive than those for erucic acid and the statistical significance and percentage of the trait variance accounted for varied among the experimental occasions. Taking the mean values of the percentage variance accounted for, these QTL accounted for 9.7, 5.4, 6.8, 9.9, 9.9, 7.2 and 5.7%, respectively for N1, N3, N4, N8, N12, N13 and N17 totaling 54.6% of the seed oil trait variance. Alleles increasing seed oil content came from both Tapidor (N1, N4 and N12) and Ningyou 7 (N3, N8, N13 and N17). The additive effect of each of these alleles was very similar, varying from 0.95 to 1.25%, and resulted in the expression of oil content, averaged over the QTL where the Tapidor allele was increasing, of 43.87% (Ningyou = 41.57%) compared to those where Ningyou 7 was the increasing allele of 43.9% (Tapidor = 41.76%).

To identify potential epistatic interactions between the QTL for oil and erucic acid, we further analysed the data using multiple QTL mapping (MQM) with cofactors for erucic acid selected by MapQTL's automatic cofactor selection routine. This was carried out for each experimental occasion separately and identified markers SA63, P8M3-295, S10M03-1-360, sR7178 and P6m5-110 plus E10HM32-140 on N1, N2, N4, N8 and N13, respectively. The cofactors were only used for the analysis if they had been identified for that particular experimental occasion. MQM mapping confirmed the position of the two main oil QTL peaks at 19 and 102–106 cM on linkage groups N8 and N13, respectively, increasing the mean LOD score over each experiment from 22.5 to 30.0 on N8, but decreasing that on N13 from 18.3 to 12. Despite the increased significance of the QTL on N8 the proportion of the trait variance accounted for was slightly reduced from 45.1 to 38.5% for N8 while the reduced significance of the QTL on N13 resulted in a reduction of over half on N13 from 30.4 to 15.6% in N13. The QTL on N1 was essentially unchanged (mean LOD = 3.0; variance accounted for = 2.4%) whereas the minor QTL detected on N2 and N4 were lost. New minor QTL were, however, detected on N15 in Wuhan 2004 and Dali 2003 and in N17 for all except Wuhan 2003 (LOD = 2.0 and 1.8, respectively). Overall, the proportion of the trait variance accounted for by these QTL after MQM analysis was reduced from 87.9 to 60%.

The effect of the two major QTL for erucic acid on seed oil content were studied using sR7178 (N8) and E10HM32-140 to P6M5-110 (N13) as cofactors for MQM analysis. There were no consistent or major effects on the LOD level or the percentage variance accounted for in any of the seed oil QTL detected i.e. N1, N3, N4, N8, N12, although those on N13 and N17 were lost.

Table 2 QTL for % seed oil in TNDH mapping population

	Position of peak QTL (cM)	Peak LOD score	Additivity[a]	% trait variance accounted for
N1				
Wuhan 2003	81.7	2.55**	0.45	6.1
Wuhan 2004	70.5	4.09***	0.59	9.6
Dali 2003	70.2	5.00***	0.71	11.9
Dali 2004	81.5	4.36***	0.72	11.1
N3				
Wuhan 2003	13.8	2.36*	-0.45	5.7
Wuhan 2004	7.1	0.94^{NS}	-0.32	2.4
Dali 2003	13.8	3.41**	-0.62	8
Dali 2004	16.8	1.89*	-0.53	5.5
N4				
Wuhan 2003	31	2.77**	0.47	6.7
Wuhan 2004	25.9	2.21*	0.47	6.3
Dali 2003	35.1	2.51**	0.53	6.6
Dali 2004	52.7	2.87**	0.52	7.6
N8				
Wuhan 2003	20.4	4.74***	−0.62	11.4
Wuhan 2004	20.4	4.65***	−0.65	11.2
Dali 2003	20.4	3.7***	−0.62	8.8
Dali 2004	20.8	3.38***	−0.62	8
N12				
Wuhan 2003	18.1	6.4***	0.72	15.7
Wuhan 2004	13.6	4.49***	0.62	10.6
Dali 2003	4.5	3.17**	0.63	9.4
Dali 2004	4.5	1.33 ^{NS}	0.41	3.7
N13				
Wuhan 2003	88.9	2.67**	−0.47	6.8
Wuhan 2004	89.7	3.9***	−0.57	9.1
Dali 2003	89.7	3.47**	−0.58	8.2
Dali 2004	89.7	3.02**	−0.57	7.2
N17				
Wuhan 2003	77.8	2.12*	−0.41	5.1
Wuhan 2004	77.9	2.3*	−0.45	5.5
Dali 2003	77.8	1.96*	−0.45	4.7
Dali 2004	78.9	3.01**	−0.60	7.6

[a] (Mean effect of Tapidor allele − mean effect of Ningyou 7 allele)/2

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$

Alignment of QTL in oilseed rape with the *A. thaliana* genome

The main reason for the development of comparative markers associated with sequences in the *A. thaliana* genome was to enable rapid assessment as to whether there are known genes for which their oilseed rape orthologues would be candidates for quantitative control of the traits being studied. We assessed the alignment to the *A. thaliana* genome of the coincident QTL on linkage groups N1, N8 and N13 for erucic acid and oil content. This was done in two ways: (1) directly; by looking up the corresponding *A. thaliana* genes for the "IGF" markers or by homology to the sequenced SSR and RFLP markers we used and (2) indirectly; by alignment of our map to that of Parkin et al. (2005), which is itself extensively aligned to the *A. thaliana* genome.

Direct alignment of linkage groups with the *A. thaliana* genome was possible for N1 and N13 by consideration only of sequenced SSR and RFLP markers and the mapped STS markers. The comparative linkage map for N1 included two sequenced RFLP markers (pW157 and pW145) and four STS markers (IGF0191b, IGF9014a, IGF9014b and IGF0557c), all of which align to *A. thaliana* chromosome 4. This alignment is consistent with that reported by Parkin et al. (2005). The comparative linkage map for N13 included three STS markers (IGF0235b, IGF0568d and IGF0117a) that aligned to *A. thaliana* chromosome 4, one sequenced RFLP marker (pW146) and one STS marker (IGF1152z) that aligned to *A. thaliana* chromosome 1, and one STS marker (IGF3165b) that aligned to *A. thaliana* chromosome 3. These results are consistent with the mosaic of alignments reported by Parkin et al. (2005) for N13. The QTL we identified on N1 and

Table 3 QTL for % erucic acid in TNDH mapping population

	Position of peak QTL (cM)	Peak LOD score	Additivity[a]	%trait variance accounted for
N1				
Wuhan 2003	73.3	3.23**	4.24	7.9
Wuhan 2004	73.3	3.17**	4.56	7.9
Dali 2003	72.3	1.85*	3.32	4.5
Dali 2004	73.3	2.46*	4.32	6.3
N2				
Wuhan 2003	115.9	1.99*	3.45	4.9
Wuhan 2004	112.3	2.25*	4.29	6.5
Dali 2003	115.9	1.86*	3.46	4.5
Dali 2004	115.9	2.51*	4.51	6.3
N6				
Wuhan 2003	84.5	2.35*	−3.61	5.7
Wuhan 2004	84.5	1.18 NS	−2.80	2.9
Dali 2003	84.5	1.84*	−3.31	4.4
Dali 2004	84.5	1.22 NS	−3.03	3.1
N8				
Wuhan 2003	19.4	19.46***	-9.57	39.1
Wuhan 2004	19.4	22.95***	−11.02	44.4
Dali 2003	19.4	23.21***	−10.58	44.1
Dali 2004	21.8	24.29***	−12.04	47.1
N10				
Wuhan 2003	14	0.86 NS	−2.65	2.8
Wuhan 2004	15	1.41 NS	−3.73	4.7
Dali 2003	18	2.15*	−4.21	6
Dali 2004	12	1.57 NS	−4.22	5.5
N13				
Wuhan 2003	109.4	13.08***	−8.59	32.7
Wuhan 2004	109.4	11.8***	−7.45	30.4
Dali 2003	109.4	11.18***	-8.34	28.3
Dali 2004	109.4	11.03***	-9.39	29.7

[a] (Mean effect of Tapidor allele − mean effect of Ningyou 7 allele)/2

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$

N13 for erucic acid content and seed oil content overlap the portion of each linkage group that can be aligned with *A. thaliana* chromosome 4.

Direct alignment of linkage group N8 to the *A. thaliana* genome was not possible as it contained only one STS marker, IGF1108c (which aligned with *A. thaliana* chromosome 1). We could, however, confirm the linkage group as N8 as it includes the SSR markers sS1702 and sR7178. These markers lie within the confidence limits of the QTL detected for erucic acid content and oil content, and enable alignment with *A. thaliana* chromosome 4. This alignment, and that which we observed between another part of N8 (containing the STS marker IGF1108c) and *A. thaliana* chromosome 1, are consistent with the alignments reported for N8 by Parkin et al. (2005).

The portion of *A. thaliana* chromosome 4 which aligns with the oil QTL and erucic QTL-containing regions of *B. napus* linkage groups N1, N8 and N13 contains a locus, *FAE1*, which encodes the fatty acid elongase (β-ketoacyl-CoA synthase) primarily responsible for the biosynthesis of very long chain fatty acids, including erucic acid, in seed oil in *A. thaliana* (James and Dooner 1990; James et al. 1995). Thus, the alignment of the QTL we identified with the *A. thaliana* genome could be conducted, and the results are consistent with the location of an appropriate candidate gene.

Discussion

Much of the quantitative genetic analysis on crops has used segregating backcross or F_2 populations. Such populations are poorly suited to the analysis of complex agronomic traits as field trials require plots of genetically uniform individuals. One solution, which has been used for the quantitative genetic analysis of the low erucic phenotype of oilseed rape variety Tapidor, is the construction of a panel of substitution lines, in which segments of the genome of one variety are introgressed into the genetic background of another variety using marker assisted selection (Burns et al. 2003). This method allows the production of large numbers of

genetically identical plants, and hence enables replicated field trials. We chose another widely used approach: chromosome doubling of haploid individuals to generate a genetically fixed doubled haploid (DH) mapping population. This has the advantage of being rapid and more cost-effective for the production of large numbers of lines with differing genetic composition.

We developed a genetic linkage map for the doubled haploid mapping population with the principal aim of integrating it with other linkage maps produced for *Brassica* populations and with the *A. thaliana* genome sequence. The former is important for correlating QTL identified in different experiments; the latter is important for efficiently exploiting the wealth of genomic and gene function information available in that species. To enhance both ease of use and the precision of QTL identification, we developed two versions of the linkage map. The high stringency linkage map consisted of 277 loci, which were selected to provide broad coverage of the linkage map and for which quality metrics indicated the most reliable scoring. This map is for use in QTL identification experiments. The low stringency linkage map included a further 74 loci with comparative power that could be positioned relative to the core map, but which gave redundant coverage of some parts of the genome and/or were likely to contain unidentifiable errors in their scoring. This latter map is used for alignment of *Brassica* linkage maps (hence alignment of QTL identified in different populations) and, by alignment with the *A. thaliana* genome, for the identification of candidate genes for the control of QTL. The inclusion of SSR and RFLP markers (particularly the "sN"-prefixed SSR markers) permitted unambiguous assignment of our linkage groups to the internationally adopted N1–N19 nomenclature.

We identified four QTL for erucic acid content of seed oil. Two of these, mapping to linkage groups N8 and N13, can be considered "major", accounting for about 45 and 30% of the variation in the population, respectively. These likely correspond to the *eru1* and *eru2* major QTL identified on N8 and N13 in a backcross population between Tapidor and a high erucic acid variety, Victor (Howell et al. 1996) and the two QTL identified in a doubled-haploid population formed between varieties Major and Stellar (Thormann et al. 1996). A study of erucic acid content in a panel of substitution lines derived from the Tapidor × Victor backcross population identified the QTL on N8, but assigned the QTL with the second greatest effect to N3 (Burns et al. 2003). However, linkage groups N3 and N13 are homoeologous, being derived from the *B. rapa* and *B. oleracea* progenitors of *B. napus*, respectively (Parkin et al. 2003). Consequently many markers reveal polymorphisms, in different populations, between either or both linkage groups, so assignment can be very difficult. Thus, the N3 assignment of Burns et al. may have been erroneous or the result of incomplete representation of the substitution lines (between 45 and 54% of the Victor genome was estimated to be represented in the panel).

The QTL on N1, N8 and N13 for erucic acid content lie in regions of the *B. napus* genome that show paralogous/homoeologous relationships to one another (Parkin et al. 2003) and align with the long arm of *A. thaliana* chromosome 4 (Parkin et al. 2005). This alignment is consistent with a *Brassica FAE1* homologue as a candidate gene for the control of erucic acid biosynthesis, as orthologous genes are likely to be present in the *B. napus* genome in the vicinity of each of these three QTL. Expression of the jojoba orthologue of *FAE1* in low erucic oilseed rape restored the high erucic phenotype, confirming the function of *FAE1* (Lassner et al. 1996). Mutation of *FAE1* orthologues in oilseed rape is associated with a loss of the corresponding enzyme activity and low erucic acid levels in seed oil (Roscoe et al. 2001). We hypothesize that the QTL on N1 is a poorly characterized minor QTL or "modifier" of the erucic trait, and may be controlled by another orthologue of *FAE1*. However, cloning and functional analysis of the gene will be required for confirmation.

We identified seven QTL for oil content of seeds. All of these were relatively minor, although accounting for ca. 55% of the variation for the trait in total. The QTL mapping to linkage groups N1, N8 and N13 coincided with QTL for erucic acid content. The coincidence of QTL for oil content and erucic acid content has been noted in previous studies. It has been suggested that the erucic acid content of seed oil is itself a major determinant of seed oil content in oilseed rape (Ecke et al. 1995); a hypothesis that has received some support (Burns et al. 2003). Our data provide further support for this hypothesis, although they do not exclude the possibility of different, but tightly linked, loci being involved. However, the control of seed oil content is clearly more complex than being dependent solely upon erucic acid content, as we have identified four reproducible QTL for oil content that do not coincide with QTL for erucic acid.

In contrast to other *B. napus* mapping populations, we developed the TN population specifically from the most genetically and phenotypically diverged varieties of oilseed rape that could be identified. This population and the linkage map that we have developed will provide the opportunity to analyze the genetic control

of a wide range of agronomic traits in oilseed rape, without the problems of genome instability associated with resynthesised *B. napus*. We report a novel methodology for the development, on a large scale, of molecular markers targeted to the oilseed rape homologues of specific genes of *A. thaliana*, and present the first extensive integrative linkage map featuring markers so developed. We have demonstrated the utility of the resource by the analysis of seed oil content and erucic acid composition. This analysis demonstrated that the two known erucic acid QTL could be identified, but also enabled the identification of two novel QTL for this trait, which were consistently identified across experiments. We also identified QTL for seed oil content; an important trait that is poorly understood. We identified seven QTL, several of which are novel and independent of erucic acid content. Further investigation of hypotheses generated using this population will be enhanced by the ability to develop targeted molecular markers. This can be done by using knowledge of the alignment of the linkage map with the *A. thaliana* genome sequence, then using the approach we report to develop markers in oilseed rape linked to orthologues of candidate genes.

Acknowledgments The TNDH population is available from the corresponding author. Marker scoring data, and the latest version of the linkage map, are available via the Web site of the senior authors (http://www.jic.bbsrc.ac.uk/staff/ian-bancroft/research_page3.htm#linkage and http://croplab.hzau.edu.cn/brassica/). We thank the EU Framework 5 Programme for financial support through INCO-DEV grant ICA4-CT-10067 and the National Grand Fundamental Research Program of Education of China under Grant No. 104171.

References

Arabidopsis Genome Initiative (2000) Analysis of the genome of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

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

Cao G, Zhu J, He C, Gao Y, Yan J, Wu P (2001) Impact of epistasis and QTL × environment interaction on the developmental behavior of plant height in rice (*Oryza sativa* L.). Theor Appl Genet 103:153–160

Cheung WY, Friesen L, Rakow GFW, Seguin-Swartz G, Landry BS (1997) A RFLP-based linkage map of mustard (*Brassica juncea* L. *Czern* and *Coss*). Theor Appl Genet 94:841–851

Ecke W, Uzunova M, Weißleder K (1995) Mapping the genome of rapeseed (*Brassica napus* L.). II. Localization of genes controlling erucic acid systhesis and seed oil content. Theor Appl Genet 91:972–977

Ferreira ME, Williams PH, Osborn TC (1994) RFLP mapping of *Brassica* using doubled haploid lines. Theor Appl Genet 89:615–621

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

Howell PM, Marshall DF, Lydiate DJ (1996) Towards developing intervarietal substitution lines in *Brassica napus* using marker-assisted selection. Genome 39:348–358

Hu JG, Vick BA (2003) Target region amplification polymorphism: a novel marker technique for plant genotyping. Plant Mol Biol Rep 21:289–294

James DWJ, Dooner HK (1990) Isolation of EMS-induced mutants in *Arabidopsis* altered in seed fatty acid composition. Theor Appl Genet 80:241–245

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

Jansen RC, Van Ooijen JW, Stam P, Lister C, Dean C (1995) Genotype by environment interaction in genetic mapping of multiple quantitative trait loci. Theor Appl Genet 91:33–37

Kaushik N, Agnihotri A (2000) GLC analysis of Indian rapeseed-mustard to study the variability of fatty acid composition. Biochem Soc Trans 28:581–583

Kresovich S, Szewc-McFadden AK, Bliek SM, McFerson JR (1995) Abundance and characterization of simple sequence repeats (SSR) loci isolated from a size-fractionated genomic library of *Brassica napus L.* (rapeseed). Theor Appl Genet 91:206–211

Lagercrantz U, Lydiate D (1996) Comparative genome mapping in *Brassica*. Genetics 144:1903–1910

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150:1217–1228

Lassner MW, Lardizabal K, Metz JG (1996) A jojoba-ketoacyl-CoA synthase cDNA complements the canola fatty acid elongation mutation in transgenic plants. Plant Cell 8:281–292

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, Jones AE, Raybould AF, Trick M, Moule CL, Edwards KJ (2002) Transferability and genome specificity of a new set of microsatellite primers among *Brassica* species of the U triangle. Mol Ecol Notes 2:7–11

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

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

Lydiate D, Sharpe A (2003) Aliging genetic maps of *Brassica napus* Using microsatellite markers. In: Plant & animal genomes XI conference, p 473

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

Meng J, Sharpe A, Bowman C, Tian Z, Qian X, Lydiate D (1996) Genetic diversity of *Brassica napus* accessions mainly from China detected with RFLP markers. Chin J Genet 23(3):221–232

Mika V, Tillmann P, Koprna R, Nerusil P, Kucera V (2003) Fast prediction of quality parameters in whole seeds of oilseed rape (*Brassica napus* L.). Plant Soil Environ 49:141–145

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

Parkin IAP, Sharpe AG, Lydiate DJ (2003) Patterns of genome duplication within the *Brassica napus* genome. Genome 46:291–303

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

Piepho HP (2000) A mixed-model approach to mapping quantitative trait loci in barley on the basis of multiple environment data. Genetics 156:2043–2050

Pires JC, Zhao JW, Schranz ME, Leon EJ, Quijada PA, Lukens LN, Osborn TC (2004) Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (*Brassicaceae*). Biol J Linn Soc 82:675–688

Rana D, van den 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

Reyna–López GE, Simpson J, Ruiz-Herrera J (1997) Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. Mol Gen Genet 253:703–710

Roscoe TJ, Lessire R, Puyaubert J, Renard M, Delseny M (2001) Mutation in the fatty acid elongase 1 gene are associated with a loss of ketoacyl-CoA synthase activity in low erucic acid rapeseed. FEBS Lett 492:107–111

Sharpe AG, Parkin IAP, Keith DJ, Lydiate DJ (1995) Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome 38:1112–1121

Szewc-McFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. Theor Appl Genet 93:534–538

Snowdon RJ, Friedt W (2004) Molecular markers in *Brassica* oilseed breeding: current status and future possibilities. Plant Breed 123:1–8

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 5:739–744

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* L. Theor Appl Genet 107:997–1002

Thormann CE, Romero J, Mantet J, Osborn TC (1996) Mapping loci controlling the concentrations of erucic and linolenic acids in seed oil of *Brassica napus* L. Theor Appl Genet 93:282–286

Uzunova MI, Ecke W, Weissleder K, Röbbelen G (1995) Mapping the genome of rapeseed (*Brassica napus* L.) I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theor Appl Genet 90:194–204

Van Ooijen JW (2004) MapQTL 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma BV, Wageningen

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Fijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lübberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP® linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Wang DL, Zhu J, Li ZK, Paterson AH (1999) Mapping of QTL with epistatic effects and QTL x environment interactions by mixed model approaches. Theor Appl Genet 99:1255–1264

Xing YZ, Tan YF, Hua JP, Sun XL, Xu CG, Zhang Q (2002) Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. Theor Appl Genet 105:248–257

Yan J, Zhu J, He C, Benmoussa M, Wu P (1998) Molecular dissection of developmental behavior of pant height in rice (*Oryza sativa* L.). Genetics 150:1257–1265

Yang YW, Lai KN, Tai PY, Li WH (1999) Rates of nucleotide substitution in Angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J Mol Evol 48:597–604

Zhao J, Becker Heiko C, Zhang D, Zhang Y, Ecke Wolfgang (2005) Oil content in a European–Chinese rapeseed population: QTL with additive and epistatic effects and their genotype–environment interactions. Crop Sci 45:51–59