Genetic linkage mapping of *Arabidopsis* lipid related genes in *Brassica napus*

Fei Chen[1], Jixiang Huang[1], Yilong Wang[1], Fei Xu[2], Xiyuang Ni[1], Jianyi Zhao[1]
[1]Crop Research Institute, Zhejiang Academy of Agricultural Sciences, 310021, Hangzhou, China; [2]College of Agriculture and Biotechnology, Zhejiang University, 310021, Hangzhou, China.

Abstract

To unravel the genetic basis of seed oil content in *B. napus*, here, we present a high-density molecular linkage map based on *Brassica* EST and BACs derived markers, which were sequence homology to corresponding Arabidopsis genes, especially for those involved in plant lipid biosynthesis or recycling pathways. A total of 374 newly developed markers were integrated into the framework of previously published SG map and formed a new map version of 477 markers with a combined map length of 1956.7 cM. Most of the newly added markers (74.9%; 280/374) are functional markers, and approximately half of them (137) exhibited high homology to 112 acyl lipid metabolism related genes in Arabidopsis. Comparative genomic analysis between B. napus and A. thaliana revealed 32 large syntenic regions. In parallel, a detailed QTL analysis for seed oil content was performed using the new map version across 11 environments and identified nine significant QTL. Co-location analysis revealed that a total of 28 acyl lipid metabolism related candidate gene loci were linked to seven of the nine QTL, indicating their potential role in genetic variability control on seed oil content.

Key words: oilseed rape (*Brassica napus* L.), oil content, lipid related genes, linkage map, QTL

Introduction

Improving seed oil content is a major goal for oilseed rape breeding. However, little is currently known about the genetic determinism of this trait since it is under polygenic control and environment sensitivity. In previous studies, by using different mapping populations, varying numbers of QTL (3 to 14 per study) for oil content in rapeseed have been mapped on 19 linkage groups (see Nesi *et al.*, 2008 and references therein). Recently, a comprehensive comparison between *B. napus* and Arabidopsis based on linkage mapping analysis using sequenced RFLP probes demonstrated that 21 conserved syntenic genomic segments were shared by *B. napus* and Arabidopsis, which could be duplicated and rearranged to represent 90% of the *B. napus* genome (Parkin *et al.*, 2005). It indicated the feasibility of using comparative genomics approach to dissect the genetic control of seed oil content in rapeseed.

In this article, we presented an updated molecular linkage map based on a large number of functional gene markers derived from *Brassica* EST, GSS and BACs, with a high sequence homology to corresponding Arabidopsis genes, especially for those involved in the plant lipid biosynthesis or recycling pathways. Comparison of this map with Arabidopsis genome was then conducted to identify syntenic regions based on homologous loci. Furthermore, a detailed QTL analysis for oil content was performed using the new map version and 11 mapping experiments (environments). Comparison of positions of QTL and candidate gene loci revealed new candidate genes loci that may influence the variation of seed oil content.

Materials and methods

Mapping population and QTL mapping: The segregation DH population from the cross Sollux"×"Gaoyou" was obtained as described in Zhao *et al.* (2005). It consisted of 282 DH lines and was named SG population. In present study, the same population was tested in additional seven environments at three locations (Hanghzou, Xi'an and Wuhan) in China. Phenotypic data over 11 environments (four were previously analyzed) were collected and used for QTL analysis by the Composite Interval Mapping program (CIM) of WinQTL cartographer 2.5 (Zeng, 1994).

Marker analysis and map update: Markers designated ZAAS, ZAASA1, ZAASA7 and miRNA were newly developed in our lab. *Brassica* EST, GSS and BAC sequences from *Brassica* Genome Gateway (http://brassica.bbsrc.ac.uk/BrassicaDB) and NCBI (http://www.ncbi.nlm.nih.gov) were randomly generated and assembled into consensus sequences. They were then aligned to the Arabidopsis genome ATH1v5. The homologous contigs were selected to design the primer pairs using the software PRIMER 3 and verifying them both in *Brassica* and Arabidopsis genomes by e-PCR. In

addition, 235 proteins/genes with known or suspected function in Arabidopsis acyl lipid metabolism pathways were selected either from the Arabidopsis Lipid Gene Database (http://www.plantbiology.msu.edu/) or kindly provided by Oil Crops Institute, CAAS. Sequences from *Brassica* database were aligned to these genes then primer pairs were designed based on mRNA, full gene or EST when homologous contigs were available.

A total of 1690 molecular markers including 1269 newly developed informative, 101 published EST-derived markers, 25 functional markers provided by Oil crops Institute, CAAS and 5 obtained from Huazhong Agricultural University, 310 public SSRs, 8 combinations of SRAP primers and 2 SCAR markers were used in present research. All primer pairs that showed polymorphisms in a screening with the two parents were applied to the population of 282 SG-lines.

Genome alignment between B. napus and A. thaliana: 280 linked markers with known sequence information were aligned with Arabidopsis genome sequences using the NCBI BLASTn program. The homologous loci of Arabidopsis ($E < 10^{-7}$) were located on a physical map of Arabidopsis using the Seqviewer program from TAIR database (http://www.arabidopsis.org/). Regions having neighboring loci that were conservatively colinear with Arabidopsis were regarded as a homologous synteny region.

Candidate gene nomination: QTL-candidate gene co-location was then carried out to search for candidate genes. When physical positions of lipid related genes fell into the confidence interval of the QTL for oil content, then co-localization of QTL and candidate genes were assumed (Keurentjes *et al.*, 2008).

Results and discussion

Functional gene Map: Using a minimum LOD score of 3.0, a total of 374 new markers were integrated into the framework of the previously published SG map (Zhao *et al.*, 2005). The updated genetic map is composed of 477 markers including 103 previously mapped SSRs and spans 1956.7 cM of the total genome with an average marker interval of 4.10 cM (Figure 1). By blastn analysis, 280 markers were shown to have homologous regions in Arabidopsis (E value $< 10^{-7}$, corresponding to 254 genes), and these homologous loci covered nearly the whole genome of Arabidopsis. The most conspicuous characteristic of the present map is that nearly half of the functional marker loci 48.6% (136 of 280) correspond to 111 orthologous Arabidopsis genes involved in the acyl lipid metabolism pathway, which comprise obvious candidate genes for seed oil content in rapeseed. Of them, 18 and 17 correspond to genes for synthesis of fatty acids in plastids and synthesis of membrance lipids in endomembrane system, respectively, which accounted for 36.0% (18 out of 50) and 28.3% (17 out of 60) of total candidate genes in these two important pathways reported. Besides, the mapped homologous gene loci included 7 of 26 genes for synthesis of membrance lipids in plastids, 7 of 38 for metabolism of acyl-lipids in mitochondria, 6 of 20 genes for synthesis and storage of oil, 21 of 164 for lipid signaling, as well as 23 representing the enzymes involved in the lipid modification, processing or degradation processes. For another 12 orthologous genes, their function was uncertain and they were classified as miscellaneous. It provides available information to reveal the genetic architecture of seed oil content in rapeseed and became a useful tool for identifying important candidates in the overlapped genomic regions between oil content QTL and orthologous lipid related candidate genes.

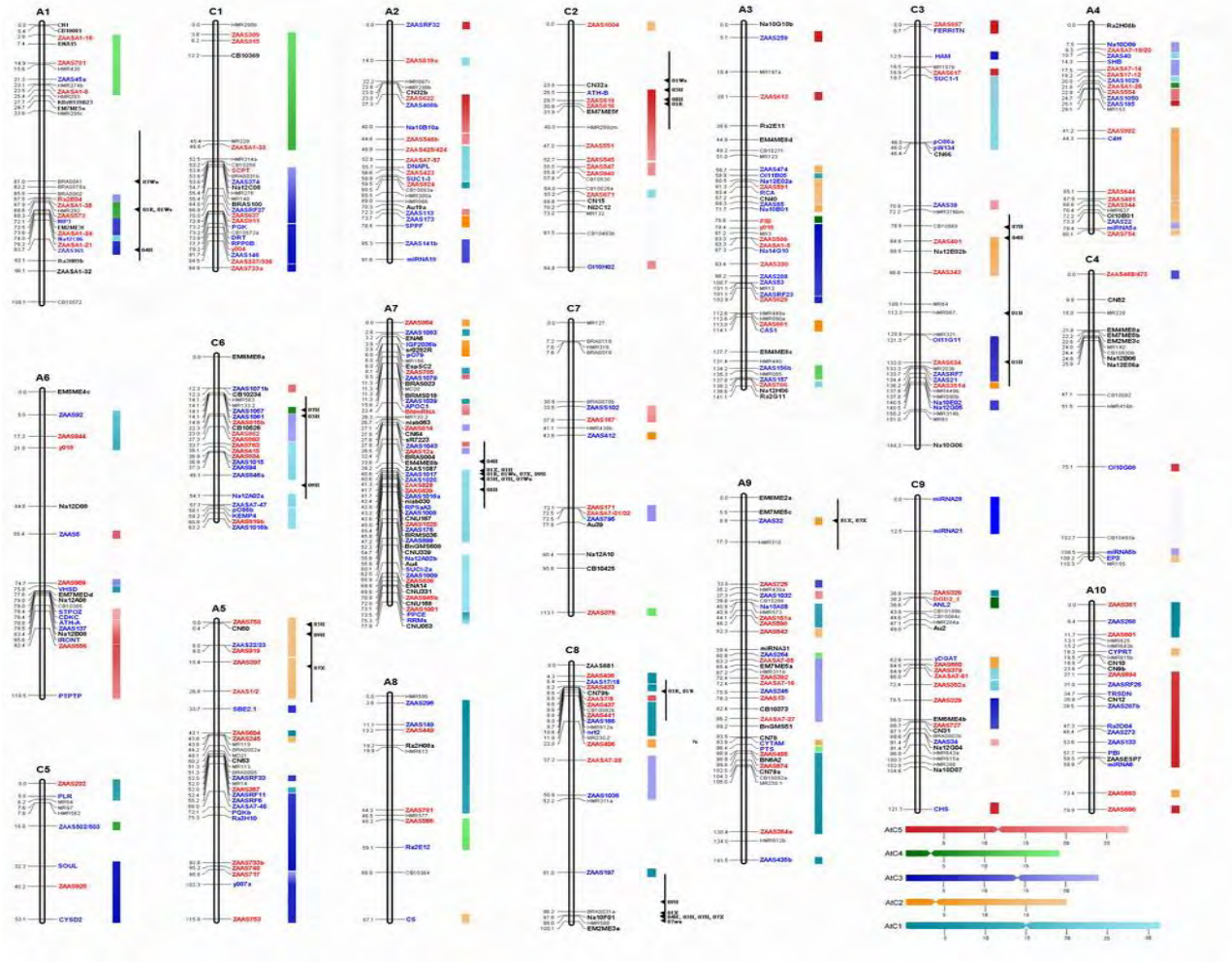


Figure 1. Linkage map of *B. napus* and comparative map with *A. thaliana*. The 19 linkage groups are represented by vertical bars. The number at the top of each linkage group, A1–A10 and C1-C9, indicates the internationally agreed chromosomal nomenclature of *B. napus*. The locus names and genetic distances are respectively listed on the right and left side of the linkage groups. Markers that are homologs to Arabidopsis lipid related genes are highlighted in bold and red; markers corresponding to other Arabidopsis genes are presented in bold and blue; the rest of newly added markers are given by bold and black; and the markers in black are originated from our previous map. The synteny regions with Arabidopsis are shown to the right of the *B. napus* linkage groups as colored vertical bars, which represent different chromosomes of Arabidopsis. Vertical lines lie on the right side of the synteny regions represent the QTL detected in this study, the different characters represent the location of experiments (see table 1).

Comparative map of B. napus and A. thaliana: Comparative genomic analysis between *B. napus* and Arabidopsis revealed multiple genome duplications/rearrangements and identified 32 large syntenic regions between these two species, each conserved chromosomal segment corresponding to Arabidopsis represented two to six times within the *B. napus* genome. In addition, our study further corroborated the previous finding of the close homology between A genome of A1 ~ A10 and corresponding C genome of C1 ~ C9 in *B. napus*. For example, the linkage groups A1 and C1 show entire-length homology, the upper half of them are homologous to the block segment (about 8.5 ~ 17.5 Mb) from AtC4, and the lower half are homologous to the top arm of AtC3. Simultaneously, the segment from AtC3 strongly conserved in A1/C1 was also co-linear with the lower half of A3/C3 and A5/C5. These results agree well with that from Parkin *et al.* (2005) and will further help to reveal the genome structure of *B. napus*.

Analysis of QTL and candidate genes for seed oil content: A detail QTL analysis for oil content based on new map version and the phenotypic survey over 11 diverse environments. In total, nine significant QTL were detected (Table 1), seven of them were located on the same linkage groups (A1,

A7, A9, C2, C3, C6 and C8) and similar genomic regions as previously reported (Zhao *et al.*, 2005). Co-location analysis revealed that a total of 28 acyl lipid metabolism related candidate gene loci were linked to seven of the nine QTL, which homology to Arabidopsis at an E-value from $2E^{-25}$ to 0 (Table 2). Notably, some of them are well-known lipid related candidate genes in Arabidopsis and happened to be located nearby or under the peak positions of major QTL region. For example, ZAAS573 and ZAASA1-24 associated with *OilA1*, ZAAS828, ZAAS839 and ZAAS1025 linked to *OilA7*, and ZAAS619 and ZAAS616 just under the peak of *OilC2*, respectively, and some other important candidates assigned with *OilC8-1*, *OilC3* and *OilC6*, indicating their potential role in genetic control on seed oil content in *B. napus*.

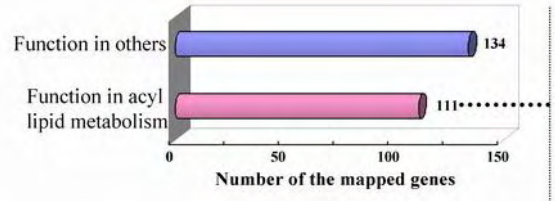
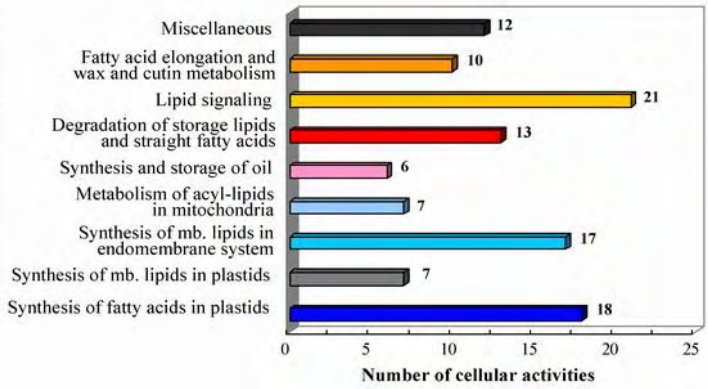


Figure 2. An overview of the homology genes mapped in present study.



Functional classification of lipid-related genes based

Table 1. QTL analyses of seed oil content in B. napus under 11 environments

Name of QTL[a]	Linkage group	LOD value	Additive effect(%)[b]	R^2%[c]	Confidential interval	Environment[d]
OilA1	A1	3.04 ~ 5.25	0.31 ~ 0.55	3.64 ~ 6.40	40.3 ~ 92.1	01R; 01We; 04H; 07Wu
OilA5	A5	2.83 ~ 4.64	-0.35 ~ -0.45	2.85 ~ 5.40	0.0 ~ 30.0	05H; 07X; 09H
OilA7	A7	3.09 ~ 20.54	-0.41 ~ -0.91	3.59 ~ 24.7	29.3 ~ 53.9	01X; 01H; 01R; 01We; 04H; 05H; 07H; 07X; 07Wu; 08H; 09H
OilA9	A9	4.65 ~ 5.09	0.39 ~ 0.40	5.97 ~ 6.12	0.3 ~ 20.5	01X; 07X
OilC2	C2	3.89 ~ 7.87	0.60 ~ 0.93	5.47 ~ 9.31	10.1 ~ 40.7	01R; 01We; 05H; 08H
OilC3	C3	3.29 ~ 6.33	-0.37 ~ -0.74	3.75 ~ 8.25	73.7 ~ 141.7	01H; 04H; 05H; 07H
OilC6	C6	2.54 ~ 3.36	-0.32 ~ -0.36	2.56 ~ 4.24	15.2 ~ 58.2	05H; 07H; 09H
OilC8-1	C8	4.14 ~ 5.18	0.47 ~ 0.61	5.01 ~ 6.00	6.0 ~ 21.9	01R; 01We
OilC8-2	C8	2.84 ~ 11.31	-0.27 ~ -0.73	3.35 ~ 14.0	81.3 ~ 100.1	01X; 01R; 01We; 04H; 05H; 07H; 07X; 07Wu; 09H

a, QTL names are abbreviations of the trait followed by its respective linkage group's number. A number 1 or 2 was added if more than one QTL were detected in one linkage group. b, additive effect is the phenotypic effect due to the substitution of a Gaoyou allele by an allele of Sollux.. c, proportion of the phenotypic variation explained by the QTL. d. the different characters represent the location of experiments, R-Reinshof, We- Weende, H-Hangzhou, X-Xi'an, Wu-Wuhan; the two numerals in front represent the year of the flowering season.

References

Keurentjes JJB, Sulpice R, Gibon Y, Steinhauser MC, Fu J, Koornneef M, Stitt M and Vreugdenhil D. 2008. Integrative analyses of genetic variation in enzyme activities of primary carbohydrate metabolism reveal distinct modes of regulation in Arabidopsis thaliana. Genome Biol. 9: R129.

Nesi N, Delourme R, Brégeon M, Falentin C and Renard M. 2008. Genetic and molecular approaches to improve nutritional value of Brassica napus L. seed. C. R. Biologies 331: 763-771.

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

Zeng ZB. 1994. Precision mapping of quantitative trait loci. Genetics 136: 1457-1468.

Zhao JY, Becker HC, Zhang DQ, Zhang YF and Ecke W. 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.