

RESEARCH ARTICLE

Open Access



# Synteny analysis of genes and distribution of loci controlling oil content and fatty acid profile based on QTL alignment map in *Brassica napus*

Nadia Raboanatahiry<sup>1,2†</sup>, Hongbo Chao<sup>1,2†</sup>, Liangxing Guo<sup>1,2</sup>, Jianping Gan<sup>2</sup>, Jun Xiang<sup>2</sup>, Mingli Yan<sup>3\*</sup>, Libin Zhang<sup>1</sup>, Longjiang Yu<sup>1</sup> and Maoteng Li<sup>1,2\*</sup>

## Abstract

**Background:** Deciphering the genetic architecture of a species is a good way to understand its evolutionary history, but also to tailor its profile for breeding elite cultivars with desirable traits. Aligning QTLs from diverse population in one map and utilizing it for comparison, but also as a basis for multiple analyses assure a stronger evidence to understand the genetic system related to a given phenotype.

**Results:** In this study, 439 genes involved in fatty acid (FA) and triacylglycerol (TAG) biosyntheses were identified in *Brassica napus*. *B. napus* genome showed mixed gene loss and insertion compared to *B. rapa* and *B. oleracea*, and C genome had more inserted genes. Identified QTLs for oil (OC-QTLs) and fatty acids (FA-QTLs) from nine reported populations were projected on the physical map of the reference genome “*Darmor-bzh*” to generate a map. Thus, 335 FA-QTLs and OC-QTLs could be highlighted and 82 QTLs were overlapping. Chromosome C3 contained 22 overlapping QTLs with all trait studied except for C18:3. In total, 218 candidate genes which were potentially involved in FA and TAG were identified in 162 QTLs confidence intervals and some of them might affect many traits. Also, 76 among these candidate genes were found inside 57 overlapping QTLs, and candidate genes for oil content were in majority (61/76 genes). Then, sixteen genes were found in overlapping QTLs involving three populations, and the remaining 60 genes were found in overlapping QTLs of two populations. Interaction network and pathway analysis of these candidate genes indicated ten genes that might have strong influence over the other genes that control fatty acids and oil formation.

**Conclusion:** The present results provided new information for genetic basis of FA and TAG formation in *B. napus*. A map including QTLs from numerous populations was built, which could serve as reference to study the genome profile of *B. napus*, and new potential genes emerged which might affect seed oil. New useful tracks were showed for the selection of population or/and selection of interesting genes for breeding improvement purpose.

**Keywords:** *Brassica napus*, QTL alignment map, Seed oil alteration, QTL, Candidate genes, Synteny, Gene network interaction, Metabolism pathway

\* Correspondence: mlyan@hnust.edu.cn; limaoteng426@mail.hust.edu.cn

†Equal contributors

<sup>3</sup>School of Life Science, Hunan University of Science and Technology, Xiangtan 411201, China

<sup>1</sup>Department of Biotechnology, College of Life Science and Technology, Key Laboratory of Molecular Biology, Huazhong University of Science and Technology, Wuhan 430074, China

Full list of author information is available at the end of the article



## Background

Dissection of the genetic architecture of a species is one of the best approach to understand its identity, evolution history and allow the understanding of the genetic network mechanism that run the entire organization [1–3]. Each gene has an important role within this organization, which might affect one phenotypic trait, or in case of pleiotropy, affect several unrelated traits [4, 5]. Hunting specific genes for agriculturally and economically valuable traits is needed. Nowadays, breeding of cultivar with high oil content and advantageous fatty acid profile has become a necessity, since the demand in oil has increased with the growing population [6, 7].

The rapeseed (*Brassica napus*) is a well-known source of vegetable oil and is the preferred oil crop for biodiesel production in Europe [8], the most important breeding goal is to increase the oil content, since 1 % increase of seed oil content is equivalent to 2.3–2.5% increase in seed yield in *B. napus* [9]. The allotetraploid *B. napus* ( $A^nA^nC^nC^n$ ,  $2n = 38$ ) was derived from hybridization of *B. rapa* ( $A^rA^r$ ,  $2n = 20$ ) and *B. oleracea* ( $C^oC^o$ ,  $2n = 18$ ). Long years of evolution and artificial selection have made the  $A^n$  and  $C^n$  genomes of *B. napus* somewhat different from the  $A^r$  genome of *B. rapa* and the  $C^o$  genome of *B. oleracea* [10]. The *Brassica* genera is closely related to the model plant *Arabidopsis thaliana*, the divergence occurred approximately 14 to 20 million years ago [11, 12]. In the evolution history of *Brassicaceae* family, *Brassica* species underwent a whole genome triplication event compared to *A. thaliana*, which promoted their speciation, and that event was followed by genome duplication and rearrangement events [13–16]. *A. thaliana* genome could be subdivided into 24 blocks and 21 among them have been conserved in *B. napus* [17, 18]. Great opportunities are opened to undertake multiple important studies for further exploitation or improvement of this crop with the release of *B. napus* genome sequence in 2014 [19]. In our knowledge, the distribution of all genes involved in oil formation in *B. napus* was not reported before.

Seed oil is mostly composed of triacylglycerol (TAG), which represents 35% of seed weight in *A. thaliana* [20, 21]. The pathway of TAG biosynthesis leading to the oil formation has been elucidated in plant, and some key genes have been identified [21–25]. Li-Beisson et al. (2013) revealed that at least, 120 enzymatic reactions and more than 600 genes encoding the proteins and regulatory factors, were involved in acyl-lipid formation in *A. thaliana* [21]. Besides, combination of multiple genes, influenced by the environment has been demonstrated to control seed oil content trait [26, 27]. Also, many studies have shown that oil content and fatty acid profile influenced each other [28–30]. Understanding the genes architecture and network that control the variation of seed oil composition

allows a better insight to get the desired profile according to the final usage. For instance, high unsaturated fatty acid (UFA) oil has been recommended for food preparation, because it allows rapid cooking time and less oil absorption [31]. Otherwise, many genes have been cloned and showed their effectiveness on oil content improvement. For example, expression of rapeseed DGAT increased oil content in *A. thaliana* [32], and BnGPDH and BnGPAT increased oil content of ~4% in transgenic seeds [33].

Quantitative trait loci (QTL) analysis is a powerful tool for genetic investigation in order to identify loci responsible for the variation of phenotypic trait. The phenotypic traits are usually valuable traits in agriculture and economy. QTL is extensively used in plant breeding. Because QTLs correlate with variation of phenotype, the corresponding loci could be amplified and consequently are expected to improve the phenotype. Multiple oil related QTLs have been discovered in different populations of *B. napus* [28, 34–49], but the number and location of QTLs in different populations varied a lot, which needs their unification in one map for comparison. Establishing a unique map combining QTLs from multiple populations would be of a great utility, which could facilitate the comparison of these QTLs. Importantly, if QTLs of multiple populations with same environment overlaps in one region, they could be defined as fixed QTLs, otherwise, they would be population specified QTLs. However, building a consensus map is challenging since the position of QTLs varies with population, and the markers used for the studies are different, so their comparison or their unification into one consensus map is rather difficult [50]. Liu et al. (2016) built a map aligning QTLs for oil content QTLs (OC-QTLs) from six populations of *B. napus* by projection on the physical map of the reference genome “*Darmor-bzh*”. One-hundred and ten QTLs could be positioned and 53 among them were overlapping [51]. Besides, the determination of QTLs coupled with the identification of associated candidate genes would permit the comprehension of these genes authority over traits [52, 53].

The purposes of the present study were as follows: (1) identify the genes involved in FA and TAG biosyntheses in *B. rapa*, *B. oleracea* and *B. napus*, and analyze the genes synteny; (2) compare the FA-QTLs and OC-QTLs from different genetic mapping populations of *B. napus* by construction of a map aligning QTLs; (3) identify the relative candidate genes in the confidential intervals of QTLs and analyze their interaction and metabolism pathway. Our study revealed a map of FA-QTLs and OC-QTLs, with diverse populations, showing fixed and specified QTLs, and additionally highlighted new potential genes that might affect seed oil content and FA composition.

## Results

### Gene synteny analysis revealed higher gene copy number in *B. napus*

In total, 439 genes related to FA and TAG biosyntheses were identified in the genome of *B. napus*, they were homologous to 110, 224 and 173 genes from *A. thaliana*, *B. rapa* and *B. oleracea*, respectively. In *B. napus*, A<sup>n</sup> and C<sup>n</sup> genomes contained 220 and 219 genes, respectively (Table 1; Additional file 1: Table S1). The genes synteny in *B. rapa*, *B. oleracea* and *B. napus* are illustrated on Fig. 1, it was found that genes were mostly located on A3 and C3 chromosomes. Obviously, the number of inherited genes greatly increased in *B. napus*, and some of them were not maintained in the same chromosome location as their parents. While observing the synteny between them, it was revealed that the genes could be lost or inserted on the genome. In fact, 27 genes were lost in total, 17 of them were lost in A<sup>n</sup> chromosomes and 10 were lost in C<sup>n</sup> chromosomes. This is the case of KASII genes Bra014202 and Bra014203 on A<sup>n</sup>8, and Bol012577 on C<sup>n</sup>6 and Bol042053 on C<sup>n</sup>7, which did not have descendant genes in *B. napus*. They thus had sequence similarity with the other KASII genes found in *B. napus*. However, seven lost genes were replaced in the other genome, i.e. genes lost in A<sup>n</sup> were replaced in C<sup>n</sup>, and reversely; so, five genes on A<sup>n</sup> genome were replaced on C<sup>n</sup> genome, and two genes on C<sup>n</sup> genome were replaced on A<sup>n</sup> genome. For instance, *B. rapa* DHLAT Bra006486 on A<sup>n</sup>3 should have been transmitted on A<sup>n</sup>3, but was replaced on C<sup>n</sup>8 in *B. napus* (BnaC08g03220D). Besides, 69 genes were inserted, of which 14 were found on A<sup>n</sup> chromosomes, and 55 were found on C<sup>n</sup> chromosomes. For example, PLA2 Bra039011 (A<sup>n</sup>7) in *B. rapa* had homologous BnaA07g01090D (A<sup>n</sup>7) and BnaC07g01540D on C<sup>n</sup>7 in *B.*

**Table 1** Distribution of genes involved in FA and oil biosynthesis in *Brassica*

Chromosomes	<i>B. rapa</i>	<i>B. napus</i>	Chromosomes	<i>B. oleracea</i>	<i>B. napus</i>
A01	21	21	C01	22	19
A02	13	15	C02	10	14
A03	38	34	C03	31	35
A04	17	11	C04	23	23
A05	31	28	C05	19	24
A06	15	17	C06	16	18
A07	21	23	C07	22	23
A08	20	16	C08	12	13
A09	35	27	C09	15	21
A10	10	12			
Scaffold	3	0	Scaffold	3	0
An	0	16	Cn	0	29
Total	224	220	Total	173	219

*napus*, which accounted for one gene insertion on C<sup>n</sup> genome.

Counting the number of lost and inserted genes, 42 additional genes were found in *B. napus* compared to *B. rapa* and *B. oleracea*. They originated from gene duplication or triplication which mostly appear in other chromosomes. For example,  $\beta$ -PDH gene Bra032361 on A<sup>n</sup>9 in *B. rapa*, had two descendant genes in *B. napus*, which were BnaA08g17570D on A<sup>n</sup>8 and BnaA09g26420D on A<sup>n</sup>9. Also, three copies of PP genes were found on C<sup>n</sup> chromosomes, BnaC01g22010D on C<sup>n</sup>1, BnaC07g31970D on C<sup>n</sup>7 and BnaCnng22260D on an unknown C<sup>n</sup> chromosome, and they were homologues of *B. rapa* PP gene Bra006653 on A<sup>n</sup>3. Note that *B. napus* PDAT gene BnaC03g53840D on C<sup>n</sup>3 had no parental homologue, this gene was homologue of *A. thaliana* PDAT gene At3g44830. These results showed higher amount of gene copy in *Brassica* compared to *A. thaliana*, and also confirmed the fact that *B. napus* genome showed mixed of gene loss and insertion compared to *B. rapa* and *B. oleracea* genomes. Moreover, distribution on chromosomes of all genes involved in oil formation in *B. napus* were highlighted in our findings.

### Identified overlapping QTLs on the reference genome

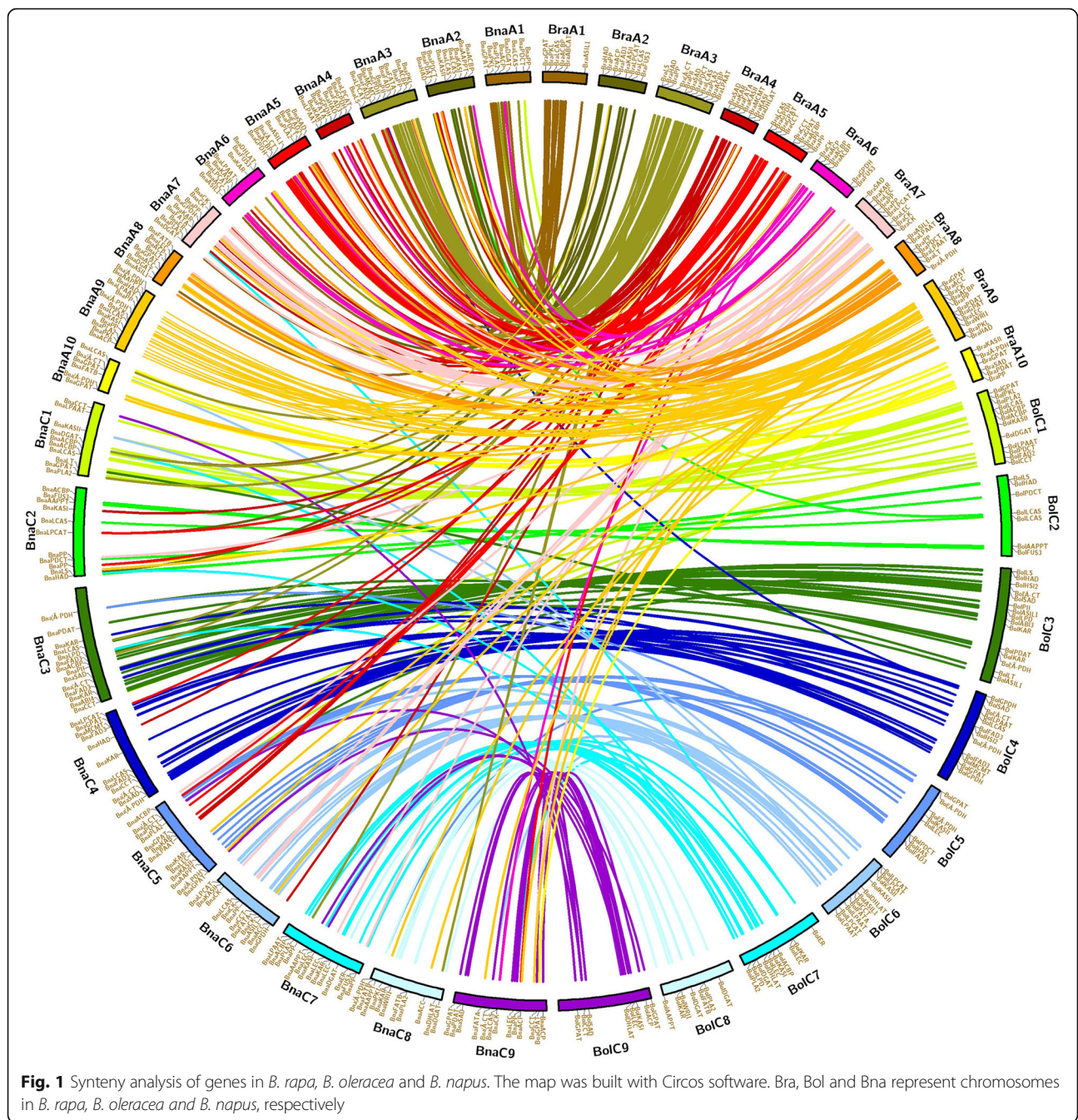
#### "Darmor-bzh"

By using of E-PCR, 217 molecular markers could be settled on the physical map of "Darmor-bzh" emphasizing 335 FA-QTLs and OC-QTLs from *DY*, *KN*, *M201*  $\times$  *M202*, *PT*, *RNSL*, *SG*, *SO*, *TN* and *Z5* populations. The detailed information about these QTLs with their respective physical location are presented on Additional file 2: Table S2. The QTL alignment map related to these QTLs is illustrated on Fig. 2 and the proportion of overlapping QTLs per chromosome are represented on Fig. 3. It was observed that 82 overlapping QTLs were found widespread in all chromosomes unless A4, A6, C1, C4 and C7 (Fig. 3, Additional file 3: Table S3). Obviously, 64 overlapping QTLs were from two populations (e.g. *qOC-A2-1-KN1* and *qOC-A2-2-TN* were overlapping on A2), 15 overlapping QTLs were from three populations (e.g. *qOC-A8-1-TN*, *qOC-A8-RNSL* and *qOC-A8-3-KN* were overlapping on A8), 3 overlapping QTLs from four populations were also observed (e.g. *qOC-A3-Z5*, *qOC-A3-DY*, *qOC-A3-3-TN* and *qOC-A3-5-KN* were overlapping on A3).

OC-QTLs overlapped the most in this study (59/82 overlapping QTLs), followed by C18:3 (five overlapping QTLs), and then C18:1, C18:2 and C22:1 (three overlapping QTLs each), and finally the saturated fatty acids and UFA (two overlapping QTLs each). Chromosome C3 hosted the highest number of overlapping QTLs (22) with all trait studied except for C18:3.

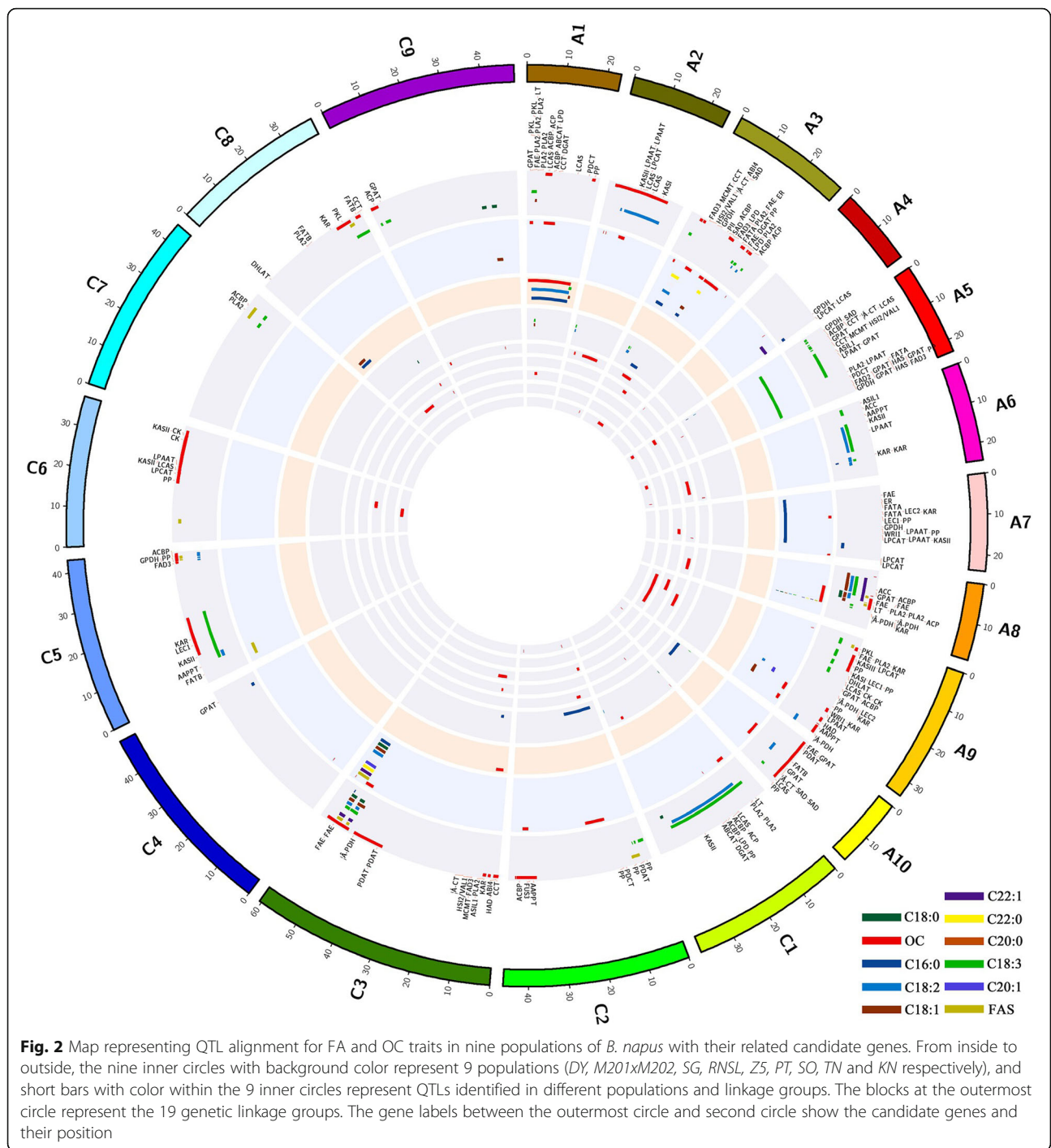
While observing the QTL positions between populations, QTLs from *TN* and *KN* populations often overlapped. *TN*





and KN populations had in common Wuhan and Shanxi, China as environments. Considering the overlapping QTLs according to their environments (*PT* was cultivated in Canada; *DY*, *RNSL* and *SO* were cultivated in Europe, *KN*, *TN*, *M201* × *M202* and *Z5* were cultivated in China, and *SG* was cultivated in Europe and China), it was unclear to delimit fixed QTL region according to these environments because overlapping QTLs were of populations with dissimilar environments, especially those with three and four populations. For example, in the region where *qOC-A1-SO*,

*qOC-A1-SG*, *qOC-A1-2-KN* and *qOC-A1-3-TN* overlapped on A1, these populations were both developed in Europe and China. However, regions could be observed in overlapping QTLs involving two populations. Thus, 43 among these 82 overlapping QTLs might be fixed QTLs for Chinese environments. No fixed QTLs for Europe were found, also because *SG* population was cultivated both in China and Europe. The remaining 39 overlapping QTLs were then of mixed population. The Canadian cultivated population *PT'* QTLs overlapping once with that of

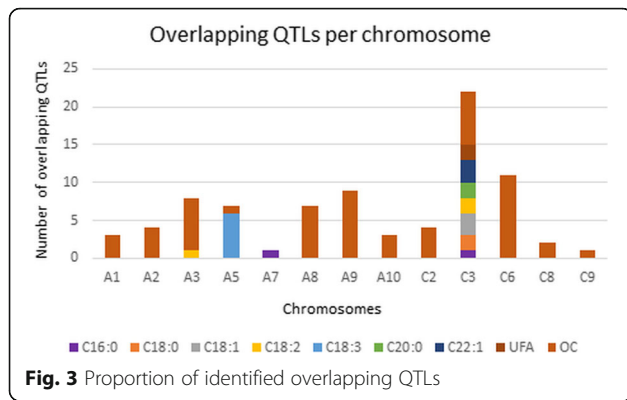


*TN* population, and appeared as the only co-localization with other populations' QTLs. The genetic architecture of this *PT* population might be very different from the others. These results confirmed that genotype and environment influenced QTLs, which in turn affected the detected overlapping QTLs of our map. Also, fixed QTL regions for particular environment could be identified with our approach (e.g. Chinese environment).

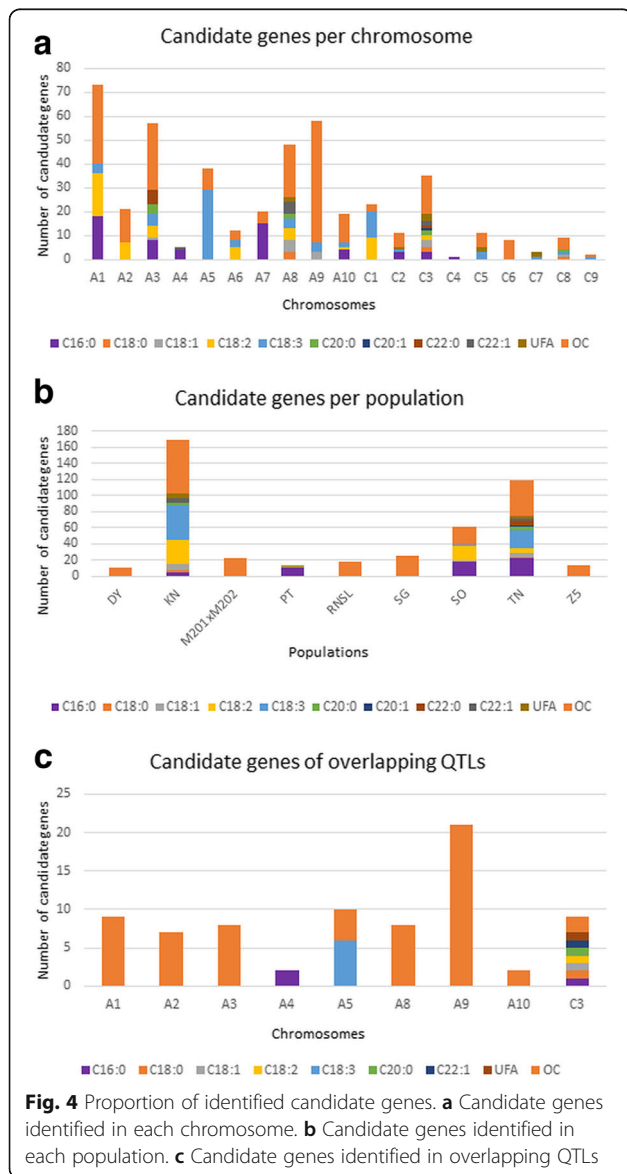
**Potential candidate genes identified in QTLs regions**

A total of 218 among the 439 genes which were mentioned above, were identified as candidate genes in 162 QTLs intervals (Additional file 4: Table S4). The proportion of candidate genes in each chromosome, in each population and in overlapping QTLs are illustrated on Fig. 4. Obviously, the highest amount of candidate genes detected was for oil content trait. Besides, it was discovered that some





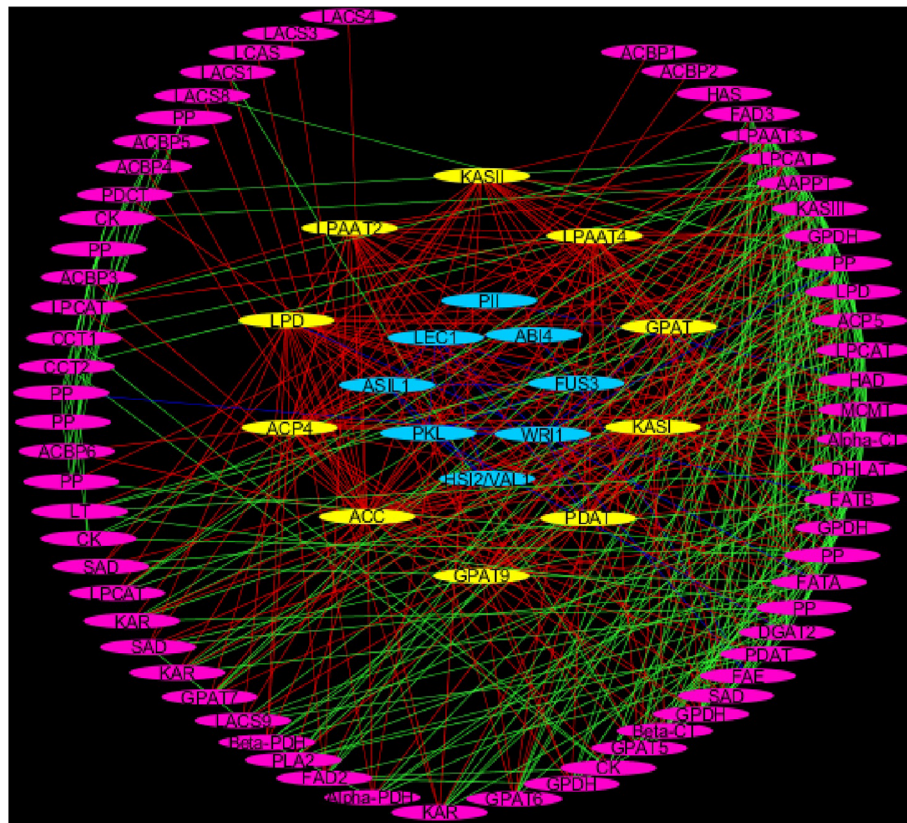
candidate genes could be found in many QTL intervals and also, they might affect more than one trait, for example, KASI gene BnaA02g24400D was located in QTL interval of C18:2-QTL (*qC18:2-A2-3-KN*) and OC-QTLs (*qOC-A2-4-KN* and *qOC-A2-1-Z5*). Considering the number of candidate genes for each trait in each chromosome, OC-QTLs had 148 candidate genes detected and 70 remaining candidate genes were for FA traits. These candidate genes were mainly detected in QTLs intervals of *TN* and *KN* populations. Otherwise, 76 among these 218 detected candidate genes were observed in 57 overlapping QTLs. They were in majority candidate genes for oil content (61/76 genes). Sixteen genes were found in overlapping QTLs involving three populations, and the remaining 60 genes were found in overlapping QTLs of two populations (Fig. 4).



The detected candidate genes varied with populations, some of these detected candidate genes might affect more than one trait and one trait might be affected by multiple genes. For instance, FAE gene BnaC03g65980D on C3 might affect all studied traits unless C18:3. Further analysis revealed that these 218 candidate genes belonged to 45 families. One trait could be affected by many gene families, as in OC traits, but the most frequent gene families for each trait were PLA2 (C16:0, C18:2, UFA and OC), FAE (C18:0, C18:1, C18:2, C20:0, C20:1, C22:0, C22:1 and UFA), ACBP and LCAS (C18:2), GPAT (C18:3), PP (C20:1), CT, FAD3, HIS2/VAL1, LPD and MCMT (C22:0), respectively. Obviously, FAE which are involved in FA elongation, emerging long chain fatty acid at the expense of C18:X, might affect all studied traits unless for C18:3. Our findings indicated candidate genes that possibly have influence over multiple traits.

**Candidate genes interaction network and metabolic pathway analyses**

In order to understand the interaction between candidate genes, we analyzed interaction network and pathway which involved them in oil formation and FA synthesis. The interaction analysis was made with STRING, and visualized with Cytoscape\_V3.2.1. Because *B. napus* and *B. oleracea* are still not available on STRING Database, we used orthologous genes in *A. thaliana* to perform the analysis (Additional file 4: Table S4). Thus, 91 genes from *A. thaliana* were used for this study. The results indicated 83 nodes and 413 edges (Fig. 5). It was surprising that the transcription factors (ABI4, ASIL1, FUS3, HSI2/VAL1, LEC1, PKL, PII, WRI1) interacted poorly with the rest of the genes (few edges connected them with the other genes). However, ten genes belonging to six families, which were ACC, ACP, GPAT, KAS, LPAAT and LPD, interacted the most with the other genes (DL ≥ 20). They might have more influence over the other genes. These ten genes were found within 57 QTLs intervals and might affect C16:0, C18:1, C18:2, C18:3, C22:1 and OC (Additional file 4: Table S4).



**Fig. 5** Candidate genes interaction network associated with fatty acid synthesis and oil formation. The analysis was made with orthologous *A. thaliana* genes by using STRING (<http://string-db.org/>) and visualized with Cytoscape\_V3.2.1. 83 nodes and 413 edges are shown. Purple nodes indicate genes involved in FA and TAG biosyntheses. Yellow nodes with their related red edges represent the genes that interact the most with the other genes ( $DL \geq 20$ ). Blue nodes with their related blue edges represent transcription factors

LPD, ACC, ACP and KAS are plastidial key enzymes which have important roles in FA biosynthesis. LPD is an enzyme that contributes to the transformation of pyruvate into acetyl-CoA by decarboxylation. ACC is an enzyme that catalyzes the carboxylation of acetyl-CoA to produce malonyl-CoA. ACP conveys the growing FA chain between enzymatic domains of fatty acid synthase (FAS) during biosynthesis. KAS are enzymes involved in FA elongation. GPAT and LPAAT are key enzymes that work in endoplasmic reticulum. GPAT catalyzes the conversion of G3P to LPA. LPAAT is an enzyme that catalyze the acylation of LPA into PA. FAE gene family which was previously found in QTL interval of all traits unless C18:3, seemed not have high interaction with the other genes. LPD, ACC, ACP and KAS were influential candidate genes for the above-mentioned traits which indicated that these traits were affected at earlier stage of FA biosynthesis. However, GPAT and LPAAT genes were also highly connected to the other genes, which indicated that traits could be affected at multiple level of the oil biosynthesis.

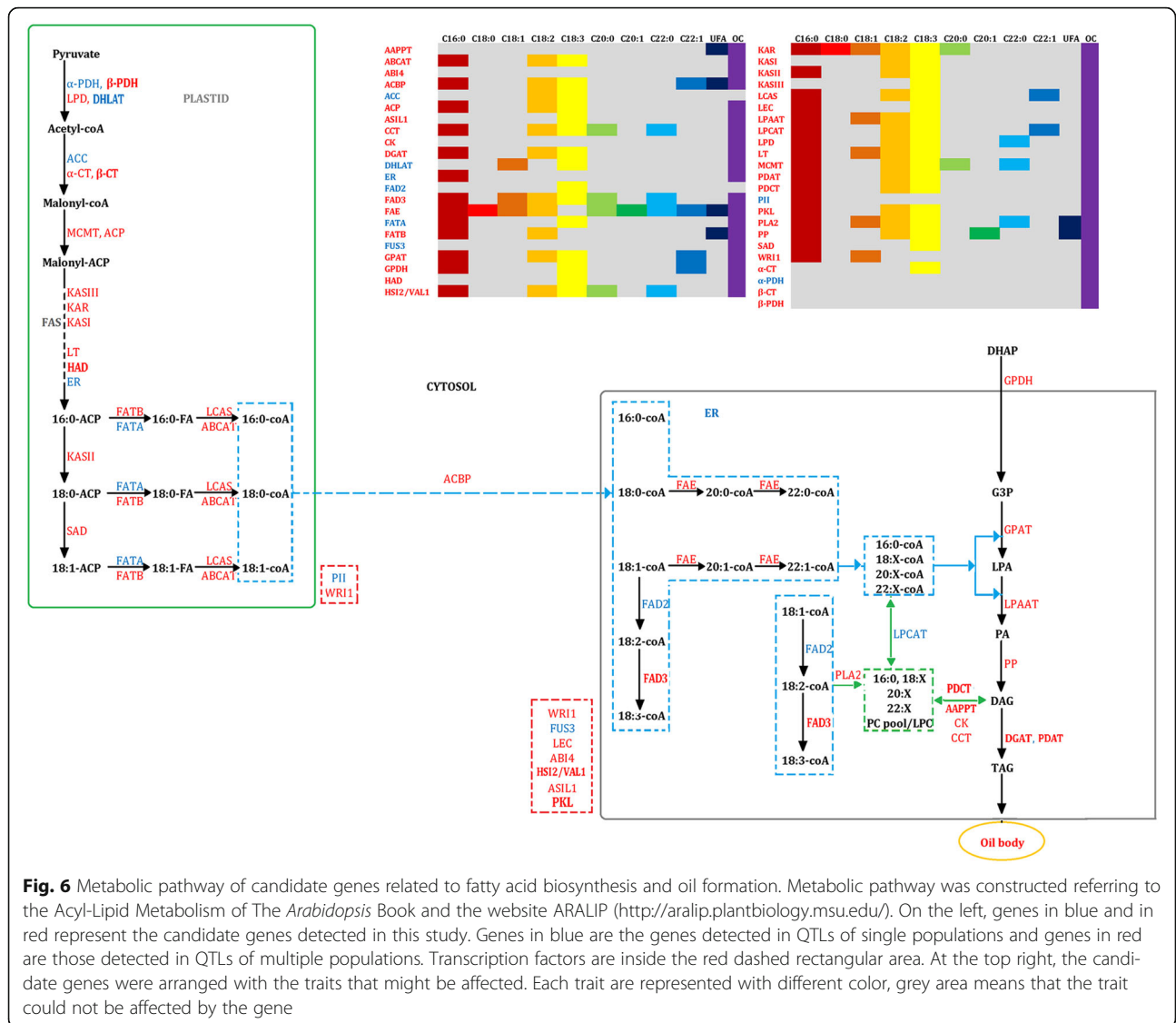
The metabolic pathway for fatty acid and TAG biosyntheses is shown on Figs. 6, 37 gene families were observed

in QTLs from multiple populations, for example, ACBP genes were seen in OC-QTLs of all nine populations. The remaining eight gene families were observed in QTLs in single population, for example, ACC genes were observed only in C18:2, C18:3 and C22:1 QTLs of KN population. All genes could affect oil content unless ACC and FAD2 genes. Most of genes could affect C16:0, C18:2 and C18:3, with 29, 24 and 32 gene families involved, respectively (Fig. 6 and Additional file 4: Table S4). Earlier, we found that genes of KASI, KASII, LPAAT, GPAT, PDAT, ACC, ACP interacted most with other genes, which might be dependent of them. These genes were found in QTLs of multiple populations unless PDAT and ACC. Moreover, they could affect oil content traits unless ACC genes. Our findings indicated that traits could be affected at multi-level of oil biosynthesis.

## Discussion

### The evolutionary history of *Brassicaceae* was reflected in the present study

While identifying the genes involved in FA biosynthesis and TAG formation, 439 genes were found in *B. napus*



and they were homologous to 110 genes of *A. thaliana*, 224 genes of *B. rapa* and 173 genes of *B. oleracea*. Multiple copies of genes belonging to unique family were perceived in both *A. thaliana* and *Brassica* species. The genes might belong to the same class, presenting a slight dissimilarity in amino acid sequence, or to different classes showing significant difference in structure (e.g. protein domain), but conserve the core structure that label them as a part of the family, like the ACBP genes [54].

It was discovered that *Brassica* species have experienced a whole genome triplication (WGT) event that occurred possibly while diverging from *Arabidopsis* [14, 16]. Thus, *Brassica* genome should have three times the size of *Arabidopsis* genome, and the number of genes and chromosomes as well should be three times higher than those of *Arabidopsis*. WGT was followed by chromosome restructuring, translocation, fusion, or recombination of

precursor chromosomes that led to the reduction of chromosome number in *Brassica* [13, 15, 16]. Consequently, pressure of this rearrangement resulted in some disorder within the genome: gene might be altered (mutation), lost (unfound), converted (e.g. location change from A<sup>n</sup> to C<sup>n</sup> in *B. napus*), chromosomal location might be dissimilar with parents (e.g. parental gene on A<sup>9</sup>2 and descendant gene on A<sup>9</sup>9) [19, 55]. Besides, duplication event also occurred, which is usually the primary explanation of new genes emergence [56, 57]. Duplicated genes resulted in function divergence and affected the evolution [58, 59]. Otherwise, hybridization followed by genome duplication in *Brassica* engendered new species like *B. napus* which is a tetraploid of *B. rapa* and *B. oleracea* [16, 60, 61]. This is called “polyploidization”. The higher amount of gene copies observed in *B. napus* was normal as the consequence of polyploidization. During this polyploidization,



chromosome rearrangement occurred in *B. napus* resulting in loss of genes compared to *B. rapa* and *B. oleracea* [19]. It was affirmed that loss of genes on A<sup>n</sup> genome might be replaced by homologous genes on C<sup>n</sup> genome, and reversely [19], but this involved seven genes only in the present study. Also, it was intriguing that more gene insertions were found in C<sup>n</sup> genome of *B. napus*. Besides, the reference genome “Darmor-bzh” which was used for the sequencing of *B. napus* was not derived from hybridization of the reference genomes used for the sequencing of *B. rapa* (Chiifu-401) [62] and *B. oleracea* (Capitata) [63], so this might explain also the blurred reason of genes distribution in *B. napus* of the present study. Additionally, *B. rapa* and *B. oleracea* are vegetables and *B. napus* is an oil crop, thus, more copy number of oil related genes were normally found in *B. napus*. Furthermore, as mentioned earlier, long years of evolution and artificial selection have resulted in difference between the parental A<sup>r</sup> genome of *B. rapa*, C<sup>o</sup> genome of *B. oleracea* and the descendant A<sup>n</sup> and C<sup>n</sup> genomes of *B. napus* [10]. Otherwise, in the present study, more gene copies were found in *B. napus*, and paralogous genes might be sub-functionalized or neofunctionalized, as these phenomena are commonly seen in newly emerged genes [64, 65].

#### Environments influenced also the QTL alignment map which revealed fixed QTLs for particular environment

In the current study, a QTL alignment map was generated according to the related molecular markers and they were aligned on the physical map of *Darmor-bzh*, QTLs detected in diverse populations had dissimilar location on the physical map of *Darmor-bzh*. It is not new that genotype and environmental characteristics influence QTLs. Zhu and Zhao (2007) emphasized that factors influencing the number of detected QTLs in one given population depends on the genetic variation between the two parents, the type

(DH or RIL) and size of the population, and the number of environments used [53]. The populations used in our study were derived from different parental lines, we both used DH and RIL, but the size of population was different and the number of environments as well (Table 2), thus, this explained the dissimilarity in QTLs detected in each population, which in consequence affected our QTLs alignment map. More genetic variation involved in the trait is one of the bases that more relative QTLs can be identified. Also, larger number of lines studied could allow higher detection of the genetic loci related to the trait variation [66]. In this study, FA-QTLs and OC-QTLs from nine populations of *B. napus*, with dissimilar environments and genotype background were used. First, looking at the parental lines of populations used in this study (Table 2), it is obvious that they have no direct relationship. In our study, we compared QTLs from these nine different populations, QTLs were for the same traits, but obtained in totally different genetic backgrounds (with no parent in common). It is a very good approach for detecting hot spot genomic regions associated with the traits, but not always accurate for validating exact QTL. These nine populations were produced from hybridization of different varieties of *B. napus*: *DY* is a hybrid of European and Korean [37], *SG* and *TN* are hybrids of European and Chinese [35, 67], *Z5*, *KN* and *M201XM202* are Chinese [42, 48, 68], *SO* is European [69] and *PT* is Canadian [70]. Genesis of variety within a species initiates in natural selection which enables species to settle into specific environmental pressures. Inherited variations occur in natural selection, individuals with suitable traits survive and reproduce better than the others, and the genetic information are inherited by their descendant. Pressures might come from biotic and/or abiotic factors [71, 72]. In selective breeding, artificial selection allows the production of new varieties according to the desired traits.

**Table 2** List of populations used to establish a map aligning QTLs for FA and OC traits

	Populations	Type	Environments	Lines	OC (%)	OC-QTLs	FA-QTLs	References
<i>Darmor-bzh</i> × <i>Yudal</i>	<i>DY</i>	SW	France	445 DH	38–54	14	–	[37]
<i>Rapid</i> × <i>NSL96/25</i>	<i>RNSL</i>	W	France, Germany	242 DH	35.7–50	10	–	[37]
<i>zy036</i> × <i>51,070</i>	<i>Z5</i>	S, W, SW	China	92 DH	30.85–51.30	12	–	[39]
<i>Sollux</i> × <i>Gaoyou</i>	<i>SG</i>	W, SW	Germany, China	282 DH	35.6–57	9	–	[40, 41]
<i>Tapidor</i> × <i>Ningyou7</i>	<i>TN</i>	S, W	China	404 DH	33.8–50.9	41	–	[44]
<i>Tapidor</i> × <i>Ningyou7</i>	<i>TN</i>	S, W, SW	China	202 DH	–	–	72	[45]
<i>KenC-8</i> × <i>N53-2</i>	<i>KN</i>	S, W, SW	China	348 DH	42–47.6	24	–	[42]
				300 DH	35–57.5	67	204	[43]
<i>M201XM202</i>	<i>M201XM202</i>	W	China	149 RIL	29.16–48.93	15	–	[48]
<i>Sansibar</i> × <i>Oase</i>	<i>SO</i>	W	Germany, Sweden	226 DH	41.2–48.6	5	16	[46]
<i>Polo</i> × <i>Topas</i>	<i>PT</i>	W	Canada	156 DH	31.5–55.5	14	131	[47]

Abbreviations: S (Spring), W (Winter), SW (Semi-winter), DH (Double haploid), RIL (Recombinant inbred lines)

Although varieties are phylogenetically divergent, overlapping QTLs could be detected, due also to the influence of environments.

Phenotypic variation of trait could be influenced by genetic and environmental factors [53], but overlapping QTLs underlined conserved regions on the genome which were responsible for trait variation. In the present results, populations which were cultivated in similar environments had more overlapping QTLs, such as *KN* and *TN* populations. Overlapping QTLs in these populations might be fixed QTLs for Chinese environment. The Canadian line also had few QTLs overlapping with the others, these QTLs were rather specific. However, the poor overlapping QTLs found in European cultivated populations were intriguing. Also, we found three overlapping QTLs in A1 and A3 regions, involving populations of dissimilar environments, these might correspond to enrichment region of associated gene variations. It is also probable that the populations share common ancestors but this needs to be verified.

In the current study, we found specified QTLs for independent populations. They could not overlap with other QTLs due probably to various factors, such as power of detection, density of genetic map, difference between parents, environments, type and size of population (Table 2). Thus, concerning these available populations, comprehensive assessment of the genetic background, or selection history of the parental lines that led to the roles of those specific QTLs, or in specific environment have not been made yet. Currently, more high density maps are being constructed and published, based on the same population. We expect in the future that the results obtained in our study would serve to assess the genetic background and evolution of rapeseed.

Otherwise, we compared our results with previously published consensus map and due to the difference of markers used, comparable results could be obtained only with those published by Wang et al. (2013). In fact, they built a consensus map based on common markers, for oil content QTL including one RIL and seven DH populations *GS/05*, *GS/12*, *DY*, *RNSL*, *Z5*, *TN* and *KN*. Six overlapping QTLs were detected: one on A1 chromosome (*KN-qOC-A1-1* overlapping with *TN-qOC-A1* and *DY-qOC-A1-2*); one on A2 (*KN-qOC-A2-3* overlapping with *DY-qOC-A2-2*), one on A8 (*KN-qOC-A8-1* overlapping with *TN-qOC-A8* and *RNSL-qOC-A8*); one on C3 (*KN-qOC-C3-2* overlapping with *RNSL-qOC-C3*); and two on C9 (*KN-qOC-C9-2* and *Z5-qOC-C9-1*, *KN-qOC-C9-3* and *Z5-qOC-C9-2*) [42]. By comparing markers interval, BRAS068 on chromosome C3 was aside of overlapping *KN-qOC-C3-2* and *RNSL-qOC-C3* of Wang et al. (2013) map and *qOC-C3-1-TN* and *qOC-C3-2-KN* of the present results. A fragile comparison with the consensus map made by Wang et al. (2013)

resulted in poor findings, due to the limitation imposed by the difference of markers. Other approaches such as using common markers might lead to more discovery.

In the present study, one locus on the C3 chromosome (53.75 Mb to 58.29 Mb) might affect six traits (C16:0, C18:0, C18:1, C18:2, C20:0, and C22:1). Any changes within this locus, harsh or beneficial, could affect these traits, it is interesting to tailor more than one desirable trait at the same time. Wang et al. (2014) investigated on genetic changes on current breeding of *B. napus*, and discovered that C genome (57.15 Mb) had extended breeding regions compared to A genome (16.80 Mb), but also C genome might have contributed to more valued alleles to produce elite traits [73]. This might explain the fact that more gene insertions were found on C genome in this study, but also the region on C3 which might affect multiple traits in our study. This region on C3 is then a favorable region to develop for rapeseed breeding. Further analysis could help into understanding of varietal characteristic of rapeseed, which is useful for the selection of population for breeding, for instance, use of the current results to compare with other populations not used in this study.

#### **New potential candidate genes were found, which might affect multiple traits**

Candidate gene investigation allows the identification of valuable genes associated to agriculturally and economically quantitative traits. Precise genetic architecture, distribution and interaction of loci that affect variation, permit the understanding of their effect on phenotypic variation [53]. In the present study, 162 QTLs underlined 218 candidate genes of *B. napus* which belonged to 45 families, they were homologous of 91 *A. thaliana* genes; and 76 among these candidate genes were found in 57 overlapping QTLs intervals. They were located in the QTL confidence intervals for C16:0, C18:0, C18:1, C18:2, C18:3, C20:0, C22:1, UFA, OC traits. In general, identification of candidate genes can be done via genome wide association studies (GWAS), linkage studies, expression studies but also it needs a prior understanding of the biological pathway. Lou et al. (2006) affirmed that combining the candidate genes with linkage studies could effectively enhance the accuracy [74].

In this study, we preselected the candidate genes by their correlation with the studied traits. The prior knowledge of the biological function of future candidate genes is important because many genes could be identified, especially if the study relies on position-dependent strategy, but the better are those which have functional consequences on the biological pathway or have close connection to the studied traits [53]. Thus, we both used function and position-dependent strategy to identify the potential candidate genes. The identified candidate genes were just putative causal genes, only experimental approaches

can validate the accuracy of these results. LT gene BnaA08g12720D was a candidate gene on A3, which might affect C18:1 and oil. The positive correlation between C18:1 and oil content has been demonstrated in multiple studies [29, 75, 76]. Oleic acid (C18:1) is an omega-9 fatty acid which composes naturally the animal and vegetable oils. Sales-Campos et al. (2012) reviewed the effect of this monounsaturated fatty acid on health, including its beneficial usage on wounds and inflammation, on regulation of blood pressure, on immunity system and on cancer healing process [77]. To take advantage of these, LT would be an ideal choice for altering C18:1 and oil at the same time, if it is proven experimentally to affect these traits. In *SO* map, Teh and Möllers (2016) identified FAD2 overlapping with QTLs for C18:1 and C18:3 on chromosome A1 by alignment with *B. rapa* [46]; however, the present study did not underline candidate genes in this QTL region. By comparing the present results with those published by Wang et al. (2015) (*TN* population only), divergence in findings was obvious. In fact, they identified 234 genes homologues of *A. thaliana* in 47 QTLs, involved in fatty acid metabolism [45]. In our study, 32 *A. thaliana* genes were found to be similar to those detected by Wang et al. (2015). Remaining 59 homologous genes were new candidate genes. Among these 32 similar candidate genes, 18 genes were potential genes for the same traits discovered in both our analysis and Wang et al. (2015) analysis (Additional file 5: Table S5). However, divergence in some *B. napus* genes were perceived, for instance, FATB-At1g08510 was also seen in our analysis, but homologue gene in *B. napus* (BnaA08g26890D) failed to be a candidate gene. Wang et al. (2015) also found FAD2, FAD3, FAE, LEC, FUS3, WRI1 and ABI genes as potential genes. Additionally, the genes ACC2, FAE1 and LPAAT were found in all FA-QTL intervals, while LEC1, LEC2, ACC2, and KASIII underlined C16:0. However, in our results, ACC2 were underlined by C18:2-QTL and C18:3-QTL only, FAE were involved in all traits unless C18:3 (but this was similar to the result of Wang et al. (2015), in which FAE fell into the CI of qC3-2 involving all traits unless C18:3). Besides, LPAAT were found in QTL interval of C16:0, C18:1, C18:2, C18:3 and OC, whereas LEC were inside QTL for C16:0, C18:3 and OC; and KASIII were found in the QTLs of C18:3 and OC traits. Otherwise, several of the candidate genes identified in our study have already demonstrated their effectiveness in enhancing oil content. As mentioned earlier, the abilities of DGAT, GPDH and GPAT to improve oil content have been demonstrated [32, 33], DGAT-BnaA08g03400D and GPAT-BnaA08g06960D both fell inside QTL for oil content in our analysis. In the present time, homologues in *B. napus* must be undergoing multiple analyses for similar or new functions discovery.

### Candidate genes could be affected at multilevel of oil formation

Focusing on network interaction and pathway analyses, genes belonging to KASI, KASII, LPAAT, GPAT, PDAT, ACC, ACP interacted most with other genes. They were found in multiple populations unless PDAT and ACC and could affect oil traits unless ACC. Similar analysis were made by Wang et al. (2015) in which direct or indirect effect of transcription factors over the genes were highlighted. In the present results, poor connection between transcription factors and candidate genes were found. As mentioned earlier, seed oil content is a trait controlled by a versatile genetic structure and also influenced by the environment [26, 27]. Association of these detected candidate genes, which obviously might depend on the ten above-mentioned key genes, underline structure that run the overall system. It has been affirmed that structure and dynamism of genetic regulatory networks influence quantitative traits [78] and genes are responsible for QTLs, affecting genetic variation of traits [79]. Because our analysis took in consideration nine populations at the same time, separated analysis of individual population might lead to different results. Also, since QTLs were dissimilar in the nine studied populations, it is possible that the genes involved in the system were different, which affected QTLs and traits in consequences. Previously, we found that some genes could affect one given trait in a population, and affect another trait in another population, this is the case of ACBP gene BnaA03g29000D which was found in QTL interval of C16:0 in *PT* population, but in oil content in *DY*, *TN* and *Z5* populations. Finally, since candidate genes interacted strongly with ten key genes of FA and TAG biosyntheses, traits were then affected at multilevel of oil formation.

### Advantages and limitations of our study

The current study aimed to identify overlapping QTLs from diverse populations of different environments background, and related potential candidate genes that might affect fatty acid profile and oil content in *B. napus*. We used function and position-based strategy to identify the candidate genes. This strategy allowed to detect QTL hotspots which maybe enrichment regions of gene variation involved in fatty acids and oil biosyntheses. This strategy also offered the advantage of eliminating genes inside QTL region which were not involved in fatty acid and oil biosyntheses. Besides, building the QTL alignment map allowed to make possible and easier the comparison of QTLs identified in diverse populations, which could be combined in one map, despite the difference of markers. Also, related candidate genes could be discovered. It was regrettable that some QTLs could not be settled on the map due to missing marker sequences. Additionally, the map helped us to verify stable QTLs, which could help us



to focus on valuable loci for fine mapping. Although stable QTLs were not confirmed yet, we could discover QTL enrichment regions which also gave us clues in genetic mechanism of close linkage of each trait or trait with trait, and discover easy variant area and conservative area. Finally, we analyzed the interaction network of candidate genes in order to understand their interaction, their influence on each other. We also built a metabolism pathway highlighting the discovered candidate genes, and the traits that they might affect. STRING and Cytoscape\_V3.2.1 offered a simple and easy way to analyze and visualize gene interaction, they are commonly used for such analysis actually, but the fact that *B. napus* and *B. oleracea* is still missing on STRING Database, so that orthologous genes in *A. thaliana* were used for the analysis, it is probable that results obtained were not accurate.

## Conclusion

In conclusion, the present study allowed to build a QTL alignment map with diverse populations which could serve as reference to study the genome profile of *B. napus*. New potential genes emerged which need experimental approach for authentication. We offered new useful tracks for the selection of population or/and selection of interesting genes for breeding improvement purpose. As perspectives, we suggest the development of functional markers based on our results. Also, since the candidate genes were detected by using of the reference genome “*Darmor-bzh*”, it would be better to test the accuracy of our results in other population.

## Methods

### Identification of *Brassica* genes involved in FA synthesis and TAG formation and gene synteny analysis

FA are synthesized in the plastid and TAG are formed in the ER. In the present study, we took in consideration the genes related to these biosyntheses for an afterward detection of potential candidate genes for oil improvement. Thus, *A. thaliana* genes related to these biosyntheses were acquired from the website ARALIP (<http://aralip.plantbiology.msu.edu/>) and TAIR ([www.arabidopsis.org](http://www.arabidopsis.org)) [80]. *Brassica* genes were identified based on homology to *A. thaliana* genes, with a score > 80, by using blastn (using *A. thaliana* nucleotides sequence) on *Brassica* Database (<http://brassicadb.org/>) [81] to get *B. rapa* and *B. oleracea* homologous genes; and browser (using *B. rapa* and *B. oleracea* gene names) on *Brassica napus* genome resource (<http://www.genoscope.cns.fr/brassicanapus/>) [19] to get *B. napus* homologous genes. The genes synteny was built with Circos software [82]: *B. napus* genes were linked to their homologous genes in *B. rapa* and *B. oleracea*.

### Identification of overlapping QTLs for FA and OC traits, and detection of potential candidate genes

OC-QTLs and FA-QTLs from nine previously reported populations were aligned into one map for comparison: *DY* (*Darmor-bzh* × *Yudal*) [37], *RNSL* (*Rapid* × *NSL96/25*) [37], *Z5* (*zy036* × *51,070*) [39], *SG* (*Sollux* × *Gaoyou*) [40, 41], *KN* (*KenC-8* × *N53-2*) [42, 43], *TN* (*Tapidor* × *Ningyou7*) [44, 45], *SO* (*Sansibar* × *Oase*) [46], *PT* (*Polo* × *Topas*) [47] and *M201* × *M202* [48]. The QTLs were projected onto the physical map of the reference genome “*Darmor-bzh*” and the position of related flanking markers were identified by using of E-PCR [83, 84]. First, the markers intervals were taken from an area less than 3 cM from the linkage map. Second, those markers’ primer sequences were acquired from related published papers. Then, by using E-PCR [83, 84], their positions on the physical map of *B. napus* “*Darmor-bzh*” could be deduced. Finally, these markers could be aligned on the physical map and the region inside two positioned markers was the QTL region. QTLs of which marker sequences were missing, or could not be placed on the corresponding chromosome were removed from this analysis. The studied traits were C16:0, C18:0, C18:1, C18:2, C18:3, C20:0, C20:1, C22:0, C22:1, combined unsaturated FA (UFA) and oil content (OC). The map was built with Circos software [82]. QTLs were renamed as “*q-trait-chromosome-population*” for uniformity, if many QTLs were detected on the same chromosome, the number of order was added to the name, e.g.: *qOC-A1-2-TN* referred to the second OC-QTL from *TN* population on chromosome A1. Overlapping QTLs were QTLs from two or more populations that were located in the same region, and potential candidate genes were genes located inside a QTL region.

### Gene interaction network and metabolism pathway analyses

In order to study the interaction between candidate genes, STRING was used (<http://string-db.org/>). STRING is a well-known database widely used to predict interactions (physical and functional) in proteins [85] which was then suitable for our study. *B. napus* and *B. oleracea* genes cannot be directly used for network analysis in STRING Database due to unavailability. Thus, interaction of candidate genes was studied by using of orthologous genes in *A. thaliana*. The orthologous genes of *A. thaliana* were submitted to STRING search using protein names and *A. thaliana* as organism. Then the resulting network was visualized with Cytoscape\_V3.2.1 [86]. The interaction was classified according to the degree layout (DL). More edges indicate more interaction with other genes. Then, potential metabolism pathway was manually constructed referring to the Acyl-Lipid Metabolism of The *Arabidopsis* Book and the website ARALIP (<http://aralip.plantbiology.msu.edu/>), gene families were placed according to their roles in oil formation [21], and corresponding gene names were summarized on Additional file 4: Table S4.

## Additional file

**Additional file 1: Table S1.** Genes involved in FA biosynthesis and TAG formation in *A. thaliana*, *B. rapa*, *B. oleracea* and *B. napus*. (XLSX 39 kb)

**Additional file 2: Table S2.** QTLs for FA and OC in nine populations, mapped to *B. napus* "Darmor-bzh" reference genome v4.1. (XLSX 33 kb)

**Additional file 3: Table S3.** Overlapping QTLs for FA and OC in nine populations, mapped to *B. napus* "Darmor-bzh" reference genome v4.1. (XLSX 21 kb)

**Additional file 4: Table S4.** Candidate genes detected in QTL intervals from nine populations, with their respective homolog in *A. thaliana*. (XLSX 31 kb)

**Additional file 5: Table S5.** Comparison with Wang et al. (2015) study, showing same *A. thaliana* orthologous genes of detected candidate genes in *B. napus*. (XLSX 10 kb)

## Abbreviations

AAPPT: Diacylglycerol Cholinephosphotransferase; ABCAT: ABC Acyl Transporter; ABI3: Homologous to the maize transcription factor Viviparous-1; ABI4: Abscisic Acid Insensitive (ABI) transcription factors; ACBP: Acyl-coA binding proteins; ACC: Acetyl-coA Carboxylase; ACP: Acyl Carrier Protein; ASIL1: Trihelix DNA Binding Family; CCT: Choline-Phosphate Cytidyltransferase; CK: Choline Kinase; DGAT: Acyl-CoA: Diacylglycerol Acyltransferase; DHLAT: Dihydrolipoamide Acetyltransferase; ER: Enoyl-ACP Reductase; FAD2: Oleate Desaturase; FAD3: Linoleate Desaturase; FAE: Fatty acid elongase; FATA: Acyl-ACP Thioesterase A; FATB: Acyl-ACP Thioesterase B; FUS3: Transcriptional factor; GPAT: Glycerol-3-Phosphate Acyltransferase; GPDH: NAD-dependent Glycerol-3-Phosphate Dehydrogenase; HAD: Hydroxyacyl-ACP Dehydrase; HAS: Holo-ACP Synthase; HSI2/VAL1: a member of a novel family of B3 domain proteins; HSL1/VAL2: a member of a novel family of B3 domain proteins; KAR: Ketoacyl-ACP Reductase; KAS1: Ketoacyl-ACP Synthase I; KASII: Ketoacyl-ACP Synthase II; KASIII: Ketoacyl-ACP Synthase III; LCAS: Long-Chain Acyl-CoA Synthetase; LEC: Transcription factor/activator; LPAAT: 1-Acylglycerol-3-Phosphate Acyltransferase; LPCAT: 1-Acylglycerol-3-Phosphocholine Acyltransferase, Lysophospholipid acyltransferase; LPD: Dihydrolipoamide Dehydrogenase; LS: Lipoate Synthase; LT: Lipoyltransferase; MCMT: Malonyl-CoA: ACP Malonyltransferase; PDAT: Phospholipid: Diacylglycerol Acyltransferase; PDCT: Phosphatidylcholine:diacylglycerol cholinephosphotransferase; PKL: a SWI / SWF nuclear-localized chromatin remodeling factor of the CHD3 group; PLA2: Phospholipase A2; PP: Phosphatidate Phosphatase; SAD: Stearoyl-ACP Desaturase; WRI1: WRINKLED1;  $\alpha$ -CT: Alpha-carboxyltransferase;  $\alpha$ -PDH: Alpha-Pyruvate Dehydrogenase;  $\beta$ -CT: Alpha-carboxyltransferase;  $\beta$ -PDH: Beta-Pyruvate Dehydrogenase

## Acknowledgments

We are thankful to Dr. Lishia Teh (Georg-August-Universität Göttingen, Göttingen, Germany), Dr. Nasir Javed (University of Manitoba, Winnipeg Manitoba, Canada) and Dr. Xiaodong Wang (Jiangsu Academy of Agricultural Sciences, Nanjing, China) for providing the marker primer sequences of *SO*, *PT* and *TN* populations.

## Funding

This research was supported by the National Basic Research Program of China (2015CB150205), the National Science Foundation of China (31671721), and the New Century Talents Support Program of the Ministry of Education of China (NCET110172).

## Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

## Authors' contributions

NR and HC collected the data, carried out the analysis and wrote the manuscript. LG, JG, JX participated in data collection and the candidate genes analysis. ML, LZ, YL and MY designed, led and coordinated the overall study. All authors have read and approved the final version of the manuscript.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Department of Biotechnology, College of Life Science and Technology, Key Laboratory of Molecular Biology, Huazhong University of Science and Technology, Wuhan 430074, China. <sup>2</sup>Hubei Collaborative Innovation Center for the Characteristic Resources Exploitation of Dabie Mountains, Huanggang Normal University, Huanggang 435599, China. <sup>3</sup>School of Life Science, Hunan University of Science and Technology, Xiangtan 411201, China.

Received: 14 June 2017 Accepted: 5 October 2017

Published online: 12 October 2017

## References

- Mackay TFC. The genetic architecture of quantitative traits. *Annu Rev Genet.* 2001;35:303–39. doi:10.1146/annurev.genet.35.102401.090633.
- Wolf JB. Genetic architecture and evolutionary constraint when the environment contains genes. *Proc Natl Acad Sci.* 2003;100(8):4655–60. doi:10.1073/pnas.0635741100.
- Hansen TF. The evolution of genetic architecture. *Annu Rev Ecol Evol Syst.* 2006;37:123–57. doi:10.1146/annurev.ecolsys.37.091305.110224.
- He XL, Zhang JZ. Toward a molecular understanding of pleiotropy. *Genetics.* 2006;173:1885–91. doi:10.1534/genetics.106.060269.
- Lobo I. Pleiotropy: one gene can affect multiple traits. *Nature Education.* 2008;1(1):10.
- Gu T. Oil, population growth, and the resource curse. North Carolina. Economics Thesis: Duke University; 2009. <https://sites.duke.edu/djepapers/files/2016/10/Gu.pdf>. Accessed 23 Nov 2014
- Lukoil. Global trends in oil & gas markets to 2025. Lukoil. 2013. [http://www.lukoil.be/pdf/Trends\\_Global\\_Oil\\_ENG.pdf](http://www.lukoil.be/pdf/Trends_Global_Oil_ENG.pdf). Accessed 23 Nov 2014.
- Boland M. Rapeseed. Agricultural Marketing Resource Center. 2012. [http://www.agmrc.org/commodities\\_products/grains\\_oilseeds/rapeseed](http://www.agmrc.org/commodities_products/grains_oilseeds/rapeseed). Accessed 17 Oct 2014
- Wang HZ. Strategy on the mid and long-term development of rapeseed variety improvement in China. *Chin J Oil Crop Sci.* 2004;26:98–101.
- Li M, Qian W, Meng J, Li Z. Construction of novel *Brassica napus* genotypes through chromosomal substitution and elimination using interloid species hybridization. *Chromosom Res.* 2004;12:417–26. doi:10.1023/B:CHRO.0000034722.66981.94.
- Yang YW, Lai KN, Tai PY, Li WH. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *J Mol Evol.* 1999;48(5):597–604. doi:10.1007/PL00006502.
- Beilstein MA, Al-Shehbaz IA, Kellogg EA. Brassicaceae phylogeny and trichome evolution. *Am J Bot.* 2006;93(4):607–19. doi:10.3732/ajb.93.4.607.
- Lagercrantz U. 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.* 1998;150:1217–28.
- Lysak MA, Koch MA, Pecinka A, Schubert I. Chromosome triplication found across the tribe Brassicaceae. *Genome Res.* 2005;15:516–25. doi:10.1101/gr.3531105.
- Cheng F, Mandáková T, Wu J, Xie Q, Lysak MA, Wang X. Deciphering the diploid ancestral genome of the Mesohexaploid *Brassica rapa*. *Plant Cell.* 2013;25(5):1541–54. doi:10.1105/tpc.113.110486.
- Cheng F, Wu J, Wang X. Genome triplication drove the diversification of *Brassica* plants. *Horticulture Research.* 2014;1:14024. doi:10.1038/hortres.2014.24.
- Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, et al. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics.* 2005;171:765–81. doi:10.1534/genetics.105.042093.
- Schranz ME, Lysak MA, Mitchell-Olds T. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends Plant Sci.* 2006;11(11). doi:10.1016/j.tplants.2006.09.002.

19. Chalhouh B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, et al. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*. 2014;345:950–3. doi:10.1126/science.1253435.
20. Kaup MT, Froese CD, Thompson JE. A role for diacylglycerol acyltransferase during leaf senescence. *Plant Physiol*. 2002;129:1–11. doi:10.1104/pp.003087.
21. Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, et al. Acyl-lipid metabolism. *The Arabidopsis Book*. 2013;11:e0161. doi:10.1199/tab.0161.
22. Ohlrogge J, Browse J. Lipid biosynthesis. *Plant Cell*. 1995;7:957–70. doi:10.1105/tpc.7.7.957.
23. Beisson F, Koo AJK, Ruuska S, Schwender J, Pollard M, Thelen JJ, et al. *Arabidopsis* genes involved in acyl lipid metabolism: a 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. *Plant Physiol*. 2003;132:681–97. doi:10.1104/pp.103.022988.
24. Baud S, Lepiniec L. Physiological and developmental regulation of seed oil production. *Prog Lipid Res*. 2010;49:235–49. doi:10.1016/j.plipres.2010.01.001.
25. Chapman KD, Ohlrogge JB. Compartmentation of triacylglycerol accumulation in plants. *J Biol Chem*. 2012;287:2288–94. doi:10.1074/jbc.R111.290072.
26. Rebetzke GJ, Pantalone VR, Burton JW, Carter Jr TE, Wilson RF. Genetic background and environment influence palmitate content of soybean seed oil. *Crop Sci*. 2001;41(6). doi:10.2135/cropsci2001.1731.
27. Shi CH, Zhang HZ, JG W, Li CT, Ren YL. Genetic and genotype × environment interaction effects analysis for erucic acid content in rapeseed (*Brassica napus* L.). *Euphytica*. 2003;130:249. doi:10.1023/A:1022867100199.
28. Ecke W, Uzunova M, Weißleder K. Mapping the genome of rapeseed (*Brassica napus* L.). II. Localization of genes controlling erucic acid synthesis and seed oil content. *Theor Appl Genet*. 1995;91:972–7. doi:10.1007/BF00223908.
29. Möllers C, Schierholt A. Genetic variation of palmitate and oil content in a winter oilseed rape doubled haploid population segregating for oleate content. *Crop Sci*. 2002;42(2). doi:10.2135/cropsci2002.0379.
30. Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, et al. A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat Genet*. 2008;40(3):367–72. doi:10.1038/ng.85.
31. Miller JF, Zimmerman DC, Vick BA. Genetic control of high oleic acid content in sunflower. *Oil Crop Sci*. 1987;27:923–6. doi:10.2135/cropsci1987.0011183X002700050019x.
32. Jako C, Kuar A, Wei Y, Zou J, Barton DL, Giblin EM, et al. Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol*. 2001;126:861–74. doi:10.1104/pp.126.2.861.
33. Liu F, Xia Y, Wu L, Fu D, Hayward A, Luo J, et al. Enhanced seed oil content by overexpressing genes related to triacylglyceride synthesis. *Gene*. 2015; 557(2):163–71. doi:10.1016/j.gene.2014.12.029.
34. Burns MJ, Barnes SR, Bowman JG, Clarke MH, Werner CP, Kearsey MJ. QTL analysis of an intervarietal set of substitution lines in *Brassica napus*: (i) seed oil content and fatty acid composition. *Heredity*. 2003;90(1):39–48. doi:10.1038/sj.hdy.6800176.
35. Qiu D, Morgan C, Shi J, Long Y, Liu J, Li R, et al. A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theor Appl Genet*. 2006;114(1):67–80. doi:10.1007/s00122-006-0411-2.
36. Zhao J, Becker HC, Zhang D, Zhang Y, Ecke W. Conditional QTL mapping of oil content in rapeseed with respect to protein content and traits related to plant development and grain yield. *Theor Appl Genet*. 2006;113:33–8. doi:10.1007/s00122-006-0267-5.
37. Delourme R, Falentin C, Huteau V, Clouet V, Horvais R, Gandon B, et al. Genetic control of oil content in oilseed rape (*Brassica napus* L.). *Theor Appl Genet*. 2006;113:1331–45. doi:10.1007/s00122-006-0386-z.
38. Chen G, Geng J, Rahman M, Liu X, Tu J, Fu T, et al. Identification of QTL for oil content, seed yield, and flowering time in oilseed rape (*Brassica napus*). *Euphytica*. 2010;175:161–74. doi:10.1007/s10681-010-0144-9.
39. Sun M, Hua W, Liu J, Huang S, Wang X, Liu G, et al. Design of new Genome- and Gene-Sourced Primers and identification of QTL for seed oil content in a specially high-oil *Brassica napus* cultivar. *PLoS One*. 2012;7(10): e47037. doi:10.1371/journal.pone.0047037.
40. Zhao J, Huang J, Chen F, Xu F, Ni X, Xu H, et al. Molecular mapping of *Arabidopsis thaliana* lipid-related orthologous genes in *Brassica napus*. *Theor Appl Genet*. 2012;124:407–21. doi:10.1007/s00122-011-1716-3.
41. Chen Y, Qi L, Zhang X, Huang J, Wang J, Chen H, et al. Characterization of the quantitative trait locus OilA1 for oil content in *Brassica napus*. *Theor Appl Genet*. 2013;126:2499–509. doi:10.1007/s00122-013-2150-5.
42. Wang X, Wang H, Long Y, Li D, Yin Y, Tian J, et al. Identification of QTLs associated with oil content in a high-oil *Brassica napus* cultivar and construction of a high-density consensus map for QTLs comparison in *B. napus*. *PLoS One*. 2013;8(12):e80569. doi:10.1371/journal.pone.0080569.
43. Chao H, Wang H, Wang X, Guo L, Gu J, Zhao W, et al. Genetic dissection of seed oil and protein content and identification of network associated with oil content in *Brassica napus*. *Scientific report*. 2017;7:46295. doi:10.1038/srep46295.
44. Jiang C, Shi J, Li R, Long Y, Wang H, Li D, et al. Quantitative trait loci that control the oil content variation of rapeseed (*Brassica napus* L.). *Theor Appl Genet*. 2014;127(4):957–68. doi:10.1007/s00122-014-2271-5.
45. Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, et al. New insights into the genetic networks affecting seed fatty acid concentrations in *Brassica napus*. *BMC Plant Biol*. 2015;15:91. doi:10.1186/s12870-015-0475-8.
46. Teh L, Möllers C. Genetic variation and inheritance of phytosterol and oil content in a doubled haploid population derived from the winter oilseed rape Sansibar × Oase cross. *Theor Appl Genet*. 2016;129(1):181–99. doi:10.1007/s00122-015-2621-y.
47. Javed N, Geng J, Tahir M, McVetty PBE, Li G, Duncan RW. Identification of QTL influencing seed oil content, fatty acid profile and days to flowering in *Brassica napus* L. *Euphytica*. 2016;207:191. doi:10.1007/s10681-015-1565-2.
48. Cheng X, Xia S, Zeng X, Gu J, Yang Y, Xu J, et al. Identification of quantitative trait loci associated with oil content and development of near isogenic lines for stable qOC-A10 in *Brassica napus* L. *Can J Plant Sci*. 2016; 96:423–32. doi:10.1139/cjps-2014-0442.
49. Huang XQ, Huang T, Hou GZ, Li L, Hou Y, YH L. Identification of QTLs for seed quality traits in rapeseed (*Brassica napus* L.) using recombinant inbred lines (RILs). *Euphytica*. 2016;210:1–16. doi:10.1007/s10681-016-1675-5.
50. Raman H, Raman R, Kilian A, Detering F, Long Y, Edwards D, et al. A consensus map of rapeseed (*Brassica napus* L.) based on diversity array technology markers: applications in genetic dissection of qualitative and quantitative traits. *BMC Genomics*. 2013;14:277. doi:10.1186/1471-2164-14-277.
51. Liu S, Fan C, Li J, Cai G, Yang Q, Wu J, et al. A genome wide association study reveals novel elite allelic variations in seed oil content of *Brassica napus*. *Theor Appl Genet*. 2016. doi:10.1007/s00122-016-2697-z.
52. Remington DL, Purugganan MD. Candidate gene, quantitative loci, and functional trait evolution in plants. *International Journal of Plant Science*. 2003;164:57–520. doi:10.1086/367812.
53. Zhu M, Zhao S. Candidate gene identification approach: progress and challenges. *Int J Biol Sci*. 2007;3(7):420–7. doi:10.7150/ijbs.3.420.
54. Raboanatahiry NH, Yin Y, Chen L, Li M. Genome-wide identification and phylogenetic analysis of kelch motif containing ACBP in *Brassica napus*. *BMC Genomics*. 2015;16:512. doi:10.1186/s12864-015-1735-6.
55. Lukens L, Zou F, Lydiate D, Parkin I, Osborn T. Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. *Genetics*. 2003;164:359–72.
56. Ohno S. Evolution by gene duplication. New York: Springer-Verlag; 1970. p. 160. doi:10.1002/tera.1420090224.
57. Zhang J. Gene duplication. *The Princeton guide to evolution* (ed. Losos J). 2013:397–405. Princeton: Princeton University Press.
58. Zhang JZ. Evolution by gene duplication: an update. *Trends in Ecology and Evolution*. 2003;18:292–8. doi:10.1016/S0169-5347(03)00033-8.
59. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. *J Genet*. 2013;92(1):155–61. doi:10.1007/s12041-013-0212-8.
60. Morinaga T. Interspecific hybridization in *Brassica*. II. The cytology of F1 hybrids of *B. cerna* and various other species with 10 chromosomes. *Japan J Bot*. 1929;4:277–89.
61. Nagaharu U. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Journal of Botany*. 1935;7:389–452.
62. Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, et al. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet*. 2011;43:1035–157.
63. Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IA, et al. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat Commun*. 2014;5:3930. doi:10.1038/ncomms4930.
64. Rastogi S, Liberles DA. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol Biol*. 2005;5:28. doi:10.1186/1471-2148-5-28.



65. Freeling M, Scanlon MJ, Fowler JE. Fractionation and subfunctionalization following genome duplications: mechanisms that drive gene content and their consequences. *Current Opinion in Genetics and Development*. 2015; 35:110–8. doi:10.1046/j.1420-9101.2003.00485.x.
66. Semagn K, Bjørnstad Å, Xu Y. The genetic dissection of quantitative traits in crops. *Electron J Biotechnol*. 2010;13(5) doi:10.2225/vol13-issue5-fulltext-14.
67. Zhao J, Becker HC, Zhang D, Zhang Y, Ecke W. Oil content in a European × Chinese rapeseed population: QTL with additive and epistatic effects and their genotype–environment interactions. *Crop Sci*. 2005;45:51–9.
68. ZY H, Wang XF, Zhan GM, Liu GH, Hua W, Wang HZ. Unusually large oil bodies are highly correlated with lower oil content in *Brassica napus*. *Plant Cell Rep*. 2009;28:541–9. doi:10.1007/s00299-008-0654-2.
69. Amar S, Becker HC, Möllers C. Genetic variation in phytosterol content of winter rapeseed (*Brassica Napus* L.) and development of NIRS calibration equations. *Plant Breed*. 2009;128:78–83. doi:10.1111/j.1439-0523.2008.01531.
70. Geng J, Javed N, McVetty PBE, Li G, Tahir M. An integrated genetic map for *Brassica napus* derived from double haploid and recombinant inbred populations. *Hereditary Genetics*. 2012;1(1):103. doi:10.4172/2161-1041.1000103.
71. Rundell RJ, Price TD. Adaptive radiation, non-adaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology and Evolution*. 2009;24:394–9. doi:10.1016/j.tree.2009.02.007.
72. Kraft NJ, Adler PB, Godoy O, James EC, Fuller S, Levine JM. Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol*. 2015;29(5):592–9. doi:10.1111/1365-2435.12345.
73. Wang N, Li F, Chen B, Xu K, Yan G, Qiao J, et al. Genome-wide investigation of genetic changes during modern breeding of *Brassica napus*. *Theor Appl Genet*. 2014;127(8):1817–29. doi:10.1007/s00122-014-2343-6.
74. Lou XY, Ma JZ, Yang MCK, Zhu J, Liu PY, Deng HW, et al. Improvement of mapping accuracy by unifying linkage and association analysis. *Genetics*. 2006;172:647–61. doi:10.1534/genetics.105.045781.
75. Geleta M, Stymne S, Bryngelsson T. Variation and inheritance of oil content and fatty acid composition in niger (*Guizotia abyssinica*). *J Food Compos Anal*. 2011;24(7):995–1003. doi:10.1016/j.jfca.2010.12.010.
76. XP M, Aryal N, JM D, JJ D. Oil content and fatty acid composition of the kernels of 31 genotypes of Chinese dwarf cherry (*Cerasus humilis* (Bge.) Sok.). *J Horticult Sci Biotechnol*. 2015;90(5):525–9. doi:10.1080/14620316.2015.11668709.
77. Sales-Campos H, Souza PR, Peghini BC, da Silva JS, Cardoso CR. An overview of the modulatory effects of oleic acid in health and disease. *Mini Review in Med Chem*. 2012;13(2):201–10. doi:10.2174/1389557511313020003.
78. Frank SA. Genetic variation of polygenic characters and the evolution of genetic degeneracy. *J Evol Biol*. 2003;16:138–42.
79. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet*. 2002;3:391–7. doi:10.1038/nrg796.
80. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res*. 2011;40:D1202–10. doi:10.1093/nar/gkr1090.
81. Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, et al. BRAD, the genetics and genomics database for *Brassica* plants. *BMC Plant Biol*. 2011;11:136. doi:10.1186/1471-2229-11-136.
82. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res Adv*. 2009. doi:10.1101/gr.092759.109.
83. Schuler GD. Sequence mapping by electronic PCR. *Genome Res*. 1997;7(5): 541–50. doi:10.1101/gr.7.5.541.
84. Rotmistrovsky K, Jang W, Schuler GD. A web server for performing electronic PCR. *Nucleic Acids Res*. 2004;32:W108–12. doi:10.1093/nar/gkh450.
85. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43:D447–52. doi:10.1093/nar/gku1003.
86. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2013;13:2498–504. doi:10.1101/gr.1239303.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
www.biomedcentral.com/submit

