

Identification of additive and epistatic effects QTLs for seed oil content in soybean based on an integrating map of two RIL populations

Yue Wang

Northeast Agricultural University

Shulin Liu

Institute of Genetics and Developmental Biology Chinese Academy of Sciences

Jiajing Wang

Northeast Agricultural University

Chang Yang

Northeast Agricultural University

Zhixi Tian

Institute of Genetics and Developmental Biology Chinese Academy of Sciences

Wen-Xia Li (■ liwenxianeau@126.com)

Northeast Agricultural University

Hailong Ning (Ininghailongneau@126.com)

Northeast Agricultural University https://orcid.org/0000-0002-5979-594X

Research article

Keywords: soybean seed oil content, additive effect QTL, epistatic effects QTL, associated recombinant inbred line, high-density integrated genetic map, candidate gene

Posted Date: July 20th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-39376/v1

License:
① ① This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract Background

Soybean seed oil has been widely used in human consumption and industrial production.

Results

In order to identify the additive and epistatic effects QTLs and QTLs by environments interactions (AE and AAE) for seed oil content in soybean, an eight-environment conjoint analysis based on two populations RIL3613 and RIL6013 with an integrating map was conducted. An new high-density integrated genetic map containing 2212 SNP markers and covering 5718.01 cM with an average distance of 2.61 cM were constructed by the combination of two linkage maps of two associated recombinant inbred line (A-RIL) populations. A total of 64 additive effect and additive × environment interaction (AE) QTL were identified on 19 chromosomes by both ICIM and IM methods, and the proportion of phenotypic variations explained (PVE) range of QTL related to oil content was 1.29–10.75%, of which 19 QTLs had overlapping marker intervals, and qOil-5-1 was identified simultaneously in both RIL populations. Compared with previous SSR positioning results, it is found 8 SNP sites within the QTL physical interval located in the SSR sites. Among them, 4 QTLs were new found. Twelve pairs of epistatic QTLs (additive × additive, AA) and QTL interactions with environments (AAE) for oil content were identified by the ICIM method, of which 3 QTLs were new found, and 2 additive effect QTLs, qOil-9-2 and qOil-15-1, linked to the other two QTLs to produce epistatic effects. A total of 5 potential candidate genes were identified based on genetic ontology and annotated information showing the relationship with seed oil content and/or fatty acid biosynthesis and metabolism.

Conclusion

These QTLs with different effects provide the good basis for molecular-assisted breeding of soybean oil content-related traits and further fine mapping of related genes.

Background

Soybean [*Glycine max* (L.) Merr.] has become one of the most primary crops in the world, high quantity oil (average 20%) that are widely used for human consumption and industrial purpose, such as cooking oils and biodiesel [1, 2]. Thus, development of increasing the soybean oil content is great significance for quality improvement.

However, oil content is quantitative traits not only controlled by multiple genes having small or large effects but also influenced by environments [3, 4]. Using molecular markers for indirect selection of important agronomic traits, and molecular marker-assisted selection (MAS) could improve the efficiency of traditional plant breeding [5, 6]. In last three decade years, numerous quantitative trait loci (QTL) for oil content in soybean seed have been reported [5, 7–13], about 327 oil QTLs listed in the Soybase Database

(http://www.soybase.org). And a few soybean oil content QTLs have been identified in multi-environment and multi-genetic backgrounds [5, 7, 14]. However, the resolution and accuracy with which QTL mapping can identify causal genetic changes is limited by the low total number of recombination events present in biparental mapping populations [2, 15]. QTL for oil content detected by different populations derived from the common parent have been less reported.

Furthermore, previous mapping studies carried out for oil trait in soybean were mainly based on the identification of additive effect QTLs, while less efforts have been made on the study of complex genetic effects such as epistasis and environment effects. Epistatic effects and QTL × environments interactions effect are all-important genetic effects [16, 17]. Epistasis, interaction between one pair of loci located in the same or different chromosomes, as an important genetic basis of complex traits [18]. Some studies had shown that epistasis significantly affected the expression of genes and genetic variation underlying soybean [17, 19–24]. Hou et al.(2014)[20] mapped oil content QTLs using simple repeat sequence (SSR) markers derived from Charleston and Dongnong594, which detected 4 pairs oil epistatic effect QTLs, and Qi et al. (2017)[12] identified epistatic interactions for seed oil content under multi-environments based on a high-density single-nucleotide polymorphisms (SNP) using the same population. The key for a better resolution of QTL architecture was the use of a genetic linkage map with relatively high marker coverage, and can predict functional genes.

In the previous results, we used SSR map to analysis the additive and epistasis (additive × additive) QTLs based on multi-environment data of two associated recombinant inbred lines (RIL) populations with common parents (RIL3613 and RIL6013) [25]. In this study, SNP mapping was used to re-identify the QTLs of oil content with the same population, the aim is to improve the detection efficiency, shorten the localization interval and candidate gene identification.

Methods

Plant materials and field design

Three soybean varieties with significant difference of oil content, Dongnong L13 (oil content 17.74%), Henong 60 (oil content 20.25%), and Heihe 36 (oil content 20.28%) were applied to construct the associated recombinant inbred line (A-RIL) populations in the current study. Dongnong L13 is a soybean variety bred by soybean research Institute of Northeast Agricultural University. Suinong 10 is the female parent and Jiunong 22 is the male parent, and the hybrid combination was prepared by sexual hybridization, and then selected by genealogy. Heihe 36 is a soybean variety bred by Heihe Branch of Heilongjiang Academy of Agricultural Sciences. Beiyi 87-9 is the female parent and Jiusan 90-66 is the male parent. A hybrid combination was prepared by irradiating F_1 air-dried seeds with 60C γ ray 0.14KGY. Henong 60 (original code name Hejiao 98-1667) is selected by the Jiamusi Branch of Heilongjiang Academy of Agricultural Sciences with Beifeng 11 as the female parent and an American dwarf variety Hobbit as the male parent by sexual hybridizing and genealogical method. Two RIL populations derived from the Dongnong L13 × Heihe 36 (RIL3613) and Dongnong L13 × Henong 60 (RIL6013) and conducted in Harbin (HRB; 45°75[®] N, 126°63[®] E) in 2008, Heilongjiang, China. The detail construction process of two RIL populations were described by Ning et al. (2018)[15]. Finally, a total of 120 and 139 lines from RIL3613 and RIL6013 were used in this research, respectively.

The parental lines and RILs were planted in 8 environments, in Keshan (KS; 48°25[®] N, 125°64[®] E) in 2013, in Harbin (HRB; 45°75[®] N, 126°63[®] E) in 2014, in Harbin and Keshan in 2015, in Acheng (AC; 45°52[®] N, 126°95[®] E), Shuangcheng (SC; 45°53[®] N, 126°32[®] E) and Harbin in 2016, in Shuangcheng (SC; 45°53[®] N, 126°32[®] E) in 2017. Details of seedling densities, fertilizers and sowing dates are shown in Supplementary Table S1. All plant materials were grown in a randomized complete block design with three replications (3 m in length, 0.70 m in apart and the seeds of an individual line were sown at 0.06 cm intervals). The field experiment was managed identically to the local soybean production.

Measurement of oil content

Seed phenotypic measurement was determined by ten mature plants randomly selecting in the middle row of per plot. Seed oil content (dry seeds, with water content of about 10%) was determined with three times by Infratec 1241 Grain Analyzer (FOSS, Sweden).

Variation analysis and heritability of phenotypic data

The significance of the difference in oil content between the parents of each population was determined by the t test, and the significance of the genotype difference between RILs and the environments was determined by ANOVA. The frequency distribution was analyzed by Microsoft Excel 2007.

For the multi-environment average value, the formula is as follows:

$$h^{2} = \frac{\sigma_{G}^{2}}{\sigma_{G}^{2} + \frac{\sigma_{GE}^{2}}{e} + \frac{\sigma_{\varepsilon}^{2}}{re}}$$

Where h^2 is broad sense heritability, σ_G^2 is the variance of genotype, σ_e^2 is the variance of error, σ_{GE}^2 indicates variance of genotype by environment effect, *r* is the number of replications and *e* is the number of environments in the study, σ_G^2 , σ_e^2 and σ_{GE}^2 were estimated using a mixed method implemented by Proc Mixed in SAS 9.2 (SAS Institute, Cary, NC, United States).

SNP genotyping and construction of genetic linkage map

By the SoySNP660K Beadchip (Beijing Boao Biotechnology Co. Ltd), three parents and 259 RIL individuals were genotyped and 600010 SNP markers were obtained. And 2212 SNP markers were obtained using the Bin function of QTL IciMapping4.0 software (www.isbreeding.net) to analyze the obtaining individual SNP data for redundant markers, removing the markers without polymorphism and with a missing rate > 10%, the partially separated markers and selecting a marker at each 100K interval (Supplementary Figure S1). Genetic maps of each population were constructed by using the MAP function of QTL IciMapping 4.0 software, divided the linkage groups according to the anchor information of markers, and arranged by

nnTwoOpt method. Using CMP to integrate the corresponding chromosomes of the two populations into one map, and finally obtained the genetic map of the A-RIL populations containing 20 linkage groups.

QTL mapping

QTL IciMapping 4.2 software was employed to analysis the additive effect QTLs and epistatic (additive × additive) effects QTLs and QTLs by environments interactions (AE and AAE) among multiple environments by performing Multi-Environment Trials (MET) module in multiple environments conjoint analysis including two methods (1) IM and ICIM-ADD: additive QTLs and QTLs by environments interactions (AE) were detected by inclusive composite interval mapping (ICIM)[26] and interval mapping (IM)[27, 28] a step in scanning represented by 1 cM, probability in stepwise regression was 0.001, the LOD threshold value used to declare significant QTL was 2.5. And (2) ICIM-EPI: epistatic (additive × additive) effects QTLs and QTLs by environments interactions (AAE) were detected by inclusive composite interval mapping (ICIM) [26]. A step in scanning represented by 5 cM, probability in stepwise regression was 0.0001, the LOD threshold value used to declare significant QTL was 5.0. The epistatic QTLS with phenotypic variability > 1% were retained. Follow the nomenclature of McCouch to name the QTL. The linkage maps were showed via MapChart 2.1 (https://www.wur.nl/en), and the chromosome length unit was Mb.

Results

Construction of genetic linkage map

A total of 600010 SNP markers have been produced using the SoySNP660K BeadChip at Beijing Boao Biotechnology Co. Ltd in the populations RIL3613 and RIL6013. A total of 2212 were selected and mapped of A-RIL populations by QTL IciMapping 4.0 software on 20 linkage groups (LG) or chromosomes (Chr) of soybean (Table 1). The SNP map of soybean covered 5718.01 cM with an average marker density of 2.61 cM (Table 1). The genetic length of each chromosome ranged from 170.98 cM on Chr_18/LG G to 587.11 cM on Chr_13/LG F. On average, each linkage group is covered by 110.60 SNP markers (Figure S1).

Phenotypic variation

Phenotypic data of seed oil content of parents and RILs in two RIL populations across 8 environments have been shown in Table 2. The phenotypic data of parents indicate a wide range of variation, with the mean values of 17.28%-21.50% and 17.28%-21.40% in the RIL3613 and RIL6013 populations, respectively. The kurtosis were recorded >1 in two environments, Harbin in 2015 (E3) and Shuangcheng in 2017 (E8) of RIL3613, in other environments in two populations, kurtosis and skewness were recorded <1. The phenotypic performance of the RIL populations was continuously distributed and relatively consistent, suggesting that the segregations of oil fit a normal distribution model, characterized by transgressive inheritance, and soybean seed oil is controlled by polygenes and are suitable for QTL mapping (Figure 2 and Table 2). The *ANOVA* results of RILs showed that genotype, environment, genotype × environment were significantly variation (*P=0.01*). The heritability (h^2) of oil content was 28.1% and 43.6% in RIL3613 and RIL6013

populations, respectively, indicating that the environment also affected the expression of oil content traits (Table 2, 3).

Additive effect QTL and additive QTL by environments interactions for oil content

A total of 64 additive effect QTLs and additive QTLs by environments interactions (AE) for oil content were simultaneously identified by ICIM and IM (MET module) for the conjoint analysis in all eight environments underlying populations RIL3613 and RIL6013 (Figure 2 and Supplementary Table S2). Including 38 additive effect QTLs for oil content mapped on excluding chr. 12 in population RIL3613, and 27 additive effect QTLs for oil content mapped on excluding chr. 12 in population RIL3613, and 27 additive effect QTLs for oil content mapped on 8 chromosomes (chrs. 3, 7, 8, 10, 11, 12, 16 and 20) in population RIL6013. All of them, the QTL of additive effect direction showed negative and positive across different genetic backgrounds and different methods, and the proportion of phenotypic variations explained (PVE) of additive QTLs and PVE of QTLs by environments interactions ranged from 1.29% to 10.75% and 0.86% to 10.66%, respectively.

We further detected a common QTL simultaneously across two genetic backgrounds and methods (Figure 2 and Table 4). The *qOil-5-1* (marker interval Gm05_4062384-Gm05_5158657) was located on chromosome 5, the LOD value ranged from 2.66 to 4.28, and the PVE by additive QTLs and PVE by environments interactions ranged from 2.09% to 2.72% and 1.77% to 2.71%, respectively. The *qOil-5-1* showed the positive additive effect in RIL6013 population, however, it showed the negative additive effect in RIL3613 population based on the different genetic backgrounds.

We found 19 regions overlapped QTLs (including *qOil-5-1*) by comparing the consequence of the different genetic backgrounds, which were mapped on chromosomes 2, 5, 9, 13, 15, 17 and 19, and 10 of them were identified in populations RIL3613 and RIL6013 (Figure 2 and Table 4). PVE of additive QTLs and PVE of QTL by environments interactions ranged from 1.57% to 9.33% and 0.86% to 7.63%, respectively. The additive effect direction of QTL appeared negative and positive across different genetic backgrounds and methods.

Epistasis effect QTLs and epistasis QTLs interaction with environment for oil content

A total of 12 pairs of epistatic QTLs (additive × additive, AA) and epistasis QTL interaction with environment (AAE) were identified, covering 9 linkage groups (chrs. 1, 3, 9, 10, 11, 12, 15, 18 and 19) across two RIL populations for the oil content by ICIM method in multiple environments conjoint analysis (Figure 3 and Table 5). And all the PVE by AA ranged from 1.61% to 3.31%. The 1 pairs QTLs showed negative epistatic effect in population RIL3613, and all the QTLs showed positive epistatic effect in population RIL6013. Several pairs epistatic effect and AAE QTLs mapped on the same chromosomes or with other chromosomes, and one QTL also could interact with multiple QTLs.

In the RIL6013 population, 11 pairs epistatic effects QTLs were identified, and 6 QTLs interacted with more than one QTL (Figure 3 and Table 5). For example, Gm01_41350513-Gm01_42850670 interacted with Gm09_585590-Gm09_1201285 and Gm15_8908864-Gm15_9850704, meanwhile, Gm09_585590-Gm09_1201285 interacted with Gm11_18740411-Gm11_34453671. The Gm09_683799-Gm09_1859079 showed epistatic effect with two QTLs (Gm09_19759328-Gm09_19958039 and Gm12_7885195-

Gm12_20168509). Moreover, the Gm15_8908864-Gm15_9850704 located on chromosome 15, which interact with 4 QTLs located on chrs. 1, 11, 12 and 15. Additionally, two significant additive effect QTLs (*qOil-9-2*, Gm09_585590-Gm09_1201285, *qOil-15-1*, Gm15_8908864-Gm15_9850704) showed epistatic interactions with other two QTLs (Table 4 and Table 5, underline).

Comparison with the QTL detected from SSR and SNP map

In present research, we using ICIM and IM of MET module with the populations RIL3613 and RIL6013 to identify QTLs associated with oil content of soybean. A total of 64 additive effect QTLs for oil content were located and compared with the previous additive QTLs detected based on SSR mapping (Wang et al. 2020) [25]. It was found that 8 SNP sites located in the 4 QTLs intervals detected by SSR map. In the interval of *qOil-D1a-2* located on chromosome 1 at 0.35-54.12 Mb, *qOil-1-1*(43.15-43.46 Mb), *qOil-1-3* (23.45-36.86 Mb) and *qOil-1-2* (4.15-4.36 Mb) located were re-identified; *qOil-11-1* at 18.74-34.45 Mb was detected in *qOil-B1-1* located at 8.90-32.19 Mb on chromosome 11; two QTL (*qOil-13-1* in 13.35-13.66 Mb and *qOil-13-2* in 10.46-13.80 Mb) were found in the interval of *qOil-F-1* (13.13-20.52 Mb) on chromosome 13; In interval of *qOil-D2-1* located at 1.75-20.11 Mb on chromosome 17, two QTL (*qOil-17-1* in *3.05-7.28* Mb and *qOil-17-2* in 4.85-5.64 Mb) were identified (Supplementary Table S3). Compared with the SSR genetic map [25], the SNP genetic map can shorten the interval for searching candidate genes and improve the detection power.

Oil candidate gene mining of common and epistasis QTLs

We detected 14 additive QTLs (including 4 SNP sites within the QTL physical interval located in the SSR map: *qOil-1-1, qOil-1-2, qOil-13-1, qOil-17-2,* and 11 common QTLs: *qOil-2-3, qOil-5-1, qOil-9-2, qOil-13-1, qOil-13-4, qOil-15-1, qOil-15-2, qOil-15-4, qOil-15-5, qOil-17-2, qOil-19-5*) and 7 epistasis QTLs (marker interval, Gm01_4091303-Gm01_5085864, Gm18_47550661-Gm18_48024037, Gm09_585590-Gm09_1201285, Gm09_19759328-Gm09_19958039, Gm10_42150935-Gm10_42851788, Gm15_8908864-Gm15_9850704, Gm19_41753854-Gm19_42163159) in two populations for the prediction of candidate genes with the physical marker distance of less than 1 Mb (Figure 2, Table 4 and 5, Supplementary Table S3, bold text). In Phytozomev12.1 website (https://phytozome.jgi.doe.gov/pz/portal.html) (Glycine Max Wm82. A2. V1), 570 genes found expressing highly in seed development process. Based on the pathway analysis via the Kyoto Encyclopedia of Genes and Genomes website (KEGG, www.kegg.jp), 5 of them are identified based on genetic ontology and annotated information showing the relationship with seed oil biosynthesis and metabolism and/or fatty acid, including *Glyma.05G049500.1, Glyma.05G053300.1, Glyma.17G067400.1 and Glyma.18G200300.1* (Table 6).

Discussion

SNP genetic map constructed based on A-RIL populations

Soybean seed oil content was an important quantitative trait, construction of genetic maps and QTLs localization have been applied in molecular MAS program with the development of molecular markers. In many previous studies, genetic maps of restriction fragment length polymorphism (RFLP) [29], amplified fragment length polymorphisms (AFLP), and/or simple sequence repeat (SSR) etc. markers [7, 30-31] were

used to map target traits. However, the coverage of these genetic maps was low, which will reduce the accuracy of QTL mapping. With the development of genome sequencing technology, millions of single nucleotide polymorphisms (SNPs) and insertion/deletions were identified. Choi et al. (2007)[32] discovered a total of 5,551 SNPs markers by the resequencing of sequence-tagged sites (STSs) developed from expressed sequence tag (EST) sequence using three mapping populations including the Minsoy × Noir 1, Minsoy × Archer as well as the Evans × PI 209332. Hyten et al. (2010)[33] established a high-throughput sovbean OTL mapping with 1.536 SNP markers using the same populations as Choi et al. (2007)[32]. Wang et al. (2016) constructed two linkage maps containing 5,728 and 4,354 bins based on 89,680 and 80,995 SNP markers, spanning a total genetic distance of 2,204.6 cM and 2,136.7 cM, with an average distance of 0.4 and 0.5 cM between neighboring bins in NJRINP and NJRI4P, respectively. Yao et al. (2020)[34] constructed a high-density genetic linkage map comprising 11,398 SNP markers on 20 linkage groups (LGs) using the specific-locus amplified fragment sequencing (SLAF-seq) method, a total of 181 recombinant inbred lines (RILs) derived from a cross between wild soybean ZYD00463 and cultivated WDD01514 were genotyped. These SNP markers encompassed 2913.78 cM of the soybean genome, with a mean distance of 0.26 cM between markers. Abundant molecular markers facilitate genome-wide analysis of soybean [35, 36], and the high-density occurrence of SNPs in the genome was more suitable for QTL localization. In this study, a 2,212 SNP markers map distributing on 20 chromosomes was established, with a length of 5718.01 cM and an average distance of 2.61 cM between adjacent markers based on A-RIL populations (RIL3613 and RIL6013) with common parent, which contributed to increased genetic diversity and identify epistatic interactions of QTLs between multiple genetic backgrounds [37]. In addition, multi-populations broken through the limitation of less recombination events in single population evaluation, using the same SNP marker could improve the accuracy of the identified QTLs.

The additive QTLs variation

Eight SNP sites within the QTL physical interval located in the SSR map and nineteen additive common QTLs were detected by two methods and in multiple backgrounds (Figure 2, Table 4 and Supplementary Table S3). Comparison of the previous studies with Soybase Datadase, 19 of 23 QTLs for the oil content had been mapped overlap in soybean with previous researches [7-9, 14, 29, 37-51] (Table 4 and Supplementary Figure S2). Among them, the same QTL that was located simultaneously in two RIL populations, *qOil-5-1* (marker interval, Gm05_4062384- Gm05_5158657) showed similar position the consensus QTL detected by Wang et al (2012)[49] in cross SD02-4-59 × A02-381100. This further confirmed the accuracy of QTLs identification in this study. The rest four QTLs (*qOil-13-1, qOil-13-4, qOil-13-5* and *qOil-17-2*) are new found.

Genetic basement of epistasis QTL and interactions with environments

Epistasis, or the interaction of a pair of loci may play an important role in the genetic of complex quantitative traits [18, 52]. If epistasis is ignored, many individual loci could not be detected, which will weaken the detection power of QTL[53]. In present study, PVE of A and AE of 64 additive QTL ranged from 1.29% to 10.75% and 0.86% to 10.66%, respectively, while those of AA and AAE effect of 12 pairs of epistatic QTL ranged from 1.61% to 3.31% (Figure 3 and Table 5). By the comparison, AA and AAE effect were less

than A and AE effect, this is similar to the study of Teng et al. (2017)[22]. Epistasis may also play an essential role in trait improvement even if epistatic variance components are low [17, 54]. In this research, two individual QTL of two pairs of epistasis effects QTLs were constructed by the intervals from the same chromosome 9 and 15, respectively, including Gm09_19759328-Gm09_19958039 interacted with Gm09_683799-Gm09_1859079, Gm15_8908864-Gm15_9850704 interacted with Gm15_2753022-Gm15_29807975, and other pairs epistasis effects QTLs were mapped on different chromosomes. One QTL also could interact with multiple QTLs. In the RIL6013 population, six QTLs interacted with more than one QTL, Gm01_41350513-Gm01_42850670, Gm09_585590-Gm09_1201285, Gm09_683799-Gm09_1859079, Gm11_18740411-Gm11_34453671, Gm12_7885195-Gm12_20168509, Gm15_8908864-Gm15_9850704 interacted with each other or other QTLs. Twelve pairs of epistatic QTLs included 14 sites, among which 11 sites contained, overlapped or crossed regions with the previous QTLs (Supplementary Figure S3) [9, 10, 29, 37-39, 47-49, 51, 55-59]. It also further illustrates the stability of these locus, and the results were similar to Yang et al. (2013)[53]. The other three sites (Gm01_4091303-Gm01_5085864, Gm01_41350513-Gm01_42850670 and Gm18_47550661-Gm18_48024037) have not been reported and can be considered as the newly discovered QTLs in this study.

Boer et al. (2002)[60] believed that the QTL interactions with environments could be regarded as a specific expression of QTL caused by year, location, temperature and other factors. Therefore, multi-environment joint analysis can be used to identify the stability of QTL and estimate the AAE effect. In this study, the PVE value by AAE effects ranged from 0.09 to 1.26% (Figure 3 and Table 5), only a pair AAE effects QTL (Gm01_4091303-Gm01_5085864~Gm19_41753854-Gm19_42163159), the PVE value of AAE effects was less than the PVE value of AA effects in the RIL3613 population, meaning that the environment played an important role, while the impact of environment were weak in remaining 11 pairs QTLs. The results showed that AA and AAE effects were important genetic effects in QTL localization [12, 61].

Relationship between additive effect QTLs and epistatic QTLs

In this study, the additive effect QTL also had epistasis effects [53, 62, 63], which should be carefully analyzed and evaluated in MAS to improve the seed oil content in soybean. Two additive effect QTLs *qOil-9-2* (marker interval, Gm09_585590-Gm09_1201285) and *qOil-15-1* (marker interval, Gm15_8908864-Gm15_9850704) showed interactions with other two QTLs (Table 4 and Table 5), such as Gm09_585590-Gm09_1201285 interacted with two QTLs Gm01_41350513-Gm01_42850670 and Gm11_18740411-Gm11_34453671, Gm15_8908864-Gm15_9850704 interacted with Gm01_41350513-Gm01_42850670, Gm09_19759328-Gm09_19958039, Gm11_18740411-Gm11_34453671, and Gm12_20168509-Gm12_7885195 interacted with Gm15_2753022-Gm15_29807975. The results showed that, should not only consider the effect of one site but also the interaction of multiple sites, not only consider the main effect of QTL, but also the additive effect, AA effects and QTL interactions with environments in the application of MAS breeding.

Oil content candidate genes prediction

We identified 3 genes that may be involved in oil anabolism through 14 additive QTLs (*qOil-1-1, qOil-1-2, qOil-2-3, qOil-5-1, qOil-9-2, qOil-13-1, qOil-13-4, qOil-15-1, qOil-15-2, qOil-15-3, qOil-15-4, qOil-15-5, qOil-17-2*

and *qOil-19-5*) with the physical marker distance of less than 1 Mb which based on pathway annotations (Table 6 and Supplementary Table S3, bold text). *Glyma.05G049500.1* was predicted to lysophospholipid acyltransferase, which was an important enzyme participating in membrane and storage lipid biosynthesis and participated in glycerolipid and glycerophospholipid metabolism pathway [64]. *Glyma.05G053300.1* was predicted to plant 4,4-dimethylsterol C-4alpha-methyl-monooxygenase, which involved in the steroid biosynthesis pathway [65], we believe that it indirectly affects the biosynthesis of fatty acids. *Glyma.17G067400.1* was predicted to diacylglycerol kinase (ATP), which is key signaling enzymes for phosphorylated diacylglycerin to produce phosphatidylic acid, and participates in the glycerophospholipid metabolism pathway, it also affect the formation of fatty acids[66]. In addition, We identified 2 genes that may be involved in oil anabolism through 7 epistatic QTLs (marker interval, Gm01_4091303-Gm01_5085864, Gm09_585590-Gm09_1201285, Gm09_19759328-Gm09_19958039, Gm10_42150935-Gm10_42851788, Gm15_8908864-Gm15_9850704, Gm18_47550661-Gm18_48024037, Gm19_41753854-Gm19_42163159) with the physical marker distance of less than 1 Mb which based on pathway annotations (Table 5 and 6). *Glyma.18G200300.1* was predicted to alcohol dehydrogenase, it is the key enzyme in alcohol fermentation, which is a flooding-response specific soybean gene expressed in root tissue, and affects soybean yield [67]. *Glyma.10G189900.1* was predicted to peroxygenase, Hanano et al. (2006)[68] study showed that peroxygenase had been purified from oat microsomes and lipid droplets, whose pathway also constitutes one branch of the so-called lipoxygenase pathway, which catalyzed the oxidation of unsaturated fatty acids (c18:2, c18:3, or c16:3) to produce corresponding fatty acid hydroperoxides. Further tests are needed to determine which of the 5 genes have a significant effect on soybean oil content, and candidate genes are targeted for further study.

Conclusion

In the A-RILs populations, we identified 8 SNP sites within the QTL physical interval located in the SSR map and 19 common additive QTLs, *qOil-5-1* simultaneously existed two RILs populations, of which 4 QTLs were new found. Moreover, a total of 12 pairs of epistatic QTLs (AA) and epistasis QTL interactions with environments (AAE) were identified, covering 9 of the 20 linkage groups, of which 3 QTL were new found. Around these QTLs, 5 potential candidate genes were identified.

Abbreviations

PVE: phenotypic variations explained ; AE: additive QTLs by environments interactions; AA: epistatic QTLs ; AAE: epistasis QTL interactions with environments

Declarations

Ethics approval and consent to participate

No ethical nor consent to participate in this research was sought, this not application in this research work.

Consent to publish

Not applicable.

Availability of data and materials

All the relevant data are all availed.

Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors gratefully acknowledge the financial support for this study provided by grants from the Key Science and Technology Project in Heilongjiang Province (SC2019ZX16B0039).

Authors' Contributions

W-XL and HN conceived and designed the experiments, YW, SL, JW, CY, ZT performed the field experiments. YW and HN analyzed and interpreted the results. YW and HN drafted the manuscript and all the authors contributed to the manuscript revision.

Acknowledgements

We are sincerely grateful to Prof Dr. Ning, Dr. Li and all the teachers in our research team for their valuable guidance in the course of this research work. To all the members of the research team, we do appreciate the moral support and the immense support we received during the period of this research work.

References

- 1. Wang J, Chen P, Wang D, Shannon G, Zeng A, Orazaly M, Wu C. Identification and mapping of stable QTL for protein content in soybean seeds. Mol Breeding. 2015;35(3):92–101.
- Li Y, Reif JC, Hong H, Li H, Liu Z, Ma Y, Li J, Tian Y, Li Y, Li W, et al. Genome-wide association mapping of QTL underlying seed oil and protein contents of a diverse panel of soybean accessions. Plant Sci. 2018;266:95–101.
- 3. Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, Schmidt ME. Seed quality QTL in a prominent soybean population. Theor Appl Genet. 2004;109(3):552–61.
- 4. Lee G, Wu X, Shannon JG, Sleper DA, Nguyen HT. Soybean. In: Kole C, editor Genome mapping and molecular breeding in plants, volume 2, oilseeds. Springer, Berlin 2007:1–3.
- 5. Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker RC. Mapping QTL for Seed Protein and Oil Content in Eight Soybean Populations. Crop Sci. 1997;37(2):370–8.
- 6. Patil G, Vuong TD, Kale S, Valliyodan B, Deshmukh R, Zhu C, Wu X, Bai Y, Yungbluth D, Lu F, et al. Dissecting genomic hotspots underlying seed protein, oil, and sucrose content in an interspecific mapping population of soybean using high-density linkage mapping. Plant Biol. 2018;16(11):1939–53.

- 7. Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, Lark KG. Genetics of Soybean Agronomic Traits: I. Comparison of Three Related Recombinant Inbred Populations. Crop Sci. 1999;39(6):1642–51.
- Pathan SM, Vuong T, Clark K, Lee J, Shannon JG, Roberts CA, Ellersieck MR, Burton JW, Cregan PB, Hyten DL, et al. Genetic Mapping and Confirmation of Quantitative Trait Loci for Seed Protein and Oil Contents and Seed Weight in Soybean. Crop Sci. 2013;53(3):765.
- 9. Han Y, Teng W, Wang Y, Zhao X, Wu L, Li D, Li W. Unconditional and conditional QTL underlying the genetic interrelationships between soybean seed isoflavone, and protein or oil contents. Plant Breeding. 2015;134(3):300–9.
- 10. Wang X, Jiang G, Green M, Scott RA, Song Q, Hyten DL, Cregan PB. Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean. Mol Genet Genomics. 2014;289(5):935–49.
- Warrington CV, Abdel-Haleem H, Hyten DL, Cregan PB, Orf JH, Killam AS, Bajjalieh N, Li Z, Boerma HR. QTL for seed protein and amino acids in the Benning × Danbaekkong soybean population, vol. 128; 2015.
- 12. Qi Z, Zhang X, Qi H, Xin D, Han X, Jiang H, Yin Z, Zhang Z, et al. Identification and validation of major QTLs and epistatic interactions for seed oil content in soybeans under multiple environments based on a high-density map. EUPHYTICA. 2017;213(8):162–75.
- Hacisalihoglu G, Burton AL, Gustin JL, Eker S, Asikli S, Heybet EH, Ozturk L, Cakmak I, Yazici A, Burkey KO, et al. Quantitative trait loci associated with soybean seed weight and composition under different phosphorus levels. J Inter Plant Biol. 2018;60(3):232–41.
- 14. Eskandari M, Cober ER, Rajcan I. Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theor Appl Genet. 2013;126(2):483–95.
- Ning H, Yuan J, Dong Q, Li W, Xue H, Wang Y, Tian Y, Li W. Identification of QTLs related to the vertical distribution and seed-set of pod number in soybean [*Glycine max* (L.) Merri]. PLOS ONE. 2018;13(4):e195830.
- 16. Wang DL, Zhu J, Li ZKL, Paterson AH. Mapping QTLs with epistatic effects and QTL × environment interactions by mixed linear model approaches. Theor Appl Genet. 1999;99(7):1255–64.
- 17. Karikari B, Li S, Bhat J, Cao Y, Kong J, Yang J, Gai J, Zhao T. Genome-Wide Detection of Major and Epistatic Effect QTLs for Seed Protein and Oil Content in Soybean Under Multiple Environments Using High-Density Bin Map. Int J Mol Sci. 2019;20(4):979–99.
- 18. Carlborg Ö, Jacobsson L, Åhgren P, Siegel P, Andersson L. Epistasis and the release of genetic variation during long-term selection. Nat Genet. 2006;38(4):418–20.
- Lark KG, Chase K, Adler F, Mansur LM, Orf JH. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. Proc Natl Acad Sci U S A. 1995;92(10):4656–60.
- 20. Hou M, Qi Z, Han X, Xin D, Jiang H, Liu C, Wu Q, Sui L, Hu G, Chen Q. QTL Mapping and Interaction Analysis of Seed Protein Content and Oil Content in Soybean. Agricultural Sciences in China. 2014;47(13):2680–9.

- 21. Wang Y, Han Y, Zhao X, Li Y, Teng W, Li D, Zhan Y, Li W. Mapping Isoflavone QTL with Main, Epistatic and QTL × Environment Effects in Recombinant Inbred Lines of Soybean. PLOS ONE. 2015;10(3):e118447.
- 22. Teng W, Zhang B, Zhang Q, Li W, Wu D, Yang H, Zhao X, Han Y, Li W. Identification of quantitative trait loci underlying seed oil content of soybean including main, epistatic and QTL × environment effects in different regions of Northeast China. Crop Pasture Science. 2017;68(7):625.
- 23. Tan R, Serven B, Collins PJ, Zhang Z, Wen Z, Boyse JF, Gu C, Chilvers MI, Diers BW, Wang D. QTL mapping and epistatic interaction analysis of field resistance to sudden death syndrome (Fusarium virguliforme) in soybean. Theor Appl Genet. 2018;131(8):1729–40.
- 24. Hina A, Cao Y, Song S, Li S, Sharmin RA, Elattar MA, Bhat JA, Zhao T. High-Resolution Mapping in Two RIL Populations Refines Major "QTL Hotspot" Regions for Seed Size and Shape in Soybean (*Glycine max* L.). Int J Mol Sci. 2020;21(3):1040.
- 25. Wang Y, Dong Q, Fang Y, Qi Z, Tian X, Song J, Wang J, Li X, Li W, Li W, et al. Identification of Quantitative Trait Loci for Seed Protein and Oil Contents in Soybean and Analysis for Epistatic and QTL × Environment Effects in Multiple Environments. Int J Agric Biol. DOI: 10.17957/IJAB/15.1464.
- 26. J W. Inclusive composite interval mapping of quantitative trait genes. Acta Genetica Sinica. 2009; 35(02):239–245.
- 27. Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. GENETICS. 1989;121(1):185–99.
- 28. Jansen RC. Interval mapping of multiple quantitative trait loci. GENETICS. 1993;135(1):205–11.
- 29. Diers BW, Keim P, Fehr WR, Shoemaker RC. RFLP analysis of soybean seed protein and oil content. Theor Appl Genet. 1992;83(5):608–12.
- 30. Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, et al. An Integrated Genetic Linkage Map of the Soybean Genome. Crop Sci. 1999;39:1464–90.
- 31. Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB. A new integrated genetic linkage map of the soybean. Theor Appl Genet. 2004;109(1):122–8.
- 32. Choi I, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon M, et al. A Soybean Transcript Map: Gene Distribution, Haplotype and Single-Nucleotide Polymorphism Analysis. GENETICS. 2007;176(1):685–96.
- 33. Hyten DL, Choi IY, Song Q, Specht JE, Carter TE, Shoemaker RC, Hwang EY, Matukumalli LK, Cregan PB. A High Density Integrated Genetic Linkage Map of Soybean and the Development of a 1536 Universal Soy Linkage Panel for Quantitative Trait Locus Mapping. Crop Sci. 2010;50(3):960–8.
- 34. Yao Y, You Q, Duan G, Ren J, Chu S, Zhao J, Li X, Zhou X, Jiao Y. Quantitative trait loci analysis of seed oil content and composition of wild and cultivated soybean. BMC Plant Biol. 2020;20(1):13–51.
- 35. Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLOS ONE. 2013;8(1):e54985.
- 36. Akond M, Liu S, Kantartzi SK, Meksem K, Bellaloui N, Lightfoot DA, Yuan J, Wang D, Anderson J, Kassem MA. A SNP genetic linkage map based on the 'Hamilton' by 'Spencer' recombinant inbred line

population identified QTL for seed isoflavone contents in soybean. Plant Breeding. 2015;134(5):580-8.

- 37. Mao T, Jiang Z, Han Y, Teng W, Zhao X, Li W. Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. Plant Breeding. 2013;132(6):630–41.
- 38. Mansur LM, Lark KG, Kross H, Oliveira A. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor Appl Genet. 1993;86(8):907–13.
- 39. Lee SH, Bailey MA, Mian MA, Carter JTE, Shipe ER, Ashley DA, Parrott WA, Hussey RS, Boerma HR. RFLP loci associated with soybean seed protein and oil content across populations and locations. Theor Appl Genet. 1996;93(5–6):649–57.
- 40. Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG. Genetic Mapping of Agronomic Traits Using Recombinant Inbred Lines of Soybean. Crop Sci. 1996;36(5):1327–36.
- 41. Fasoula VA, Harris DK, Boerma HR. Validation and Designation of Quantitative Trait Loci for Seed Protein, Seed Oil, and Seed Weight from Two Soybean Populations. Crop Sci. 2004;44(4):1218–25.
- 42. Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, Schmidt ME. Seed quality QTL in a prominent soybean population. Theor Appl Genet. 2004;109(3):552–61.
- 43. Kabelka EA, Diers BW, Fehr WR, LeRoy AR, Baianu IC, You T, Neece DJ, Nelson RL. Putative Alleles for Increased Yield from Soybean Plant Introductions. Crop Sci. 2004;44(3):784–91.
- 44. Reinprecht Y, Poysa VW, Yu K, Rajcan I, Ablett GR, Pauls KP. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. GENOME. 2006;49(12):1510–27.
- 45. Chen Q, Zhang Z, Liu C, Xin D, Shan D, Qiu H, Shan C. QTL Analysis of Major Agronomic Traits in Soybean. Agricultural Sciences in China. 2007;6(4):399–405.
- 46. Kim H, Kim Y, Kim S, Son B, Choi Y, Jum-Soon Kang YPYC. Analysis of Quantitative Trait Loci (QTLs) for Seed Size and Fatty Acid Composition Using Recombinant Inbred Lines in Soybean. Journal of Life Science. 2010;20(8):1186–92.
- 47. Qi Z, Wu Q, Han X, Sun Y, Du X, Liu C, Jiang H, Hu G, Chen Q. Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. EUPHYTICA. 2011;179(3):499–514.
- 48. Zhaoming Q, Xue H, Yanan S, Qiong W, Dapeng S, Xiangyu DU, Chunyan L, Hongwei J, Guohua H, Qingshan C. An Integrated Quantitative Trait Locus Map of Oil Content in Soybean, *Glycine max* (L.) Merr., Generated Using a Meta-Analysis Method for Mining Genes. Agricultural Sciences in China. 2011;10(11):1681–92.
- 49. Wang X, Jiang G, Green M, Scott RA, Hyten DL, Cregan PB. Quantitative trait locus analysis of saturated fatty acids in a population of recombinant inbred lines of soybean. Mol Breeding. 2012;30(2):1163–79.
- 50. Wang X, Jiang G, Green M, Scott RA, Song Q, Hyten DL, Cregan PB. Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean. Mol Genet Genomics. 2014;289(5):935–49.
- 51. Akond M, Liu S, Boney M, Kantartzi SK, Meksem K, Bellaloui N, Lightfoot DA, Kassem MA. Identification of Quantitative Trait Loci (QTL) Underlying Protein, Oil, and Five Major Fatty Acids' Contents in Soybean. American Journal of Plant Sciences. 2014;05(01):158–67.

- 52. Yang C, Tang D, Qu J, Zhang L, Zhang L, Chen Z, Liu J. Genetic mapping of QTL for the sizes of eight consecutive leaves below the tassel in maize (Zea mays L.). Theor Appl Genet. 2016;129(11):2191–209.
- 53. Yang Z, Xin D, Liu C, Jiang H, Han X, Sun Y, Qi Z, Hu G, Chen Q. Identification of QTLs for seed and pod traits in soybean and analysis for additive effects and epistatic effects of QTLs among multiple environments. Mol Genet Genomics. 2013;288(12):651–67.
- 54. Jannink J, Moreau L, Charmet G, Charcosset A. Overview of QTL detection in plants and tests for synergistic epistatic interactions. GENETICA. 2009;136(2):225–36.
- 55. CHEN Q. QTL Analysis of Major Agronomic Traits in Soybean. Ag Sci in China. 2007;6(4):399-405.
- 56. Leite DC, Pinheiro JB, Campos JB, Di Mauro AO, Unêda-Trevisoli SH. QTL mapping of soybean oil content for marker-assisted selection in plant breeding program. Genetics and molecular research. 2016; 15(1).
- 57. Pathan SM, Vuong T, Clark K, Lee J, Shannon JG, Roberts CA, Ellersieck MR, Burton JW, Cregan PB, Hyten DL, et al. Genetic Mapping and Confirmation of Quantitative Trait Loci for Seed Protein and Oil Contents and Seed Weight in Soybean. Crop Sci. 2013;53(3):765–74.
- 58. Reinprecht Y, Poysa VW, Yu K, Rajcan I, Ablett GR, Pauls KP. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. GENOME. 2006;49(12):1510–27.
- 59. Shibata M, Takayama K, Ujiie A, Yamada T, Abe J, Kitamura K. Genetic relationship between lipid content and linolenic acid concentration in soybean seeds. Breeding Sci. 2008;58(4):361–6.
- 60. Boer MP, Ter Braak CJ, Jansen RC. A penalized likelihood method for mapping epistatic quantitative trait Loci with one-dimensional genome searches. GENETICS. 2002;162(2):951–60.
- 61. Cao Y, Li S, Chen G, Wang Y, Bhat JA, Karikari B, Kong J, Gai J, Zhao T. Deciphering the Genetic Architecture of Plant Height in Soybean Using Two RIL Populations Sharing a Common M8206 Parent. Plants. 2019;8(10):373.
- 62. JIANG Z, HAN Y, TENG W, LI Y, ZHAO X, ZHANG Z, MAN W, LI W. Impact of epistasis and QTL × environmental interaction on the mass filling rate during seed development of soybean. Genet Res. 2012;94(02):63–71.
- 63. Qi Z, Hou M, Han X, Liu C, Jiang H, Xin D, Hu G, Chen Q. Identification of quantitative trait loci (QTLs) for seed protein concentration in soybean and analysis for additive effects and epistatic effects of QTLs under multiple environments. Plant Breeding. 2014;133(4):499–507.
- 64. Jasieniecka-Gazarkiewicz K, Lager I, Carlsson AS, Gutbrod K, Peisker H, Dörmann P, Stymne S, Banaś A. Acyl-CoA:lysophosphatidylethanolamine acyltransferase activity affects growth. Plant Physiol. 2017;174(2):986–98.
- 65. Rahier A. Dissecting the sterol C-4 demethylation process in higher plants. From structures and genes to catalytic mechanism. Steroids. 2011;76(4):340–52.
- 66. Escobar-Sepúlveda HF, Trejo-Téllez LI, Pérez-Rodríguez P, Hidalgo-Contreras JV, Gómez-Merino FC. Diacylglycerol Kinases Are Widespread in Higher Plants and Display Inducible Gene Expression in Response to Beneficial Elements, Metal, and Metalloid Ions. Front Plant Sci. 2017; 08.

- 67. Komatsu S, Thibaut D, Hiraga S, Kato M, Chiba M, Hashiguchi A, Tougou M, Shimamura S, Yasue H. Characterization of a novel flooding stress-responsive alcohol dehydrogenase expressed in soybean roots. Plant Mol Biol. 2011;77(3):309–22.
- Hanano A, Burcklen M, Flenet M, Ivancich A, Louwagie M, Garin J, Blée E. Plant Seed Peroxygenase Is an Original Heme-oxygenase with an EF-hand Calcium Binding Motif. J Biol Chem. 2006;281(44):33140-51.

Tables

 $\label{eq:Table 1} \textbf{Table 1} \text{ Distribution of markers and their properties on the linkage groups}$

Chr./LG	Number of SNP markers	Length (cM)	Average marker density (cM)
1/D1a	160	246.45	1.55
2/D1b	108	380.61	3.56
3/N	50	180.18	3.68
4/C1	148	352.57	2.76
5/A1	65	210.05	3.28
6/C2	129	335.84	2.62
7/M	165	300.52	1.83
8/A2	74	213.22	2.92
9/K	132	285.28	2.18
10/O	130	229.63	1.78
11/B1	48	186.86	3.98
12/H	67	267.44	4.05
13/F	254	587.11	2.32
14/B2	115	316.77	2.78
15/E	83	380.58	4.64
16/J	173	267.26	1.55
17/D2	72	289.22	4.07
18/G	62	170.98	2.80
19/L	59	321.54	5.54
20/I	118	195.90	1.67
Total	2212	5718.01	2.61

Table 2 Summarization of oil content of parents and RILs in two populations under eight environments

Population	${\tt Environment}^1$	Parents Average (%)			RIL l			
				Range	Average	Std^2	Skew ³	Kurt ⁴
				(%)	(%)			
RIL3613		Dongnong L13	Heihe 36					
	E1	17.90	20.10	14.40-20.13	17.05	0.94	0.35	0.43
	E2	18.10	20.30	16.73-22.39	19.49	0.95	0.10	0.45
	E3	17.28	19.80	16.45-21.76	20.15	0.46	-0.19	2.61
	E4	18.25	20.15	17.19-21.88	19.39	0.82	-0.06	0.06
	E5	17.28	19.88	17.16-21.14	20.10	0.55	-0.79	0.81
	E6	17.98	20.53	17.57-21.21	20.00	0.52	-0.41	0.28
	E7	18.26	20.90	17.61-21.74	20.66	0.39	-0.43	0.40
	E8	19.90	21.50	19.10-22.00	21.32	0.34	-0.42	1.17
RIL6013		Dongnong L13	Henong 60					
	E1	17.90	20.18	15.19-21.89	17.03	0.85	0.17	-0.64
	E2	18.10	20.22	17.89-22.37	20.37	1.03	-0.21	-0.57
	E3	17.28	20.92	16.54-21.92	20.38	0.51	0.03	0.68
	E4	18.25	20.78	17.72-21.83	19.63	0.85	-0.07	0.23
	E5	17.28	20.38	16.70-22.25	20.24	0.59	-0.35	0.82
	E6	17.98	20.24	17.19-22.31	20.77	0.60	-0.25	0.26
	E7	18.26	20.78	17.51-21.78	20.26	0.62	-0.45	0.24
	E8	19.90	21.40	20.10-22.40	21.28	0.39	-0.34	0.88

1 E1-E8 represent the 8 environments (2013 Keshan, 2014 Harbin, 2015 Harbin, 2015 Keshan, 2016 Acheng, 2016 Shuangcheng, 2016 Harbin, 2017 Shuangcheng), respectively

 $\ensuremath{\mathbf{2}}$ Std represent standard deviation of the phenotypic trait

 $\ensuremath{\mathsf{3}}$ Skew represent skewness of the phenotypic trait

 $\ensuremath{\mathbf{4}}$ Kurt represent kurtosis of the phenotypic trait

Table 3 Joint ANOVA of oil content of two RILs and heritability

Population	Source	DF	SS	MS	F	$\Pr > F$
RIL3613	Environment	7	3961.16	565.880	1074.36	<.0001
	Genotype	119	201.29	1.692	3.21	<.0001
	Genotype \times Environment	777	995.10	1.281	2.43	<.0001
	Error		8.66	0.541		
	h^{21}	0.281				
RIL6013	Environment		4792.95	684.707	1235.17	<.0001
	Genotype	138	328.59	2.381	4.30	<.0001
	Genotype \times Environment	925	1262.57	1.365	2.46	<.0001
	Error	16	9.62	0.601		
	h ²¹	0.436				

1 h^2 represent broad-sense heritability

Table 4 Additive effects and additive by environments interactions (AE) QTLs simultaneous mapping for oil content in two RIL populations

QTL	Left Marker	Right Marker	Chr^1	Population	Located by ICIM and IM methods ⁷				
					LOD ²	PVE(%) ³	PVE(A) ⁴ (%)	PVE(AE) ⁵ (%)	Add^6
qOil-2-3	Gm02_7564870	Gm02_7653076	2	RIL3613	2.94, 2.51	6.58, 8.65	0.88, 1.02	5.71, 7.63	0.033, 0.031
qOil-2-4	Gm02_6350994	Gm02_9755950		RIL6013	3.20, 6.65	2.46, 3.75	0.07, 0.35	2.39, 3.40	0.014, 0.040
qOil-5-1	Gm05_4062384	Gm05_5158657	5	RIL3613	3.25, 2.95	2.09, 2.72	0, 0.01	2.09, 2.71	-0.002, -0.003
				RIL6013	2.66, 4.28	2.11, 2.23	0.34, 0.23	1.77, 2.00	0.031, 0.032
<u>qOil-9-2</u>	<u>Gm09_585590</u>	Gm09_1201285	9	RIL6013	3.46, 4.26	2.63, 2.05	0.63, 0.30	2.00, 1.74	-0.042, -0.037
q0i1-9-3	Gm09_683799	Gm09_1859079		RIL3613	3.83, 3.09	5.24, 6.63	1.07, 1.14	4.16, 5.48	0.037, 0.033
q0il-13-3	Gm13_16968208	Gm13_21164873	13	RIL3613	4.67, 4.62	6.91, 9.33	0.07, 0.05	6.84, 9.28	0.009, 0.007
q0il-13-4	Gm13_21066582	Gm13_21695191		RIL6013	4.09, 4.70	4.25, 2.95	0.36, 0.20	3.89, 2.75	0.034, 0.032
q0il-13-5	Gm13_21695191	Gm13_23818871		RIL3613	2.76, 3.01	3.15, 4.44	0.08, 0.02	3.07, 4.42	0.01, 0.004
<u>qOil-15-1</u>	Gm15_8908864	<u>Gm15_9850704</u>	15	RIL6013	2.76, 3.68	1.81, 1.57	0.26, 0.17	1.55, 1.40	-0.026, -0.027
q0il-15-2	Gm15_7654010	Gm15_8235391		RIL3613	2.82, 2.91	4.81, 6.59	0.20, 0.32	4.61, 6.27	0.016, 0.018
q0il-15-3	Gm15_7569225	Gm15_7654010		RIL6013	3.73, 3.40	3.09, 1.83	1.33, 0.88	1.75, 0.95	0.067, 0.069
q0il-15-4	Gm15_9750834	Gm15_9970727		RIL3613	3.16, 3.26	5.51, 7.54	0.17, 0.29	5.34, 7.26	0.015, 0.016
q0il-15-5	Gm15_9970727	Gm15_10064172		RIL6013	3.06, 3.70	2.45, 2.06	0.77, 0.57	1.67, 1.48	0.047, 0.052
q0il-17-1	Gm17_3048712	Gm17_7282795	17	RIL3613	3.54, 3.50	5.44, 7.57	0.27, 0.36	5.17, 7.20	-0.02, -0.019
q0il-17-2	Gm17_4850258	Gm17_5639017		RIL6013	3.37, 3.25	3.23, 2.04	1.71, 1.02	1.53, 1.02	0.071, 0.070
q0il-19-1	Gm19_23094757	Gm19_47475093	19	RIL3613	3.37, 3.38	6.76, 9.17	2.10, 2.78	4.66, 6.39	-0.057, -0.056
q0il-19-2	Gm19_17827436	Gm19_38866655		RIL3613	3.77, 4.18	5.02, 7.08	0.02, 0.01	5.00, 7.07	-0.005, -0.002
qOil-19-4	Gm19_2179973	Gm19_38732396		RIL6013	3.48, 4.94	2.21, 2.41	0.05, 0.04	2.16, 2.37	0.012, 0.014
q0il-19-5	Gm19_34458684	Gm19_35275610		RIL6013	3.09, 3.51	2.35, 1.91	1.37, 1.05	0.98, 0.86	0.062, 0.069

1 Chr represent the chromosome

2 LOD is log of odd score

3 PVE represent the total phenotypic contribution of the additive effect and additive QTLs by environments interactions (AE) QTLs

 $4\ \text{PVE}(\text{A})$ is the phenotypic contribution of the additive effect QTL

5 PVE(AE) represent the phenotypic variation by additive effects and environment interactions (AE) QTL

 $6 \; \text{Add}$ represent the additive effect of the QTL

7 The first value from the ICIM method and the second value from IM method at the current scanning position

Table 5 Identification of epistatic effect QTLs (AA) and epistatic QTLs by environments interactions (AAE) for oil content in multi-genetic backgrounds and multi-environments

							RIL3613									
Chr^1	Marker Interval	Chr^1	Marker Interval	LOD ²	PVE ³	PVE (A	PVE	AA^6	Epistatic QTLs by environments interactions (AAE) ⁷							
	1		2		(%)	A) ⁴ (%)	¹ (%) (AAE) ⁵		AAE	AAE	AAE	AAE	AAE	AAE	AAE	AAE
							(%)		_01	_02	_03	_04	_05	_06	_07	_08
1	Gm01_4091303-	19	Gm19_41753854-	6.01	1.97	0.71	1.26	-0.16	-0.359	-0.286	0.136	0.089	0.139	0.106	0.091	0.086
	Gm01_5085864		Gm19_42163159													
	RIL6013															
1	Gm01_41350513-	9	<u>Gm09 585590-</u>	5.22	2.39	2.10	0.30	0.11	0.054	0.037	0.028	0.055	-0.057	0.044	-0.093	-0.068
	Gm01_42850670		Gm09_1201285													
9	Gm09_19759328-	9	Gm09_683799-	5.63	2.33	1.85	0.48	0.11	0.069	0.090	-0.001	-0.003	-0.065	0.044	-0.080	-0.054
	Gm09_19958039		Gm09_1859079													
9	Gm09_1398911-	10	Gm10_42150935-	5.01	2.54	1.67	0.87	0.10	0.134	0.094	0.031	0.049	-0.062	-0.095	-0.103	-0.049
	Gm09_20250790		Gm10_42851788													
9	Gm09_585590-	11	Gm11_18740411-	5.49	2.37	2.01	0.36	0.11	0.028	0.081	0.033	0.002	-0.075	0.090	-0.092	-0.067
	Gm09_1201285		Gm11_34453671													
9	Gm09_683799-	12	Gm12_7885195-	5.54	2.30	1.77	0.54	0.11	0.069	0.083	0.003	-0.019	-0.063	0.063	-0.085	-0.051
	Gm09_1859079		Gm12_20168509													
1	Gm01_41350513-	15	Gm15_8908864-	5.61	2.91	1.94	0.97	0.11	0.042	0.142	0.027	-0.022	-0.089	0.082	-0.091	-0.091
	Gm01_42850670		Gm15_9850704													
9	Gm09_19759328-	15	Gm15_8908864-	6.13	3.31	2.14	1.17	0.11	0.099	0.149	0.005	-0.045	-0.096	0.056	-0.086	-0.083
	Gm09_19958039		Gm15_9850704													
11	Gm11_18740411-	15	Gm15_8908864-	6.05	3.10	2.14	0.97	0.11	0.038	0.157	0.037	-0.022	-0.096	0.063	-0.093	-0.084
	Gm11_34453671		Gm15_9850704													
12	Gm12_7885195-	15	Gm15_8908864-	6.04	3.29	2.10	1.19	0.11	0.100	0.142	0.007	-0.060	-0.094	0.076	-0.091	-0.080
	Gm12_20168509		Gm15_9850704													
15	Gm15_8908864-	15	Gm15_2753022-	5.38	2.91	1.75	1.16	0.11	0.098	0.161	-0.009	-0.053	-0.083	0.042	-0.074	-0.083
	Gm15_9850704		Gm15_29807975													
3	Gm03_32549208-	18	Gm18_47550661-	5.19	1.61	1.53	0.09	0.11	0.002	0.148	0.042	0.001	-0.063	-0.005	-0.049	-0.077
	Gm03_33991515		Gm18_48024037													

1 Chr represent the chromosome

2 LOD is log of odd score

3 PVE represent the total phenotypic contribution of the additive effect and additive QTLs by environments interactions (AE) QTLs

 $4\ \text{PVE}(\text{AA})$ is the phenotypic contribution of the epistatic effect QTLs (AA)

5 PVE(AAE) represent the phenotypic variation by epistatic QTLs by environments interactions (AAE)

6 AA represent the epistatic effect of the QTL

7 AAE _01, AAE _2, AAE _3, AAE _4, AAE _5, AAE _6, AAE _7, AAE _8 represent the epistatic QTLs by environments interactions (AAE) contributed by multiple environments interactions among 8 environments (2013 Keshan, 2014 Harbin, 2015 Harbin, 2015 Keshan, 2016 Acheng, 2016 Shuangcheng, 2016 Harbin, 2017 Shuangcheng), respectively

Table 6 Details of 5 genes annotated in the KEGG database

Gene name ¹	Chromosome	Position	KO number	Annotation
Glyma.05G049500.1	chr05	44188234423186	K13519	LPT1; lysophospholipid acyltransferase [EC:2.3.1.51 2.3.1.23 2.3.1]
Glyma.05G053300.1	chr05	48105494814012	K14423	SMO1; plant 4,4-dimethylsterol C-4alpha-methyl-monooxygenase [EC:1.14.18.10]
Glyma.10G189900.1	chr10	4228989242291691	K17991	PXG; peroxygenase [EC:1.11.2.3]
Glyma.17G067400.1	chr17	52116435218864	K00901	dgkA; diacylglycerol kinase (ATP) [EC:2.7.1.107]
Glyma.18G200300.1	chr18	4796905147975187	K00001	E1.1.1.1; alcohol dehydrogenase [EC:1.1.1.1]

 $1 \mbox{ Indicates the gene which correlates with the QTL }$

Figures



Figure 1

Frequency distribution oil content. E1-E8 represent 8 environments (2013 Keshan, 2014 Harbin, 2015 Harbin, 2015 Keshan, 2016 Acheng, 2016 Shuangcheng, 2016 Harbin, 2017 Shuangcheng), respectively



Figure 2

Additive QTLs of oil content in RIL3613 (red bars) and RIL6013 (blue bars) Chromosomes genetic groups were arranged tandemly as a circle. The green bar is the QTL simultaneously in two RIL populations. The elongation is the common QTLs for two RIL populations.



Figure 3

Epistatic interactions of oil content QTLs in RIL3613 (red lines) and RIL6013 (green lines) populations. Chromosomes genetic groups were arranged tandemly as a circle.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

• TableS3ComparisonwitnSSRmap.docx

- TableS2QTLofAandAEmappingforoilcontent.docx
- TableS1Fieldmanagementmethodsandsoiltexture.docx
- FigureS3.docx
- FigureS2.docx
- FigureS1.docx