

Identification and validation of major QTLs and epistatic interactions for seed oil content in soybeans under multiple environments based on a high-density map

Qi Zhaoming · Zhang Xiaoying · Qi Huidong · Xin Dawei · Han Xue ·
Jiang Hongwei · Yin Zhengong · Zhang Zhanguo · Zhang Jinzhu ·
Zhu Rongsheng · Hu Zhenbang · Liu Chunyan · Wu Xiaoxia · Chen Qingshan ·
Che Daidi

Received: 19 February 2017 / Accepted: 26 June 2017 / Published online: 7 July 2017
© Springer Science+Business Media B.V. 2017

Abstract The seed oil content in soybean is an important trait that drives successful soybean quality. A recombination of inbred lines derived from a cross between the ‘Charleston’ and ‘Dongnong594’ cultivars was planted in one location (HRB) across 13 years and two locations (HXL and JMS) across 6 years in China (25 environments in total), and the genetic effects were partitioned into additive main effects, epistatic main effects and their environmental interaction effects using composite interval mapping and inclusive composite interval mapping models based on a high-density genetic map. Twelve main-effect quantitative trait loci (QTLs) were identified on

chromosomes Ch3, Ch4, Ch6, Ch13, Ch15, Ch17, Ch18, Ch19 and Ch20 and detected in more than two environments. Among the intervals of the main-effect QTLs, eleven candidate genes were screened for their involvement in seed oil content and/or fatty acid biosynthesis and metabolism processes based on gene ontology and annotation information, moreover, the main effect QTL were identified in a wild soybean chromosome segment substitution lines population, the phenotype and the QTL fragment showed the identity relationship significantly. Furthermore, an analysis of epistatic interactions showed that four epistatic QTL pairs were detected, and they could explain approximately 70% of the phenotypic variation interaction with environments in total. The additive main-effect QTLs and epistatic QTL pairs contributed to high phenotypic variation under multiple environments, and the results were also validated and corroborated with previous research, indicating

Qi Zhaoming, Zhang Jinzhu and Qi Huidong have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10681-017-1952-y) contains supplementary material, which is available to authorized users.

Q. Zhaoming · Z. Xiaoying · Q. Huidong ·
X. Dawei · H. Xue · J. Hongwei · Z. Zhanguo ·
Z. Jinzhu · Z. Rongsheng · H. Zhenbang ·
L. Chunyan · W. Xiaoxia · C. Qingshan (✉) ·
C. Daidi (✉)
Northeast Agricultural University,
Harbin 150030, Heilongjiang, People’s Republic of China
e-mail: qshchen@126.com

C. Daidi
e-mail: daidiche@163.com

Q. Zhaoming
e-mail: qizhaoming1860@126.com

H. Xue · J. Hongwei · L. Chunyan
Crop Research and Breeding Center of Land-Reclamation
of Heilongjiang Province, Harbin 150090, Heilongjiang,
People’s Republic of China

Y. Zhengong
Crop Breeding Institute, Heilongjiang Academy of
Agricultural Sciences, Harbin 150086, Heilongjiang,
People’s Republic of China

that marker-assisted selection can be used to improve soybean oil content and that the candidate genes can also be used as a foundation data set for research on gene functions.

Keywords Soybean seed oil content · Major QTL mapping · Epistatic interactions · Candidate gene · QTL validation

Introduction

Soybean (*Glycine max* [L.] Merr.) is a crop with substantial economic value, which represented 61 percent of the world's oilseed production, soybean oil could be used into food products such as margarine, salad dressings and cooking oils, and industrial products such as plastics and biodiesel fuel (<http://www.soystats.com>) (Wilson 2004, 2008). Thus, development of increasing the soybean oil content is important for commercial production.

Soybean oil content is quantitatively inherited in soybeans and controlled by multiple genes by classical genetic studies (Wilcox 1985). As the development of molecular technology and quantitative genetics, numerous of soybean oil content quantitative trait loci (QTLs) have been reported over the past two decades, and there were over 188 QTLs listed in the USDA Soybean Genome Database (SoyBase, <http://www.soybase.org>). The mapped populations primarily included segregation populations, recombinant inbred lines (RILs) and backcross populations, and the main mapping methods included the analyses of variance (ANOVAs; Lander et al. 1987), interval mapping (IM; Lander and Botstein 1989; Haley and Knott 1992; Jansen 1993), composite interval mapping (CIM; Rodolphe and Lefort 1993; Zeng 1994), and multiple interval mapping (MIM; Kao et al. 1999). Among the published QTLs, some of them showed 'hot regions' that were identified four or more times at the same or similar intervals in different studies, and these regions include Gm05: 35.2–40.8 Mb, Gm09: 40.3–46.8 Mb, Gm12: 34.1–40.6 Mb, Gm14: 33.8–49.2 Mb, Gm15: 0.8–13.9 Mb, Gm18: 51.6–59.8 Mb, Gm19: 32.9–48.0 Mb and Gm20: 23.5–34.6 Mb (<http://www.soybase.org>).

Epistatic interaction effect between genes and their interaction with environments were two important genetic factors, which make a substantial contribution

to the variation observed in complex traits (Carlborg and Haley 2004; Würschum et al. 2013). Most traits controlled by major genes are less affected by these two factors, while most quantitative traits are greatly affected by either one of them or both (Xu and Crouch 2008). Till now, many researches have been conducted to identify the QTL \times environmental effects in various crops by these methods in recent years, such as tomato (Eshed and Zamir 1996; Chapman et al. 2012), rice (Liu et al. 2007; Wang et al. 2012b), corn (Wang and Guo 2010; Yang et al. 2015), soybean (Korir et al. 2011; Zhang et al. 2015; Qi et al. 2016), wheat (Mohan et al. 2009; Deng et al. 2015), common bean (González et al. 2015), cotton (Jia et al. 2014), chickpea (Bajaj et al. 2015), rapeseed (Xu et al. 2015), peanut (Wang et al. 2015), rapeseed (Luo et al. 2014), asparagus bean (Xu et al. 2013) and groundnut (Ravi et al. 2011). These studies showed that QTL \times environment interactions greatly influenced the genetic architecture. Thus, it is meaningful to study the genetic characteristics of soybean oil content across different environments.

The identification of genes/QTLs and the development of higher-density molecular genetic maps have facilitated the study of complex quantitative traits, and the analysis of genes have influenced the related traits into individual Mendelian factors (Xing et al. 2012). Till now, few reports were reported focus on mapping soybean oil content QTLs under 25 environments or more based on a high-density genetic maps. Therefore, in this study, a RIL population was sequenced previously by specific-length amplified fragment sequencing (SLAF-seq) and a high density genetic map has been constructed (Qi et al. 2014a, b), with the population was tested for more than 10 years and employing multiple mapping methods to (1) identify the major QTLs of oil content and environment specific QTL, (2) mining the related genes of major QTLs, and (3) analyze epistatic QTL pairs across years (or environments) to better understand the implications of these QTLs when breeding for soybean oil content.

Materials and methods

QTL mapping population and experimental design

A total of 147 $F_{2:10}$ – $F_{2:22}$ populations derived from RILs were advanced by single-seed descent from crosses between the American cultivars 'Charleston'

(♀) (Cooper et al. 1995) and Chinese lines ‘Dongnong594’ (♂) (Developed by Northeast Agricultural University, Harbin, Heilongjiang, China). The RIL population was constructed by methods from Chen et al. (2005). A total of 147 RILs were planted in Harbin (HRB; longitude 126°38'E, latitude 45°45'N) from 2002 to 2014; Hongxinglong (HXL; longitude 129°55'–134°35'E, latitude 45°35'–47°17'N) from 2007 to 2013; and Jiamusi (JMS; longitude 130°21'E, latitude 46°49'N) from 2007 to 2013.

The plants were arranged in a randomized complete block experimental design involving single-row plots (1 m length and 0.5 m width), with two and three replicates in 2002–2007 and 2008–2012, respectively. The timing and frequency of cultural management procedures used for the trials were in accordance with normal production practices for the respective environments.

QTL validation population

A set of 194 wild soybean CSSLs population (Suinong 14, a cultivated variety in northeast of China, as the recurrent parent; and ZYD00006, a wild soybean variety, as the donor parent) had been constructed in previously study (Xin et al. 2016), which including 190 substituted segment covered 82.55% genome of wild soybean. The CSSLs and their parents were grown in a randomized complete block design with three replications in Harbin (HRB; longitude 126°38'E, latitude 45°45'N) in 2014 year. The other field management was the same as QTL mapping population.

Measurement of soybean oil content

Seeds from five randomly harvested plants were taken per plot for each line to analysis of seed oil content. Seed oil phenotypic data was analyzed using the FOSS Infratec TM1241 Grain Analyzer (FOSS Tecator AB, Hoeganaes, Sweden) at the 13% moisture basis. The seed oil content for each RIL line was averaged from the data of two or three plots replicates.

Map construction

The linkage map of this RIL population was previously constructed by specific-length amplified fragment sequencing (SLAF-seq) (Qi et al. 2014a, b). The high-density genetic map included 5308 markers on 20

linkage groups and was 2,655.68 cM in length, with an average distance of 0.5 cM between adjacent markers.

QTL mapping

Windows QTL Cartographer Ver. 2.5 (Wang et al. 2012a) was employed to independently analyze the QTL under each environment using a CIM model. IciMapping V 3.3 (<http://www.isbreeding.net/software/?type=detail&id=12>) was also used for joint analyses of QTLs to identify the major loci. These two models were used with a mapping step of 2.0 cM, an LOD significance threshold determined by 1000 permutation tests, and significance set at the $P = 0.05$ level. The naming of QTL was followed according to QTL-nomenclature described by McCouch (1997). All of the favorable alleles were derived from the parent ‘Charleston’ as indicated by the positive values of the additive effect and the parent ‘Dongnong 594’, as indicated by the negative values of the additive effects.

The MCIM model of QTL Network V2.0 (Yang et al. 2007; Yang et al. 2008) was used to analyze the additive QTL effects, additive \times additive epistatic QTL pairs, additive QTL \times environment effects, and epistatic QTL pair \times environment interactions for seed oil content data of all 25 environments. The critical F value of the MCIM model was calculated with 10,000 permutation tests. The QTL effects were estimated using the Markov chain Monte Carlo method with 20,000 Gibbs sampler iterations and candidate interval selection and putative QTL detection, and the QTL effects were calculated with an experiment-wise type-I error under $\alpha = 0.001$ (Wang et al. 1994; Yang et al. 2007; Xing et al. 2012).

Results

Phenotypic variation in the RIL population in 25 environments

Phenotypic value of seed oil content of RILs population in 25 environments across 13 years was shown in Table 1. The differences between the two parents were significant in most of the 25 environments, and Mean values of ‘Charleston’ and ‘Dongnong 594’ were

Table 1 Phenotype of oil content of the population and parents in 25 environments across 14 years

Year & environment	Charleston (%)	DongNong 594 (%)	RIL lines						ANOVA
			Max (%)	Min (%)	Mean (%)	SD	Kurtosis	Skewness	P value
2002HRB	19.63	21.83	24.33	16.68	21.45	1.38	0.23	-0.34	0.4471
2003HRB	19.27	19.68	22.61	17.9	20.44	0.89	0.10	-0.04	0.5395
2004HRB	19.03	19.26	21.39	17.22	19.3	0.74	-0.13	-0.01	0.5641
2005HRB	19.11	19.44	24.15	14.15	19.28	1.44	0.80	0.31	0.114
2006HRB	20.30	21.20	22.1	18.9	20.84	0.58	0.27	-0.54	0.1845
2007JMS	20.65	20.98	22.76	18.64	20.99	0.78	0.42	-0.39	0.1856
2007HXL	21.09	21.63	25.88	18.99	22.39	1.35	-0.09	0.02	0.7031
2007HRB	20.97	21.26	23.35	19.49	21.46	0.81	-0.09	0.05	0.1879
2008HXL	21.80	19.89	22.76	18.64	20.93	0.79	-0.29	0.29	0.6263
2008JMS	21.77	20.73	24.13	19.14	22.08	1.05	-0.73	0.40	0.0012
2008HRB	21.89	20.31	23.57	18.64	21.6	0.93	-0.44	0.08	0.3447
2009HXL	21.24	18.96	22.28	17.85	20.42	0.67	-0.59	1.46	0.4312
2009JMS	20.93	20.07	21.98	19.16	20.93	0.47	-0.70	0.91	0.0004
2009HRB	21.36	20.26	21.79	19.08	20.71	0.55	-0.74	0.56	0.0000
2010HXL	21.73	21.42	22.36	20.14	21.56	0.39	1.35	-0.69	0.0045
2010JMS	21.51	21.49	23.2	19.09	21.49	0.48	4.84	-0.80	0.0003
2010HRB	21.21	21.42	22.03	19.73	21.37	0.41	2.76	-1.27	0.0000
2011HXL	19.78	20.99	22.34	20.17	21.21	0.43	-0.05	-0.26	0.1510
2011JMS	20.56	21.14	22.55	20.57	21.41	0.36	1.12	0.63	0.0069
2011HRB	19.98	21.40	22.51	20.01	21.33	0.40	0.55	-0.26	0.9668
2012HXL	19.58	19.69	20.86	16.42	19.10	0.81	0.54	-0.53	0.0111
2012JMS	21.29	21.49	22.45	18.58	21.02	0.57	2.51	-0.66	0.0329
2012HRB	20.73	21.16	22.74	18.02	20.55	0.67	2.10	-0.41	0.2144
2013HRB	21.25	21.51	23.19	19.70	21.45	0.57	0.72	-0.02	0.7530
2014HRB	21.00	21.20	22.10	19.30	21.07	0.46	1.22	-0.60	0.0784
Average	20.71	20.74	22.78	18.65	20.98				

20.71 and 20.74%, respectively. The range of variation was narrow for ‘Charleston’ (19.03–21.89%) but was wider for ‘Dongnong 594’ (18.96–22.22%). The phenotypic data of the RIL populations also showed wide ranges of variation in the 25 environments. Some environments in this population had higher and others had lower mean of the trait as compared to the parents. Most skewness and kurtosis values were <1.00. In HRB location, the oil coefficient of variation ranged from 0.40 to 1.44; in HXL location, the oil content of RILs ranged from 0.36 to 1.05; in JMS location, the ranged from 0.39 to 1.35. An ANOVA analysis was employed to detect the year and location interaction effects of the phenotypic data, and most of the effects showed a significant interaction ($P = 0.05$).

Overall, the phenotypic performance of the RIL populations was continuously distributed and relatively consistent in 25 environments, suggesting that the segregations of this trait fit a normal distribution model.

Main effects of the stable additive QTL analysis

An environment-universal QTL (QTL_U) was defined as a QTL that was associated with the trait in at least one macro-environment that corresponded to the combination of year and location (Palomeque et al. 2009; Mao et al. 2013), in this research, the main-effect QTLs associated with the oil content were identified in more than two environments or

Table 2 Main-effect QTL mapping based on multiple environments

QTL	Environment	Chr	Analysis method	QTL position	Interval	L-marker	R-marker	LOD	ADD	R ² (%)
QOil3-2	2007HXL	3	CIM	60.4	58.2–65.9	Mark874678	Mark839399	3.11	0.47	9.0
	2010HRB	3	CIM	67.7	65.1–69.7	Mark883955	Mark890734	3.25	0.15	8.0
QOil4-1	2010HXL	4	CIM	0.0	0–5.6	Mark771298	Mark775719	4.81	–0.19	12.0
	2013HRB	4	CIM	2.4	1.9–3.4	Mark823650	Mark766640	3.20	–0.20	7.0
QOil6-1	2012JMS	6	ICIM	30.0	29.9–30.2	Mark450142	Mark493782	3.60	–0.21	10.0
	2012JMS	6	CIM	30.4	29.2–31.4	Mark461434	Mark450512	3.19	–0.19	7.0
QOil13-5	2014HRB	13	CIM	115.4	114.1–117.2	Mark93009	Mark136617	3.19	–0.14	7.0
	2012HXL	13	CIM	116.4	112.5–123.5	Mark94256	Mark134620	6.22	–0.44	16.0
QOil15-1	2006HRB	15	CIM	4.0	2.6–6.4	Mark34302	Mark75332	3.42	0.20	9.0
	2006HRB	15	ICIM	5.0	4.0–5.758	Mark53701	Mark39547	2.93	0.16	10.0
QOil17-5	2011HXL	17	CIM	130.1	129.5–132.6	Mark1396229	Mark1431798	3.08	0.14	8.0
	2013HRB	17	CIM	135.3	130.5–136.7	Mark1396229	Mark1422265	2.70	0.16	6.0
QOil17-6	2013HRB	17	ICIM	137.0	136.7–137.7	Mark1422265	Mark1431709	2.51	0.17	8.0
	2014HRB	17	ICIM	137.0	136.7–137.7	Mark1422265	Mark1431709	3.58	0.17	11.0
QOil17-7	2012HRB	17	CIM	140.1	138.1–141.9	Mark1407275	Mark1385955	2.98	0.22	8.0
	2012JMS	17	CIM	140.1	138.1–141.6	Mark1407275	Mark1385955	3.15	0.16	7.0
QOil17-8	2013HRB	17	CIM	143.2	142.6–145.7	Mark1385955	Mark1393366	2.78	0.15	6.0
	2012HRB	17	CIM	145.2	142.0–146.1	Mark1385955	Mark1393366	3.91	0.25	10.0
	2012JMS	17	CIM	145.2	142.6–146.4	Mark1385955	Mark1393366	3.13	0.16	7.0
QOil18-1	2007JMS	18	CIM	0.0	0–5.7	Mark179802	Mark190765	3.22	–0.25	8.5
	2004HRB	18	ICIM	1.0	0–1.38	Mark179802	Mark142708	2.71	–0.18	9.0
QOil19-2	2007HRB	19	CIM	5.1	4.6–7	Mark931755	Mark954052	2.86	–0.22	6.0
	2011JMS	19	CIM	8.3	4.6–10.6	Mark931755	Mark899805	3.72	–0.12	10.0
QOil20-6	2010HXL	20	CIM	76.3	72.2–78.5	Mark1167645	Mark1144580	3.06	0.12	7.0
	2008HRB	20	ICIM	79.0	74.2–79.0	Mark1184737	Mark1144580	2.67	0.25	8.0
	2008HRB	20	CIM	79.0	78.3–83.2	Mark1184737	Mark1179016	3.04	0.25	8.0

Chr represent the chromosome, *L*-marker represent the left marker of the interval, *R*-marker represent the right marker of the interval, *ADD* represent the additive effect of the QTL and *R*² represent the phenotypic contribution of the QTL

detected by two methods. Seven of them (QOil3-2, QOil4-1, QOil13-5, QOil17-5, QOil17-7, QOil17-8 and QOil19-2) were identified on chromosomes 3, 4, 13, 17 and 19 by the CIM method. The LOD ranged from 2.70 to 6.22 with significance threshold determined by 1000 permutation test ($P < 0.05$), QOil17-8 was identified in three environments (2012HRB, 2012JMS and 2013HRB). QOil17-6 was mapped on chromosome 17 by the ICIM method in both 2013HRB and 2014HRB, and the main-effect QTL could explain approximately 10% of the phenotypic variation for the seed oil content. Four main effects QTL (QOil6-1, QOil15-1, QOil18-1 and QOil20-6) were identified by both CIM and ICIM methods and the LOD ranged from

2.67 to 3.60. QOil20-6 on chromosome 20 was identified in 2008HRB by two methods and also detected at the similar genome region in 2008HRB and 2010HXL by CIM method (Table 2), so above 12 QTLs was defined as main effect QTLs for soybean oil content in this research.

Environment-specific QTL (QTL_{SP}) was defined to be associated with the oil content that was significant only in one environment (Palomeque et al. 2009; Mao et al. 2013). In HRB location across 13 years, thirty-six QTLs were detected and the LOD ranged from 2.51 to 4.52. Among these QTLs, four environment-universal QTLs were found and also contributed to main effect QTLs, QOil15-1, QOil17-6 and QOil17-8 could explain approximately 10% of the phenotypic contribution for

the seed oil content; thirty-two environment-specific QTLs were detected which located on Gm02, Gm 03, Gm 04, Gm 05, Gm 08, Gm10, Gm 11, Gm 12, Gm 13, Gm 14, Gm 15, Gm 16, Gm 17, Gm 18, Gm 19 and Gm20, and seven of them also contributed to main effect QTLs. In JMS location across 6 years, fourteen QTLs were detected and the LOD ranged from 2.51 to 3.60. Among these QTLs, one environment-universal QTL was found and also contributed to main effect QTLs, it could explain 10% of the phenotypic contribution for the seed oil content, thirteen environment-specific QTL were detected which located on Gm 02, Gm 06, Gm 08, Gm 012, Gm13, Gm 17, Gm 18 and Gm19, and four of them also contributed to main effect QTLs. In HXL location across 6 years, twenty-seven QTLs were detected and the LOD ranged from 2.57 to 6.22. Among these QTLs, no environment-universal QTL were found, all these QTLs were environment-specific QTLs which located on Gm 02, Gm 03, Gm 04, Gm 05, Gm 07, Gm 09, Gm10, Gm 11, Gm 13, Gm 15, Gm 16, Gm 17, Gm 18 and Gm20, and five of them also contributed to main effect QTLs, QOil4-1 and QOil3-5 could explain 11.78 and 16.16% of the phenotypic contribution for the seed oil content (Supplement Table 1).

Candidate gene mining of the main-effect QTLs

By comparison of the marker physical position of the main-effect QTLs and sequencing soybean genome (Williams 82, version 289), 1489 candidate genes were screened out from the gene annotation data set of soybean (http://phytozome.jgi.doe.gov/pz/portal.html#?info?alias=Org_Gmax), (Supplemental Table 2). Based on the blast with Kyoto Encyclopedia of Genes and Genomes website (KEGG, www.kegg.jp), eleven of them showed the relationship with seed oil content and/or fatty acid biosynthesis and metabolism based on the gene ontology and annotation information, including one gene in QOil3-2, seven genes in QOil4-1, two genes in QOil6-1 and one gene in QOil13-5 (Supplemental Table 3). These candidate genes functions have already been studied in the previously research. Glyma03g18661 was screened from QOil3-2 and predicted as Beta-ketoacyl synthase related gene, which was a key enzyme for fatty acid synthesis (Garwin et al. 1980). Glyma04g02271 was screened from QOil4-1 and predicted as sterol methyltransferase 2, this gene involved in the steroid biosynthesis (Shi et al.1996), researchers found it may influenced on fatty acid indirectly (Rakotomanga et al. 2005; Merah et al. 2012).

Glyma04g05110 and Glyma04g05131 were screened from QOil4-1 and both of them were predicted as fatty acid desaturase (*FAD*) gene, which contribute to seed linolenic acid levels (Bilyeu et al. 2003). Glyma04g03590 was screened from QOil4-1 and predicted as phospholipid glycerol acyltransferase, which involve in the ether lipid/glycerophospholipid metabolism process, it also verified influence the fatty acid synthesis (Heath et al. 1994). Glyma04g03740 was screened from QOil4-1 and predicted as allene oxide synthase, which involved in the process of alpha-linolenic acid metabolism and specialized for the metabolism of fatty acid hydro-peroxides (Song et al. 1993). Glyma04g06110 was screened from QOil4-1 and predicted as 3-ketoacyl-CoA synthase 10 (*KCS10*), which may encodes a fatty acid elongase 3-ketoacyl-CoA synthase and influence the fatty acid indirectly (Todd et al. 1999). Glyma04g04961 was screened from QOil4-1 and predicted as alpha beta-hydrolases superfamily protein and this gene involved in the process of glycerophospholipid metabolism, this kind of enzymes were collected in involved in ESTHER database (Lenfant et al. 2013). Glyma06g17625 and Glyma06g17640 were screened from QOil6-1 and both of them were predicted as fatty acyl-ACP thioesterases B (*FatB*), which encode a group of enzymes displayed high specificity towards 12:0-ACP which correlates with the composition of the oil present in the seeds of that species (Salas and Ohlrogge 2002). Glyma13g28860 was screened from QOil13-5 and was predicted as cytochrome P450, family 710, subfamily A, polypeptide 1 which involved in the process of steroid biosynthesis, and this kind of gene influence the fatty acid indirectly (Muerhoff et al. 1992) All of the details on the gene annotation and expression information are shown in Supplemental Table 3.

Identification of epistatic effect QTLs and epistatic QTL interactions with the environment

With the analyzing the oil content data set of all 25 environments by QTLnetwork software, four epistatic QTL pairs were identified. One epistatic QTL pair, which was located on Gm01 (70.5 cM) and Gm18 (0.2 cM), showed additive epistatic effects and interactions with the environment (AAE), the phenotypic variation explained by the epistatic QTLs was 7.41%, and the phenotypic variation explained by epistasis \times environment interactions was 15.26%. Furthermore, this QTL pair showed negative values for epistatic effects in 2004HRB, 2010HRB, 2011HRB, 2014HRB,

2007JMS and 2010JMS environment. The second epistatic QTL pair, located on Gm03 (5.1 cM) with the marker interval of Mark859393-Mark855624 and Gm15 (0.1 cM) with the marker interval of Mark18771-Mark23202, showed additive epistatic effects and interactions with the environment. The phenotypic variation explained by the epistatic QTLs was 7.01%, and the phenotypic variation explained by the additive epistasis \times environment interactions was 14.55%, and this QTL pair presented positive values for the epistatic effects and positive epistatic effects in the 2003HRB, 2006HRB 2007HRB and 2010HRB environments, although it showed the negative values of 2012HRB; moreover, it also showed epistatic interaction in HRB environment rather than the other location. The third epistatic QTLs pairs were located on Gm04 (0.1 cM) with the marker interval of Mark771298-Mark756945 and Gm08 (30.1 cM) with the marker interval of Mark1312817-Mark1334259, showed additive epistatic effects and interactions with the environment. The phenotypic variation explained by the epistatic QTL was 8.16%, and the phenotypic variation explained by the epistasis \times environment interactions was 18.24%. This QTL pair had negative epistatic effects in the 2008JMS, 2009JMS and 2012JMS environments, although it had a positive epistatic effect in 2011JMS. It showed epistatic interaction in JMS environment rather than the other location. The fourth epistatic QTLs pairs were located on Gm06 (30.2 cM) with the marker interval of Mark450142-Mark493782 and Gm07 (0.1 cM) with the marker interval of Mark582982-Mark532996, showed additive epistatic effects and interactions with the environment. The phenotypic variation explained by the epistatic QTL was 10.17%, and the phenotypic variation explained by the epistasis \times environment interactions was 22.05%. This QTL pair had negative epistatic effects in the 2007HXL, 2009 HXL and 2012HXL environments, although it had a positive epistatic effect in 2011HXL, it also showed epistatic interaction in HXL environment rather than the other location (Table 3). Furthermore, one of QTL pair contributed over 20% of the phenotypic variation, these four QTL pairs contributed around 70% of phenotypic variation interaction with environments in total.

QTL validation

To identify the effect of QTLs mapping results, a CSSLs population was used. We screened out the

CSSLs with main effect QTL fragment, environment-specific QTL fragment and non-QTL fragment. Distribution of CSSLs seed oil content was showed in Fig. 1. The seed oil content of the CSSLs with main effect QTL fragment was range from 21 to 22.76%. CSSL-158 with the QOil3-2 showed the highest oil content (22.76%) in CSSLs. CSSL-77 brings the QOil17-5, QOil17-6, QOil17-7 and QOil17-8 together, which the seed oil content was 21.73%. CSSL-52 with QOil6-1 and CSSL-91 with QOil13-5 showed the seed oil content were 22.15 and 22.33%, respectively. The seed oil content of CSSLs with environment-specific QTL fragment were range from 20.7 to 22.38%. The CSSLs with main effect QTL fragment and most environment-specific QTL fragment showed the significant higher than the Suinong14 (the recurrent parent of CSSLs, average of seed oil content was 20.20%) and the CSSLs without any QTL fragment (seed oil content of non-QTL fragment CSSLs was average at 20.20%).

Discussion

High-density genetic map could identify more recombination events in a population and will increase accuracy of QTL mapping (Xie et al. 2010). In soybean, many genetic linkage maps have been published based on restriction fragment length polymorphism (RFLP) markers (Keim et al. 1990), RFLP, isozyme, morphological, and biochemical markers (Shoemaker and Olson 1993; Lark et al. 1993), simple sequence repeat (SSR) markers (Cregan et al. 1999) and genetic map of integrated SSR, RFLP, random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), etc. markers (Song et al. 2004). With the advances in genome sequencing technology, few high-density genetic map have been constructed, Choi et al. (2007) constructed a soybean transcript map with single nucleotide polymorphisms (SNPs). Hyten et al. (2010) built a high-density integrated genetic linkage map "Universal Soy Linkage Panel" (USLP 1.0) with 1536 SNPs. Lee et al. (2015) constructed a high-density genetic map with SNPs based on the Infinium BARCSoySNP6 K BeadChip array. Song et al. (2016) have constructed a high resolution genetic linkage maps based on new soybean genome sequence assembly data Glyma1.01. In this study, a high-density soybean genetic map was

Table 3 Estimated epistatic effects (AA) and environmental (AAE) interaction of QTLs for soybean seed oil content across all environments

Ch_i	Interval_i	Position_i	Ch_j	Interval_j	Position_j	AA ^a	AAE1 ^b	AAE2 ^b	AAE3 ^b	AAE4 ^b	AAE5 ^b	AAE6 ^b
Gm01	Mark586999- Mark588867	70.5	Gm18	Mark179802-Mark142708	0.2	-0.2503**	NS	NS	-0.1867*	NS	NS	NS
Gm03	Mark859393- Mark855624	5.1	Gm15	Mark18771-Mark23202	0.1	0.2635**	NS	0.1922*	NS	NS	0.2339*	0.0728*
Gm04	Mark771298- Mark756945	0.1	Gm08	Mark1312817- Mark1334259	30.1	-0.2642**	NS	NS	NS	NS	NS	NS
Gm06	Mark450142- Mark493782	30.2	Gm07	Mark582982-Mark532996	0.1	-0.3468**	NS	NS	NS	NS	NS	NS
AAE7 ^b	AAE8 ^b	AAE9 ^b	AAE10 ^b	AAE11 ^b	AAE12 ^b	AAE13 ^b	AAE14 ^b	AAE15 ^b	AAE16 ^b	AAE17 ^b	AAE18 ^b	AAE19 ^b
NS	NS	-0.1732*	-0.1356*	NS	NS	-0.1735*	NS	NS	NS	NS	NS	NS
NS	NS	0.0642*	NS	-0.0641*	NS	NS	NS	NS	NS	NS	NS	NS
NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
NS	NS	NS	NS	NS	NS	NS	-0.2180*	NS	-0.1729*	NS	NS	NS
AAE18 ^b	AAE19 ^b	AAE20 ^b	AAE21 ^b	AAE22 ^b	AAE23 ^b	AAE24 ^b	AAE25 ^b	AAE26 ^b	AAE27 ^b	AAE28 ^b	AAE29 ^b	AAE30 ^b
NS	NS	-0.1132*	NS	NS	-0.1261*	NS	NS	NS	7.41	15.2635	NS	NS
NS	NS	NS	NS	NS	NS	NS	NS	NS	7.01	14.5492	NS	NS
NS	NS	NS	-0.0102*	-0.1045*	NS	0.0149*	-0.5849**	NS	8.16	18.2393	NS	NS
0.0075*	-0.2236*	NS	NS	NS	NS	NS	NS	NS	10.17	22.0463	NS	NS

^a The estimated additive by additive epistatic effect

^b Epistatic effects by environment interaction effect: E1, 2002HRB; E2, 2003HRB; E3, 2004HRB; E4, 2005HRB; E5, 2006HRB; E6, 2007HRB; E7, 2008HRB; E8, 2009HRB; E9, 2010HRB; E10, 2011HRB; E11, 2012 HRB; E12: 2013 HRB; E13: 2014 HRB; E14: 2007HXL; E15, 2008HXL; E16, 2009HXL; E17, 2010HXL; E18, 2011HXL; E19, 2012HXL; E20, 2007JMS; E21, 2008JMS; E22, 2009JMS; E23, 2010JMS; E24, 2011JMS and E25, 2012JMS

^c Phenotypic variation explained by epistatic QTL

^d Phenotypic variation explained by epistasis × environment (AAE) interactions

* Significant at P = 0.05, ** Significant at P = 0.01 and NS represent no significance, respectively

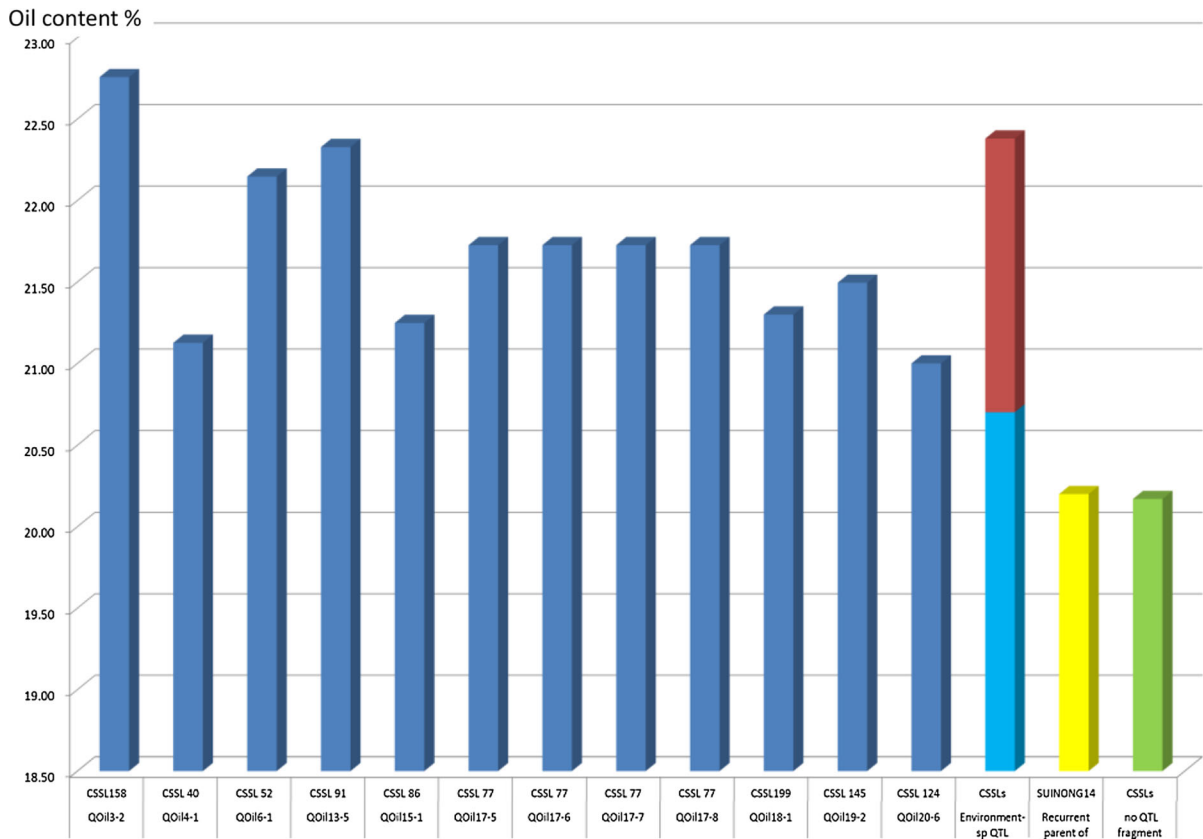


Fig. 1 Distribution of CSSLs population oil content. *Dark blue* represent the CSSL with main effect QTLs, *white blue* and *red* represent the border of CSSL with environment-specific QTLs,

yellow represent the recurrent parent (Suinong14) of CSSLs population, *green* represent the CSSLs without any QTL fragment. (Color figure online)

used, and it was created by SLAF-seq that included 5308 markers on 20 linkage groups and was 2,655.68 cM in length, with an average distance of 0.5 cM between adjacent markers (Qi et al. 2014a, b), furthermore, the RILs were crossed between an American cultivars ‘Charleston’ and a Chinese lines ‘Dongnong594’ as distant hybridization, Charleston was semi-dwarf lines and selected from HC74-634RE × HC78-676 (Cooper et al. 1995), and ‘Dongnong594’ was selected from Dongnong42 serials, Northeast Agricultural University in China. In order to reduce the errors caused by the environment, the RILs were planted in HRB (from 2002 to 2014), JMS (from 2007 to 2013) and HXL (from 2007 to 2013), and each of the locations or years was different from a statistical perspective. As described by Jansen et al. (1995), the accuracy of the QTL position and effects can be evaluated if the data are collected in different locations and years from a statistical perspective.

Twelve main-effect QTLs were identified, they contributed the seed oil content significantly. Comparison of the result with previously study with the same population and similar data set based on the SSR marker linkage map (Qi et al. 2014a, b), there were no overlap for the main effect QTL with previously research, only few matches with environment specific QTL, but the QTLs detected in this study showed much accuracy with significantly much markers and shorter average maker distance. Furthermore, six of them had a good match with those detected in previously research. QOil4-1 showed similar position and bigger interval than the consensus QTL detected by Qi et al. (2011) and a main effect QTL detected by Wang et al. (2014a, b), although the interval was bigger than the result of meta-analysis, but the peak position was closed. QOil6-1 was identified by both the CIM and ICIM methods and it mapped with a smaller confidence interval at the central part of a QTL, which was identified under 4 environments by

Hyten et al. (2004). QOil13-5 located on Gm03 from 3.1 to 3.3 Mb had been identified as the three QTLs Seed oil 37-7, Seed oil 37-8 and Seed oil 38-4 by Eskandari et al. (2013) and Rossi et al. (2013), respectively. QOil15-1 showed a smaller genome region as detected by Li et al. (2011). QOil19-2 was detected as a hot region by Mansur et al. (1996), Orf et al. 1999, Hyten et al. (2004), Qi et al. (2011). QOil20-6 was identified by both the CIM and ICIM methods in 2008HRB and had the same QTL peak, which were also identified by Diers et al. (1992), Sebolt et al. (2000), Tajuddin et al. (2003), Shibata et al. (2008), Qi et al. (2011), Specht et al. (2001), Wang et al. (2014a, b) in the classic population, this genomic region was also a very hot region. As the soybean genome of Williams 82 Glyma1.01 was sequenced completely and published by Schmutz et al. (2010), which make the identification of QTLs or genes much easier. Based on the release of sequence data and bioinformatics tools, 1489 candidate genes were screened from the main-effect QTLs, and eleven of them showed the relationship with seed oil content and/or fatty acid biosynthesis and metabolism based on the gene ontology and annotation information and their homologous or similar genes have been researched for the functional analysis or validation in the other plant. All of the gene annotation details and expression information are shown in Supplemental Table 3. To validate the effect of main effect QTLs, a wild soybean CSSLs population was used. We found seed oil content of CSSLs with main effect QTL and environment-specific QTL were range from 21 to 22.76% and 20.7 to 22.38%, respectively. Which were significant higher than the recurrent parent of CSSLs and the CSSLs without any QTL fragments. The consistency with validation population and previously research indicated that these QTL regions were “hot regions” that should be studied in further research on genes and marker-assisted selection (MAS) and the candidate genes could be the further research target.

The phenotype data correlation analysis was conducted between oil content and protein content, plant height, seeds length and width, seeds and pods weight per plant, seeds weight per plant, 100-seeds weight etc. agronomic traits of the RIL population for all environments, results showed there were significance negative relations between oil content and protein content in most of environments (Supplemental Table 4, except two environmental data showed minor

positive relation), the result were identified with previously important research (Brim and Burton 1979; Brummer et al. 1997; Wilcox 1998; Cober and Voldeng 2000). Hwang et al. (2014) conducted a genome-wide association study of seed protein and oil content in soybean, six SNPs had a significant association with both protein and oil and showed negative relation between protein and oil. However, based on the genetic map and phenotype data under multiple environments of this RIL, main-effect QTLs of seed protein content (Qi et al. 2016) and oil content (this research) have been identified, we found the a main effect QTL located on 24.6–24.8 Mb of Gm20, the positive allele was from “Charleston”, which showed the positive relationship for oil and protein, furthermore, this region also inside in a hot region identified by Bolon et al. (2010), who using Soy GeneChip and transcriptome sequencing technology association analysis identified the candidate genes for soybean protein and oil content. Moreover, there were only six genes in this interval, but no evidence showed they involved in fatty acid biosynthesis or metabolize, it could be a further research point.

Epistasis effects and environmental factors make a substantial contribution to variation in complex traits, for quantitative trait, the phenotype of a given genotype cannot be predicted by the sum of its component single locus effects (Carlborg and Haley 2004). Previously studies showed the mechanism of epistasis of complex traits is much more complicated comparison with the individual gene effects and might decrease the individual QTLs effect (Zhao et al. 2008; Palomeque et al. 2010; Korir et al. 2011). The genetic architecture of a complex trait was influenced by the main gene and also can be largely attributed to epistasis and gene \times environment interactions (Allard 1996), additive–additive epistatic effects of QTL pairs and random terms for environmental effects, additive–environmental interaction effects and additive–additive–environmental interaction effects were the key factor (Yang et al. 2005, 2007, 2008). Several recent studies have been conducted to demonstrate epistasis and QTL \times environmental interaction effects (Ma et al. 2007; Xing et al. 2012; Wu et al. 2012; Liu et al. 2016). In this study, four QTL pairs were screened, and they could explain approximately 70% of the phenotypic variation, which is a high value because some pairs could explain around 20% of the phenotypic

variation of seed oil content. These QTL pairs showed additive epistatic effects and interactions with the environment which detected under serious type-I errors below $\alpha = 0.001$. Epistasis is common and can cause cryptic genetic variation in quantitative traits and can produce small additive effects and generate a loss of heritability and lack of replication (Mackay et al. 2009; Mackay 2014). The identification of epistatic interactions between genes and/or QTLs is a valuable starting point for a more thorough understanding of these genetic networks (Carlborg and Haley 2004). Future breeding strategies with to increase seed oil content should consider not only the main effect of QTLs, but also exploring epistasis effects more thoroughly and their interactions with the environments.

Acknowledgements This study was conducted in the Key Laboratory of Soybean Biology in Chinese Ministry of Education in Heilongjiang Province and financially supported by National Key R & D Program of the Ministry of Science and Technology of China (2016YFD0100500, 2016YFD0100300, 2016YFD0100201-21), the Research Fund for the national natural science foundation of China (31471516, 31401465, 31400074, 31501332), the China Post Doctoral Project (2015M581419), “Dongnongxuezheshi project (To Chen QS)” “Young Talent project (To Qi ZM, 518062)” of Northeast agriculture university, SIPT Project of Northeast Agricultural University (201610224146), Heilongjiang Funds for Distinguished Young Scientists (JC2016004) and the outstanding academic leaders projects of Harbin (2015RQXXJ018).

Author’s contribution Qi ZM, Qi HD and Zhang JZ contributed equally to this research. Che DD and Chen QS designed and conducted field tests and drafted the manuscript. Qi HD, Zhang XY, Xin DW and Yin ZG provided field management and protein test data. Qi ZM, Jiang HW, Zhang ZG, Zhu RS, Liu CY, Han X, Zhang JZ and Hu ZB provided the data mapping analysis, Qi ZM, Chen QS and Che DD designed the experiment, organized manuscript. All authors read and approved the final manuscript. Authors state that the experiments comply with the current laws of the country in which they were performed.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Allard RW (1996) Genetic basis of the evolution of adaptedness in plants. *Euphytica* 92:1–11. doi:[10.1007/BF00022822](https://doi.org/10.1007/BF00022822)
- Bajaj D, Upadhyaya HD, Khan Y, Das S, Badoni S, Shree T, Kumar V, Tripathi S, Gowda CL, Singh S, Sharma S, Tyagi AK, Chattopdhyay D, Parida SK (2015) A combinatorial approach of comprehensive QTL-based comparative genome mapping and transcript profiling identified a seed weight-regulating candidate gene in chickpea. *Sci Rep* 5:9264. doi:[10.1038/srep09264](https://doi.org/10.1038/srep09264)
- Bilyeu KD, Palavalli L, Slep DA, Beuselinck PR (2003) Three microsomal omega-3-fatty acid desaturase genes contribute to soybean linolenic acid levels. *Crop Sci* 43(5):1833–1838
- Bolon YT, Joseph B, Cannon SB, Graham MA, Diers BW, Farmer AD, May GD, Muehlbauer GJ, Specht JE, Tu ZJ, Weeks N, Xu WW, Shoemaker RC, Vance CP (2010) Complementary genetic and genomic approaches help characterize the linkage group I seed protein QTL in soybean. *BMC Plant Biol* 10:41
- Brim CA, Burton JW (1979) Recurrent selection in soybeans: II: selection for increased percent protein in seeds. *Crop Sci* 19:494–498
- Brunner EC, Graef GL, Orf J, Wilcox JR, Shoemaker RC (1997) Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Sci* 37(2):370–378
- Carlborg Ö, Haley CS (2004) Epistasis: too often neglected in complex trait studies? *Nat Rev Genet* 5:618–625
- Chapman NH, Bonnet J, Grivet L, Lynn J, Graham N, Smith R, Sun GP, Walley PG, Poole M, Causse M, King GJ, Baxter C, Seymour GB (2012) High-resolution mapping of a fruit firmness-related quantitative trait locus in tomato reveals epistatic interaction associated with a complex combination locus. *Plant Physiol* 159:1644–1657
- Chen QS, Zhang ZC, Liu CY, Wang WQ, Li WB (2005) Construction and analysis of soybean genetic map using recombinant inbred line of Charleston × dongnong 594. *Sci Agric Sin* 38:1312–1316
- Choi IY, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, Hwang EY, Yi SI, Young ND, Shoemaker RC, van Tassell CP, Specht JE, Cregan PB (2007) A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics* 176:685–696
- Cober ER, Voldeng HD (2000) Developing high-protein, high-yield soybean populations and lines. *Crop Sci* 40:39–42
- Cooper RL, Martin RJ, St Martin SK, Calip-DuBois A, Fioritto RJ, Schmitthenner AF (1995) Registration of ‘Charleston’ soybean. *Crop Sci* 35(2):593–593
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. *Crop Sci* 39:1464–1490. doi:[10.2135/cropsci1999.3951464x](https://doi.org/10.2135/cropsci1999.3951464x)
- Deng Z, Hu S, Chen F, Li W, Chen J, Sun C, Zhang Y, Wang S, Song X, Tian J (2015) Genetic dissection of interaction between wheat protein and starch using three mapping populations. *Mol Breed* 35:1–9. doi:[10.1007/s11032-015-0216-6](https://doi.org/10.1007/s11032-015-0216-6)
- Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. *Theor Appl Genet* 83:608–612
- Eshed Y, Zamir D (1996) Less-than-additive epistatic interactions of quantitative trait loci in tomato. *Genetics* 143:1807–1817
- Eskandari M, Cober ER, Rajcan I (2013) Genetic control of soybean seed oil: II. QTL and genes that increase oil

- concentration without decreasing protein or with increased seed yield. *Theor Appl Genet* 126:1677–1687
- Garwin JL, Klages AL, Cronan JE (1980) Beta-ketoacyl-acyl carrier protein synthase II of *Escherichia coli*. Evidence for function in the thermal regulation of fatty acid synthesis. *J Biol Chem* 255(8):3263–3265
- González AM, Yuste-Lisbona FJ, Rodiño AP, De Ron AM, Capel C, García-Alcázar M, Lozano R, Santalla M (2015) Uncovering the genetic architecture of *Colletotrichum lindemuthianum* resistance through QTL mapping and epistatic interaction analysis in common bean. *Front Plant Sci* 6:141. doi:10.3389/fpls.2015.00141
- Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315–324. doi:10.1038/hdy.1992.131
- Heath RJ, Jackowski S, Rock CO (1994) Guanosine tetraphosphate inhibition of fatty acid and phospholipid synthesis in *Escherichia coli* is relieved by overexpression of glycerol-3-phosphate acyltransferase (plsB). *J Biol Chem* 269(42):26584–26590
- Hwang EY, Song Q, Jia G, Specht JE, Hyten DL, Costa J, Cregan PB (2014) A genome-wide association study of seed protein and oil content in soybean. *BMC Genom* 15(1):1
- Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, Schmidt E (2004) Seed quality QTL in a prominent soybean population. *Theor Appl Genet* 109:552–561
- Hyten DL, Choi I, Song Q, Specht JE, Carter TE, Shoemaker RC, Hwang E, Matukumalli LK, Cregan PB (2010) 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* 50:960–968
- Jansen RC (1993) Interval mapping of multiple quantitative trait loci. *Genetics* 135:205–211
- Jansen RC, Van Ooijen JW, Stam P, Lister C, Dean C (1995) Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. *Theor Appl Genet* 91:33–37. doi:10.1007/BF00220855
- Jia Y, Sun X, Sun J, Pan Z, Wang X, He S, Xiao S, Shi W, Zhou Z, Pang B, Wang L, Liu J, Ma J, Du X, Zhu J (2014) Association mapping for epistasis and environmental interaction of yield traits in 323 cotton cultivars under 9 different environments. *PLoS ONE* 9:e95882. doi:10.1371/journal.pone.0095882
- Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. *Genetics* 152:1203–1216
- Keim P, Diers BW, Olson TC, Shoemaker RC (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* 126:735–742
- Korir PC, Qi B, Wang Y, Zhao T, Yu D, Chen S, Gai J (2011) A study on relative importance of additive, epistasis and unmapped QTL for aluminium tolerance at seedling stage in soybean. *Plant Breed* 130:551–562. doi:10.1111/j.1439-0523.2011.01862.x
- Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181. doi:10.1016/0888-7543(87)90010-3
- Lark KG, Weisemann JM, Matthews BF, Palmer R, Chase K, Macalma T (1993) A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: ‘Minsoy’ and ‘Noir 1’. *Theor Appl Genet* 86:901–906. doi:10.1007/BF00211039
- Lee S, Freewalt KR, McHale LK, Song QJ, Jun TH, Michel AP, Dorrance AE, Rouf MA (2015) A high-resolution genetic linkage map of soybean based on 357 recombinant inbred lines genotyped with BARCSoySNP6K. *Mol Breed* 35(2):1–7
- Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chantonnat A (2013) ESTHER, the database of the α/β -hydrolyase fold superfamily of proteins: tools to explore diversity of functions. *Nucl Acids Res* 41(D1):D423–D429
- Li H, Zhao T, Wang Y, Yu D, Chen S, Zhou R, Gai J (2011) Genetic structure composed of additive QTL, epistatic QTL pairs and collective unmapped minor QTL conferring oil content and fatty acid components of soybeans. *Euphytica* 182(1):117–132
- Liu G, Yang J, Xu H, Zhu J (2007) Influence of epistasis and QTL \times environment interaction on heading date of rice (*Oryza sativa* L.). *J Genet Genom* 34:608–615. doi:10.1016/S1673-8527(07)60069-1
- Liu YT, Li CY, Shi XX, Feng H, Wang YG (2016) Identification of QTLs with additive, epistatic, and QTL \times environment interaction effects for the bolting trait in *Brassica rapa* L. *Euphytica*. doi:10.1007/s10681-016-1710-6
- Luo YX, Luo CY, Du DZ, Fu Z, Yao YM, Xu CC, Zhang HS (2014) Quantitative trait analysis of flowering time in spring rapeseed (*B. napus* L.). *Euphytica* 200(3):321–335
- Ma XQ, Tang JH, Teng WT, Yan JB, Meng YJ, Li JS (2007) Epistatic interaction is an important genetic basis of grain yield and its components in maize. *Mol Breed* 20(1):41–51
- Mao T, Jiang Z, Han Y, Teng W, Zhao X, Li W (2013) Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. *Plant Breed* 132(6):630–641
- Mackay TF (2014) Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat Rev Genet* 15:22–33. doi:10.1038/nrg3627
- Mackay TF, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* 10:565–577. doi:10.1038/nrg2612
- Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci* 36(5):1327–1336
- McCouch SR, Cho YG, Yano M, Paul E, Blinstrub M, Morishima H, Kinoshita T (1997) Report on QTL nomenclature. *Rice Genet Newsl* 14:11–13
- Merah O, Langlade N, Alignan M, Roche J, Pouilly N, Lippi Y, Vear F, Cerny M, Bouniols A, Mouloungui Z, Vincourt P (2012) Genetic analysis of phytosterol content in sunflower seeds. *Theor Appl Genet* 125:1589–1601
- Mohan A, Kulwal P, Singh R, Kumar V, Mir RR, Kumar J, Prasad M, Balyan HS, Gupta PK (2009) Genome-wide

- QTL analysis for pre-harvest sprouting tolerance in bread wheat. *Euphytica* 168:319–329. doi:[10.1007/s10681-009-9935-2](https://doi.org/10.1007/s10681-009-9935-2)
- Muerhoff AS, Griffin KJ, Johnson EF (1992) The peroxisome proliferator-activated receptor mediates the induction of CYP4A6, a cytochrome P450 fatty acid omega-hydroxylase, by clofibrilic acid. *J Biol Chem* 267(27):19051–19053
- Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, Lark KG (1999) Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. *Crop Sci* 39(6):1642–1651
- Palomeque L, Liu LJ, Li W, Hedges B, Cober ER, Rajcan I (2009) QTL in mega-environments: I. Universal and specific seed yield QTL detected in a population derived from a cross of high-yielding adapted × high-yielding exotic soybean lines. *Theor Appl Genet* 119:417–427. doi:[10.1007/s00122-009-1049-7](https://doi.org/10.1007/s00122-009-1049-7)
- Palomeque L, Liu LJ, Li W, Hedges BR, Cober ER, Smid MP, Lukens L, Rajcan I (2010) Validation of mega-environment universal and specific QTL associated with seed yield and agronomic traits in soybeans. *Theor Appl Genet* 120:997–1003. doi:[10.1007/s00122-009-1227-7](https://doi.org/10.1007/s00122-009-1227-7)
- Qi Z, Wu Q, Han X, Sun Y, Du X, Liu C, Jiang H, Hu G, Chen Q (2011) Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. *Euphytica* 179(3):499–514
- Qi ZM, Huang L, Zhu RS, Xin DW, Liu CY, Han X, Jiang HW, Hong WG, Hu GH, Zheng HK, Chen QS (2014a) A high-density genetic map for soybean based on specific length amplified fragment sequencing. *PLoS ONE* 9(8):e104871. doi:[10.1371/journal.pone.0104871](https://doi.org/10.1371/journal.pone.0104871)
- Qi ZM, Han X, Hou M, Xin DW, Wang ZY, Zhu RS, Hu ZB, Jiang HW, Li CD, Liu CY, Hu GH, Chen QS (2014b) QTL analysis of soybean oil content under 17 environments. *Can J Plant Sci* 94:245–261
- Qi ZM, Pan JB, Han X, Qi HD, Xin DW, Wei Li, Mao XR, Wang ZY, Jiang HW, Liu CY, Hu ZB, Hu GH, Zhu RS, Chen QS (2016) Identification of major QTLs and epistatic interactions for seed protein concentration in soybean under multiple environments based on a high-density map. *Mol Breed* 36:55. doi:[10.1007/s11032-016-0475-x](https://doi.org/10.1007/s11032-016-0475-x)
- Rakotomanga M, Saint-Pierre-Chazale M, Loiseau PM (2005) Alteration of fatty acid and sterol metabolism in miltefosine-resistant *Leishmania donovani* promastigotes and consequences for drug-membrane interactions. *Antimicrob Agents Chemother* 49(7):2677–2686
- Ravi K, Vadez V, Isobe S, Mir R, Guo Y, Nigam S, Gowda M, Radhakrishnan T, Bertoli D, Knapp S, Varshney R (2011) Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). *Theor Appl Genet* 122:1119–1132
- Rodolphe F, Lefort M (1993) A multi-marker model for detecting chromosomal segments displaying QTL activity. *Genetics* 134:1277–1288
- Rossi ME, Orf JH, Liu LJ, Dong Z, Rajcan I (2013) Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian × Chinese crosses. *Theor Appl Genet* 126:1809–1823
- Salas JJ, Ohlrogge JB (2002) Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Arch Biochem Biophys* 403(1):25–34
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183
- Sebolt AM, Shoemaker RC, Diers BW (2000) Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Sci* 40(5):1438–1444
- Shi J, Gonzales RA, Bhattacharyya MK (1996) Identification and Characterization of an S-Adenosyl-L-methionine:-Sterol-C-methyltransferase cDNA from Soybean. *J Biol Chem* 271(16):9384–9389
- Shibata M, Takayama K, Ujiie A, Yamada T, Abe J, Kitamura K (2008) Genetic relationship between lipid content and linolenic acid concentration in soybean seeds. *Breed Sci* 58(4):361–366
- Shoemaker RC, Olson TC (1993) Molecular linkage map of soybean [*Glycine max* L.] Merr]. In: Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Publisher, New York, pp 131–138
- Song WC, Funk CD, Brash AR (1993) Molecular cloning of an allene oxide synthase: a cytochrome P450 specialized for the metabolism of fatty acid hydroperoxides. *Proc Natl Acad Sci* 90(18):8519–8523
- Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109:122–128
- Song QJ, Jenkins J, Jia GF, Hyten DL, Pantalone V, Jackson SA, Schmutz J, Cregan PB (2016) Construction of high resolution genetic linkage maps to improve the soybean genome sequence assembly Glyma1.01. *BMC Genom* 17(1):1
- Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, Germann M, Orf JH, Lark KG (2001) Soybean response to water: a QTL analysis of drought tolerance. *Crop Sci* 41:493–509
- Tajuddin T, Satoshi W, Naoki Y, Kyuya H (2003) Analysis of quantitative trait loci for protein and lipid contents in soybean seeds using recombinant inbred lines. *Breed Sci* 53:133–140
- Todd J, Post-Beittenmiller D, Jaworski JG (1999) KCS1 encodes a fatty acid elongase 3-ketoacyl-CoA synthase affecting wax biosynthesis in *Arabidopsis thaliana*. *Plant J* 17(2):119–130
- Wang F, Guo CY (2010) Molecular mapping and identification of quantitative trait loci for yield components in rapeseed (*Brassica napus* L.). *Hereditas* (Beijing) 32(3):271–277
- Wang CS, Rutledge JJ, Gianola D (1994) Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genet Sel Evol* 26:91–115
- Wang SC, Basten CJ, Zeng ZB (2012a) Windows QTL cartographer 2.5 department of statistics, North Carolina State

- University, Raleigh, N.C. <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>. Accessed 01 Aug 2012
- Wang Z, Chen Z, Cheng J, Lai Y, Wang J et al (2012b) QTL analysis of Na⁺ and K⁺ concentrations in roots and shoots under different levels of NaCl stress in rice (*Oryza sativa* L.). *PLoS ONE* 7:e51202
- Wang L, Cheng JP, Lai YY, Du WL, Huang X, Wang ZF, Zhang HS (2014a) Identification of QTLs with additive, epistatic and QTL × development interaction effects for seed dormancy in rice. *Planta* 239:411–420
- Wang X, Jiang GL, Green M, Scott RA, Song Q, Hyten DL, Cregan PB (2014b) Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean. *Mol Genet Genom* 289(5):935–949
- Wang ML, Khera P, Pandey MK, Wang H, Qiao L, Feng S, Tonniss B, Barkley NA, Pinnow D, Holbrook CC, Culbreath AK, Varshney RK, Guo B (2015) Genetic mapping of QTLs controlling fatty acids provided insights into the genetic control of fatty acid synthesis pathway in peanut (*Arachis hypogaea* L.). *PLoS ONE* 10(4):e0119454
- Wilcox JR (1985) Breeding soybeans for improved oil quantity and quality. R. Shibles, ed. In: *Proc 3rd World Soybean Res. Con.* Westview Press, Boulder, CO, pp 380–386
- Wilcox JR (1998) Increasing seed protein in soybean with eight cycles of recurrent selection. *Crop Sci* 38(6):1536–1540
- Wilson RF (2004) Seed composition. In: Boerma HR, Specht JE (eds) *Soybeans: improvement, production and uses*, 3rd edn. ASA, CSSA, SSA, Madison, pp 621–677
- Wilson RF (2008) Soybean: market driven research needs. In: Stacey G (ed) *Genetics and genomics of soybean*. Springer, New York, pp 3–15
- Wu X, Chang X, Jing R (2012) Genetic insight into yield-associated traits of wheat grown in multiple rain-fed environments. *PLoS ONE* 7:e31249
- Würschum T, Maurer HP, Dreyer F, Reif JC (2013) Effect of inter- and intragenic epistasis on the heritability of oil content in rapeseed (*Brassica napus* L.). *Theor Appl Genet* 126(2):435–441
- Xie W, Feng Q, Yu H, Huang X, Zhao Q, Xing Y, Yu S, Han B, Zhang Q (2010) Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. *Proc Natl Acad Sci USA* 107:10578–10583
- Xin DW, Qi ZM, Jiang HW, Hu ZB, Zhu RS, Hu JH, Han HY, Hu GH, Liu CY, Chen QS (2016) QTL location and epistatic effect analysis of 100-Seed weight using wild soybean (*Glycine soja* Sieb. & Zucc.) chromosome segment substitution lines. *PLoS ONE* 11(3):e0149380
- Xing GN, Zhou B, Wang YF, Zhao TJ, Yu DY, Chen SY, Gai JY (2012) Genetic components and major QTL confer resistance to bean pyralid (*Lamprosema indicata* Fabricius) under multiple environments in four RIL populations of soybean. *Theor Appl Genet* 125:859–875
- Xu YB, Crouch JH (2008) Marker-assisted selection in plant breeding: from publication to practice. *Crop Sci* 48:391–407
- Xu P, Wu XH, Wang BG, Hu TT, Lu ZF, Liu YH, Qin DH, Wang S, Li GJ (2013) QTL mapping and epistatic interaction analysis in asparagus bean for several characterized and novel horticulturally important traits. *BMC Genet* 14:4
- Xu JF, Long Y, Wu JG, Xu HM, Zhao ZG, Wen J, Meng JL, Shi CH (2015) QTL identification on two genetic systems for rapeseed glucosinolate and erucic acid contents over two seasons. *Euphytica*. 205(3):647
- Yang J, Zhu J (2005) Methods for predicting superior genotypes under multiple environments based on QTL effects. *Theor Appl Genet* 110:1268–1274
- Yang J, Zhu J, Williams RW (2007) Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics* 23:1527–1536
- Yang J, Hu C, Hu H, Yu R, Xia Z, Ye X, Zhu J (2008) QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics* 24:721–723
- Yang C, Tang D, Zhang L, Liu J, Rong T (2015) Identification of QTL for ear row number and two-ranked versus many-ranked ear in maize across four environments. *Euphytica* 5:1–15
- Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457–1468
- Zhang YH, Liu MF, He JB, Wang YF, Xing GN, Li Y, Yang SP, Zhao TJ, Gai JY (2015) Marker-assisted breeding for transgressive seed protein content in soybean [*Glycine max* (L.) Merr.]. *Theor Appl Genet* 128:1061–1072
- Zhao G, Wang J, Han Y, Teng W, Sun G, Li W (2008) Identification of QTL underlying the resistance of soybean to pod bore *Leguminivora glycinivorella* (Mats.) obratzsov, and correlation with plant, pod and seed traits. *Euphytica* 164:275–282