RESEARCH ARTICLE Legume Research- An International Journal, Volume 46 Issue 1: 18-24 (January 2023)

Genetic Analysis and Quantitative Trait Locus Mapping using the Major Gene Plus Polygene Model for Soybean [*Glycine max* (L.) Merr.] Main Quality Traits

Dan Yao[1#], Qiuzhu Zhao[1#], Tingting Li[1], Jiaxin Wang[1], Lixue Wang[1], Yixuan Liu[1], Wenyuan Hao[2], Huijing Liu[1]

10.18805/LRF-705

ABSTRACT

Background: Soybean is an important oil and cash crop, soybean protein is the main source of protein for human consumption. Soybean oil is also major edible oil that is rich in unsaturated fatty acids and is more conducive to human long-term consumption than animal oil. An important research direction of current soybean breeding is improvement of the protein and fat content using marker assisted selection and many reports of associated quantitative trait locus (QTL) mapping have been documented. Until to now, a total of 241 QTLs associated with protein content and 315 related to fat content have been collected on the SoyBase website.

Methods: In this study, 236 F_2 generation plants and a derivative group were constructed by using Jiyu50 and Jinong18, obtained from Jilin Province. Combining three years of phenotypic and molecular detection data, using the ICIM method, one-dimensional scanning detected 24 QTLs related to protein and fat content, two-dimensional ICIM analysis of $F_{2:3}$ families detected 7 pairs of epistatic QTLs associated with fat content.

Result: Based on a mixed model for major genes and polygenes, the C model was determined the optimal genetic model for protein and fat content, with genetic rates of multiple genes of 71.15% and 79.15%, respectively. In the detection of high-oil molecular markers, the detection coincidence degree was Sat_238 > Satt 100> Satt 150 > Satt 636> Sat_287 = Sat_342, while this was Satt150 = Sat_342> Satt100 > Satt 636> Sat_287> Sat_238 for high-protein markers. The QTL localization results showed that 9 microeffect QTLs related to protein content and 7 microeffect QTLs related to fat content were detected. Isolation analysis results were generally similar to those of QTL mapping. One stable QTL associated with protein content and two stable QTLs associated with fat content were identified as being of some application value in soybean molecular marker-assisted breeding.

Key words: Fat content, *Glycine max*, Main gene, Multiple gene, Protein content, Quantitative trait loci.

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] is an important oil and cash crop, with a grain protein content of about 40% and a fat content of 20%. Soybean protein is the main source of protein for human consumption and contains eight essential amino acids (Li *et al.* 1986). Soybean oil is also a major edible oil that is rich in unsaturated fatty acids and is more conducive to human long-term consumption than animal oil (Lin 2013). An important research direction of current soybean breeding is an improvement of the protein and fat content and many reports of associated quantitative trait locus (QTL) mapping have been documented. For example, Nine QTLs associated with fat content and eight related to protein content were mapped (Diers *et al.* 1992), while in 1996, six Restriction fragment length polymorphism (RFLP) sites associated with soybean oil content and thirteen RFLP sites associated with protein content were mapped (Lee *et al.* 1996). One QTL relevant to protein content and two relevant to fat content were detected following a cross between Essex × Williams and mapped to M, L, A1 and D2 linkage groups (Chapman *et al.* 2003). QTLs related to soybean protein and fat content were located in B2, D1a, N, E and other linkage groups using $F_{2:10}$ inbred lines from a Charleston × Dong nong 594 cross (Chen *et al.* 2007). Four QTLs related to protein content were detected at two sites by a Qi huang 26 × smooth leather cross which formed 170 F_2 segregation populations; the QTLs were distributed in D2, E and K linkage groups (Lin *et al.* 2010). Twenty five QTLs related to protein content were detected in the combined inbred lines of Charleston × Dongnong 594 and distributed in A1, B1,

[1]Department of College of Life Sciences, Jilin Agricultural University, Changchun 130118, China.
[2]Jilin Provincial Academy of Agricultural Sciences, Jilin, Changchun 130118, China.
[#]These authors contributed equally to this work.

Corresponding Authors: Huijing Liu and Wenyuan Hao, Department of College of Life Sciences, Jilin Agricultural University, Changchun 130118, China; Jilin Provincial Academy of Agricultural Sciences, Jilin, Changchun 130118, China.
Email: huijingliu@jlau.edu.cn; wenyuan_h@163.com

How to cite this article: Yao, D., Zhao, Q., Li, T., Wang, J., Wang, L., Liu, Y., Hao, W. and Liu, H. (2023). Genetic Analysis and Quantitative Trait Locus Mapping Using the Major Gene Plus Polygene Model for Soybean [*Glycine max* (L.) Merr.] Main Quality Traits. Legume Research. 46(1): 18-24. doi: 10.18805/LRF-705.

Submitted: 20-06-2022 **Accepted:** 19-09-2022 **Online:** 31-10-2022

C2, D1a, D1b, E, J, L and N linkage groups (Shan *et al.* 2011). Seven QTLs associated with protein content and six related to fat content were detected (Pathan *et al.* 2013). Two QTLs associated with protein content were detected with a phenotypic variation of 1.92%~2.03% and two related to fat content demonstrated phenotypic variations of 2.56%-6.98% (Li *et al.* 2015). Five QTLs associated with protein content and nine related to fat content were detected (Zhu *et al.* 2017). Until now, a total of 241 QTLs associated with protein content and 315 related to fat content have been collected on the SoyBase website (http://soybase.org/). Of these, seven Simple Sequence Repeats (SSRs) related to protein content (Satt 100, T155-1, Satt 578, Satt 277, A 454-1, A 144-1 and Satt 567) were mapped more than twice and distributed in C2, A1, C1, C2, I, E and M linkage groups, while nine SSRs related to fat content (Satt 277, T 155-1, Satt 100, A 111-1, A 063-1, Satt 468, A 566-2, Satt 270 and Satt 398) were also mapped more than twice and distributed in C2, A1, C2, A2, C1, D1a, H, I and L linkage groups (Hou *et al.* 2014).

As can be seen from the above examples, many QTLs are known to be related to soybean protein and fat content and they are scattered throughout most of the linkage groups. However, the number and location of QTLs detected in populations of different genetic backgrounds differ significantly (Zhen *et al.* 2011). Therefore, the analysis of QTL stability among different generations will help improve the efficiency of soybean protein and the fat content selection and accelerate the breeding process (Shen *et al.* 2001). In this study, QTL mapping of soybean protein and fat content was carried out using the ICIM method of QTL IciMapping v3.0 software and F_2 and $F_{2:3}$, $F_{3:4}$ derived populations were obtained by the hybridization of soybean Ji Yu 50 and Jinong 18. The genetic development of soybean protein and fat content was analyzed using the major gene and polygene mixed genetic model. This will provide the theoretical basis for the selection of molecular markers in the breeding of high-quality soybean varieties.

MATERIALS AND METHODS

Construction of soybean segregation group

In this study, soybean variety Jiyu 50 (female parent, Ji Shen bean 2001015) was hybridized with Jinong 18 (male parent, Ji Shen bean 2006) in the experimental field of Jilin Agricultural University in the summer of 2013 to obtain the F_0 generation. In October 2015, an F_1 individual plant with 450 seeds was obtained. In October 2016, 236 F_2 individual plants were obtained. In October 2017, F_3 strains were obtained from 236 individual plants from the F_2 generation. In October 2018, $F_{3:4}$ strains were obtained from the F_3 strains.

Primers

According to the soybean public genetic map published (Song *et al.* 2004), 380 pairs of SSR primers were primarily confirmed from the soybean database SoyBase (http://soybase.agron.iastate.edu) and synthesized by Changchun KuMei Co., Ltd. (Changchun, China).

Genetic analysis methods

Genetic analysis of quality traits was carried out using the primary gene and polygene mixed genetics model with the five generation joint separation analysis method (P_1, F_1, P_2, F_2 and $F_{2:3}$) (Wang *et al.* 1998).

Soybean protein and fat content

The soybean protein and fat content of Jiyu 50, Jinong 18, the F_2 generation and their derivative groups was determined by a near infrared grain quality analyzer (BUCHI NIRLab N-200 MCS 100) (Switzerland) between 2016~2018.

DNA extraction and SSR analyses

Genomic DNA of the soybean was isolated from the leaf tissue by the CTAB method (Wang and Fang, 2002; Li *et al.*, 2017). Polymerase chain reaction (PCR) was performed in a 15 µL volume containing 0.6 µL genomic DNA (50 ng/µL), 0.6 µL dNTP mixtures (10 mM), 0.6 µL SSR primer (25 µM), 1.5 µL 10X PCR buffer (contain Mg^{2+}), 0.15 µL Taq polymerase (5 U/µL) and 10.95 µL double-distilled water. The PCR conditions were 4 min at 94°C; followed by 35-40 cycles of 45 s at 94°C, 30 s at 50°C and 30 s at 72°C; then 8 min at 72°C (Nazima *et al.*, 2018). After amplification, the PCR products were mixed with a loading buffer, denatured for 5 min at 94°C and kept at 0°C. The denatured PCR products were separated on 8% (w/v) denaturing polyacrylamide gel and visualized by silver staining (Sanguinetti *et al.*, 1994).

Construction of the molecular genetic map

PCR amplification banding patterns identical to those of the male parent were recorded as “1”, those identical to the female parent were recorded as “2”, heterozygous banding patterns as “3” and missing banding patterns as “–”. Mapmaker Exp 3.0 software (The Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Cambridge, MA, USA) was used for map construction with the “Group” command to perform interlocking analysis and grouping of markers (Wang 2009). If the number of linkage markers was less than 8, the “Compare” command was used to sort and if the number was greater than 8, the “Ripple” command was used to sort (Schneider *et al.* 1997, Liu *et al.* 2000). The error detection level was set at 1% and the recombination rate was converted to genetic distance (cM) using the Kosambi function. Win QTL Cart 2.5 software (Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA) was used to draw the genetic linkage map.

Mapping quantitative trait loci

The Inclusive Composite Interval Mapping algorithm was applied to determine the protein and fat content of the F_2 and F_3 segregation groups by QTL IciMapping v3.0 software (Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing, China). The scanning step was 1.0 cM. The probabilistic levels of stepwise regression variables were 0.01 and 0.02, respectively. LOD≥2.0

(LR = 9.66) was used as the threshold for QTL mapping and effect estimation (Zhang *et al.* 2008).

RESULTS AND DISCUSSION

Soybean protein and fat content in F_2 population

The protein and fat content was shown to vary greatly between the parents and to have a near normal distribution and a wide distribution frequency in F_2 population (Fig 1); this was typical of a quantitative genetic model. Protein content variation was 37.29%~44.50% and the average value was 40.90% in F_2. The protein content is the separation of mid parent, biased towards the female parent. Regarding fat content in the F_2 group, the variation was 17.31%~23.34% and the average value was 19.73% (Table 1). The fat content is also the separation of mid parent, biased towards the female parent. Super parent isolated single plants were detected in the offspring for both protein and fat content.

Genetic analysis of soybean protein and fat content

Based on analysis of the major gene plus polygene genetic model, the likelihood function and the Akaike information criterion (AIC) value of the protein and fat content under different genetic models were obtained by the IECM algorithm. According to the principle of minimum AIC value, C and E-2 were preliminarily determined as alternative models for protein content, C and E-1 as alternative models for fat content.Further fitness test results showed that the C model (polygenic genetic model) was the most suitable model for protein and fat content (Supplementary Table 1 and Table 2) and the polygenic heritability was 71.15% and 79.15%, respectively.

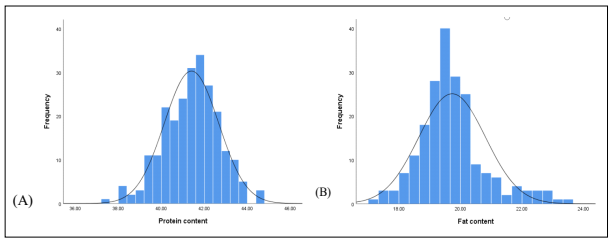

Fig 1: Histogram of soybean protein and fat content in F_2. (A) Frequency distribution of soybean protein content (B) Frequency distribution of soybean fat content.

QTL mapping of soybean protein and fat content

In this study, 380 pairs of SSR primers were used to screen the polymorphic primers between the parents. The results showed that 118 pairs of primers showed polymorphism in the parents and the polymorphism rate was 31.05%. SSR-PCR was carried out on 236 individual plants of F_2 isolated population using polymorphic primers. Finally, a SSR linked genetic map containing 102 markers was constructed. Fourteen QTLs related to protein content were detected, which were distributed in six linkage groups, including 4 (C2), 6 (A1), 12 (G), 13 (C1), 17 (M) and 22 (F). Among them, four QTLs had a phenotypic variation of more than 20% and one was stable in 2 years. Ten QTLs related to fat content were detected that were distributed in five linkage groups, including 1 (A1), 4 (C2), 12 (G), 17 (M) and 22 (F). Among them, one stable QTL in three continuous years was detected in the 12 (G) linkage group Sat_287~Sat_342 marker interval and one stable QTL in two continuous years was detected in the (C2) linkage group Satt100~Sat_238 marker interval. Additionally, three major QTLs related to soybean fat content were also detected (Supplementary Table 3 and Fig 2).

Table 1: Distribution of soybean protein and fat content in parent plants and F_2.

Traits	Jiyu 50	Jinong 18	Average value	Standard variance	Max	Min	Skewness	Kurtosis
Protein content %	42.45	37.52	41.41	1.28	44.50	37.29	-0.29	0.11
Fat content %	19.58	23.79	19.73	1.09	23.34	17.31	1.07	1.83

Supplementary Table 1: Suitability test for protein content genetic models.

Model	Generations	U_1^2	U_2^2	U_3^2	${}_nW^2$	D_n
C	P_1	0.00 (0.99)	0.00 (0.98)	0.01 (0.86)	0.24 (>0.05)	0.13 (>0.05)
	F_1	1.26 (0.34)	0.89 (0.21)	0.33 (0.45)	0.22 (>0.05)	0.15 (>0.05)
	P_2	0.00 (1.00)	0.06 (0.81)	0.08 (0.77)	0.14 (>0.05)	0.08 (>0.05)
	F_2	1.68 (0.12)	3.65 (0.09)	2.12 (0.48)	0.22*	0.09
	$F_{2:3}$	3.15 (0.18)	4.12*	2.97 (0.53)	0.16*	0.04 (>0.05)
E-2	P_1	0.03 (0.86)	0.07 (0.69)	0.34 (0.56)	0.33 (>0.05)	0.17 (>0.05)
	F_1	1.51 (0.22)	1.17 (0.45)	0.19 (0.68)	0.39*	0.21 (>0.05)
	P_2	0.07 (0.81)	0.03 (0.95)	0.21 (0.59)	0.28 (>0.05)	0.16 (>0.05)
	F_2	3.97*	4.59*	3.12 (0.06)	0.23*	0.12*
	$F_{2:3}$	7.51**	2.59 (0.12)	4.78*	0.34 *	0.06 (>0.05)

Note: U_1^2, U_2^2, U_3^2: The homogeneity test statistic, the numbers in parentheses are the corresponding probabilities; ${}_nW^2$: Smirnov test statistic; D_n: Kolmogorov test statistic; *: Significant difference at 0.05; **: Significant difference at 0.01 level.

Marker-assisted selection of soybean protein and fat content

High fat and high protein molecular markers were detected using six stable SSR markers (Satt 100, Sat_287, Satt 150, Satt 636, Sat_238 and Sat_342) related to soybean protein and fat content in 108 soybean materials from Biotechnology Center of Jilin Agricultural University. The detection coincidence degree of high-oil molecular markers was (high to low): Sat_238 (95.12%) > Satt 100 (87.80%) > Satt150 (75.61%) > Satt636 (71.95%) > Sat_287 = Sat_342 (68.29%). The detection coincidence degree of high-protein molecular markers was (high to low): Satt150 = Sat_342 (84.00%) > Satt 100 (77.33%) > Satt636 (70.67%) > Sat_287 (52%) > Sat_238 (50.33%) (Table 2 and Fig 3).

Supplementary Table 2: Suitability test for fat content genetic models.

Model	Generations	U_1^2	U_2^2	U_3^2	$_nW^2$	D_n
C	P_1	0.00 (0.99)	0.02 (0.95)	0.44 (0.46)	0.12 (>0.05)	0.08 (>0.05)
	F_1	0.85 (0.31)	1.17 (0.28)	1.51 (0.22)	0.14 (>0.05)	0.07 (>0.05)
	P_2	0.06 (0.81)	0.08 (0.77)	0.03(0.86)	0.21 (>0.05)	0.11 (>0.05)
	F_2	2.42 (0.12)	4.83*	0.71 (0.56)	0.15*	0.06 (>0.05)
	$F_{2:3}$	0.25 (0.62)	0.00 (0.95)	3.96*	0.04 (>0.05)	0.02 (>0.05)
E-1	P_1	0.00 (1.00)	0.00 (0.96)	0.36 (0.78)	0.19 (>0.05)	0.10 (>0.05)
	F_1	2.14 (0.31)	1.21 (0.27)	1.23 (0.64)	0.16 (>0.05)	0.09 (>0.05)
	P_2	0.07 (0.79)	0.07 (0.68)	0.05 (0.72)	0.18 (>0.05)	0.12 (>0.05)
	F_2	5.04*	4.01*	0.50 (0.48)	0.17*	0.11*
	$F_{2:3}$	2.41 (0.12)	2.62 (0.19)	4.22*	0.14*	0.08 (>0.05)

Note: U_1^2, U_2^2, U_3^2: The homogeneity test statistic, the numbers in parentheses are the corresponding probabilities; $_nW^2$: Smirnov test statistic; D_n: Kolmogorov test statistic; *: Significant difference at 0.05; **: Significant difference at 0.01 level.

Supplementary Table 3: QTLs of soybean protein and fat content detected by ICIM-ADD.

Traits	Population	Linkage group	QTL	Marker flanking	QTL position (Front/Back)/cM	LOD value	AE	PVE(%)
Protein	F_2 generation	6(A1)	*qPC-6-1*	Sat_368~ Sat_369	0.9/1.5	2.10	-0.79	4.74
			qPC-6-2	Sat_369~ Satt411	3.5/6.7	2.22	0.86	5.53
		12(G)	*qPC-12-1*	Sat_140~ Satt570	13.0/45.0	2.13	1.10	8.76
			qPC-12-2	Satt199~ Sat_143	20.8/6.7	2.30	-1.61	6.66
			qPC-12-3	Sat_143~ Sat_064	11.3/32.9	4.80	-1.86	24.49
		13(C1)	*qPC-13-1*	Satt190~ Satt396	1.0/9.7	2.32	-0.79	5.04
	$F_{2:3}$ families	17(M)	*qPC-17-1*	Satt636~ Satt150	17.0/17.6	2.74	0.24	8.97
			qPC-17-2	Sat_121~ Satt308	35.9/18.8	2.88	-0.01	5.85
		22(F)	*qPC-22-1*	Satt 505~ Sct-199	36.0/18.7	2.87	-0.01	5.97
			qPC-22-2	Sct-199~ Sat_074	10.3/20.0	2.40	-0.01	2.36
	$F_{3:4}$ families	4(C2)	*qPC-4-1*	Satt 100~ Sat_238	0.1/32.0	2.17	0.70	9.23
		12(G)	*qPC-12-1*	Satt 690~ Sat_140	17.3/56	2.56	0.21	36.65
		17(M)	*qPC-17-1*	Satt 636~ Satt150	15.0/19.6	2.80	-0.93	25.25
		22(F)	*qPC-22-1*	Satt 595~ Sat_298	18.1/16.5	2.97	-0.93	25.30
Fat	F_2 generation	4(C2)	*qOC-4-1*	Satt 100~ Sat_238	11.4/20.7	2.20	0.86	5.20
		12(G)	*qOC-12-1*	Sat_287~ Sat_342	18.0/36.9	3.45	-0.82	31.78
	$F_{2:3}$ families	1(A1)	*qOC-1-1*	Satt 593~ Satt599	19.0/29.5	2.31	0.42	7.05
		4(C2)	*qOC-4-1*	Satt 100~ Sat_238	16.4/15.7	2.11	0.45	9.31
		12(G)	*qOC-12-1*	Sat_287~ Sat_342	15.0/39.9	2.01	-0.74	20.34
		17(M)	*qOC-17-1*	Satt636~ Satt150	12.0/22.6	2.27	-0.39	5.42
			qOC-17-2	Sat_121~ Satt308	27.9/26.8	2.54	-0.21	3.65
		22(F)	*qOC-22-1*	Satt 505~ Sct_199	27/27.4	2.52	-0.21	3.72
			qOC-22-2	Satt 595~ Sat_298	21.1/13.5	2.46	-0.40	5.98
	$F_{3:4}$ families	12(G)	*qOC-12-1*	Sat_287~ Sat_342	13.0/41.9	3.64	-1.22	37.60

Note: Linkage group in capital letters in parentheses represents the public genetic map of the linkage group number; QTL position (front /rear): “front” means the first marker distance away from the marker interval, “rear” means the second marker distance away from the marker interval. A new QTL nomenclature was established by the American Soybean Genetic Association in 2004.

Comparison of major gene plus polygene model analysis with QTL mapping for soybean main quality traits

The quantitative analysis of genetic traits and molecular marker loci can be carried out on the genetic traits of quantitative traits and the results of the two analyses can be used to confirm each other. The inheritance of protein content was shown to follow a polygene genetic model according to the results of the fifth generation (P_1, F_1, P_2, F_2 and $F_{2:3}$) populations. Therefore, it should be possible to locate several QTLs with LODs of similar sizes. QTL mapping showed that nine minor QTLS with similar LOD sizes were detected in soybean F_2 and $F_{2:3}$ populations (phenotypic variation rate <10%) and only one major QTL was detected (phenotypic variation rate >10%) protein content. The inheritance of fat content was also shown to follow a polygene genetic model according to the results of the fifth generation (P_1, F_1, P_2, F_2 and $F_{2:3}$) populations. In the F_2 and $F_{2:3}$ populations, seven minor QTLs with similar LOD sizes were detected and only two major QTLs. In general, the results of the model analysis are similar to those of QTL mapping. Our findings are also consistent with those reported by Xu (2006) and Wang (2001), but differ to those reported by Zheng *et al*. (2007). This may be because the isolation analysis method can only detect genes with strong effects in QTL mapping analysis, while other genes are classified as micro-polygenes. Thus, the number of major genes detected in QTL mapping usually exceeds the number of major genes detected by model analysis, which is consistent with the findings of Wang (2000) and Xu (2006). Additionally, because the population used in our study was the early F_2 and $F_{2:3}$ populations after hybridization, not the stable RIL population, the genetic parameters were less relative to the RIL population and the F_2 data used for statistical analysis did not represent average values.

Table 2: High fat marker selection screening in 108 soybean materials using six SSR markers.

Markers	Numbers of high oil-type plants	Coincidence detection%	Numbers of high protein-type plants	Coincidence detection %
Sat_238	78	95.12	34	50.33
Satt 100	72	87.80	58	77.33
Satt 150	62	75.61	63	84.00
Satt 636	59	71.95	53	70.67
Sat_287	56	68.29	39	52.00
Sat_342	56	68.29	63	84.00

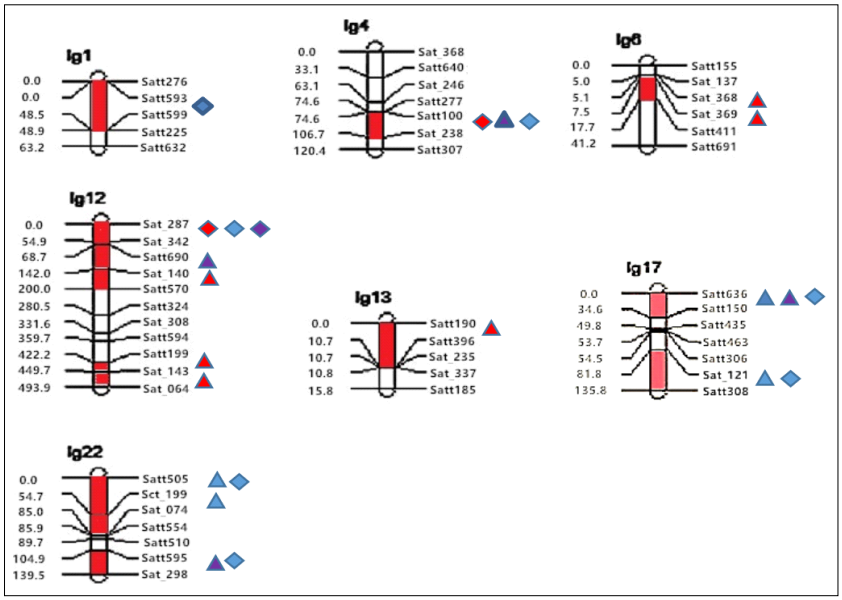

Fig 2: Location of additivity effect QTLs on linkage groups.
Note: Red triangles represent QTL of protein content in F_2 generation.
Blue triangles represent QTL of protein content in $F_{2:3}$ families.
Purple triangles represent QTL of protein content in $F_{3:4}$ families.
Red diamonds represent QTL of fat content in F_2 generation.
Blue diamonds represent QTL of fat content in $F_{2:3}$ families.
Purple diamonds represent QTL of fat content in $F_{3:4}$ families.

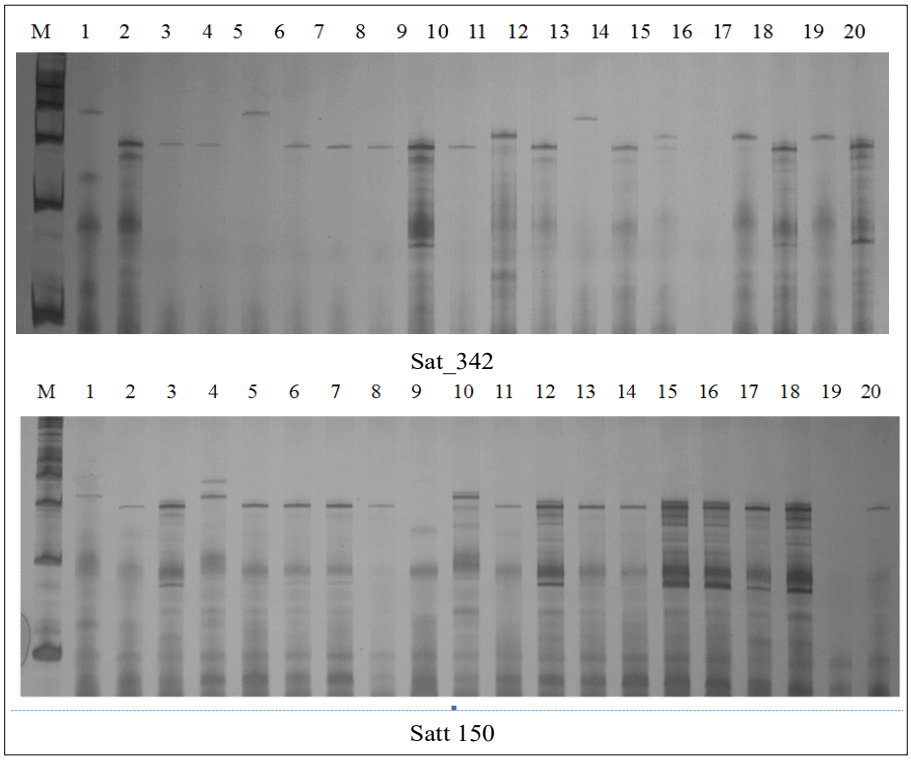

Fig 3: Electrophoresis using Sat_342 and Satt150 primers of soybean materials.
Note: M: Marker; 1: Jiyu 50; 2: Jinong 18; 3–20: selected soybean materials.

Thus, the experimental results were affected by the environment, so should be verified by extending the parental analysis.

Marker-assisted selection of soybean main quality traits

Using the six pairs of SSR markers, which were localized and stable in relation to soybean protein and fat content, 108 soybean seed resources were analyzed for high oil and high protein. The detection coincidence degree of SSR markers exceeded 50%, with a maximum of 95.12%. We found that the Satt100 marker was closely linked to the fat content and was identified as a marker of repeated positioning more than twice, which is consistent with previous studies (Hou *et al.* 2014). The marker related to fat content was stable under different genetic background conditions, indicating that it could be used in marker-assisted selection breeding for high oil and protein content in soybean. The detection coincidence degree was high (>90%) for Sat_238, but was lower (<90%) for the other five markers (Satt100, Satt150, Sat_342, Satt636 and Sat_287). This could be because the protein and fat content are quantitative traits controlled by multigenes. Therefore, future work should further investigate markers that are closely linked with protein and fat content (Yang *et al.* 2008). Additionally, although we identified QTLs that were stable in different generations, we did not verify the stability under different environmental conditions. Further testing and verification is under way to confirm the selection effect of these molecular markers.

CONCLUSION

The inheritance of protein content was shown to follow a polygene genetic model according to the results of the fifth generation (P_1, F_1, P_2, F_2 and $F_{2:3}$) populations. QTL mapping showed that nine minor QTLS with similar LOD sizes were detected in soybean F_2 and $F_{2:3}$ populations (phenotypic variation rate <10%) and only one major QTL was detected (phenotypic variation rate >10%). In the F_2 and $F_{2:3}$ populations, seven minor QTLs with similar LOD sizes were detected and only two major QTLs. In general, the results of the model analysis are similar to those of QTL mapping.

Using the six pairs of SSR markers, which were localized and stable in relation to soybean protein and fat content, 108 soybean seed resources were analyzed for high oil and high protein. The detection coincidence degree of SSR markers exceeded 50%, with a maximum of 95.12%. We found that the Satt100 marker was closely linked to the fat content and was identified as a marker of repeated positioning more than twice. The marker related to fat content was stable under different genetic background conditions, indicating that it could be used in marker-assisted selection breeding for high oil and protein content in soybean.

ACKNOWLEDGEMENT

This work was partially supported by Department of Science and Technology of Jilin Province (20210202006NC). We would like to thank all teachers and graduate students of Jilin Agricultural University Center for Biotechnology for their

support and help with this research. We thank Sarah Williams, PhD, from Liwen Bianji, Edanz Group China (www.liwenbianji.cn), for editing the English text of a draft of this manuscript.

Conflict of interest: None.

REFERENCES

Chapman, A., Pantalone, V.R., Ustun, A., Allen, F.L., Landau-Ellis, D., Rigiano, R.N. and Gresshoff, P.M. (2003). Quantitative trait loci for agronomic and seed qualit traits in an F_2 and $F_{4:6}$ soybean population. Euphytica. 129: 387-393. DOI: 10.10 23/A:1022282726117.

Chen, Q.S., Zhang, Z.C., Liu, C.Y., Xin, D.W., Shan, D.P., Qiu, H.M. and Shan, C.Y. (2007). QTL analysis of major agronomic traits in soybean. Scientia Agricultura Sinica. 1: 41-47.

Diers, B.W., Keim, P., Fehr, W.R. and Shoemarker, R.C. (1992). RFLP analysis of soybean seed protein and oil content. Theoretical and Applied Genetics. 83(5): 608-612. DOI: 10.1007/BF00226905.

Hou, M., Qi, Z.M., Han, X., Xin, D.W., Jiang, H.W., Liu, C.Y., Wu, Q., Sui, L.L., Hu, G.H. and Chen, Q.S. (2014). QTL mapping and interaction analysis of seed protein content and oil content in soybean. Scientia Agricultura Sinica. 47(13): 2680-2689.

Lee, S.H., Bailey, M.A., Mian, M.A.R., CarterJr, T.E., Shipe, E.R., Ashley, D.A., Parrott, W.A., Hussey, R.S. and Boerma, H.R. (1996). RFLP loci associated with soybean seed protein and oil content across populations and lotions. Theoretical and Applied Genetics. 93: 649-657. DOI: 10.1007/bf00224058.

Li, F.H., Chang, R.Z., Shu, S.Z., Chang, B.Y., Chen, Z.P. and Zuo, J.W. (1986). The analysis of protein content and amino acid composition in cultivation, wild, semi-wild soybean. Soybean Science. 1: 65-72.

Lin, J. (2013). Unsaturated fatty physiologic function and its application progress. Guangdong Chemical Industry. 40(6): 92-93.

Lin, Y.H., Zhang, L.W., Li, W., Zhang, L.F. and Xu, R. (2010). QTLs mapping related to protein content of soybean. Soybean Science. 29(2): 207-209.

Liu, F., Zhuang, B.C., Zhang, J.S. and Chen, S.Y. (2000). Construction and analysis of soybean genetic map. Acta Genetica Sinica. 27(11): 1018-1026.

Li, W., Sun, M.M., Zhao, D., Zhang, J.Y., Cao, G.L., Han, Y.P., Li, W.B. and Teng, W.L. (2015). Stability analysis of QTL associated with soybean protein and oil content across different generations. Soybean Science. 34(1): 42-45.

Pathan, S.M., Vuong, T., Clark, K., Lee, J.D., Shannon, J.G., Roberts, C.A., Ellersieckd, M.R. *et al.* (2013). Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean. Crop Science. 53(3): 765-774. DOI: 10.2135/cropsci2012.03.0153.

Sanguinetti, C.J., Neto, E.D. and Simpson, A.J.G. (1994). RAPD silver staining and recovery of pcr products separated on polyacrylamide gels. Biotechniques. 17: 914-921. DOI: 10.1002/bmc.1130080616.

Schneider, K.A., Brothers, M.E. and Kelly, J.D. (1997). Marker assisted selection to improve drought resistance in common bean. Crop Science. 37: 51-60. DOI: 10.2135/cropsci 1997.0011183X003700010008x.

Shan, D.P., Liu, H.Y., Jiang, H.W., Dong, X.H., Chen, Q.S. and Hu, G.H. (2011). QTL analysis of soybean protein content using two methods in 5 different environments. Chinese Journal of Oil Crop Sciences. 33(1): 9-14.

Shen, X.L., Yuan, Y.L., Guo, W.Z., Zhu, X.F., Zhang, T.Z. and Gao, J.S. (2001). Genetic stability of a major QTL for fiber strength and its marker-assisted selection in upland cotton. High Technology Letters. 11: 13-16.

Song, Q.J., Marek, L.F., Shoemaker, R.C., Concibido, V.C., Delannay, X., Specht, J.E. and Cregan, P.B. (2004). A new integrated genetic linkage map of the soybean. Theoretical and Applied Genetics. 109(1): 122-128. DOI: 10.1007/s00 122-004-1602-3.

Wang, J.K. and Gai, J.Y. (1998). Identification of major gene and polygene mixed inheritance model of quantitative traits by using joint analysis of P-(1), F-(1), P-(2), F -(2) and F-(2:3) generations. Acta. Agron. Sin. 24: 651-659.

Wang, X.J. (2000). Inheritance, stability analysis and QTL mapping of yield related traits in soybean. Ph.D. diss. Chinese Academy of Agricultural Sciences, Beijing. (Abstract in English).

Wang, G. L. and Fang, H.J. (2002). Plant Gene Engineering Principle and Technology. 2nd Ed. Science Press, Beijing.

Wang, J.K. (2009). Inclusive composite interval mapping of quantitative trait genes. Acta Agronomica Sinica. 35(2): 239-245.

Wang, X.Z. (2008). Inheritance, stability analysis and QTL mapping of yield related traits in soybean. Doctoral Dissertation, Chinese Academy of Agricultural Sciences, Beijing.

Wang, Y.J. (2001). Establishment and adjustment of RIL population and its application to map construction, mapping genes resistant to SMV and QTL analysis of agronomic and quality traits in soybeans. Doctoral Dissertation, Nanjing Agricultural University.

Xu, P. (2006). Inheritance analysis and mapping QTLs related to protein and oil content of soybeans. Doctoral dissertation, Nanjing Agricultural University.

Yang, Z., Liu, L. J., Gao, M.J., Pu, G.F. and Zhang, L. (2008). Molecular marker assistant selection for high oil QTL in soybean. Soybean Science. 27: 921-924.

Zhen-Yu, G.E., Liu, X.B., Liu, B.H. *et al.* (2011). QTL Mapping of Protein and Oil Content in Soybean[J]. Soybean Science.

Zheng, Y.Z., Gai, J.Y., Zhou, R.B., Tian, S.J., Lu, W.G. and Li, W.D. (2007). Inheritance of fat and fatty acid composition contents in soybean. Soybean Science. 6: 801-806.

Zhang, L., Li, H., Li, Z. and Wang, J. (2008). Interactions between markers can be caused by the dominance effect of quantitative trait loci. Genetics. 180(2): 1177-1190. DOI: 10.1534/genetics.108.092122.

Zhu, M.Y., Han, F.X., Sun, J.M., Yan, S.R. and Yang, H. (2017). Detection of Soybean protein and oil content QTL with backcross introgression lines. Journal of Plant Genetic Resources. 18(6): 1204-1209. DOI: 10.13430/j.cnki. jpgr.2017.06.022.