



# 基于高密度大豆 SNP 遗传图谱的蛋白质和脂肪含量 QTL 定位及基因注释

刘家伶, 张 君, 韩 笑, 李雪莹, 姚 丹, 宋 阳, 张 卓, 杨松楠

(吉林农业大学 农学院, 吉林 长春 130118)

**摘要:** 为构建脂肪和蛋白含量上存在显著差异的大豆高密度遗传图谱, 并对其基于 SNP 标记的 QTL 定位及定位区间内的基因注释, 本研究以吉农 45 为母本、绥农 76 为父本杂交获得的 F<sub>2</sub> 代分离群体为基础, 通过靶向测序基因型分型技术获得 SNP 标记, 构建高密度遗传图谱, 结合群体籽粒脂肪和蛋白含量, 采用复合区间作图法 (CIM) 和完备区间作图法 (ICIM) 进行 QTL 定位, 并根据 SNP 标记提供的染色体物理位置信息确定 QTL 所在的物理区间, 利用 Soybase 和 Phytozome 数据库对区间进行基因挖掘、注释和筛选。结果表明: 共获得 2 399 个多态性分子标记, 构建了总图距为 1 020.5 cM、标记间平均距离为 0.61 cM 的遗传图谱。CIM 法共检测到 4 个 QTL, 其中 2 个蛋白质含量位点分别为 qPro-11-1 和 qPro-11-2, 2 个脂肪含量位点分别为 qOil-7-1 和 qOil-19-1, 贡献率为 1.31% ~ 21.69%。ICIM 法共检测到 3 个 QTL, 其中 2 个蛋白质含量位点分别为 qPro-11-3 和 qPro-14-1, 1 个脂肪含量位点为 qOil-16-1, 贡献率为 8.41% ~ 17.83%。通过基因注释, 在 6 个物理区间内分别筛选到 26, 29, 6, 1, 124 和 36 个与脂肪和蛋白贮藏、油脂生物合成、油分代谢过程、脂肪酸生物合成、种子发育和种子萌发相关的基因。在定位的 6 个不同位点中, 与脂肪含量相关的 QTL 分别位于 7, 16 和 19 号染色体, 与蛋白质含量相关的 QTL 分别位于 11 和 14 号染色体。有 4 个位点与前人定位结果相似, qPro-11-2 和 qPro-14-1 为新位点。

**关键词:** SNP; 蛋白质; 脂肪; 遗传图谱; QTL 定位; GO 注释

## Protein and Fat Content QTL Mapping and Gene Annotation Based on High-density Soybean SNP Genetic Map

LIU Jia-ling, ZHANG Jun, HAN Xiao, LI Xue-ying, YAO Dan, SONG Yang, ZHANG Zhuo, YANG Song-nan

(Agricultural College, Jilin Agricultural University, Changchun 130118, China)

**Abstract:** In order to construct a high-density genetic map of soybean with significant differences in fat and protein content, and carry out QTL mapping based on SNP markers and gene annotation within the mapping interval to find new loci, this study was based on the F<sub>2</sub> segregation population obtained by crossing Jinong 45 as female parent and Suinong 76 as male parent, SNP markers were obtained by targeted sequencing genotyping technique to construct a high-density genetic map. Combined with seed fat and protein content, composite interval mapping (CIM) method and inclusive composite interval mapping (ICIM) method were used to map QTL, and the physical interval of QTL was determined according to the chromosome physical location information provided by SNP markers. Soybase and Phytozome databases were used to mine, annotate and screen genes in the interval. A total of 2 399 polymorphic molecular markers were obtained, and the genetic maps with 1 020.5 cM of a total map distance and 0.61 cM of average distance between markers were constructed. A total of 4 QTL were detected by CIM method, of which 2 protein content loci were qPro-11-1 and qPro-11-2, and 2 fat content loci were qOil-7-1 and qOil-19-1, with a contribution rate of 1.31% – 21.69%. A total of 3 QTL were detected by ICIM method, of which two protein content loci were qPro-11-3 and qPro-14-1, and one fat content locus was qOil-16-1, with a contribution rate of 8.41% – 17.83%. Through gene annotation, 26, 29, 6, 1, 124 and 36 genes related to fat and protein storage, oil biosynthesis, oil metabolism, fatty acid biosynthesis, seed development and seed germination were screened in 6 physical intervals, respectively. Among the 6 different loci, adipose QTL was located on chromosomes 7, 16 and 19, and protein QTL was located on chromosomes 11 and 14, respectively. qPro-11-2 and qPro-11-3 were the same loci located by the two methods. Four loci were similar to the results of previous loci, and qPro-11-2 and qPro-14-1 were new loci.

**Keywords:** SNP; Protein; Fat; Genetic map; QTL mapping; Go annotation

收稿日期: 2020-12-04

基金项目: 吉林省教育厅科学研究项目 (JKKH20210213KJ); 国家科技重大专项 (2020ZX08009 – 15B)。

第一作者: 刘家伶 (1996—), 女, 在读硕士, 主要从事作物遗传育种研究。E-mail: 1326987198@qq.com。

通讯作者: 杨松楠 (1989—), 女, 博士, 副教授, 主要从事大豆遗传育种研究。E-mail: ysn785620774@126.com。

大豆 (*Glycine max* L.) 由于含有丰富的脂肪 (15% ~ 25%) 和蛋白质 (35% ~ 50%), 在农业和工业领域有广泛的应用<sup>[1-2]</sup>, 提高大豆籽粒中脂肪和蛋白质的含量已经成为一个重要的品质育种目标<sup>[3-4]</sup>。大豆脂肪和蛋白质含量是非常复杂的数量性状, 由多个基因控制, 并受到环境条件的影响<sup>[5-6]</sup>。然而, 在分子标记辅助选择育种中, 与蛋白和脂肪含量相关基因的分离和应用研究较少<sup>[7-9]</sup>。而基于分子标记的蛋白和脂肪含量数量性状定位的研究对选育优质的大豆品种具有重要意义。

连锁图谱是遗传研究的重要组成部分, 已被应用于生物技术领域, 包括数量性状位点分析、基因定位和克隆以及分子标记辅助选择育种<sup>[10]</sup>。高密度遗传图谱的构建已经成为精细定位的重要手段<sup>[11]</sup>。Qi 等<sup>[12]</sup>报告了 1 份利用重组自交系 (RILs) 群体的大豆高密度遗传图谱, 包括 20 个连锁群的 5 308 个标记, 图谱长度为 2 655.68 cM, 相邻标记间的平均距离为 0.5 cM。朱明月等<sup>[13]</sup>从 718 对引物中筛选出 106 个多态性标记, 通过 Icimapping 4.1 构建了 1 张共 106 个 SSR 标记, 总长 534.63 cM, 平均遗传距离 5.04 cM 的遗传连锁图谱。

在过去的几十年中, 许多研究者使用分子标记识别大豆农艺性状的 QTL<sup>[14]</sup>。到目前为止, 在 Soybase (<http://soybase.org>) 上已经报道了许多控制大豆农艺性状的 QTL, 包括 334 个与大豆脂肪含量相关的 QTL 和 313 个与蛋白含量相关的 QTL。其中, 一些 QTL 的遗传位置非常接近 (<10 cM), 表明它们可能是相同的 QTL<sup>[15]</sup>。Sonah 等<sup>[16]</sup>使用全基因组关联分析 (GWAS) 在 5, 15 和 20 号染色体上定位到了与大豆籽粒脂肪和蛋白质含量相关的 QTL。Hyten 等<sup>[17]</sup>在 6, 7 和 19 号染色体上定位到了与脂肪含量相关的 QTL。盛英华等<sup>[18]</sup>以低亚麻酸材料 FA10 为父本, 分别与中豆 32 和中豆 43 杂交得到两个 F<sub>2</sub> 分离群体, 在两个群体间分别定位到 39 和 18 个与脂肪酸相关的 QTL, 并且在 16 号染色体上定位到了与脂肪含量相关的 QTL。Zhang 等<sup>[19]</sup>和 Warrington 等<sup>[20]</sup>均在第 14 (B2) 号染色体上检测到了与蛋白含量相关的 QTL。Wang 等<sup>[21]</sup>利用复合区间作图法在 11 号染色体上检测到了与大豆籽粒蛋白含量相关的 QTL。前人基于分离群体的连锁分析的研究结果表明, 检测到大量的与大豆蛋白和脂肪含量相关的 QTL, 在大豆 20 条染色体上均有分布。

前人对大豆籽粒脂肪含量和蛋白含量相关的

QTL 研究很多, 大部分的研究基于 SSR 标记, 但由于性状的复杂性, 可重复检验的位点很少, 且前人对相关基因挖掘的研究也较少。本研究以东北地区的优良品种吉农 45 和绥农 76 为亲本, 构建了 F<sub>2</sub> 代分离群体。并以此分离群体为材料, 对大豆籽粒蛋白和脂肪含量进行基于 SNP 标记的 QTL 定位分析, 并对定位区间内相关基因进行基因注释和筛选, 检测重复位点并定位新位点, 为分子标记辅助育种、QTL 的精细定位和相关基因克隆奠定理论基础。

## 1 材料与方法

### 1.1 材料

高油品种吉农 45 (蛋白含量为 37.28%, 脂肪含量 21.66%); 高蛋白品种绥农 76 (蛋白含量 46.78%, 脂肪含量 16.86%)。

### 1.2 方法

1.2.1 杂交群体获得 于 2017 年在吉林农业大学教学科研基地种植吉农 45 和绥农 76, 7 月中旬以吉农 45 为母本、绥农 76 为父本进行杂交, 10 月份收取杂种 F<sub>1</sub> 代, 并到海南加代, 获得 F<sub>2</sub> 代种子, 2018 年 4 月加代种子种植于吉林农业大学试验田, 共得到 F<sub>2</sub> 代群体 143 株, 成熟后单株收获, 用于品质性状分析。

1.2.2 脂肪和蛋白含量测定 使用近红外光谱仪 (Buchi 公司 N-500) 测量蛋白质和脂肪含量, 每株 3 次重复, 取平均值。

1.2.3 DNA 提取及 SNP 测序 采集 143 株 F<sub>2</sub> 代个体及亲本苗期幼嫩叶片, 使用植物基因组 DNA 提取试剂盒 (DP305-03, 北京天根生化科技有限公司) 提取 DNA, 采用靶向测序基因型分型 (Genotyping by Targeted Sequencing, GBTS) 技术进行 SNP 多态性分析 (由石家庄博瑞迪生物技术有限公司完成), 得到 SNP 原始数据。

1.2.4 遗传图谱构建 通过比较亲本间基因型的差异, 删掉 SNP 多态位点原始数据中基因型相同的位点。再利用 Joinmap 4.0, 根据亲本和群体个体脂肪和蛋白质含量表型数据进一步筛选并绘制遗传图谱。

1.2.5 QTL 定位 构建遗传连锁图谱后, 使用 WinQTL Cart 2.5 进行 CIM 作图, 使用 QTL IciMapping 4.2.53 进行 ICIM 作图, 以 LOD = 2.5 为阈值进行 QTL 定位。

1.2.6 基因注释 对已获得的 QTL 进行基因注释。以 Glyma.Wm82.a2.v1 为基因模型,从 Soybase 数据库 (<http://www.soybase.org>) 中获取参考基因组的信息,根据 Soybase (<http://www.soybase.org>) 和 Phytozome (<http://phytozome.jgi.doe.gov>) 中获得的基因注释信息进行基因预测。然后,利用 WEGO2.0 在线软件(<http://wego.genomics.org.cn/document>)对 QTL 区间的 GO 注释信息进行分子功能、生物过程和细胞成分的分类和注释。

1.3 数据分析

使用 Excel 2010 进行数据整理,使用 DPS v7.05 数据处理软件进行数据的统计分析。

2 结果与分析

2.1 表型数据分析

由表 1 可知,F<sub>2</sub> 的蛋白质含量为 34.33% ~ 48.27%,脂肪含量为 18.28% ~ 26.03%,脂肪含量的变异系数为 7.35%,大于蛋白含量的变异系数 6.04%。而 F<sub>2</sub> 的脂肪和蛋白含量的最大值均高于

亲本吉农 45 和绥农 76。在 F<sub>2</sub> 中两个目标性状的偏度和峰度的绝对值均小于 1,表明这些性状的分离符合正态分布模型,具有多基因控制的数量性状的遗传特征。

2.2 遗传图谱构建

对两个亲本材料及 F<sub>2</sub> 代群体基因多态性的检测得到的 3 万个 SNP 多态性位点数据,通过比较亲本间基因型的差异,删掉基因型相同的位点,筛选得 3 966 个 SNP,利用 Joinmap 4.0 进一步筛选出 2 399 个用于构建遗传连锁图谱的 SNP 分子标记。遗传图谱总图距为 1 020.5 cM,20 条染色体上的平均标记数为 120 个,染色体平均长度为 51.03 cM。其中,第 3 号染色体上的分子标记数最多,共 238 个;第 11 号染色体图距最长,为 73.7 cM;第 20 号染色体上分子标记数最少,共 21 个;第 17 号染色体图距最短,为 25.3 cM。20 号染色体的 SNP 标记间距离为 0.14 ~ 2.54 cM,平均距离 0.62 cM。标记间最大距离在第 11 号染色体上,为 25.24 cM;最小距离在第 5 号染色体上,为 2.31 cM(图 1,表 2)。

表 1 亲本材料及 F<sub>2</sub> 代群体脂肪、蛋白含量统计结果

Table 1 Descriptive statistics of fat and protein content of the parents and F<sub>2</sub> population

性状 Character	吉农 45 Jinong 45	绥农 76 Suinong 76	F <sub>2</sub>						
			最大值 Max.	最小值 Min.	均值 Average	标准差 SD	变异系数 CV/%	偏度 Skew	峰度 Kurtosis
蛋白含量 Protein content/%	37.92	46.78	48.27	34.33	40.11	0.02	6.04	0.53	0.17
脂肪含量 Fat content/%	21.66	16.86	25.03	18.28	22.15	0.01	7.35	-0.41	0.31

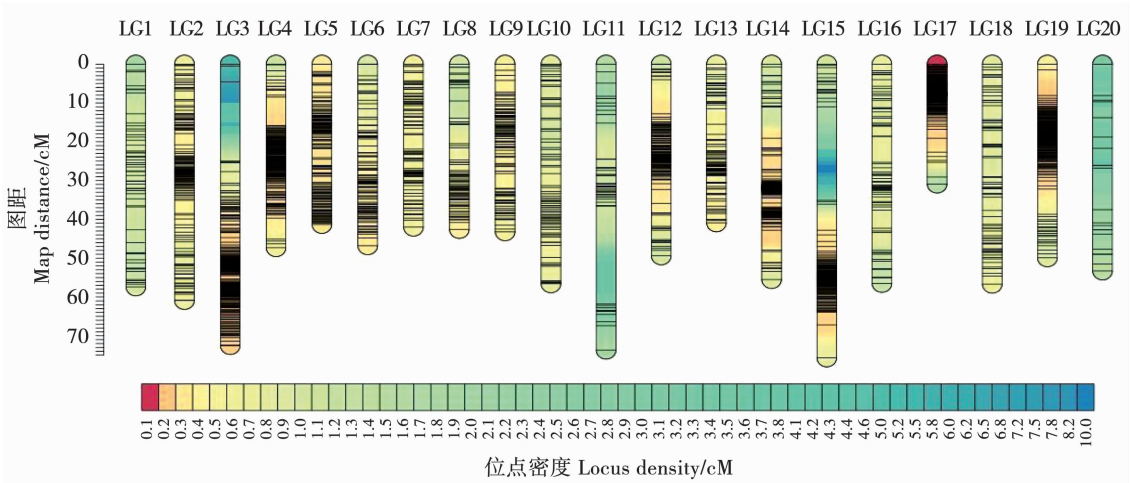


图 1 基于 SNP 分子标记构建的 F<sub>2</sub> 代群体遗传图谱

Fig. 1 F<sub>2</sub> population genetic map based on SNP molecular markers

表 2 F<sub>2</sub>代群体染色体的 SNP 标记数量及遗传距离统计

Table 2 SNP molecular marker number and genetic distance statistical data of the F<sub>2</sub> population on chromosomes

染色体 Chromosome	SNP 分子标记数 Number of SNP markers	染色体长度 Length of chromosome /cM	标记间平均距离 Average distance between markers/cM	标记间最大距离 Maximum distance between markers/cM
1	54	57.4	1.06	5.85
2	143	60.8	0.43	3.45
3	238	72.5	0.30	20.34
4	189	37.9	0.20	7.39
5	149	41.0	0.28	2.31
6	109	46.8	0.43	2.96
7	83	42.0	0.51	3.7
8	70	40.8	0.58	4.85
9	106	42.5	0.40	3.79
10	81	56.6	0.70	3.69
11	47	73.7	1.57	25.24
12	145	49.4	0.34	8.74
13	85	40.9	0.48	4.02
14	177	52.7	0.30	8.02
15	170	67.2	0.40	24.15
16	72	55.1	0.77	6.77
17	180	25.3	0.14	3.83
18	77	56.7	0.74	4.93
19	203	47.9	0.24	6.59
20	21	53.3	2.54	10.75
总计 Total	2399	1020.5	—	—
平均 Mean	120	51.03	0.62	8.07

2.3 QTL 定位分析

以 LOD2.5 为标准,采用 CIM 法共检测到 4 个 QTL,其中包括两个与大豆籽粒蛋白含量相关的 QTL,均在 11 号染色体上,分别为 qPro-11-1 和 qPro-11-2,qPro-11-2 的 LOD 值最高,为 7.963 1,其贡献率为 27.3%;检测到两个与大豆籽粒脂肪含量相关的 QTL,位于 7 号染色体和 19 号染色体上,分别为 qOil-7-1 和 qOil-19-1,其中 qOil-7-1 的 LOD 值最高,为 4.399 6,贡献率为 4.13%(表 3 和图 2)。

以 LOD2.5 为标准,ICIM 法共检测到 3 个 QTL,其中 2 个与大豆籽粒蛋白含量相关,在 11 号和 14 号染色体上,分别为 qPro-11-3 和 qPro-14-1,

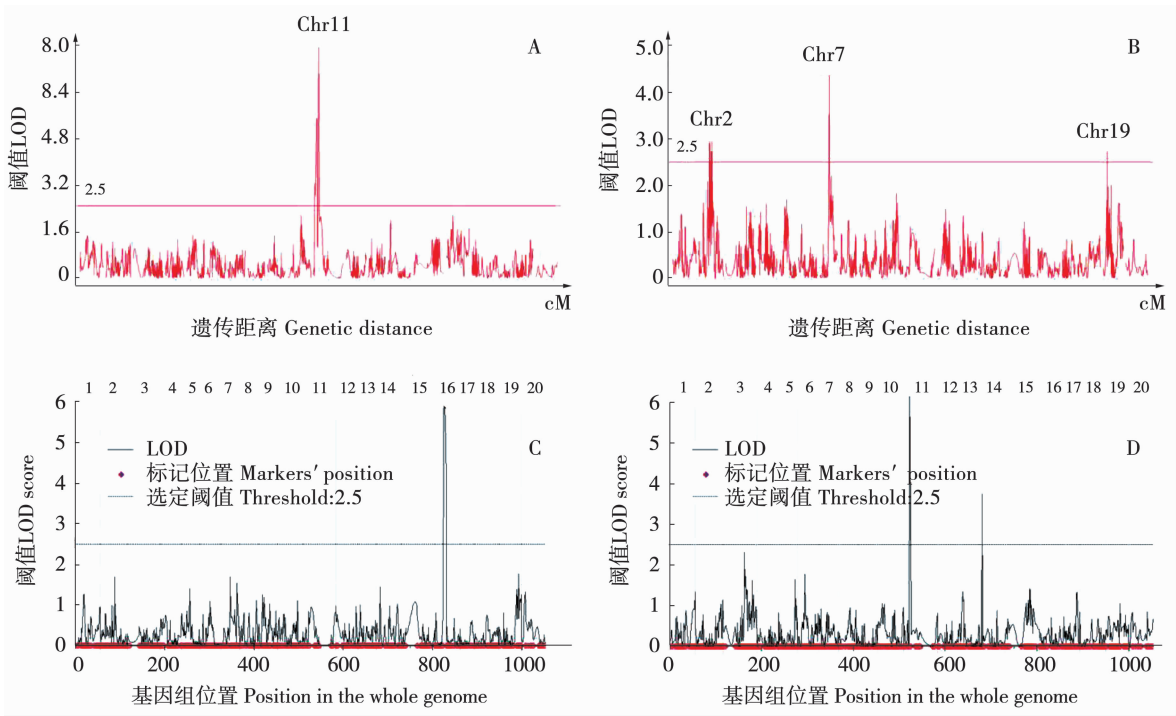
其中 qPro-11-3 的 LOD 值最高,为 6.149 5,贡献率为 12.43%;检测到 1 个与大豆籽粒脂肪含量相关的 QTL,在 16 号染色体上,为 qOil-16-1,LOD 值为 5.9106,贡献率为 17.83%(表 3 和图 2)。

综上所述,CIM 法和 ICIM 法均在 11 号染色体上检测到与大豆籽粒蛋白含量相关的 QTL,并且与蛋白含量相关的 QTL 大多集中在 11 号染色体上。依照定位区间可知 qPro-11-2 与 qPro-11-3 为同一 QTL,取两者的交集为其新区间,即 qPro-11-3,下文记为 qPro-11-2。有 4 个位点与前人定位结果相似,qPro-11-2 和 qPro-14-1 为新位点。

表 3 利用 ICIM 和 CIM 法在 F<sub>2</sub> 代群体中检测到的籽粒脂肪和蛋白含量 QTL 位置及其效应

Table 3 The position and effect of QTL for seed fat and protein content detected in the F<sub>2</sub> population by ICIM and CIM

方法 Method	QTL	染色体 Chromosome	置信区间 Confidence interval/cM	定位区间 Mapping interval /bp	QTL 位置 QTL Position/cM	阈值 Lod	加性效应 Additive effect	显性效应 Dominant effect	贡献率 R <sup>2</sup> /%	参考文献 References
CIM	qPro-11-1	11	6.3 – 9.1	3070668 – 5011653	7.91	5.5201	0.0124	– 0.0159	21.69	[22][23]
	qPro-11-2	11	11.9 – 13.6	33237625 – 33761671	12.91	7.9631	0.0158	– 0.0128	27.30	
	qOil-7-1	7	22.1 – 23.9	3780293 – 5414292	22.61	4.3996	– 0.0077	– 0.0083	4.13	[22][24]
	qOil-19-1	19	10.1 – 10.5	13420344 – 33678446	10.31	2.7166	0.0059	0.0067	1.31	[25][26]
ICIM	qPro-11-3	11	11.50 – 13.50	33237625 – 33757938	13	6.1495	0.0117	– 0.0084	12.43	
	qPro-14-1	14	5.50 – 6.50	2095056 – 2194692	6	3.7433	– 0.0111	– 0.0066	8.41	
	qOil-16-1	16	18.50 – 24.50	26263392 – 37149918	21	5.9060	– 0.0077	0.0068	17.83	[5][22][27]



A 和 B 分别为 CIM 法获得的与蛋白质和脂肪含量相关的位点;C 和 D 分别为 ICIM 法获得的与脂肪和蛋白质含量相关的位点。

A and B were the protein and fat content related loci obtained by CIM method respectively ; C and D were the fat and protein content related loci obtained by ICIM method respectively.

图 2 F<sub>2</sub> 代群体蛋白和脂肪含量 QTL 分布情况

Fig. 2 The distribution of QTL for protein and fat content detected in the F<sub>2</sub> population

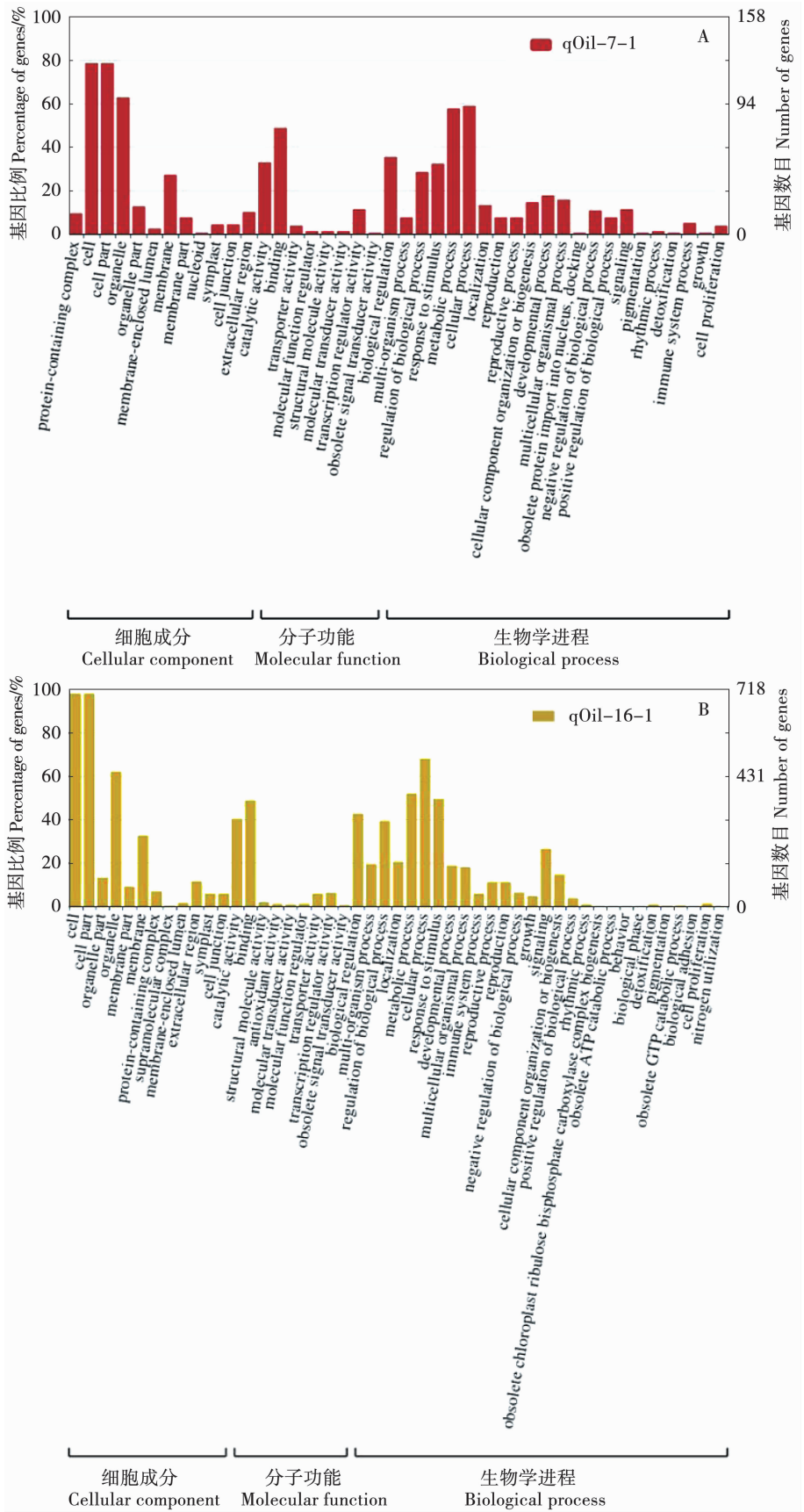
2.4 QTL 候选基因的预测

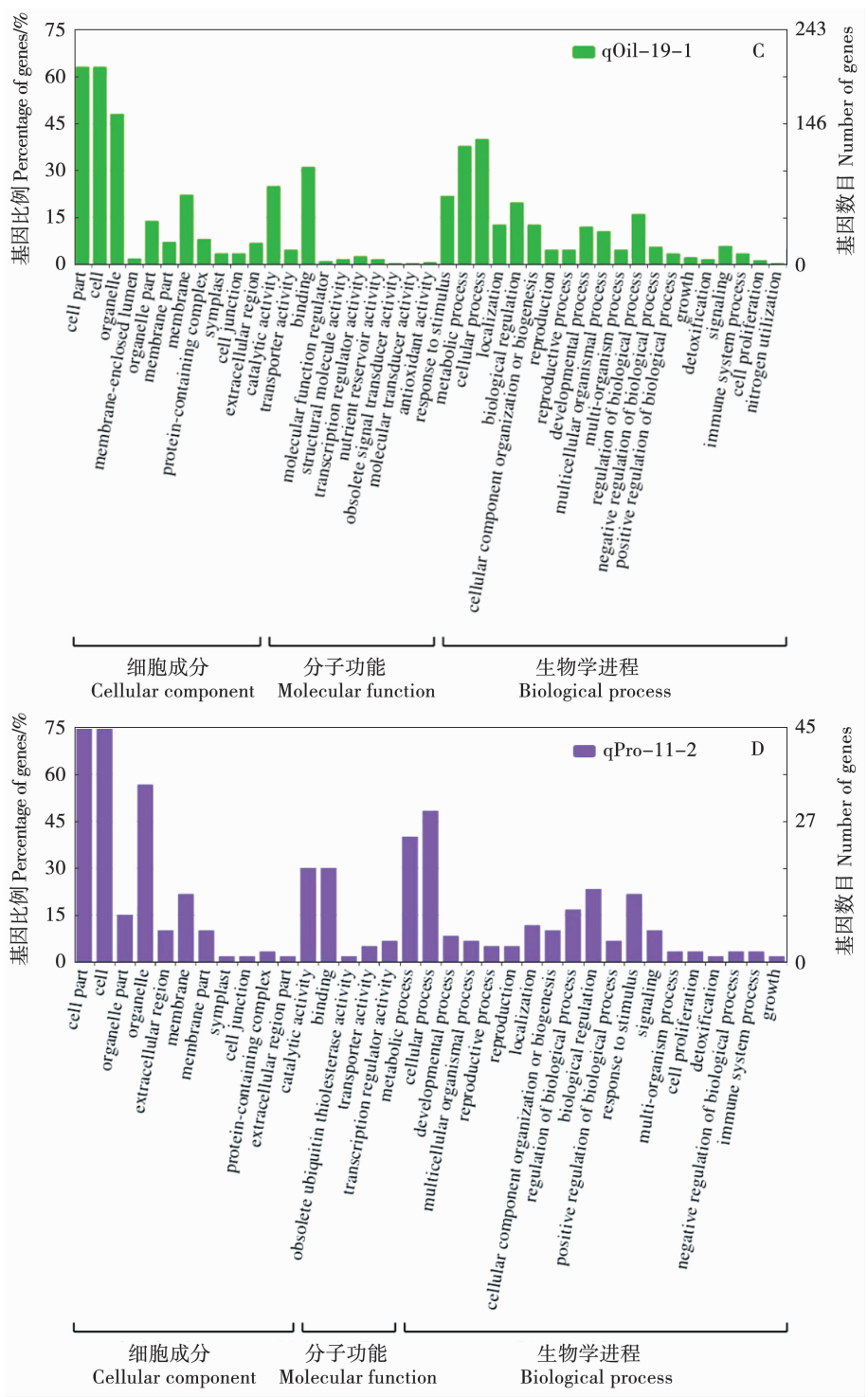
以 Glyma. Wm82. a2. v1 为参考基因组,对 qOil-7-1 (3780293 – 5414292 bp)、qPro-11-1 (3070668 – 5011653 bp)、qPro-11-2 (33237625 – 33757938 bp)、qPro-14-1 (2095056 – 2194692 bp)、qOil-16-1 (26263392 – 37149918 bp) 和 qOil-19-1 (13420344 – 33678446 bp) 位点的物理区间内的 158,247,60,14,958 和 325 个模式基因进行分类,并分别对它们的物理位置的细胞成分、分子功能和生物过程功能进

行注释。参考 GO 注释信息,在 6 个物理区间内分别筛选到 26,29,6,1,124 和 36 个与脂肪和蛋白贮藏、油脂生物合成、油分代谢过程、脂肪酸生物合成、种子发育和种子萌发相关的基因(图 3)。

两个新位点的相关基因注释如表 4 所示,位点内基因多与转运、蛋白质加工和降解相关,因而位点影响籽粒蛋白含量的原因可能与其参与了贮藏蛋白合成的相关过程有关。







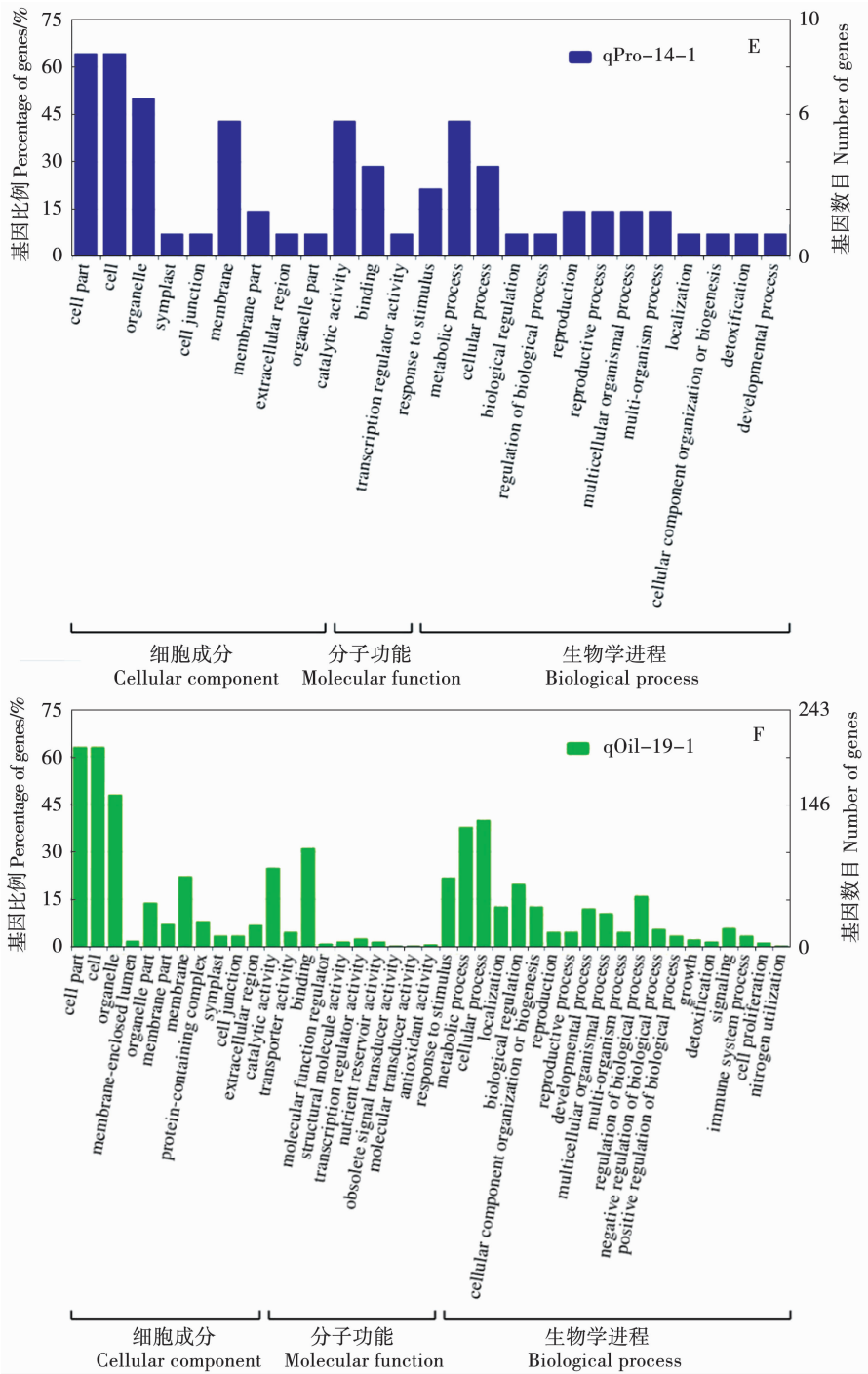


图 3 Go 基因注释信息

Fig. 3 Go gene annotation information

表 4 基于新 QTL 注释到的与籽粒蛋白和脂肪含量相关的基因

Table 4 Genes annotated with the new QTL related to seed protein and fat content

基因 ID Gene ID	拟南芥同源基因 ID Homolog ID in Arabidopsis	基因本体论的生物进程描述 Gene ontology biological process descriptions
Glyma. 11G238100	AT4G39820. 1	生物进程 biological process
		细胞阳离子稳态;半胱氨酸生物合成过程;二价金属离子转运;硫代葡萄糖苷生物合成过程;吡啶乙酸生物合成过程;硫酸盐转运;跨膜转运;转运
Glyma. 11G238400	AT5G10180. 1	cellular cation homeostasis; cysteine biosynthetic process; divalent metal ion transport; glucosinolate biosynthetic process; indoleacetic acid biosynthetic process; sulfate transport; transmembrane transport; transport



续表 4

基因 ID Gene ID	拟南芥同源基因 ID Homolog ID in <i>Arabidopsis</i>	基因本体论的生物进程描述 Gene ontology biological process descriptions
<i>Glyma. 11G238500</i>	AT5G10180. 1	细胞阳离子稳态;半胱氨酸生物合成过程;二价金属离子转运;硫代葡萄糖苷生物合成过程;吲哚乙酸生物合成过程;硫酸盐转运;跨膜转运;转运 cellular cation homeostasis; cysteine biosynthetic process; divalent metal ion transport; glucosinolate biosynthetic process; indoleacetic acid biosynthetic process; sulfate transport; transmembrane transport; transport
<i>Glyma. 11G239800</i>	AT4G39850. 1	脂肪酰辅酶 A 转运;脂肪酸 β 氧化;参与过氧化物酶体组成;种子萌发的正调控;蛋白酶体泛素依赖的蛋白质分解代谢过程;蛋白酶体核心复合物组装;对错误折叠蛋白质的反应;毒素分解代谢过程;跨膜转运;运输 fatty-acyl-CoA transport; fatty acid beta-oxidation; peroxisome organization; positive regulation of seed germination; proteasomal ubiquitin-dependent protein catabolic process; proteasome core complex assembly; response to misfolded protein;toxin catabolic process; transmembrane transport; transport
<i>Glyma. 11G242900</i>	AT4G40080. 1	生物过程;硝酸盐转运;硝酸盐相关反应 biological process; nitrate transport; response to nitrate
<i>Glyma. 14G029000</i>	AT3G59110. 1	蛋白质自磷酸化;蛋白质磷酸化 protein autophosphorylation; protein phosphorylation

3 讨 论

本研究以高油品种吉农 45 和高蛋白品种绥农 76 为亲本构建 F<sub>2</sub> 代分离群体,后代中出现了脂肪含量和蛋白含量的最大值均大于双亲的现象。其可能原因是双亲籽粒中控制脂肪和蛋白含量的位点是不同的,从而导致后代出现超亲分离现象。

本研究使用 SNP 分子标记,与 SSR 分子标记相比,其具有在全基因组中高密度分布的特点,而且能够提供标记所在的物理位置;此外,高密度遗传图谱相较于低密度遗传图谱更适用于 QTL 定位,在提高 QTL 检测效率的同时还能提供更紧密的遗传标记。

目前,在大豆的不同群体中,已经发现了许多脂肪和蛋白质含量相关的 QTL 基因。在本研究中, F<sub>2</sub> 群体中检测到 6 个不同的控制脂肪和蛋白质含量的 QTL。其中 3 个与脂肪含量相关,分别为 qOil-7-1 (3780293 – 5414292 bp)、qOil-16-1 (26263392 – 37149918 bp) 和 qOil-19-1 (13420344 – 33678446 bp); 3 个与蛋白含量相关,分别为 qPro-11-1 (3070668 – 5011653 bp)、qPro-11-2 (33237625 – 33757938 bp) 和 qPro-14-1 (2095056 – 2194692 bp)。并且 ICIM 和 CIM 两种方法均在 11 号染色体上检测到了与蛋白含量相关的 QTL。将已发表的脂肪和蛋白质含量的 QTL/基因与本研究检测到的 QTL/基因进行比较,这 6 个 QTL 中,有 4 个 QTL 与前人报道的控制大豆籽粒蛋白和脂肪含量相关的 QTL 具有相同或相近的标记或标记区间。其中,qOil-7-1 所在区段与 Mao 等<sup>[22]</sup>和 Han 等<sup>[23]</sup>定位的油分含量 QTL 相一致,并且较 Mao 等<sup>[22]</sup>的区段缩短约 6.7 Mb,

与 Han 等<sup>[23]</sup>定位的区段在 7 号染色体 4559602 – 5414292 bp 处重合;qPro-11-1 所在区段与 Mao 等<sup>[22]</sup>和 Brummer 等<sup>[24]</sup>定位的控制蛋白含量 QTL 的区段有所重合,三者重合区域在 11 号染色体 4260934 – 5114461 bp 处;qOil-16-1 定位区间是一个稳定的 QTL 位点,分别在 Wang 等<sup>[4]</sup>、Mao 等<sup>[22]</sup>、Lee 等<sup>[27]</sup>定位的油分含量相关 QTL 结果中出现,其重合区域位于 16 号染色体 30404629 – 31474289 bp 处;qOil-19-1 则与 Fasoula 等<sup>[25]</sup>与 Kim 等<sup>[26]</sup>的籽粒油分含量 QTL 定位结果所在区域重合,且其区间跨幅较两者结果缩小约 1.5 Mb。此外,在本研究中检测到的 qPro-11-2、qPro-14-1 所在的区段在已发表相关文献中并未有相关位点的报道,因而 qPro-11-2、qPro-14-1 属于新发现的位点。

功能预测结果显示,与蛋白含量相关的两个新 QTL 区段主要参与了蛋白质的转运、加工和代谢过程,影响籽粒蛋白含量的原因可能是这些过程参与了籽粒贮藏蛋白的相关代谢通路,但其注释结果还需要进一步验证。并且 qPro-11-2 在两种定位方法下的贡献率均大于 12%,具有重要的潜在利用价值,今后应对其稳定性等进行深入分析,以期为大豆分子辅助育种提供高效准确的分子标记,并为相关基因的克隆奠定基础。

4 结 论

本研究选用亲本之间蛋白和脂肪含量有显著差异的两品种杂交构建 F<sub>2</sub> 分离群体,利用 SNP 分子标记构建了总图距 1 020.5 cM、标记间平均距 0.61 cM 的遗传图谱。采用复合区间作图法(CIM)和完备区间作图法(ICIM)进行 QTL 定位,CIM 法共

检测到4个QTL,其中包括2个蛋白位点和2个脂肪位点;ICIM法共检测到3个QTL,其中包括2个蛋白位点和1个脂肪位点,且qPro-11-2和qPro-11-3为两种方法定位的相同位点。在定位的6个不同位点中,脂肪QTL位于7、16和19号染色体,蛋白QTL位于11和14号染色体。有4个位点与前人定位结果相似,qPro-11-2和qPro-14-1为新位点。与蛋白含量相关的两个新QTL区段主要参与了蛋白质的转运、加工和代谢过程。

## 参考文献

[1] Nichols D M, Glover K D, Carlson S R, et al. Fine mapping of a seed protein QTL on soybean linkage group I and its correlated effects on agronomic traits[J]. Crop Science, 2006, 46(2): 834-839.

[2] Hurburgh C R. Long-term soybean composition patterns and their effect on processing[J]. Journal of the American Oil Chemists' Society, 1994, 71(12): 1425-1427.

[3] Li Z, Orf J H, Bajjalieh N, et al. QTL for seed protein and amino acids in the Benning × Danbaekkong soybean population[J]. Theoretical & Applied Genetics International Journal of Breeding Research & Cell Genetics, 2015, 128(5): 839-850.

[4] Wang J, Chen P, Wang D, et al. Identification and mapping of stable QTL for protein content in soybean seeds[J]. Molecular Breeding, 2015, 35(3): 92.

[5] Burton J W. Quantitative genetics: Results relevant to soybean breeding [M]// Wilcox J R. ed. Soybeans: Improvement, production, and uses, 2nd edn. ASA, CSSA, and SSSA, Madison, 1987: 211-242.

[6] Wilcox J R. Increasing seed protein in soybean with eight cycles of recurrent selection[J]. Crop Science, 1998, 38(6): 1536-1540.

[7] Wang X, Jiang G L, Green M, et al. Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean[J]. Molecular Genetics & Genomics, 2014, 289(5): 935-949.

[8] Zhang Y H, Liu M F, He J B, et al. Marker-assisted breeding for transgressive seed protein content in soybean [*Glycine max* (L.) Merr.] [J]. Theoretical & Applied Genetics, 2015, 128(6): 1061-1072.

[9] Jun T H, Van K, Kim M Y, et al. Association analysis using SSR markers to find QTL for seed protein content in soybean[J]. Euphytica, 2008, 162(2): 179-191.

[10] Ma J Q, Huang L, Ma C L, et al. Large-scale SNP discovery and genotyping for constructing a high-density genetic map of tea plant using specific-locus amplified fragment sequencing (SLAF-seq) [J]. PLoS One, 2015, 10(2): e0121093.

[11] Liebhard R, Koller B, Gianfranceschi L, et al. Creating a saturated reference map for the apple (*Malus × Domestica* Borkh.) genome[J]. Theoretical & Applied Genetics, 2003, 106(8): 1497-1508.

[12] Qi Z M, Huang L, Zhu R S, et al. A high-density genetic map for soybean based on specific length amplified fragment sequencing [J]. PLoS One, 2014, 9(8): e104871.

[13] 朱明月. 利用回交导入系群体定位大豆蛋白质、脂肪含量及脂肪酸含量 QTL[D]. 北京: 中国农业科学院, 2017. (Zhu M Y.

Identification of quantitative trait loci controlling oil concentration and fatty acid components with backcross introgression lines in soybean[D]. Beijing: Chinese Academy of Agricultural Sciences, 2017.)

[14] Carpentieri-Pipolo V, Pipolo A E, Abdel-Haleem H, et al. Identification of QTLs associated with limited leaf hydraulic conductance in soybean [J]. Euphytica, 2012, 186(3): 679-686.

[15] Pathan S M, Vuong T, Clark K, et al. Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean[J]. Crop Science, 2013, 53(3): 765-774.

[16] Sonah H, O'Donoghue L S, Cober E, et al. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean[J]. Plant Biotechnology Journal, 2014, 13(2): 211-221.

[17] Hyten D L, Pantalone V R, Sams C E, et al. Seed quality QTL in a prominent soybean population [J]. Theoretical & Applied Genetics, 2004, 109(3): 552-561.

[18] 盛英华, 张延瑞, 戴亚楠, 等. 不同群体中大豆脂肪酸组分 QTL 定位研究[J]. 中国油料作物学报, 2020(5): 796-806. (Sheng Y H, Zhang Y R, Dai Y N, et al. QTL mapping of fatty acids in different F<sub>2</sub> populations in soybean[J]. Chinese Journal of Oil Crops, 2002(5): 796-806.)

[19] Zhang W K, Wang Y J, Luo G Z, et al. QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers[J]. Theoretical and Applied Genetics, 2004, 108(6): 1131-1139.

[20] Warrington C V, Abdel-Haleem H, Hyten D L, et al. QTL for seed protein and amino acids in the Benning × Danbaekkong soybean population[J]. Theoretical & Applied Genetics, 2015, 128(5): 839-850.

[21] Wang X, Jiang G L, Green M, et al. Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean[J]. Molecular Genetics & Genomics, 2014, 289(5): 935-949.

[22] Mao T T, Jiang Z F, Han Y P, et al. Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments[J]. Plant Breeding, 2013, 132(6): 630-641.

[23] Han Y P, Teng W, Wang Y, et al. Unconditional and conditional QTL underlying the genetic interrelationships between soybean seed isoflavone, and protein or oil contents[J]. Plant Breeding, 2015, 134(3): 300-309.

[24] Brummer E C, Graef G L, Orf J, et al. Mapping QTL for seed protein and oil content in eight soybean populations [J]. Crop Science, 1997, 37(2): 370-378.

[25] Fasoula V A, Harris D K, Roger B H. Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations[J]. Crop Science, 2004, 44(4): 1218-1225.

[26] Kim H K, Kim Y C, Kim S T, et al. Analysis of quantitative trait loci (QTLs) for seed size and fatty acid composition using recombinant inbred lines in soybean[J]. Journal of Life Science, 2010, 20(8): 1186-1192.

[27] Lee S H, Bailey M A, Mian M A R, et al. RFLP loci associated with soybean seed protein and oil content across populations and locations[J]. Theoretical and Applied Genetics, 1996, 93(5-6): 649-657.