

大豆高蛋白种质中引 1106 蛋白质含量的 QTL 分析

王 婉^{1,4}, 韩德志², 闫洪睿², 栾晓燕³, 王 俊¹, 邱丽娟⁴

(¹ 长江大学农学院, 湖北荆州 434025; ² 黑龙江省农业科学院黑河分院, 黑河 164300; ³ 黑龙江省农业科学院大豆研究所, 哈尔滨 150086;

⁴ 农作物基因资源与遗传改良国家重大科学工程 / 农业部北京大豆生物学重点实验室 / 中国农业科学院作物科学研究所, 北京 100081)

摘要: 大豆籽粒蛋白质含量由多基因控制且易受环境条件的影响, 发掘高蛋白基因是促进大豆优质分子育种的重要手段。本研究选用综合性状优良、蛋白质含量较低的黑河 50 为轮回亲本, 以引进的高蛋白种质中引 1106 为供体亲本, 构建了由 384 个家系组成的回交高代群体。利用近红外光谱仪测定回交群体的蛋白质含量, 使用 SSR 分子标记技术鉴定回交群体 BC₁F₆ 基因型, 通过 QTL ICIMapping4.1 的区间作图法 (IM-ADD) 和完备区间作图法 (ICIM-ADD) 定位蛋白含量 QTL, 共获得 9 个蛋白含量 QTL, 其中 IM 定位到 7 个蛋白含量 QTL, 而 ICIM 定位到 3 个蛋白含量 QTL, 两种方法同时在 8 号染色体上定位到一个 QTL (*qPro-8-1*), 该 QTL 两侧分子标记是 SSR₅₀ 和 SSR₅₁, 可解释表型变异分别是 2.26% 和 7.85%, 定位区间物理距离大小为 218.71 kb, 该 QTL 尚未见报道, 是一个与蛋白质含量相关的新 QTL 位点, 为高蛋白大豆品种选育提供了材料和理论依据。

关键词: 蛋白质含量; QTL 定位; 回交群体; 区间作图法; 完备区间作图法

QTL Mapping of Soybean Protein Content from High-Protein Soybean ZhongYin1106

WANG Wan^{1,4}, HAN De-zhi², YAN Hong-rui², LUAN Xiao-yan³, WANG Jun¹, QIU Li-juan⁴

(¹ College of Agriculture, Yangtze University, Hubei Jingzhou 434025; ² Soybean Research Center, Heihe Institute, Heilongjiang

Academy of Agricultural Sciences, Heihe 164300; ³ Soybean Research Institute, Heilongjiang Academy of Agricultural Science,

Harbin 150086; ⁴ National Key Facility for Crop Gene Resources and Genetic Improvement/Key Laboratory of Soybean

Biology in Beijing (MOA)/Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081)

Abstract: The protein content in soybean is genetically controlled by multiple genes. Identification of the quantitative trait loci (QTL) associating with this trait will benefit for a practice use in soybean breeding. In this study, a backcross population consisting of 384 lines was derived from a cross of elite cultivar Heihe50 as recurrent parent and high protein accession ZhongYin1106 as donor parent. The protein content of parents and the offspring segregants was detected by Near-infrared spectrometer, and genotyping was conducted with SSR molecular markers. A genetic linkage map was constructed by software QTL ICIMapping4.1. Nine QTL associating with protein content were revealed, including seven QTL detecting by interval mapping (IM) and 3 QTLs detecting by inclusive composite interval mapping (ICIM). One QTL *qPro-8-1*, which was located in

收稿日期: 2019-06-06 修回日期: 2019-06-12 网络出版日期: 2019-06-13

URL: <http://doi.org/10.13430/j.cnki.jpgr.20190606003>

第一作者研究方向为大豆基因挖掘, E-mail: sxnydxsk1201ww@163.com; 韩德志为共同第一作者

通信作者: 邱丽娟, 研究方向为大豆基因资源挖掘与利用, E-mail: qiulijuan@caas.cn

王俊, 研究方向为大豆分子生物学与遗传育种, E-mail: wangjagri@yangtzeu.edu.cn

基金项目: 中国农科院创新工程基金项目; 黑龙江省农业科学院院级科研项目 (2017XQ04); 湿地生态与农业利用教育部工程中心开放基金 (KF201808)

Foundation project: Innovation Project of Chinese Academy of Agricultural Sciences, Scientific Research Project of Heilongjiang Academy of Agricultural Sciences (2017XQ04), Open Fund of Engineering Research Center of Ecology and Agricultural Use of Wetland, Ministry of Education (KF201808)

a 218.71kb interval of markers SSR_50 and SSR_51, can explain 2.26% and 7.85% phenotype variance using both IM and ICIM mapping methods, respectively. The locus *qPro-8-1* is genetically independent of other QTL of protein content. Thus, this work provided elite materials and theoretical foundation in molecular breeding for high-protein soybean varieties.

Key words: protein content; QTL mapping; backcross population; interval mapping; inclusive composite interval mapping

大豆 (*Glycine max* (L.) Merr.) 是当今世界植物蛋白和食用植物油的重要来源, 营养价值很高, 籽粒中不仅含有丰富的蛋白质, 还含有人和动物的 8 种必需氨基酸, 而且氨基酸全面且配比较为合理^[1]。在增加大豆产量的基础上, 提高大豆籽粒蛋白质含量一直是国内外大豆育种研究的重要方向^[2]。黑龙江省是我国重要的大豆商品粮基地, 种植大豆具有比较优势, 种植面积约占全国的 1/3, 但大豆蛋白质含量偏低^[3-4]。利用高蛋白品种对黑龙江省本地优异品种进行蛋白质含量改良可以在保证产量的同时提高大豆的品质。

大豆蛋白质含量是多基因控制的数量性状, 易受环境影响。数量遗传学的发展、现代分子标记技术的完善, 为数量性状的遗传规律提供了更有效的方法, 使发掘有价值的主效基因并进行克隆成为可能。继 1988 年第一张大豆品种 RFLP 图谱问世以来^[5], 国内外学者利用不同的群体 (F_2 , BC 群体、RIL 等) 和不同类型的分子标记 (RFLP、AFLP、SSR 等) 构建出数张不同密度的大豆遗传图谱^[6-7], Keim 等^[8]1990 年首次报道大豆 QTL 定位后, 相继出现了众多与数量性状相关的 QTL 定位的报道。目前为止, 已经有许多关于大豆蛋白 QTL 研究的报道^[9-10], 并可在专业网站 (www.soybase.com) 查到相关定位结果。Sebolt 等^[11]利用两个野生豆 (A81-356022 × PI468916) 构建 BC₃ 群体, 发现野生豆的蛋白质含量与 20 号染色体 (I 连锁群) 紧密连锁的标记显著相关, 可解释的表型变异为 41%~65%。Lee 等^[12]利用两个群体 (Young × PI416937 和 PI97100 × Coker237) F_4 和 F_2 , 通过单因素方差分析和区间作图法在 15 (E 连锁群)、18 (G 连锁群)、12 (H 连锁群) 和 9 (K 连锁群) 号染色体定位了蛋白质含量 QTL, 其中有 1 个 QTL 在两个群体中均与蛋白含量相关, 可解释的表型变异率分别为 4.9% 和 7.4%。Liang 等^[13]利用两年的蛋白含量 (Jindou23 × Huibuzhi, F_{13} RIL) 进行 QTL 定位, 共检测到 6 个蛋白含量 QTL, 分布于 6 (C2 连锁群)、12 (H 连锁群)、14 (B2 连锁群) 和 18 (G

连锁群) 号染色体上, 两年同时定位到 1 个 QTL (Satt557~Satt100), 可解释的表型变异分别为 5.89% 和 6.34%。Jun 等^[9]利用 96 份种质资源和 150 个 SSR 标记, 通过关联分析, 定位到 11 个大豆蛋白的 QTL, 其中 16 号染色体 (J 连锁群) 的 Satt431 和 7 号染色体 (M 连锁群) 的 Satt551 是两个新的蛋白 QTL。Li 等^[14]利用 159 份大豆材料, 通过 GLM 法仅在第 1 号染色体上检测到 1 个与高蛋白含量关联的 QTL, 可解释 5.94% 表型变异。Patil 等^[15]利用 Williams 82 和 PI4834608 构建的 RIL 群体检测到 5 个大豆蛋白含量相关的 QTL, 位于 6、8、13、19 和 20 号染色体上, 可解释 4.6%~19.6% 的表型变异。截至 2019 年 2 月 Soybase 数据库公布的大豆蛋白质含量相关 QTL 为 246 个 (5 个全基因组关联分析获得), 这些 QTL 分布在 20 条染色体, 其中 20 号染色体上检测到的蛋白含量 QTL 数目最多, 为 25 个, 可解释的表型变异范围为 0.07%~65%, 表型变异率大于 10% 的 QTL 有 12 个; 其次是 6 号和 9 号染色体, QTL 数目均为 17 个, 可解释的表型变异范围分别为 0.05%~27.6% 和 0.07%~24.4%。

虽然国内外学者对大豆蛋白含量 QTL 做了大量的研究, 但精细定位鲜见报道, 其中 20 号染色体 (I 连锁群) 蛋白含量 QTL 在公共图谱上最小的遗传距离为 0.54 cM。黎志康^[16]提出利用性状优良的高代回交群体有望育成有应用价值的新品系, 从而将 QTL 定位与育种实践紧密结合起来。本研究以黑龙江省低蛋白高产品种黑河 50 和欧洲引进的高蛋白种质中引 1106 杂交衍生出的 384 个高代回交群体为研究对象, 结合 SSR 分子标记数据和表型数据, 用 QTL ICIMapping4.1 的区间作图法 (IM) 和完备区间作图法 (ICIM), 对大豆蛋白质含量进行 QTL 定位, 以期获得稳定的、与大豆蛋白质含量性状相关的基因, 为分子标记辅助育种和基因挖掘提供依据。

1 材料与方法

1.1 试验群体

以 2008 年黑龙江省农业科学院黑河分院选育

的具有早熟、抗倒、高产、低蛋白品种黑河 50 为母本,欧洲引进的高蛋白种质中引 1106 为父本配置杂交组合,2014 年在 F_2 中选择 1 株高蛋白(蛋白含量为 45.79%)单株,与亲本黑河 50 进行回交,获得 BC_1F_1 ,通过单粒传法加代,获得由 384 个株系组成的回交群体。

1.2 田间试验设计

BC_1F_4 (2016 年)和 BC_1F_6 (2017 年)播种于黑河市黑龙江省农业科学院黑河分院实验基地,均采用 3 m 行长,行距为 0.60 m,株距为 0.05 m, BC_1F_5 (2016 年冬)播种于三亚市中国农业科学院作物科学研究所海南试验站,行距为 0.35 m,株距为 0.06 m,四周均设保护行。全生育期田间管理同一般大田栽培,植株成熟后从每行选择正常一致的 5 株进行收获。

1.3 蛋白质含量测定

每株选择完整的大豆籽粒,用德国 Bruker 公司生产的傅立叶变换近红外光谱仪测定大豆籽粒的蛋白质含量^[17]。每份样品扫描 3 次,收集近红外吸收光谱信息,用已构建好的大豆蛋白质干基模型在软件 OPUS 中分析样品光谱数据得到蛋白质含量数据,以 3 次平均值代表每份样品的蛋白质含量,5 株的平均值作为该株系的蛋白质含量。因亲本中引 1106 在黑河地区生育期过长,收获的种子无法测定蛋白质含量,所以回交群体种植在黑河地区时,中引 1106 的蛋白质含量数据缺失。

1.4 基因组 DNA 提取

2016 年在 BC_1F_6 群体针对 384 个家系分别取不同单株叶片等量混合,采用经典的 CTAB 法^[18]提取基因组 DNA,同时提取亲本的基因组 DNA。使用 1% 的琼脂糖凝胶电泳检测 DNA 质量,并用 Nanodrop One 检测所提 DNA 的浓度,稀释成所用 DNA 浓度 20 ng/ μ L, -20℃ 保存备用。

1.5 SSR 标记选择及基因型鉴定

用实验室保存的分布于 20 个染色体的 872 对 SSR^[19]引物进行多态性筛选,选择多态性高、扩增性强、条带清晰的 SSR 引物,SSR 引物由北京博迈德生物技术有限公司合成。

PCR 反应采用 20 μ L 反应体系,反应液组成为 5 μ L 的 DNA 模板(20 ng/ μ L),2 μ L 10× Easy Taq Buffer(含 Mg^{2+}),1.5 μ L 2.5 mmol/L dNTPs,8.3 μ L 的 ddH₂O,0.2 μ L Easy Taq 酶(5 U/ μ L),3 μ L 2 μ mol/L SSR 引物(正反向引物混合)。PCR 反应程序为 95℃ 预变性 5 min,95℃ 变性 30 s,55℃ (SSR 引

物)退火 30 s,72℃ 延伸 30 s,循环 35 次,最后 72℃ 延伸 5 min,12℃ 保存。扩增产物中加入 6 μ L 6× loading buffer(0.125 g 的溴酚蓝,0.125 g 的二甲苯青,49 mL 的甲酰胺,1 mL 0.5 mol/L, pH8.0 的 EDTA),95℃ 变性 3 min,取出后迅速置于冰水混合物中冷却。

用 6% 变性聚丙烯酰胺凝胶,取 1 μ L 变性后的 PCR 产物上样,恒定功率 100 W,电泳 1~2 h,最后进行银染显色^[20]。将与轮回亲本黑河 50 的条带相同的记为 A,与供体亲本中引 1106 的条带相同的记为 B,杂合的带型记为 H,缺失或模糊不清的带型记为 -。

1.6 数据统计分析

利用 SPSS19.0 和 Microsoft Excel 2013 计算群体蛋白质含量的最大值、最小值、平均值、标准差、偏度、峰度、变异系数和相关系数,利用 ggplot2(R 语言程序包)中的 vioplot 程序分析定位区间边界各个基因型对应的蛋白质含量分布。

1.7 QTL 定位

利用 QTL ICIMapping4.1 软件^[21]中的 map 模块进行遗传图谱构建,根据 Churchill 等^[22]的方法,Kosambi 函数计算图距,模拟运算 1000 次,设定 QTL 的显著性水平为 0.05,利用 BIP 功能中的区间作图法(IM-ADD)和完备区间作图法(ICIM-ADD)进行蛋白质含量 QTL 定位。

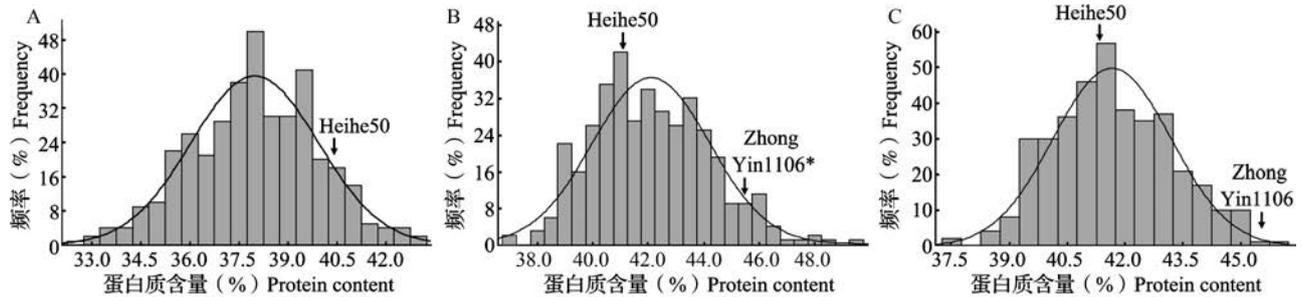
2 结果与分析

2.1 亲本及 BC_1 群体的蛋白质含量分布

轮回亲本黑河 50 和供体亲本中引 1106 的蛋白质含量存在差异,黑河 50 的蛋白质含量为 41.27%,中引 1106 的蛋白质含量为 45.50%,黑河 50 的蛋白质含量在不同年份间均低于中引 1106 的蛋白质含量。群体内 384 个家系间的蛋白质含量变异丰富(表 1)。从表中可以看出, BC_1F_4 家系蛋白质含量的变异幅度是 32.67%~42.93%,变异系数为 5.11%; BC_1F_5 家系蛋白质含量的变异幅度是 37.01%~49.55%,变异系数为 5.01%; BC_1F_6 家系蛋白质含量的变异幅度 37.59%~46.04%,变异系数为 3.67%。回交群体的 3 个世代中均存在明显的超亲分离,且超亲分离比随着环境的变化有所不同,以 BC_1F_5 群体的超亲家系数最多,超亲品系为选育高蛋白品种提供了基础材料。对 BC_1F_4 、 BC_1F_5 和 BC_1F_6 蛋白质含量的频率分布和偏度、峰度进行分析,发现其偏度值较小,偏度值为负值(呈现左偏分布),且峰度较窄,蛋白质含量呈现正态分布(图 1),可用于蛋白质含量的 QTL 定位。

表 1 BC₁F₄、BC₁F₅ 和 BC₁F₆ 3 个世代群体的蛋白质含量统计分析Table 1 Statistical analysis of protein content in three populations of BC₁F₄, BC₁F₅ and BC₁F₆

年份 Year	群体 Population	最小值 (%) Max.	最大值 (%) Min.	平均值 ± 标准差 Mean ± SD	峰度 Kurtosis	偏度 Skewness	变异系数 (%) CV
2016	BC ₁ F ₄	32.67	42.93	38.00 ± 1.94	-0.20	-0.08	5.11
2016 冬	BC ₁ F ₅	37.01	49.55	42.08 ± 2.11	0.39	-0.01	5.01
2017	BC ₁ F ₆	37.59	46.04	41.65 ± 1.53	0.17	-0.34	3.67

图 1 BC₁F₄ 家系 (A)、BC₁F₅ 家系 (B) 和 BC₁F₆ 家系 (C) 的蛋白质含量频率分布Fig.1 Frequency distribution of protein content in BC₁F₄ families (A), BC₁F₅ families (B) and BC₁F₆ families (C)

相关性分析(表 2)表明,世代间蛋白质含量存在一定的相关性,BC₁F₄、BC₁F₅ 和 BC₁F₆ 群体蛋白含量均呈现极显著正相关,相关系数最高达 0.430(表 2)。这证明尽管不同世代蛋白质含量变异幅度不同,但变化趋势相似。

2.2 SSR 标记的筛选

利用 872 对 SSR 引物扩增亲本黑河 50 和中引 1106 的基因组 DNA,共检测出 365 对多态性 SSR 标记(图 2),占引物总数的 41.86%。

表 2 3 个回交世代蛋白质含量的相关性

Table 2 Correlation coefficients among three backcross populations

群体 Population	BC ₁ F ₄	BC ₁ F ₅	BC ₁ F ₆
BC ₁ F ₄	1	<0.0001	<0.0001
BC ₁ F ₅	0.364**	1	<0.0001
BC ₁ F ₆	0.315**	0.430**	1

** 代表在 0.01 水平下显著

** Represent significance at the 0.01 probability levels

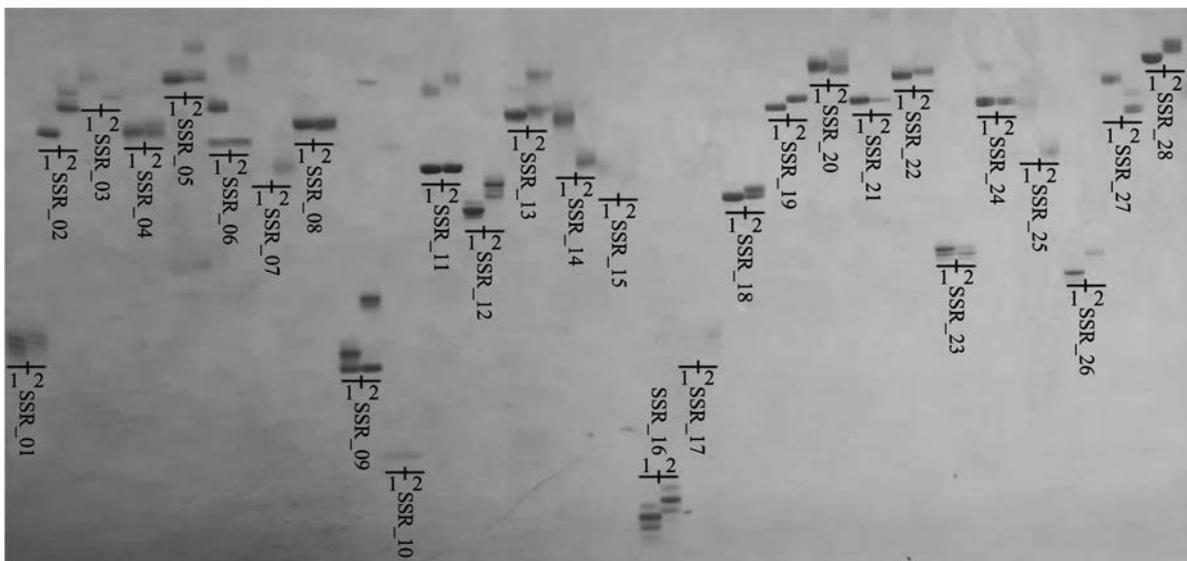


图 2 28 个 SSR 引物在黑河 50 (1) 和中引 1106 (2) 多态性筛选结果

Fig.2 Polymorphism screening results of 28 SSR markers between Heihe50 (1) and ZhongYin1106 (2)

2.3 大豆蛋白质含量 QTL 分析

以筛选出的多态性 SSR 引物为基础构建连锁群,用 ICIMapping4.1 设定 QTL 的显著性水平为 0.05,进行 1000 次模拟运算获得 QTL 显著的 LOD 阈值,区间作图法和完备区间作图法的阈值分别为 2.34 和 2.24。两种方法在 3 条染色体上共定位到 9 个蛋白含量相关的 QTL,同时定位的 QTL 有 1 个

(*qPro-8-1*),该 QTL 位于 SSR_50~SSR_51,标记间的物理距离为 218.71 kb。

用区间作图法 (IM-ADD) 在 BC₁F₆ 群体中共检测到 7 个与蛋白质相关的 QTL (图 3),分别位于 5、6 和 8 号染色体上,其中 8 号染色体上 SSR_63~SSR_64 的表型贡献率最高,为 6.07%。加性效应均为负值,加性效应来自中引 1106 的贡献(表 4)。

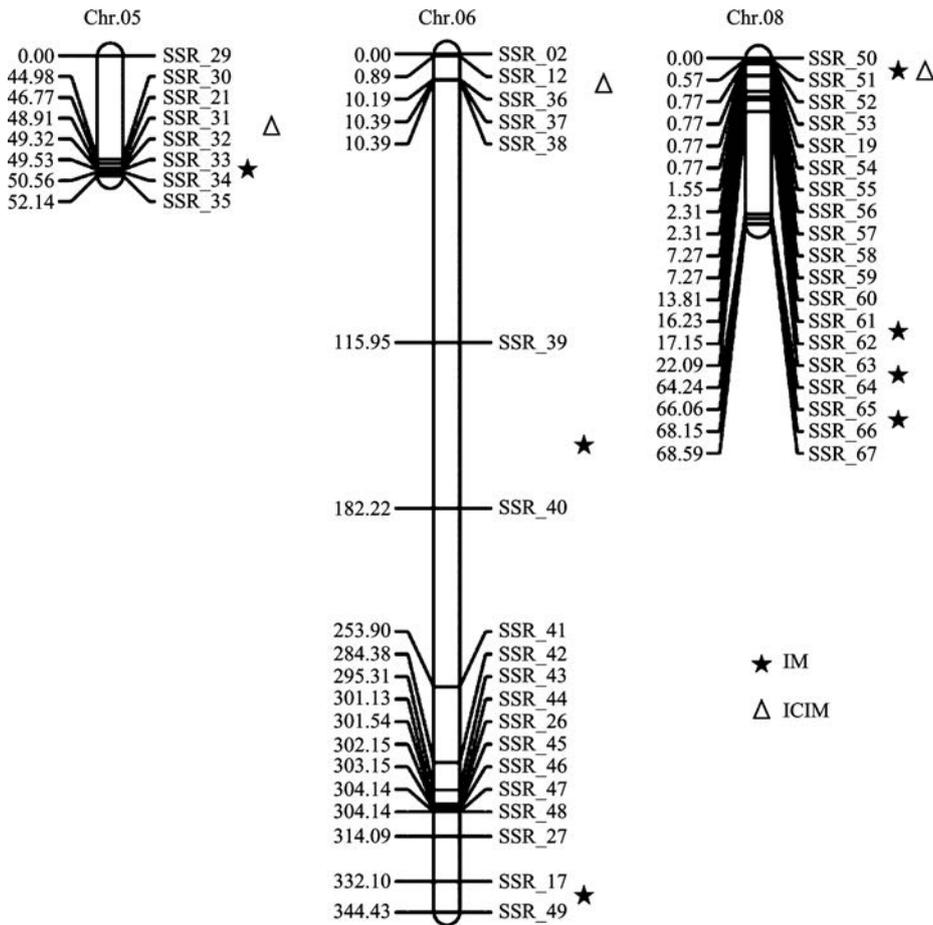


图 3 检测到的 QTL 在连锁群中的分布

Fig.3 Distribution of QTL for protein content on linkage groups

表 4 IM 模型 BC₁F₆ 群体蛋白含量 QTL 的染色体位置和参数

Table 4 Chromosomal locations and parameters associated with the quantitative trait loci (QTL) of protein contents in the BC₁F₆ population by IM

QTL 名称 Name of QTL	染色体 Chromosome	标记区间 Marker interval	LOD	表型变异率 (%) PVE	加性效应 Additive effective
<i>qPro-5-1</i>	5	SSR_33~SSR_34	8.38	2.17	-0.58
<i>qPro-6-1</i>	6	SSR_39~SSR_40	2.72	5.98	-0.99
<i>qPro-6-2</i>	6	SSR_17~SSR_49	3.12	0.83	-0.34
<i>qPro-8-1</i>	8	SSR_50~SSR_51	9.02	2.26	-0.54
<i>qPro-8-2</i>	8	SSR_61~SSR_62	4.33	1.13	-0.38
<i>qPro-8-3</i>	8	SSR_63~SSR_64	6.14	6.07	-0.95
<i>qPro-8-4</i>	8	SSR_65~SSR_66	4.41	1.23	-0.48

用完备区间作图法 (ICIM-ADD) 在 BC_1F_6 群体中共检测到 3 个与蛋白质相关的 QTL (图 3), 分别位于 5、6 和 8 号染色体上, 其中 $SSR_{50}\sim SSR_{51}$ 的表型贡献率最大, 为 7.85%。加性效应均为负值, 加性效应来自中引 1106 的贡献 (表 5)。

2.4 $qPro-8-1$ 位点两侧基因型与蛋白质含量的差异性分析

统计 BC_1F_6 群体在 $qPro-8-1$ 两侧标记 SSR_{50} 和 SSR_{51} 处基因型和蛋白质含量。标记 SSR_{50} 处, 具有黑河 50 基因型 (A) 和中引 1106 基因型

(B) 的家系分别有 266 个和 95 个, 蛋白质含量变化范围分别为 37.00%~48.26% 和 39.06%~49.55%。标记 SSR_{51} 处, 具有黑河 50 基因型 (A) 和中引 1106 基因型 (B) 的家系分别有 266 个和 92 个, 蛋白质含量变化范围分别为 37.23%~48.26% 和 39.06%~49.55%。用 R 语言对基因型为 A 和 B 时的蛋白质含量分布进行差异显著性检验 (图 4), 结果在标记 SSR_{50} 和 SSR_{51} 处, 基因型为 A 和 B 时的蛋白质含量分布均具有极显著差异, 表明 $qPro-8-1$ 真实存在。

表 5 ICIM 模型 BC_1F_6 群体蛋白含量 QTL 的染色体位置和参数

Table 5 Chromosomal locations and parameters associated with the quantitative trait loci (QTL) of protein contents in the BC_1F_6 population by ICIM

QTL 名称 Name of QTL	染色体 Chr.	标记区间 Marker interval	LOD	表型变异率 (%) PVE	加性效应 Additive effective
$qPro-5-2$	5	$SSR_{31}\sim SSR_{32}$	4.09	6.79	-0.39
$qPro-6-3$	6	$SSR_{12}\sim SSR_{36}$	2.32	4.27	-0.30
$qPro-8-1$	8	$SSR_{50}\sim SSR_{51}$	4.81	7.85	-0.38

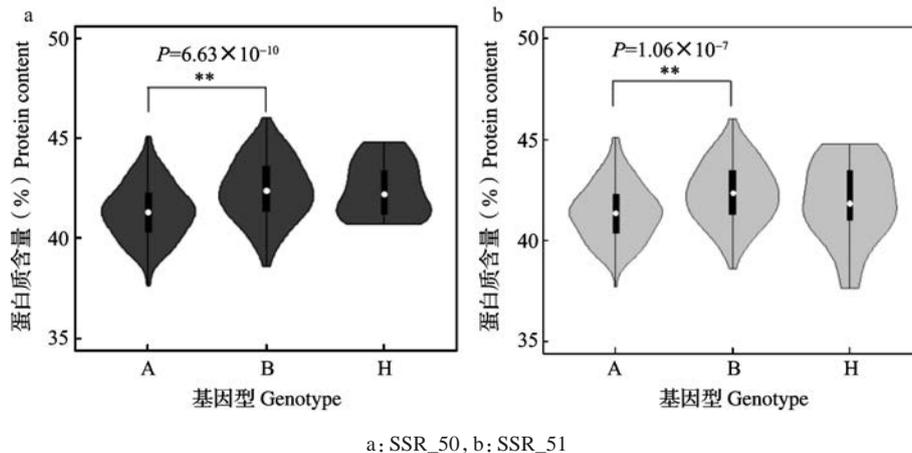


图 4 $qPro-8-1$ 两侧标记 SSR_{50} 、 SSR_{51} 、不同基因型的蛋白含量分布及差异显著性测验

Fig.4 Distribution of protein content and difference significance test on SSR_{50} and SSR_{51}

3 讨论

3.1 BC_1 回交群体的优势

通常用于大豆基因定位研究的群体包括重组自交系 (RIL)、 F_2 、近等基因系等, 与这些群体类型相比, 回交群体具有如下优点: 经过回交和多次自交后, 家系间的遗传背景差异越来越小, 供体背景对目标性状 QTL 效应检测的影响较小, QTL 检测更为准确, 更容易检测到新的位点; 在回交的基础上, 后代含有大量轮回亲本的遗传信息, 增加了微效 QTL 的几率^[23-25]。回交群体结合分子标记技术, 还可以快速改良现有品种的某种特定性状, 快速获得相关

的近等基因系, 成为有应用价值的新品系。许多学者利用回交群体对大豆数量性状进行了 QTL 研究, 包括耐旱^[26]、产量^[27-28]、油酸^[29]、蛋白质和油分含量^[30-32]等性状相关的 QTL, 并且有研究学者^[11]通过回交群体验证不同遗传背景下 QTL 的稳定性。本研究利用回交群体对蛋白质含量进行 QTL 定位, 在保持轮回亲本黑河 50 优势的基础上, 将供体亲本中引 1106 的高蛋白特性导入后代中, 检测到 9 个蛋白 QTL 且检测出 1 个新的蛋白含量 QTL 位点。

3.2 SSR 分子标记的优点

分子标记是随着分子生物学的发展而发展起来

的一种新型的鉴定方法,具有多态性高、检测手段简单迅速、无基因多效性、能够明确辨别等位基因、实验重复性好的特点^[33]。目前作物遗传多样性研究中应用较广的分子标记技术有 RAPD、RFLP、AFLP 和 SSR 等,其中 SSR 标记具有较高的多态性、共显性分离、位点专业化、标记覆盖整个基因组且分布均匀、DNA 样本用量少等优点,在小麦、水稻、玉米、大豆^[34]、高粱、大麦等主要作物中广泛用来构建连锁图谱、进行遗传变异和分子标记等研究。自 Akkaya 等^[35]首次研究大豆基因组中微卫星 DNA (SSR 标记),证明 SSR 标记大量存在且具有丰富的多态性,之后更多的学者利用 SSR 标记构建大豆的遗传连锁图谱,随着大豆参考基因组的公布^[36],设计了大量的 SSR 引物序列并被公开,目前设计出来的大豆基因组 SSR 引物超过 3 万个,为基因定位提供了便利。本研究中仅筛选了分布于 20 条染色体的 872 个 SSR 引物的多态性,进行蛋白含量 QTL 定位。

3.3 大豆蛋白 QTL 定位结果比较

本研究中 IM 作图法定位到的 QTL 数目比 ICIM 作图法多,且 IM 作图法得到的 QTL 在 8 号染色体上形成 QTL 簇,可能与区间作图法本身具有的缺点有关,区间作图法假定每条染色体上至多有一个 QTL,当实际情况不相符时,例如一条染色体上有 2 个 QTL,若 2 个 QTL 的作用方向相反,往往检测不到;若作用方向相同,2 个 QTL 间可能会出现一个“幻影”QTL,且 QTL 的置信区间较大,一般在 10~30 cM 之间,会造成待估 QTL 位置与效应值的偏差^[37];而 ICIM 作图法简化了控制背景遗传变异的过程^[38],有较低的抽样误差,较高的作图效率,有 QTL 的区域 ICIM 有显著的 LOD 值,没有 QTL 的区域 LOD 接近于 0,ICIM 对作图参数有着很好的稳健性,同时也很容易推广到上位性作图^[39-41]。

本研究采用 QTL ICIMapping4.1 的 ICIM 和 IM 两种方法共定位到 9 个蛋白含量 QTL,分布在 5、6 和 8 号染色体上,分别有 2、3 和 4 个蛋白含量 QTL,可解释的表型变异率为 0.83%~7.85%,其中 ICIM 作图法检测到 3 个 QTL,位于 5、6 和 8 号染色体,但 IM 作图法也在 5、6 和 8 号染色体检测出 7 个蛋白含量 QTL,其中 2 种方法同时检测到的 QTL 有 1 个 (*qPro-8-1*),位于 SSR₅₀~SSR₅₁ 之间,区间大小为 218.71 kb,加性效应来自父本中引 1106 的贡献,可解释的遗传变异为 7.85% 和 2.26%,可能是一个控制大豆蛋白质含量的主效应 QTL,由于区间内还存在较多基因,还需利用分离群体进一步缩

小区间,分析出候选基因,用于分子标记育种中,加快改善黑龙江省蛋白质大豆品种含量低的现状。

对定位到的 QTL 与前人研究结果进行比较,发现 9 个蛋白含量 QTL 中有 6 个尚未见报道,其中 5 号染色体上的蛋白含量 QTL (*qPro-5-1* 和 *qPro-5-2*) 与 Mao 等^[42]和 Jun 等^[9]在该染色体上检测到的两个蛋白含量 QTL (*QSPA1_1* 和 *Seed protein 41-1*) 物理距离均相差 1M bp 左右。8 号染色体上的 3 个蛋白含量 QTL (*qPro-8-1*、*qPro-8-2* 和 *qPro-8-3*) 与 Lu 等^[43]在该染色体上检测到的蛋白 QTL (*Seed protein 34-4*) 距离相近,物理距离相差 0.49~3.3 Mbp。其他 3 个蛋白含量 QTL (*qPro-6-1*、*qPro-6-3* 和 *qPro-8-4*) 与前人研究的 QTL 部分重叠,其中 *qPro-6-1* 与 Pathan 等^[44]和 Mao 等^[42]在该染色体上定位到的蛋白含量 QTL (*cqSeed protein-012*、*Seed protein 36-7* 和 *Seed protein 36-8*) 区间部分重叠,*qPro-6-3* 与 Kabelka 等^[45]和 Rossi 等^[46]在 6 号染色体上检测到的 QTL (*Seed protein 21-3* 和 *Seed protein 35-1*) 部分重叠,8 号染色体上的蛋白 QTL (*qPro-8-4*) 与 Brummer 等^[47]在该染色体上定位到的蛋白 QTL (*Seed protein 3-1*) 部分重叠,重叠的蛋白含量 QTL 是否为同一个 QTL,还需通过实验进行验证。

参考文献

- [1] Sharma S, Kaur M, Goyal R, Gill B S. Physical characteristics and nutritional composition of some new soybean (*Glycine max* (L.) Merrill) genotypes. *Journal of Food Science and Technology*, 2014, 51 (3): 551-557
- [2] Helms T C, Orf J H. Protein, oil, and yield of soybean lines selected for increased protein. *Crop Science*, 1998, 38 (3): 707-711
- [3] 罗珊,康玉凡,濮绍京,张丽,李振华,肖伶俐,李永华. 黑河地区 55 份大豆品种资源农艺性状和营养成分的聚类分析. *大豆科学*, 2009, 28 (3): 421-425
Luo S, Kang Y F, Pu S J, Zhang L, Li Z H, Xiao L L, Li Y H. Cluster analysis of agronomic characters and nutrition composition of 55 soybean accessions from Heihe area. *Soybean Science*, 2009, 28 (3): 421-425
- [4] Liu X, Jin J, Wang G, Herbert S J. Soybean yield physiology and development of high-yielding practices in Northeast China. *Field Crops Research*, 2008, 105 (3): 157-171
- [5] Apuya N R, Frazier B L, Keim P, Roth E J, Lark K G. Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merrill. *Theoretical and Applied Genetics*, 1988, 75 (6): 889-901
- [6] 杨喆,关荣霞,王跃强,刘章雄,常汝镇,王曙明,邱丽娟. 大豆遗传图谱的构建和若干农艺性状的 QTL 定位分析. *植物遗传资源学报*, 2004, 5 (4): 309-314
Yang Z, Guan R X, Wang Y Q, Liu Z Q, Chang R Z, Wang

- S M, Qiu L J. Construction of genetic map and QTL analysis for some agronomic traits in soybean. *Journal of Plant Genetic Resources*, 2004, 5(4): 309-314
- [7] 刘硕, 罗玲, 刘章雄, 关荣霞, 邱丽娟. 大豆蛋白质含量 QTL 的“整合”及 Overview 分析. *大豆科学*, 2011, 30(1): 1-14
- Liu S, Luo L, Liu Z X, Guan R X, Qiu L J. Integration of QTLs related to soybean protein content and “Qualification” of them by overview method. *Soybean Science*, 2011, 30(1): 1-14
- [8] Keim P, Diers B W, Shoemaker R C. Genetic analysis of soybean hard seedness with molecular markers. *Theoretical and Applied Genetics*, 1990, 79(4): 465-469
- [9] Jun T H, Van K J, Kim M Y, Lee S H, Walker D R. Association analysis using SSR markers to find QTL for seed protein content in soybean. *Euphytica*, 2008, 162(2): 179-191
- [10] Hwang E Y, Song Q J, Jia G F, Specht J E, Hyten D L, Costa J, Cregan P B. A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics*, 2014, 15: 1-12
- [11] Sebolt A M, Shoemaker R C, Diers B W. Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Science*, 2000, 40(5): 1438-1444
- [12] Lee S H, Bailey M A, Mian M A R, Carter Jr T E, Shipe E R, Ashley D A, Parrott W A, Hussey R S, Boerma H, R. RFLP loci associated with soybean seed protein and oil content across populations and locations. *Theoretical and Applied Genetics*, 1996, 93(5-6): 649-657
- [13] Liang H Z, Yu Y L, Wang S F, Lian Y, Wang T F, Wei Y L, Gong P T, Liu X Y, Fang X J, Zhang M C. QTL mapping of isoflavone, oil and protein content in soybean (*Glycine max* L. Merr.). *Agricultural Science in China*, 2010, 9(8): 1108-1116
- [14] Li Y H, Smulders M J M, Chang R Z, Qiu L J. Genetic diversity and association mapping in a collection of selected Chinese soybean accessions based on SSR marker analysis. *Conservation Genetics*, 2011, 12(5): 1145-1157
- [15] Patil G, Vuong T D, Kale S, Valliyodan B, Deshmukh R, Zhu C S, Wu X L, Bai Y H, Dennis Y B, Lu F, Kumpatla S, Shannon J G, Varshney R K, Nguyen H T. Dissecting genomic hotspots underlying seed protein, oil, and sucrose content in an interspecific mapping population of soybean using high-density linkage mapping. *Plant Biotechnology Journal*, 2018, 16(11): 1939-1953
- [16] 黎志康. 我国水稻分子育种计划的策略. *分子植物育种*, 2005, 3(5): 603-608
- Li Z K. Strategies for molecular rice breeding in China. *Molecular Plant Breeding*, 2005, 3(5): 603-608
- [17] Mark H L. Qualitative near-infrared reflectance analysis using mahalanobis distances. *Analytical Chemistry*, 1985, 57(7): 1449-1456
- [18] Doyle J J, Doyle J L. Isolation of plant DNA from fresh tissue. *Focus*, 1990, 12(13): 39-40
- [19] Song Q J, Jia G F, Zhu Y L, Grant D, Nelson R T, Hwang E Y, Hyten D L, Cregan P B. Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. *Crop Science*, 2010, 50(5): 1950-1960
- [20] Trigiano R N, Caetano-Anolles G. Laboratory exercises on DNA amplification fingerprinting for evaluating the molecular diversity of horticultural species. *HortTechnology*, 1998, 8(3): 413-423
- [21] Meng L, Li H H, Zhang L Y, Wang J K. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop Journal*, 2015, 3(3): 269-283
- [22] Churchill G A, Doerge R W. Empirical threshold values for quantitative trait mapping. *Genetics*, 1994, 138(3): 963-971
- [23] Tanksley S D, Nelson J C. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics*, 1996, 92(2): 191-203
- [24] Kim K S, Diers B W, Hyten D L, Rouf Mian M A, Shannon J G, Nelson R L. Identification of positive yield QTL alleles from exotic soybean germplasm in two backcross populations. *Theoretical and Applied Genetics*, 2012, 125(6): 1353-1369
- [25] 朱月明, 韩粉霞, 孙君明, 闫淑荣, 杨华. 利用回交导入系群体定位大豆蛋白质含量与脂肪含量 QTL. *植物遗传资源学报*, 2017, 18(6): 1204-1209
- Zhu Y M, Han F X, Sun J M, Yan S R, Yang H. Detection of soybean protein and oil content QTL with backcross introgression lines. *Journal of Plant Genetic Resources*, 2017, 18(6): 1204-1209
- [26] 李灿东, 蒋洪蔚, 刘春燕, 郭泰, 王志新, 吴秀红, 郑伟, 邱鹏程, 张闻博, 宋英博, 栾奕娜, 陈庆山, 胡国华. 大豆耐旱选择群体 QTL 定位. *作物学报*, 2011, 37(4): 603-611
- Li C D, Jiang H W, Liu C Y, Guo T, Wang Z X, Wu X H, Zheng W, Qiu P C, Zhang W B, Song Y B, Luan Y N, Chen Q S, Hu G H. QTL identification of drought tolerance to soybean in selection population. *Acta Agronomica Sinica*, 2011, 37(4): 603-611
- [27] Wang D, Graef G L, Procopiuk A M, Diers B W. Identification of putative QTL that underlie yield in interspecific soybean backcross populations. *Theoretical and Applied Genetics*, 2004, 108(3): 458-467
- [28] Guzman P S, Diers B W, Neece D J, Martin S K S, LeRoy A R, Grau C R, Hughes T J, Nelson R L. QTL associated with yield in three backcross-derived populations of soybean. *Crop Science*, 2006, 47(1): 111-122
- [29] 邹筱, 韩粉霞, 陈明阳, 孙君明, 南金平, 闫淑荣, 杨华. 大豆脂肪酸主要组分含量 QTL 定位. *作物学报*, 2014, 40(9): 1595-1603
- Zou X, Han F X, Chen M Y, Sun J M, Nan J P, Yan S R, Yang H. Quantitative trait loci associated with major fatty acid components in soybean. *Acta Agronomica Sinica*, 2014, 40(9): 1595-1603
- [30] 刘顺湖, 周瑞宝, 喻德跃, 陈受宜, 盖钧镒. 大豆蛋白质有关性状的 QTL 定位. *作物学报*, 2009, 35(12): 2139-2149
- Liu S H, Zhou R B, Yu D Y, Chen S Y, Gai J Y. QTL mapping of protein related traits in soybean [*Glycine max*(L.) Merr.]. *Acta Agronomica Sinica*, 2009, 35(12): 2139-2149
- [31] Ma Z Z, Sun D J, Jiang H W, Fan D M, Wang J Z, Zeng Q L, Liu C Y, Li T, Hu G H, Chen Q S. Genotyping and QTL mapping of protein content with wild soybean backcross introgressive lines. *Chinese Journal of Oil Crop Sciences*, 2014, 36(3): 316-322
- [32] Ning X, Wu D P, Zhan Y H, Liu Y, Sun M Y, Zhao X, Teng W

