# 大豆优异种质蛋白质和脂肪含量QTL挖掘

祁 航<sup>1,2</sup>, 王 婉<sup>2</sup>, 唐晓飞<sup>3</sup>, 薛永国<sup>3</sup>, 曹 旦<sup>3</sup>, 刘鑫磊<sup>3</sup>, 栾晓燕<sup>3</sup>, 杜吉到<sup>1</sup>, 邱丽娟<sup>1,2</sup> (<sup>1</sup>黑龙江八-农垦大学农学院, 大庆 163319; <sup>2</sup>农作物基因资源与遗传改良国家重大科学工程/农业农村部种质资源利用重点实验室/中国 农业科学院作物科学研究所, 北京 100081; <sup>3</sup>黑龙江农业科学院大豆研究所, 哈尔滨 150086)

摘要:大豆是全球重要的豆科作物,是人们食用和饲用主要的蛋白质和油脂来源。大豆蛋白质和脂肪含量受多基因控制 且易被环境等因素影响,发掘高蛋白和脂肪基因位点对于定向培育大豆新品种具有重要意义。本研究选用由高产优质品种黑 农84与蛋白含量较高品系京河4号杂交构建的880个家系组成的遗传分离群体,利用中豆芯一号 SNP标记和SSR 分子标记技 术鉴定F<sub>2</sub>群体的基因型,结合蛋白质、脂肪含量表型,通过QTL IciMapping4.2 的完备区间作图法(ICIM-ADD)定位获得2个蛋 白QTL和2个脂肪QTL。qPro\_11\_1和qOil\_11\_1,位于分子标记SSR\_11\_1087与SSR\_11\_1090之间,区间大小为126.27 kb,遗 传贡献率分别为4.05%和3.23%,共有注释基因5个。qPro\_14\_1和qOil\_14\_1,位于SSR\_14\_0421与SSR\_14\_0429之间,区间 大小为246.09kb,遗传贡献率分别为4.67%和7.13%,共有注释基因15个。本研究定位的稳定的蛋白质含量位点和脂肪含量 新位点,为大豆高蛋白和高油分子标记辅助选择育种和基因图位克隆奠定了基础。

关键词:大豆;蛋白质含量;脂肪含量;SSR;QTL定位

# QTL Mining of Protein and Oil Content in Elite Soybean Germplasm

QI Hang<sup>1,2</sup>, WANG Wan<sup>2</sup>, TANG Xiaofei<sup>3</sup>, XUE Yongguo<sup>3</sup>, CAO Dan<sup>3</sup>, LIU Xinlei<sup>3</sup>, LUAN Xiaoyan<sup>3</sup>, DU Jidao<sup>1</sup>, QIU Lijuan<sup>1,2</sup>

(<sup>1</sup>Agronomy of College, Heilongjiang Bayi Agricultural University, Daqing 163319;<sup>2</sup>National Key Facility for Crop Gene Resources and Genetic Improvement / Key Laboratory of Crop Germplasm Utilization, Ministry of Agriculture and Rural Affairs / Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081;<sup>3</sup>Soybean Research Institute, Heilongjiang Academy Agricultural of Science, Harbin 150086)

Abstract: Soybean is an important legume crop in the world, and serves as the main donor of protein and oil for human diet and animal feeding. The protein and oil content are genetically controlled by multiple genes, and are largely environment-dependent. It is of great significance to explore high protein and oil loci in targeted breeding of new soybean varieties. In this study, a genetic segregation population including 880 families was generated by Heinong 84 (high yield and quality) crossed with Jinghe 4 (high protein content), and subjected for the protein and oil content quantification. The genotyping was conducted using ZDX1 SNP array and SSR molecular markers, combine. Two protein QTL and two oil QTL were identified through the complete interval mapping method (ICIM-ADD) using QTL IciMapping 4.2. *qPro\_11\_1* and *qOil\_11\_1* were co-localized in an interval of 126.27 kb between the molecular markers SSR\_11\_1087 and SSR\_11\_1090, showing the genetic contribution rate of 4.05% and 3.23% respectively, as well as five annotation genes in the region. *qPro\_14\_1* and *qOil\_14\_1* were co-localized between the molecular markers SSR\_14\_0421 and SSR\_14\_0429, with the interval size of 246.09 kb with 15 annotation genes, showing the genetic contribution rate of 4.67% and 7.13%

收稿日期: 2023-03-23 修回日期: 2023-04-19 网络出版日期: 2023-05-19

URL: https://doi.org/10.13430/j.cnki.jpgr.20230323002

第一作者研究方向为大豆优异基因挖掘,E-mail: 515073095@qq.com

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

基金项目:中国农业科学院创新工程

Foundation project: The Innovation Project of Chinese Academy of Agricultural Sciences

respectively. Collectively, this study identified the QTL loci on protein content and oil content, thus laying a foundation for the marker-assisted selection breeding and gene map-based cloning of soybean with high protein and high oil content.

Key words: soybean; protein content; oil content; SSR; QTL mapping

大豆[Glycine max (L.) Merr.]是一种重要的经 济农作物,是人类膳食中植物蛋白和食用油的主要 来源之一,籽粒蛋白质和脂肪对人类消费和工业应 用具有很大价值<sup>[1-2]</sup>。籽粒蛋白质是决定大豆营养 价值和经济价值的关键因素<sup>[3]</sup>,大豆蛋白质和脂肪 含量是由多个基因控制并易受环境影响的复杂数 量性状<sup>[4]</sup>。黑龙江为我国主要的大豆商品粮基地, 播种面积占全国的70%,其地理位置及气候条件的 特殊性,使得黑龙江的黑土地大豆种植相比于其他 省市具有优势。对黑龙江省大豆的品质分析结果 表明,大豆蛋白质和脂肪含量与南方等地区相比存 在一定的差距,具体表现是蛋白质含量偏低,蛋脂 总量偏低<sup>[5]</sup>。因此,利用黑龙江大豆品种挖掘大豆 蛋白和脂肪含量相关的QTL/基因,对培育高蛋白、 高油含量新品种具有重要的理论和实践意义。

随着数量遗传学的发展,大豆蛋白质、脂肪含 量QTL/基因可采用多种分子标记技术进行鉴定, 包括限制性片段长度多态性(RFLP, restriction fragment length polymorphism)、扩增片段长度多态 性(AFLP, amplified fragment length polymorphism) 和简单重复序列多态性(SSR, simple sequence repeat polymorphism)等<sup>[6-8]</sup>。1988年Apuya等<sup>[9]</sup>利用RFLP 标记进行连锁分析,构建了第一张大豆品种遗传图 谱。Keim 等<sup>[10-11]</sup>利用驯化大豆与野生大豆杂交的 F2分离群体构建了一个大豆分子标记遗传连锁图 谱,确定了遗传标记与数量性状位点间的关联,并 首次利用RFLP标记对大豆进行QTL定位。此后, 更多的学者利用不同的群体(F,、RIL、BC)和分子标 记(RAPD、AFLP、SSR)构建了密度不同的遗传图 谱<sup>[12-14]</sup>。相对于 RFLP 等标记多态性低的缺点, Akkaya 等<sup>[15]</sup>证实了 SSR 标记在大豆基因组中的高 多态性。1999年, Cregan<sup>[16]</sup>报道了606个SSR基因 座的开发,使大豆连锁图谱的标记密度进一步增 加。Song 等<sup>[17]</sup>构建了一个完整的大豆遗传公共图 谱,其包含1849个标记,其中1015个SSR、709个 RFLP、73个RAPD。SSR分子标记技术成熟且易操 作,SSR分子标记为大豆种质系统的遗传分析提供 了实用有效的工具。

蛋白质和脂肪是大豆品质育种的重要目标,国

内外已有大量的研究,并有专门的学术共享网站 (www.Soybase.com)。Diers 等<sup>[18]</sup>利用栽培与野生 大豆杂交形成的种间群体,首次报道了与大豆蛋白 质和脂肪含量相关的QTL。Lee等<sup>[19]</sup>利用两个群体 的F<sub>4</sub>和F<sub>2</sub>衍生系,通过单因素方差分析和区间作图 法分别在E、G、H、K4个连锁群检测到蛋白质QTL, 可解释的表型变异为4.9%~7.4%,其中在E、G、H3 个连锁群同时检测到与脂肪相关的QTL。Sebolt 等<sup>[20]</sup>利用野生大豆构建的BC,群体,在连锁群I上检 测到蛋白质主效QTL且与蛋白质含量显著相关,可 解释的表型变异为41%~65%。Jun 等<sup>[21]</sup>、Qi 等<sup>[22]</sup>利 用SSR标记,通过关联分析和复合区间方法定位, 定位到多个大豆蛋白质QTL,可解释的表型变异为 4.4%~14.5%。Yesudas 等<sup>[23]</sup>通过 SSR 标记的方差分 析和复合区间映射在第18及20号染色体上发现种 子油分QTL,且在第20号染色体还发现1个种子蛋 白质QTL。Panthee等<sup>[24]</sup>利用RIL群体通过SSR标 记的单因素方差分析和复合区间作图定位到3个种 子脂肪QTL,分布在第1、10和12号染色体。目前 SoyBase数据库公布的大豆蛋白质和脂肪含量相关 QTL数量分别为241和315个,其中主要分布在第 15和20号染色体,可解释的遗传贡献率为0.07%~ 65%。虽然大豆蛋白质和脂肪含量相关QTL鉴定 已有众多报道,但应用于分子育种的QTL较少且区 间较大,仍需要进一步挖掘QTL/基因。

本研究利用黑龙江省高产优质品种黑农 84 为 母本,高蛋白品系京河4号为父本,构建包含 880 个 家系的F<sub>2</sub>群体,结合 SSR 分子标记基因型数据和群 体的表型数据,采用QTL IciMapping4.2完备区间作 图法(ICIM-ADD),定位与蛋白质和脂肪含量相关 的QTL位点,为改善大豆品质性状和分子标记辅助 育种及基因挖掘提供理论依据。

# 1 材料与方法

#### 1.1 群体构建

以黑龙江省农业科学院培育的优质多抗、高产品种黑农84(蛋白质含量43.26%,脂肪含量20.41%)为母本,黑龙江省农业科学院黑河分院培育的高蛋白品系京河4号(蛋白质含量45.21%,脂

肪含量19.45%)为父本。黑农84与京河4号杂交, 通过单粒传法进行加代,构建包含880个家系的F2 遗传分离群体。分别于2019、2020年在黑龙江省农 科院试验基地春播种植F2、F2;3群体,均采用3m行 长,行距为0.60m,株距为0.06m。全生育期田间管 理同一般大田管理,植株成熟后从每行选择正常一 致的5株进行收获。

#### 1.2 表型数据测定

选择表面干净、无裂纹、菌斑、颗粒完整的大豆 种子样品约30~50粒,利用德国Bruker公司生产的 傅立叶变换近红外光谱仪<sup>[25]</sup>对群体样品扫描,采集 样品的近红外吸收光谱信息,用已构建好的大豆蛋 白质、脂肪干基模型,在软件OPUS中分析样品光谱 数据得到蛋白质和脂肪含量数据。F<sub>2</sub>群体每个单株 测定3次重复,取平均值代表该样品的蛋白质和脂 肪含量,F<sub>2:3</sub>群体每个株系取5个单株,每株测定3 次,取5个单株的平均值代表样品的蛋白质和脂肪 含量。

利用 Microsoft Excel 2019 和 Minitab17 分析群体最大值、最小值、平均数、标准差、偏度、峰度、变异系数和相关系数,并绘制蛋白质、脂肪含量分布图,利用R语言对区间两端标记基因型与蛋白质、脂肪含量进行差异性分析。

#### 1.3 基因组 DNA 提取

2019年取黑农 84 与京河 4 号的 F<sub>2</sub>群体 880 个家 系的三出复叶的等量叶片,采用经典的 CTAB 法<sup>[26]</sup> 提取叶片基因组 DNA,同时提取亲本的基因组 DNA。用1%的琼脂糖凝胶电泳和紫外分光光度计 检测 DNA 的质量和浓度。

#### 1.4 中豆芯一号芯片检测

试验所用的中豆芯一号芯片为200K SNP芯片,采用商业试剂盒(天根植物基因组DNA试剂盒, DP305)从大豆幼叶中提取基因组DNA。本研究使 用基于Illumina<sup>®</sup>平台开发的ZDX1 SNP阵列作为分 型工具,使用GenomeStudio软件获取 SNP基因型, 对分型信号进行测试和调整,获得亲本及F<sub>2</sub>群体的 基因型。

#### 1.5 SSR标记选择及基因型鉴定

基于实验室200K芯片QTL定位结果,选择SSR标记进行QTL定位验证及区间缩小。在大豆数据库SoyBase(https://www.soybase.org)网站上选择QTL区间内部及附近的SSR标记,11、14号染色体上共计49对SSR标记(表1),根据引物序列信息,由北京博迈德生物技术有限公司合成。利用合成的SSR标记对黑农84和京河4号进行多态性筛选,选择具有多态性的SSR标记对F<sub>2</sub>群体进行基因型鉴定。

引物序列信息
<b>计物</b> 序列信息

#### Table 1 SSR primer sequence information

标记名称	上游引物序列(5'-3')	下游引物序列(5'-3')
Marker name	Upper primer sequence(5'-3')	Lower primer sequence(5'-3')
11_1040	TGGTTATTTTTGTGGGACATTG	TACTGTTTGAACTTGACTTACTTTTGT
11_1046	TTTGTTGACTTCAATGAATTTGCT	TCCACATATTGTGTTTTGAAGAAA
11_1050	TGCTACCAGTCCAAGCCTTT	AGGGTGTGAAGGGGAAAAAC
11_1052	GATCTCCAAATTAAGCCTTGAAC	TGATGGAAACCTAGCAAGCA
11_1056	TGGGAGATGTCCAATTTTCA	CCTTTTGTCATATGCTACACACG
11_1058	TTTTCTTCTGCACACGTGATT	TTCACAACCATATCCTATGTCCA
11_1059	TTTCAATTTTGCGAGGACAA	TTTAACAAGACAAGAAATGCAAAAA
11_1061	ATCCAGGTAACAACATGGCA	CATGCAAAAGCAAGCCCTAT
11_1063	GCATCCATTCACATAAAATTGG	CGTAAAAAGTCGTTCAAAACCA
11_1064	TTGCGAACAAAATGCAAAAG	TCACAGCGGATTATTTGGTG
11_1067	TGTCCAGCCATGCATTTTAC	GGATGTTCAAATTAAAAATTCAGAAA
11_1068	CACCATCTCACAACTTGAGCA	GCTGATTCCTTGGCAAACAT
11_1070	TCCAATGAAGGATTGGTGTG	GCGGGGGGTGAAACAGTAGTA
11_1071	CCCAAAGATGGGTCTTCTGA	TGTTACCAAAAGAAAGTCACCTTATG
11_1072	TTGGCCCACCTTGATAACTC	TGAAGGTTGAGGATGCATTG
11_1073	AAGCTCAATTTGAACACATAAAAATG	CATCTTAGGCTCCCGGATCT
11_1074	GTTTTCTCCATCACCTCCCA	CCTTGGACGAACATTGGAGT
11_1075	CAAGATACTAATCTCTGCTGAGGC	ATCTTGCCTTGTGCATTGTG

标记名称	上游引物序列(5'-3')	下游引物序列(5'-3')
Marker name	Upper primer sequence(5'-3')	Lower primer sequence(5'-3')
11_1076	GCACGAATTGAAAAAGTTTAAGTTG	CGAGCTGCTACCAGTTTTTGT
11_1077	GAGGCACCGACAAAGTTGAT	CAGTTGATTCGCGCTTAATTT
11_1078	TGAAAGAGTATTTGAAATCGAAGAGA	TCAATGAGGAAGGATCAAGACA
11_1079	CCCTCAAACAATTGGATAAAAGA	GTGGACACTTCATTAAAAGTAGGTT
11_1080	GCGCGGAAGAAGAAGTTTA	CAGAACACTACTCAATGTCATCGTT
11_1081	CCAACTAAATAAGCTAGATCCCATC	AAAAATGGCAAGGAAAAGTGA
11_1082	TTAAGGCTGAGGTCCCCTTT	TCACAGGGTGGTGTTGAAGA
11_1083	AATCACGTGTCAAAATTATTATATGC	TCAAATTAAGTCATCAATGGCA
11_1084	TCGTGCCTTAATTGTGTTGG	CAAATTGTTTTGCCGCTCTT
11_1087	TGGATCGATCTTGGTACGATA	CGACAAATGTGATACGGGAC
11_1089	GCTTTTATTAATGCCAGGGG	TTTCCCTTTTCCTTCTCCGT
11_1090	AAGTGACTTATTAGCCCGATCA	CAGGCCTGCCTGTAAACAAT
11_1091	CCACTTCACAAGTGTGGAAAAA	TGGGTCAAGCAAAAATTTAACA
14_0409	TTGCAGAGTCCACAAACTGC	TCACAAGAAGTGGCCAAAAC
14_0411	GATTCAACTCCCACAAACACA	CTCTGCGAACATTCAACACC
14_0413	GAAGGGAAATGGTTTGGGAT	ACCACTTCCAACGGCACTAC
14_0415	GTGTAATAGTTGTCAAAGTTATACCGA	TGCAGCAATTTTACTGGGTG
14_0417	CGGGCTATCTTCCTCCTCTC	TGCATGAGGTTGTTGTAGACG
14_0419	GAGACGAGAAAAGTGGACGAA	TGAATTGGATTCTGTGTGCAT
14_0421	ACATCTTTGCCATCTCAAACC	CGTCTCAACTCATCTCAATGTCA
14_0422	ATGGCCACTTGTGAAGTTCC	TGGGTGCTTTTACCGAACTC
14_0423	GTGCTATTGGCACTACCCGT	TTTGCATCCCACACAATACAA
14_0425	TCACTGTATTAGATCCAATGCTCAA	TGGATTGAAACGCTACAATGAAT
14_0427	CAATAATCAAAGTTAGGTATCATGAGC	CGTGCTCCTGAATAGACAAGG
14_0429	CAGCCTCAAGAATCCACACA	AGCGCGTGAGAAGTTGTTTC
14_0431	TCAATTGCGTTCAAACACTCA	TCCACATCGAAATTTCTTCAAA
14_0432	GAAGCTTATTTGCTGGCTGG	AACCTGCATGTTGCACAATC
14_0433	TGGACCAACATCGTCACCTA	TTCCAATTTTCCTGTTTCGG
14_0434	GTTGGTTACCCCACGGTATG	TACCATTAGCACGCGTTTTG
14_0435	TCCGGATGGACTTGACTCAT	TGCATGTCAAGCAAGTTGAG
14_0436	GCTTATCATACACGTATAAAGGTTTGC	TCATTGAGAAGGAGTTGTGCT
14_0437	GGCTTATTATTCCTTCTTTTGTCA	AAGTGTATCTTCTCAAATTTTGTCCA
14_0438	TGCTTGGCACTTGTACATCC	GGGAAAACTGAAAATCCCCA
14_0439	CTGAAATGCTTCCCAGTGGT	CTCAAAAGGGGGTTCCATCAA
14_0441	CGGATGATGTTGAAGTCACG	GAACCTCCGGAAAGTTTCTT
14_0442	GCTTGACAAAGGGGGAAAA	CAGGTTTTTAAGAATGAAAACGAAA
14_0443	CTATGTTTAGGGACCGCAGC	TCCGTTGTCAACCTGAAAGTC
14_0444	GGGTGAGGTCTTCGTTGTTC	ACTAGAACTTGAGGCCACGG
14_0445	TGAGCTTTTATCACCCCCAA	CAATTGAAAAGTTTAACAAATTGAGA
14_0446	TGAAAGATGACTAACGACGACAA	TGCGTGTGAAGGGAAATAAA
14_0447	TGATCAAGATTAATCCCTACCACC	CCCTTCATATTGTGTTTTTTCCA

表1(续)

黑色加粗的标记为筛选的多态性SSR标记

The markers in bold black font are screened polymorphic SSR markers

PCR反应采用10 µL反应体系:2 µL的DNA模板(20 ng/µL),1 µL 10×Easy Taq Buffer(含Mg<sup>2+</sup>), 0.8 µL 2.5 mmol/L dNTPs, 5.1 µL的ddH<sub>2</sub>O, 0.1 µL Easy Taq 酶(5 U/µL),1µL 2µmol/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 ℃退火30 s, 72 ℃延伸30 s,循环 34次;最后72 ℃延伸5 min, 12 ℃保存。扩增产物 中加入3 µL 6×loading buffer(终体积50 mL:0.125 g 嗅酚蓝,0.125 g二甲苯青, 49 mL甲酚胺,1 mL 0.5 mol/L pH8.0 的 EDTA),取出 PCR 产物迅速置于4 ℃冰箱 保存。

通过6%非变性聚丙烯酰胺凝胶电泳检测F<sub>2</sub>群体 SSR标记的扩增产物,取1μL加入6×loading buffer 的PCR产物点样,恒定功率100W,电泳约1h,最后 进行银染显色<sup>[27]</sup>。其中与母本黑农84条带相同的 记为A,与父本京河4条带相同的记为B,杂合带型 记为H,缺失或模糊的带型记为-。

#### 1.6 QTL定位

结合检测的表型数据与200K芯片和SSR标记 鉴定的群体基因型数据。使用QTL IciMapping4.2<sup>[28]</sup> 软件中的\*map模块进行遗传图谱构建,根据Churchill 等<sup>[29]</sup>的方法,模拟运算1000次,设定QTL的显著性 水平为0.05,并用完备区间作图法(ICIM)对大豆蛋 白质、脂肪含量进行QTL定位分析,蛋白质含量QTL 命名前缀为qPro,脂肪含量QTL命名前缀为qOil。

#### 1.7 基因功能注释及表达量查询

利用 Phytozome v13(https://phytozome-next.jgi.

Tabla	2	Statistical analysis of materia and all content in Heineng 94× linghe
Table	4	Statistical analysis of protein and on content in memory 84~Jinghe 4

doe.gov)网站, *Glycine max Wm82.a4.v1* 基因组,输入 基因号检索基因的功能注释。在大豆数据库 SoyBase(https://www.soybase.org)网站, Expression 工具栏查询,以Williams82为参考基因组,选择 seed 10DAF(DAF, day after flowering)、seed 14DAF、seed 21DAF、seed 25DAF、seed 28DAF、seed 35DAF和 seed 42DAF种子发育时期,输入基因号查询基因的 表达量,使用R语言进行基因表达量统计作图。

### 2 结果与分析

#### 2.1 亲本及群体的蛋白质、脂肪含量分布

亲本黑农 84 和京河 4 号的蛋白质、脂肪含量 存 在 差 异,2019 年 黑农 84 的 蛋 白 质 含 量 为 43.26%,脂肪含量为 20.41%,京河 4 号的蛋白质含 量为 45.21%,脂肪含量为 19.45%。880 个家系间 的蛋白质和脂肪含量存在较大差异(表 2),2019 年的蛋白质含量变异幅度为 37.71%~50.61%,变 异系数为 4.54%,脂肪含量变异幅度为 11.17%~ 22.42%,变异系数为 6.04%;2020 年的蛋白质含量 变异幅度为 36.79%~49.89%,变异系数为 5.90%, 脂肪含量变异幅度为 15.37%~21.60%,变异系数为 6.30%。两个世代的群体中均存在明显的超亲分 离,对其进行频率分布和偏度、峰度分析,蛋白质 和脂肪含量均呈现正态分布(图 1),可用于蛋白 质和脂肪含量的QTL定位。

	左八	团件01	古河 4 日	分	离群体 Segregating	g population	
Traits	平份 Year	黑权 84 Heinong84	录河4号 Jinghe4	平均值±标准差 Mean±SD	变异系数(%) CV	峰度 Kurtosis	偏度 Skewness
蛋白质含量(%)	2019	43.26±1.04	45.21±1.06	44.20±0.07	4.54	-0.02	0.08
Protein content	2020	41.59±2.09	43.49±1.36	43.77±0.11	5.90	-0.22	-0.33
脂肪含量(%)	2019	20.41±0.93	19.45±1.20	19.54±0.04	6.04	3.18	-0.77
Oil content	2020	20.11±0.84	19.47±0.53	18.65±1.17	6.30	-0.42	-0.05

相关性分析结果显示(表3),两个世代间蛋白质 含量和脂肪含量均存在一定的相关性。F<sub>2</sub>与F<sub>2:3</sub>群 体蛋白质含量相关性系数为0.32,呈极显著的正相 关;脂肪含量相关性系数为0.34,呈极显著的正相关。 F<sub>2</sub>群体的蛋白质和脂肪含量及F<sub>2:3</sub>群体间的相关性 系数均为负数,说明同一世代间蛋白质和脂肪含量 呈现负相关性。这表明尽管不同世代的蛋白质和脂 肪含量变异幅度不同,但整体的变化趋势一致。

#### 2.2 基于200K芯片的蛋白质含量QTL定位

利用F<sub>2</sub>群体880个单株的基因型,结合蛋白质 和脂肪含量表型数据,利用QTLIciMapping4.2完备 区间作图法(ICIM-ADD)进行相关QTL定位,未定 位到脂肪含量QTL,在第11和14号染色体定位到2 个蛋白质含量QTL:*qPro\_1*和*qPro\_2*。其中*qPro\_1* 区间大小为75.72 kb,包含2个基因;*qPro\_2*区间大 小为411.79 kb,包含26个基因(表4)。





性

Table 3	Correlation of	f protein and	l oil content	between	F <sub>2</sub> and	<b>F</b> <sub>2:3</sub>
---------	----------------	---------------	---------------	---------	--------------------	-------------------------

#*/ <del>/</del>	44-14	F <sub>2</sub>		F <sub>2:3</sub>	
和刊4 Population	在 Traits	蛋白质含量 Protein content	脂肪含量 Oil content	蛋白质含量 Protein content	脂肪含量 Oil content
	疋白氏へ早	1	On content		On content
P <sub>2</sub>	蛋日质含重	1			
	脂肪含量	-0.09	1		
F <sub>2:3</sub>	蛋白质含量	0.32***	-0.32***	1	
	脂肪含量	-0.34***	0.34***	-0.91***	1

\*\*\*代表在0.001水平下显著

\*\*\*represent significance at the 0.001 probability levels

#### 表4 200K芯片鉴定的蛋白质含量QTLs

#### Table 4 Protein content QTLs identified by 200K chip

QTL 名称 QTL name	染色体 Chr.	左侧标记 Left marker	右侧标记 Right marker	LOD值 LOD	遗传贡 献率(%)	加性效应 Additive	区间大小(kb) Interval	基因数量 Number of
					PVE	effect		genes
qPro_1	11	Gm11_15669676	Gm11_15593956	2.83	1.53	0.25	75.72	2
qPro_2	14	Gm14_7329914	Gm14_7741706	3.41	2.18	-0.34	411.79	26

-表明加性效应为负值,增效基因来源于京河4号,下同

- indicates that additive effect is negative, synergistic gene from Jinghe 4, the same as below

#### 2.3 蛋白质和脂肪含量QTL精细定位

为进一步分析蛋白质含量位点 qPro\_1和 qPro\_ 2,分别选择两个区间内部及左右两侧共49对 SSR 引物扩增母本黑农 84 和父本京河4号的 DNA 进行 引物多态性筛选,在第11、14号染色体各检测到11

#### 对多态性SSR标记(图2)。

以筛选出的多态性SSR标记鉴定群体基因型,结 合 $F_2$ 与 $F_{2,3}$ 群体的表型数据,利用QTL IciMapping4.2 构建遗传图谱定位QTL位点。结果表明, $F_2$ 与 $F_{2,3}$ 两个世代重复定位到蛋白质和脂肪含量QTL各两 个,分别为 qPro\_11\_1、qPro\_14\_1 和 qOil\_11\_1、 qOil\_14\_1(图3)。qPro\_11\_1 和 qOil\_11\_1 长度为 126.27 kb,遗传贡献率分别为4.05%和3.23%,均包 含5个基因,且qPro\_11\_1与qPro\_1发生重叠,重叠 区间大小为66.37 kb。qPro\_14\_1和qOil\_14\_1长度 为246.09 kb,遗传贡献率分别为4.67%和7.13%,均 包含15个基因,qPro\_14\_1包含于 qPro\_2 内(表6)。

### 2.4 *qPro\_11\_1*不同基因型的蛋白质和脂肪含量差 异显著性分析

F<sub>2</sub>群体中,11\_1090标记A基因型(与黑农84 同)蛋白质含量为44.36%±2.17%、B基因型(与京河 4号同)蛋白质含量为43.74%±2.01%,A相对于B增 加了0.62%;A基因型脂肪含量为19.51%±1.25%、B 基因型脂肪含量为19.79%±1.15%,A相对于B降低 了0.28%。11 1087标记A基因型蛋白质含量为

表6 蛋白质和脂肪含量OTLs

44.42%±2.17%、B基因型蛋白质含量为43.67% ±2.02%,A相对于B增加了0.75%;A基因型脂肪含量为19.47%±1.26%、B基因型脂肪含量为19.80% ±1.18%,A相对于B降低了0.33%(图4)。



 1:黑农84;2:京河4号
 1: Heinong 84;2: Jinghe 4
 图 2 部分 SSR标记在黑农84和京河4号多态性筛选结果
 Fig. 2 Polymorphism of partial SSR markers between Heinong 84 and Jinghe 4



a:蛋白质和脂肪QTL定位遗传图谱,染色体左侧绿色标记为qPro\_1和qPro\_2两端标记,红色标记为qPro\_11\_1、qOil\_11\_1、qPro\_14\_1和 qOil\_14\_1两端标记;b:qPro\_11\_1和qOil\_11\_1定位LOD图;c:qPro\_14\_1和qOil\_14\_1定位LOD图

a: Protein and oil content QTL mapping genetic map, *qPro\_1* and *qPro\_2* are marked in green on the left side of chromosome, and *qPro\_11\_1*, *qOil\_11\_1*, *qPro\_14\_1* and *qOil\_14\_1* are marked in red; b: Locate LODs in *qPro\_11\_1* and *qOil\_11\_1*; c: Locate LODs in *qPro\_14\_1* and *qOil\_14\_1* **图 3** Chr.11和Chr.14蛋白质和脂肪含量QTL定位

Fig. 3 Protein and oil content QTL mapping in Chr.11 and Chr.14

Table 6 Prot	ein and oil con	tent QTLs						
QTL名称 QTL name	左侧标记 Left marker	右侧标记 Right marker	物理位置(bp) Position	区间大小(kb) Interval	LOD	遗传贡 献率(%) PVE	加性效应 Additive effect	基因数量 Number of genes
qPro_11_1	11_1090	11_1087	15534064~15660330	126.27	4.51	4.05	0.74	5
qOil_11_1	11_1090	11_1087	15534064~15660330	126.27	3.90	3.23	-0.33	5
qPro_14_1	14_0421	14_0429	7468386~7714532	246.09	3.31	4.67	-0.61	15
qOil_14_1	14_0421	14_0429	7468386~7714532	246.09	6.78	7.13	0.39	15



\*、\*\*、\*\*\*、\*\*\*\*分别表示在0.05、0.01、0.001、0.0001水平差异显著

a, b: 11\_1090 marker; c, d: 11\_1087 marker; A: The same genotype as Heinong84; B: The same genotype as Jinghe4, the same as below;
\*, \*\*, \*\*\* and \*\*\*\* indicated significant differences at 0.05, 0.01, 0.001 and 0.0001 levels respectively

图4 Chr.11定位QTL的蛋白质和脂肪含量差异显著性测验



 $F_{2:3}$ 群体中,11\_1090标记A基因型蛋白质含量 为44.24%±2.78%、B基因型蛋白质含量为42.84% ±2.51%,A相对于B增加了1.40%;A基因型脂肪含 量为18.40%±1.31%、B基因型脂肪含量为19.02% ±1.07%,A相对于B降低了0.62%。11\_1087标记A 基因型蛋白质含量为44.34%±2.75%、B基因型蛋白 质含量为42.75%±2.54%,A相对于B增加了1.59%; A基因型脂肪含量为18.37%±1.31%、B基因型脂肪 含量为19.05%±1.08%,A相对于B降低了 0.68%(图4)。

结果表明两个世代的两个位点的不同基因型 在蛋白质和脂肪含量分布上均具有显著性差异,表 明该标记与蛋白质和脂肪含量存在密切关联。

#### 2.5 qPro\_14\_1 不同基因型的蛋白质和脂肪含量差 异显著性分析

 $F_2$ 群体中,14\_0421标记A基因型蛋白质含量 为44.10%±2.06%、B基因型蛋白质含量为44.77% ±2.07%,A相对于B降低了0.67%;A基因型脂肪含量 为19.62%±1.23%、B基因型脂肪含量为19.18% ±1.56%,A相对于B增加了0.44%。14\_0429标记A基 因型蛋白质含量为44.16%±2.01%、B基因型蛋白质含 量为44.74%±2.01%,A相对于B降低了0.58%;A基因 型脂肪含量为19.62%±1.18%、B基因型脂肪含量为19.17%±1.08%,A相对于B增加了0.45%(图5)。

 $F_{2,3}$ 群体中,14\_0421标记A基因型蛋白质含量 为43.50%±2.75%、B基因型蛋白质含量为44.67% ±2.19%,A相对于B降低了1.17%;A基因型脂肪含 量为18.91%±1.22%、B基因型脂肪含量为18.13% ±1.06%,A相对于B增加了0.78%。14\_0429标记A 基因型蛋白质含量为43.57%±2.66%、B基因型蛋 白质含量为44.46%±2.24%,A相对于B降低了 0.89%;A基因型脂肪含量为18.83%±1.22%、B基因 型脂肪含量为18.24%±1.08%,A相对于B增加了 0.59%(图5)。

结果表明两个世代的两个位点不同基因型的 蛋白质和脂肪含量分布均具有显著性差异,表明该 标记与蛋白质和脂肪含量存在密切关联。

#### 2.6 候选区间基因功能注释及表达量分析

通过SoyBase(https://www.soybase.org)数据库 查询 qPro\_11\_1和 qPro\_14\_1区间内共20个基因, 利用 Phytozome v13(https://phytozome-next.jgi.doe. gov)检索其中15个基因具有功能注释(表7),根据 基因功能注释推测3个基因与蛋白质含量相关, Glyma. 11g164300具有蛋白磷酸酶2C功能, Glyma. 14g085600 和 Glyma. 14g085800 参与半胱氨酸生物合成。利用 SoyBase 数据库对基因在种子发育不同时期的表达量进行查询分析(图6),表达量较高的有3个: Glyma. 11g164700、Glyma. 14g08580

0、Glyma. 14g086000。 推测 Glyma. 11g164300、 Glyma. 11g164700、Glyma. 14g085600、Glyma. 14g085 800、Glyma. 14g086000可能为候选基因,但需要进 一步的研究验证。



表7 候选区间内基齿功能注	料
---------------	---

 Table 7
 Gene function annotation in candidate interval

基因名称 Gene name	功能注释 Functional annotation	基因名称 Gene name	功能注释 Functional annotation
Glyma.11g164300	蛋白磷酸酶2C相关	Glyma.14g085300	未知
Glyma.11g164400	未知	Glyma.14g085400	富含亮氨酸重复序列受体样蛋白激酶
Glyma.11g164500	[酰基载体蛋白]S-丙二酰转移酶/MCAT	Glyma.14g085500	WRKY转录因子相关
Glyma.11g164600	葡萄糖基/葡萄糖醛酸转移	Glyma.14g085600	半胱氨酸蛋白酶家族C1相关
Glyma.11g164700	肉桂酰辅酶A还原酶样蛋白	Glyma.14g085700	EF-钙结合域蛋白
Glyma.14g084800	蛋白核糖核酸酶P1	Glyma.14g085800	半胱氨酸蛋白酶家族C1相关
Glyma.14g084900	未知	Glyma.14g085900	甘露糖-对-多糖醇利用缺陷1 LEC35 相关
Glyma.14g085000	RAN结合蛋白9相关	Glyma.14g086000	F-box蛋白 SKIP2
Glyma.14g085100	丝氨酸棕榈酰转移酶1	Glyma.14g086100	未知
Glyma.14g085200	驱动蛋白	Glyma.14g086200	未知



Fig. 6 Relative expression levels of seeds at the different developmental stages in the candidate interval of SoyBase website

## 3 讨论

#### 3.1 群体亲本材料的选择

黑龙江省农业科学院利用分子标记辅助育种, 通过杂交、回交等方法对多个目标基因进行选择, 选育了遗传背景一致且农艺性状优良的品种黑农 84<sup>[30]</sup>。黑农84聚集了亲本的优点,遗传基础更加丰 富,是优质高产的黑龙江大豆品种。京河4号为黑 龙江省农业科学院黑河分院选育的优质高蛋白大 豆品系,由黑河50与中引1106杂交获得,连续3年 蛋白质含量均稳定高于45%。利用黑农84和京河4 号杂交构建遗传群体进行蛋白质和脂肪含量QTL 定位,可以为筛选优质高产且高蛋白的大豆材料奠 定基础。

#### 3.2 分子标记的应用

提高大豆种子蛋白质和脂肪含量是大豆育种的主要目标之一,分子标记辅助选择育种为高品质大豆育种提供了一个新的途径<sup>[31]</sup>。分子标记技术的发展为分子育种提供了多种类型的分子标记,例如 RAPD、RFLP、AFLP、SSR 和 SNPs (Single nucleotide polymorphism)等<sup>[32-33]</sup>。其中 SSR 已被提倡为强大的分子标记系统,是指示重要农艺性状和育种领域应用最为广泛的一项技术,具有高重复性、共显性遗传和广泛的基因组覆盖等特点<sup>[34-35]</sup>。

目前SSR分子标记在水稻、小麦、玉米、大麦、大豆 等主要作物中广泛应用于遗传图谱的构建。 Sharma等<sup>[36]</sup>、Jlassi等<sup>[37]</sup>、Farooqi等<sup>[38]</sup>利用SSR标记 定位到较多与品质和抗逆性相关的QTL及等位基 因。同时随着SNPs的出现,使得SNP芯片广泛应 用于大豆、水稻、小麦等作物的遗传多样性研究、高 密度遗传图谱构建以及许多复杂数量性状相关的 基因座或基因鉴定<sup>[39-40]</sup>。例如Nsabiyera等<sup>[41]</sup>、 Kim<sup>[42]</sup>、Phansak等<sup>[43]</sup>利用SNP芯片鉴定到稳定的品 质和抗性QTL及等位基因。本研究同时利用SSR 标记和SNP芯片高效快速地鉴定到2个稳定的蛋白 质含量QTL,1个稳定的脂肪含量QTL及1个新的 脂肪含量QTL,为进一步基因挖掘提供了基础,有 助于培育大豆高蛋白高油新种质。

#### 3.3 与前人定位QTL的比较

目前大豆 SoyBase(http://www.soybase.org/)数 据库中在第11、14号染色体上定位的蛋白质含量 QTL 位点各有8个,脂肪含量QTL 数量分别为4和 19。Gai等<sup>[44]</sup>检测到两个蛋白质含量QTL,这两个 QTL 均与 $qPro_{11}$ 发生重叠,大小为126.27 kb。 Akond 等<sup>[45]</sup>检测到的蛋白质含量QTL,长度为 6.37 Mb,与 $qPro_{14}$ 1重叠,大小为246 kb。Diers 等<sup>[46]</sup>利用  $F_{2,3}$ 家系鉴定到含油量显著位点,长度为 1.56 Mb,与qOil 14 1位置最近,距离0.62 Mb。Qi 等<sup>[47]</sup>检测到的脂肪QTL大小为12.14 Mb,与qOil\_ 11\_1产生11 kb的重叠。数量性状容易受环境因素 的影响,但在不同年份、不同地点和不同材料间多 次定位到同一QTL,说明该QTL是可稳定遗传的 QTL位点<sup>[48-49]</sup>。本研究定位的qPro\_11\_1、qPro\_14\_ 1及qOil\_11\_1与已发表的不同遗传背景的QTL重 叠,说明是稳定主效的QTL,而qOil\_14\_1为1个新 的脂肪QTL位点,可用于下一步的精细定位和图位 克隆等研究。

#### 3.4 候选基因预测

根据 SoyBase (https://www.soybase.org/)数据 库查询的基因注释推测, Glyma. 14g085600和 Glyma. 14g085800作为半胱氨酸蛋白酶参与生物过 程,半胱氨酸为含硫氨基酸,是大豆蛋白质所需的 必需氨基酸。研究表明,在大豆中表达玉米醇溶蛋 白,半胱氨酸和蛋氨酸含量分别提高了 2.36%和 3.08%,影响大豆种子的蛋白质结构<sup>[50]</sup>。大豆蛋白 质在籽粒发育过程中积累,并易受环境因素影响, Glyma. 11g164300蛋白磷酸酶在生长发育和环境胁 迫响应中起重要作用,可能调控大豆籽粒蛋白质含 量<sup>[51]</sup>。因此推测上述基因可能参与大豆蛋白质合 成与积累,但仍需进一步实验鉴定与验证。

#### 参考文献

- Huang J, Ma Q, Cai Z, Xia Q, Li S, Jia J, Chu L, Lian T, Nian H, Cheng Y. Identification and mapping of stable QTLs for seed oil and protein content in soybean [*Glycine max* (L.) Merr.]. Journal of Agricultural and Food Chemistry, 2020, 68 (23): 6448-6460
- [2] Li X, Wang P, Zhang K, Liu S, Qi Z, Fang Y, Wang Y, Tian X, Song J, Wang J, Yang C, Sun X, Tian Z, Li W X, Ning H. Fine mapping QTL and mining genes for protein content in soybean by the combination of linkage and association analysis. Theoretical and Applied Genetics, 2021, 134: 1095-1122
- [3] Patil G, Mian R, Vuong T, Pantalone V, Song Q, Chen P, Shannon G J, Carter T C, Nguyen H T. Molecular mapping and genomics of soybean seed protein: A review and perspective for the future. Theoretical and Applied Genetics, 2017, 130: 1975-1991
- [4] Pathan S M, Vuong T, Clark K, Lee J D, Shannon J G, Roberts C A, Ellersieck M R, Burton J W, Cregan P B, Hyten D L, Nguyen H T, Sleper D A. Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean. Crop Science, 2013, 53 (3): 765-774
- [5] Liu X B, Jin J, Wang G H, Herbert S J. Soybean yield physiology and development of high-yielding practices in

Northeast China. Field Crops Research, 2008, 105 (3) : 157-171

- [6] Kumar S P J, Susmita C, Sripathy K V, Agarwal D K, Pal G, Singh A N, Kumar S, Rai A K, Gandara J S. Molecular characterization and genetic diversity studies of Indian soybean (*Glycine max* (L.) Merr.) cultivars using SSR markers. Molecular Biology Reports, 2022, 49: 2129-2140
- Xia Z, Tsubokura Y, Hoshi M, Hanawa M, Yano C, Okamura K, Ahmed T A, Anai T, Watanabe S, Hayashi M, Kawai T, Hossain K G, Masaki H, Asai K, Yamanaka N, Kubo N, Kadowaki K, Nagamura Y, Yano M, Sasaki T, Harada K. An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using a single F<sub>2</sub> population. DNA Research, 2007, 14(6): 257-269
- [8] Chugh V, Kaur D, Purwar S, Kaushik P, Sharma V, Kumar H, Rai A, Singh C M, Kamaluddin, Dubey R B. Applications of molecular markers for developing abiotic-stress-resilient oilseed crops. Life Science, 2023, 13(1): 88
- [9] 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: 889-901
- [10] Keim P S, Diers B W, Olson T C, Shoemaker R C. RFLP mapping in soybean: Association between marker loci and variation in quantitative traits. Genetics, 1990, 126: 735-742
- [11] Keim P, Diers B W, Shoemaker R C. Genetic analysis of soybean hard seededness with molecular markers. Theoretical and Applied Genetics, 1990, 79: 465-469
- [12] 邱丽娟.利用RAPD标记鉴定大豆种质.作物学报,1997,23
   (4):408-417
   Qiu L J. Identification of soybean germplasm by RAPD markers. Acta Agronomica Sinica, 1997, 23(4):408-417
- [13] Young W P, Schupp J M, Keim P. DNA methylation and AFLP marker distribution in the soybean genome. Theoretical and Applied Genetics, 1999, 99: 785-792
- [14] Mulato B M, Moller M, Zucchi M I, Quecini V, Pinheiro J B. Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. Publications of the Astrophysics Branch Ottawa, 2010, 45: 276-283
- [15] Akkaya M S, Bhagwat A A, Cregan P B. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics, 1992, 132(4): 1131-1139
- [16] Cregan P B, Jarvik T, Bush A L, Shoemaker R C, Lark K G, Kahler A L, Kaya N, Vantoai T T, Lohnes D G, Chung J. An integrated genetic linkage map of the soybean genome. Crop Science, 1999, 39(5): 1464-1490
- [17] Song Q J, Marek L F, Shoemaker R C, Lark K G, Concibido V C, Delannay X, Specht J E, Cregan P B. A new integrated genetic linkage map of the soybean. Theoretical and Applied Genetics, 2004, 109: 122-128
- [18] Diers B W, Shoemaker R C. Restriction fragment length polymorphism analysis of soybean fatty acid content. Journal

of the American Oil Chemists Society, 1992, 69: 1242-1244

- [19] Lee S H, Bailey M A, Mian M A R, Carter 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: 649-657
- [20] 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
- [21] Jun T H, Van K, 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: 179-191
- [22] Qi Z, Hou M, Han X, Liu Ch Y, Jiang H W, Xin A W, Hu G H, Chen Q S. Identification of quantitative trait loci (QTLs) for seed protein concentration in soybean and analysis for additive effects and epistatic effects of QTLs under multiple environments. Plant Breeding, 2014, 133(4): 499-507
- [23] Yesudas C R, Bashir R, Geisler M B, Lightfoot D A. Identification of germplasm with stacked QTL underlying seed traits in an inbred soybean population from cultivars Essex and Forrest. Molecular Breeding, 2013, 31: 693-703
- Panthee D R, Pantalone V R, West D R, Saxton A M, Sams C
   E. Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. Crop Science, 2005, 45(5): 2015-2022
- [25] Zhu Z Y, Chen S B, Wu X Y, Xing C R, Yuan J. Determination of soybean routine quality parameters using near-infrared spectroscopy. Food Sciences and Nutrition, 2018, 6(4): 1109-1118
- [26] Allen G C, Flores-Vergara M A, Krasynanski S, Thompson W F. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nature Protocols, 2006, 1(5): 2320-2325
- [27] Huang L, Deng X H, Li R H, Xia Y S, Bai G H, Siddique K H M, Guo P G. A fast silver staining protocol enabling simple and efficient detection of SSR markers using a non-denaturing polyacrylamide gel. Journal of Visualized Experiments, 2018 (134): e57192
- [28] 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. The Crop Journal, 2015, 3(3): 269-283
- [29] Churchill G A, Doerge R W. Empirical threshold values for quantitative trait mapping. Genetics, 1994, 138(3): 963-971
- [30] 栾晓燕,刘鑫磊,薛永国,马岩松,王家军,张必弦,于佰 双,王广金.多抗高产大豆新品种黑农84的选育研究.大豆 科学,2018,37(6):839-842

Luan X Y, Liu X L, Xue Y G, Ma Y S, Wang J J, Zhang B X, Yu B S, Wang G J. Innovating soybean Heinong 84 with high quality and multi-resistance by molecular marker gene polymerization. Soybean Science, 2018, 37(6): 839-842

- [31] Zhang Y H, Liu M F, He J B, Wang Y F, Xing G N, Li Y, Yang S P, Zhao T J, Gai J Y. Marker-assisted breeding for transgressive seed protein content in soybean [*Glycine max* (L.) Merr.]. Theoretical and Applied Genetics, 2015, 128: 1061-1072
- [32] Kage U, Kumar A, Dhokane D, Karre S, Kushalappa A C. Functional molecular markers for crop improvement. Critical Reviews in Biotechnology, 2016, 36(5): 917-930
- [33] Bevan M W, Uauy C. Genomics reveals new landscapes for crop improvement. Genome Biology, 2013, 14(6): 1-11
- [34] Azizi M M F, Lau H Y, Abu-Bakar N. Integration of advanced technologies for plant variety and cultivar identification. Journal of Biosciences, 2021, 46: 1-20
- [35] Hwang T A E Y, Sayama T, Takahashi M, Takada Y, Nakamoto Y, Funatsuki H, Hisano H, Sasamoto S, Sato S, Tabata S, Kono I, Hoshi M, Hanawa M, Yano C, Xia Z, Harada K, Kitamura K, Ishimoto M. High-density integrated linkage map based on SSR markers in soybean. DNA Research, 2009, 16(4): 213-225
- [36] Sharma M, Gangurde S S, Salgotra R K, Kumar B, Singh A K, Pandey M K. Genetic mapping for grain quality and yield-attributed traits in Basmati rice using SSR-based genetic map. Journal of Biosciences, 2021, 46(3): 50
- [37] Jlassi I, Bnejdi F, Saadoun M, Hajji A, Mansouri D, Ben-Attia M, EI-Gazzah M, EI-Bok S. SSR markers and seed quality traits revealed genetic diversity in durum wheat (*Triticum durum* Desf.). Molecular Biology Reports, 2021, 48 (4): 3185-3193
- [38] Farooqi M Q U, Sa K J, Hong T K, Lee J K. Bulk segregant analysis (BSA) for improving cold stress resistance in maize using SSR markers. Genetics and Molecular Research, 2016, 15(4): 4210-4238
- [39] Chen H D, Xie W B, He H, Yu H H, Chen W, Li J, Yu R B, Yao Y, Zhang W H, He Y Q, Tang X Y, Zhou F S, Deng X W, Zhang Q F. A high-density SNP genotyping array for rice biology and molecular breeding. Molecular Plant, 2014, 7 (3): 541-553
- [40] Winfield M O, Allen A M, Burridge A J, Barker G A, Benbow H R, Wilkinson P A, Coghill J, Waterfall C, Davassi A, Scopes G, Pirani A, Webster T, Brew F, Bloor C, King J, West C, Griffiths S, King L, Bentley A R, Edwards K J. High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotechnology Journal, 2016, 14(5): 1195-1206
- [41] Nsabiyera V, Baranwal D, Qureshi N, Kay P, Forrest K, Valarik M, Dolezel J, Hayden M J, Bariana H S, Bansal U K. Fine mapping of Lr49 using 90K SNP chip array and flowsorted chromosome sequencing in wheat. Frontiers of Plant Science, 2020, 10: 1787
- [42] Kim S M. Identification of novel recessive gene xa44 (t) conferring resistance to bacterial blight races in rice by QTL linkage analysis using an SNP chip. Theoretical and Applied

Genetics, 2018, 131: 2733-2743

- [43] Phansak P, Soonsuwon W, Hyten D L, Song Q J, Cregan P B, Graef G L, Specht J E. Multi-population selective genotyping to identify soybean [*Glycine max* (L.) Merr.] seed protein and oil QTLs. Genes Genomes Genetics, 2016, 6(6): 1635-1648
- [44] Gai J Y, Wang Y J, Wu X L, Chen S Y. A comparative study on segregation analysis and QTL mapping of quantitative traits in plants-with a case in soybean. Frontiers of Agriculture in China, 2007, 1: 1-7
- [45] Akond M, Liu S, Boney M, Kantartzi S K, Kassem M A. Identification of quantitative trait loci (QTL) underlying protein, oil, and five major fatty acids' contents in soybean. American Journal of Plant Sciences, 2014, 5: 158-167
- [46] Diers B W, Keim P, Fehr W R, Shoemaker R C. RFLP analysis of soybean seed protein and oil content. Theoretical and Applied Genetics, 1992, 83: 608-612
- [47] Qi Z M, Wu Q, Han X, Sun Y N, Du X Y, Liu C Y, Jiang H W, Hu G H, Chen Q S. Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. Euphytica, 2011, 179: 499-514
- [48] 武阳春, 郭兵福, 谷勇哲, 栾晓燕, 邱红梅, 刘鑫磊, 李海

燕, 邱丽娟. 大豆蛋白含量新位点 qPRO-19-1 的定位. 植物 遗传资源学报, 2021, 22(1): 139-148

Wu Y C, Guo B F, Gu Y Z, Luan X Y, Qiu H M, Liu X L, Li H Y, Qiu L J. Localization of a new protein content locus *qPRO-19-1* in soybean. Journal of Plant Genetic Resources, 2021, 22(1): 139-148

- [49] 王婉, 韩德志, 闫洪睿, 栾晓燕, 王俊, 邱丽娟. 大豆高蛋白 种质中引 1106蛋白质含量的QTL分析. 植物遗传资源学报, 2020, 21(1): 130-138
  Wang W, Han D Z, Yan H R, Luan X Y, Wang J, Qiu L J. QTL analysis of protein content in soybean high-protein germplasm for citation 1106. Journal of Plant Genetic Resources, 2020, 21(1): 130-138
- [50] Kim W S, Krishnan H B. Impact of co-expression of maize 11 and 18 kDa δ-zeins and 27 kDa γ-zein in transgenic soybeans on protein body structure and sulfur amino acid content. Plant Science, 2019, 280: 340-347
- [51] Bai G, Yang D H, Zhao Y, Ha S, Yang F, Ma J, Gao X S, Wang Z M, Zhu J K. Interactions between soybean ABA receptors and type 2C protein phosphatases. Plant Molecular Biology, 2013, 83: 651-664