



2022 年大豆抗病性、产量和品质相关性状分子标记研究进展

韩英鹏, 杨振红

(东北农业大学 大豆研究所/大豆生物学教育部重点实验室/农业农村部东北大豆生物学与遗传育种重点实验室, 黑龙江 哈尔滨 150030)

摘要: 分子标记的应用对解析作物性状的分子遗传基础至关重要。当下, 单核苷酸多态性 (Single Nucleotide Polymorphism, SNP) 是研究作物性状相关位点的重要分子标记, 可通过连锁分析和关联分析应用于大豆育种工作。相较传统杂交选育手段, 分子标记辅助育种 (Marker-Assisted Selection, MAS) 以杂交育种为基础, 凭借分子标记与目标性状基因紧密连锁的特点, 通过检测分子标记即可高效快速检测目的基因的筛选策略, 成为备受关注的育种方式。随着大豆各组学学科的兴起, 基因组学研究与其他组学之间的联合分析使分子标记的应用范围得到进一步延伸, 成为大豆现代育种工作中的关键部分, 使所筛选的遗传位点被有效利用。此外, 随着生物技术的不断更新, 以往被忽略的 microRNA (non-coding single stranded RNA, miRNA) 和长链非编码 RNA (long non-coding RNA, lncRNA) 重新进入育种家的视野, 使其与分子标记相关联的目的性状基因组区域能够被更好识别。为了详细梳理分子标记在大豆育种中的应用范围, 本综述汇总 2022 年大豆多种抗病性、产量和品质相关性状分子标记的研究进展, 旨在为大豆分子标记辅助育种工作提供参考。

关键词: 大豆; 分子标记; MAS; 农艺性状

Research Progress of Molecular Markers Related to Soybean Disease Resistance, Yield and Quality Traits in 2022

HAN Yingpeng, YANG Zhenhong

(Soybean Research Institute, Northeast Agricultural University/Key Laboratory of Soybean Biology in Chinese Ministry of Education/Northeast Key Laboratory of Soybean Biology, Genetics and Breeding, Ministry of Agriculture and Rural Affairs, Harbin 150030, China)

Abstract: The application of molecular markers is crucial for the analysis of molecular genetic basis of crop traits. Currently, Single Nucleotide Polymorphism (SNP) is an important molecular marker in the study of crop trait related sites, which is applied in soybean breeding through linkage analysis and association analysis. Compared with traditional cross-breeding breeding methods, Marker-Assisted Selection (MAS) is based on cross-breeding and is closely linked with target trait genes. The screening strategy of efficient and rapid detection of target genes by detecting molecular markers has become a popular breeding method. With the rise of soybean group science, the joint analysis between genomic research and other omics has further extended the application range of molecular markers, and become a key part of modern soybean breeding, so that the selected genetic loci can be effectively used. In addition, with the continuous update of biotechnology, microRNA (non-coding single stranded RNA, miRNA) and long non-coding RNA (lncRNA), which were neglected in the past, have re-entered the field of vision of breeders. The genomic regions of target traits associated with molecular markers can be better identified. In order to sort out the application scope of molecular markers in soybean breeding in detail, this review summarizes the research progress of molecular markers related to disease resistance, yield and quality of soybean in 2022, hoping to provide reference information for soybean researchers in soybean molecular marker assisted breeding.

Keywords: soybean; molecular marker; MAS; agronomic traits

DNA 分子标记指 DNA 水平的遗传标记, 可反映物种整个基因组的遗传多样性。随着现代分子技术不断发展, DNA 分子标记主要分为 3 类^[1]。第一类是以分子杂交为核心的分子标记, 包括限制性片段长度多态性标记 (Restriction Fragment Length Polymorphism, RFLP 标记)、DNA 指纹技术 (DNA fingerprinting) 和原位杂交 (in situ hybridization)。第二类是基于 DNA 片段 PCR 扩增的分子标记, 如 AFLP (Amplified Fragment Length Polymorphism) 和 SSR (Simple Sequence Repeat)。第

三类为高通量测序的分子标记, 主要包括 SNP 标记^[2] 和表达序列标签 (Expressed Sequences Tags, EST 标记), 其中 SNP 标记为目前常用的分子标记。

分子标记辅助育种以育种年限短、效率高、准确性强和简便性等特点, 逐渐代替传统育种, 成为大豆育种研究的热点^[3]。近年来随着基因组技术的发展和 DNA 标记的大量开发, 高精密度的大豆遗传图谱不断被构建和完善, 目前已鉴定出大量与重要性状相关的数量性状核苷酸 (Quantitative Trait Nucleotides, QTN) 位点, 其中部分 QTN 已在大豆标

收稿日期: 2023-05-26

基金项目: 黑龙江省重点研发项目 (JD22A015); 财政部、农业农村部现代农业产业技术体系建设专项资金项目 (CARS-04-PS04); 黑龙江省重点基金项目 (ZD2022C002); 国家自然科学基金面上项目 (31971967)。

第一作者: 韩英鹏 (1978—), 博士, 教授, 主要从事大豆基因组学及分子辅助育种研究。E-mail: hyp234286@aliyun.com。

记辅助育种中得到应用。Jun 等^[4]利用大豆蚜虫易感品种 WYANDOT 和抗性 PI 567301B 杂交培育的 203 个 F_{7,9} 重组自交系 (Reconstituted Inbred Lines, RIL) 和 516 个多态性 SNP 标记构建了全基因组分子连锁图谱,并通过 SSR 标记确定 1 个新的主要基因,位于已鉴定抗性基因 RAG2 基因的附近,间接证明抗性品种 PI 567301B 为抗异种病类型。Shin 等^[5]利用标记辅助选择的反复回交方法,成功在品种 Ohsuzu 中引入对 SMV 菌株 C 和 D 具有抗性的 *Rsv3* 基因,接菌实验结果显示,含有接近 *Rsv3* 基因位点的分子标记对的品系显示出对两种 SMV 菌株的抗性。变体棉子糖合酶 3 (RAFFINOSE SYNTHASE 3,RS3) 是大豆种子碳水化合物表型的关键贡献者,当 RS2 突变等位基因存在时,超低低聚糖 (Raffinose Family of Oligosaccharides, RFO) 表型有所改善,利用直接标记辅助选择的方法对大豆的豆粕性状的改良起到重要作用^[6]。充分了解大豆重要性状的相关 QTL (Quantitative Trait Locus) 位点和分子标记,及时捕捉大豆分子标记辅助育种方面的动态,有助于明确大豆分子育种的方向以及提高育种效率和准确性。本综述汇总了 2022 年大豆多种抗病性、产量和品质相关性状分子标记的主要研究进展,对大豆分子标记辅助育种的前景进行展望,期望在前人研究的基础上,为大豆的育种工作提供一定参考。

1 国际相关研究动态

国际范围内,2022 年大豆分子标记方面的研究主要集中在大豆抗病性、产量和品质性状,这些研究结果大多采用连锁分析、全基因组关联分析 (Genome-Wide Association Study, GWAS) 和混池分组分析法 (Bulked Segregant Analysis, BSA),结合大量 SNP 标记,获得与目标农艺性状相关的数量性状 QTL 区域以及 QTN 位点。随后采用竞争性等位基因特异性 PCR (Kompetitive Allele-Specific PCR, KASP) 技术对筛选后的优良位点进行精细定位,鉴定出与性状相关的候选基因。此项技术流程的运用不仅缩短了大豆的育种年限,同时也保证了大豆育种工作的准确性和可信度。

1.1 国际大豆抗病性状研究

1.1.1 大豆疫霉根腐病 过去十年中,通过 GWAS 研究确定了 70 多个与大豆疫霉部分抗性 (Partial Resistance, PR) 相关的 QTL。但是,其中多数对抗性水平影响小,表现为针对单个表型变量或者单个分离株,从而限制了相关 QTL 位点在实际育种计划中的使用。De 等^[7]将 357 个大豆种质的抗疫霉表型与已鉴定的 70 多个 QTL 位点进行联合分析,采用 GWAS 技术进行基因分型,鉴定出具有相对重大

影响的新型大豆疫霉抗性 QTL,其中位于 15 号染色体上的 1 个大豆疫霉抗性 QTL 位点跨越了 -500 kb 染色体区域 (Chr15:36,395,777-36,895,690),包括 31 个 SNP 变异位点,分布于 5 个基因 (*Glyma. 15G217100*、*Glyma. 15G217200*、*Glyma. 15G217300*、*Glyma. 15G217350* 和 *Glyma. 15G217400*) 上下游范围内,*Glyma. 15G217100* 启动子上的唯一 SNP 变异使携带 R 等位基因品系抗病性提高了 60%。Karhoffd 等^[8]采用 RNA-seq 技术,分析了 (OX20-8 和 PI 427106、OX20-8 和 PI 427105B) 后代所构建成的两组近等基因系 (Near-Isogenic Line, NIL),定位于 18 号染色体的 *QDRL-18* 位点的范围由 1 853 kb (981 ~ 2 833 kb) 缩减到 732 kb 区域 (1 713 ~ 2 445 kb),此位点可以解释高达 45% 的表型变异 (Phenotypic Variation, PV),促进分子标记辅助育种的同时增强对 *Phytophthora sojae* 的抗性。经该位点预测的 82 个基因中有 1 个新抗疫霉基因 *Glyma. 18G026900* 被确定,该基因在携带抗性等位基因的接种 NILs 中下调,在携带易感等位基因的接种 NILs 中上调,大豆疫酶根腐病的研究未曾中断,目前为止,30 个抗大豆疫酶根腐病 (Resistance to *Phytophthora Sojae*, Rps) 基因被鉴定,其物理位置广泛分布于大豆 10 条染色体上,其中 3 号染色体上含有 14 个 Rps 基因,其次为 18 号和 13 号染色体,这些 Rps 基因是否存在基因复制事件和抗性的类别都有待研究,否则无法有效提供能够提高品种抗性的合适模型^[9]。

1.1.2 大豆锈病 大豆锈病是大豆生长期阶段最重要的叶面病害之一,与大豆锈病相关的 7 个 Rpp 抗性基因位点已被报道^[10],但真菌种群间的广泛致病型变异为识别与大豆锈病 (Soybean Rust, SBR) 抗性相关的其他基因和位点增加了难度。为此,在 2018—2015 年美国东南部的田间条件下,Walker 等^[11]对来自日本、印度尼西亚和越南的 191 种大豆植物引种和来自其他国家的 65 种引种植物对 *Phakopsora pachyrhizi* 的抗性进行了筛选,使用 SoySNP50K Infinium BeadChipMcCaghey 进行 GWAS 分析,鉴定出与 SBR 抗性相关的 7 条染色体上的 8 个基因组区域,其中有 6 个未被报道的基因组区域,包含了多种与 SBR 抗性相关的 Rpp 位点。Chanchu 等^[12]利用清迈 5 号和 SUKHOTHAI 2 杂交所产生的 108 个重组自交系群体进行抗 SBR 分析,1 个单一的抗 SBR 位点 QSBR18.1 被定为在 18 号染色体上的标记 T001855631M 和 SC21_3420 之间,覆盖了 212 kb 区域,包含 *RPP4* 基因,为品系提供了 21.31% ~ 35.09% 抗性, QTL 区域所包含的 *GLYMA. 18G226250*、*GLYMA. 18G226300* 和 *GLYMA. 18G226500* (GM18: 51,753,139-51,832,726) 与

RPP4-B 基因所在区域 (GM18: 51, 678, 981-51, 792,957 bp)重叠,被视为 SBR 抗性的新候选基因,解析了 SBR 抗性的新遗传作用基础,极大地促进了分子标记辅助育种的应用。

1.2 国际大豆产量性状研究

1.2.1 大豆百粒重 籽粒大小和形状是大豆产量形成的重要决定因素,解析它们的遗传基础和分子机制对大豆产量性状的改善具有积极作用。通过 GWAS 分析和基因分型测序 (Genotype-By-Sequencing, GBS)方法评估了 200 个国内外种质和 32 000 个高质量 SNPs,鉴定出 4 个与百粒重性状相关的 SNPs,分别位于 17 和 18 号染色体,对该性状次要等位基因的影响幅度最高达 - 0. 85^[13]。Ayalew 等^[14]利用两年 541 个大豆种质与 36 257 个 SNPs 进行 GWAS 分析,发现 19 个与种子产量相关的 SNPs,分布于 9 号、15 号、16 号和 17 号染色体,其中 9 号染色体上的 ss715604501 (4203661)和 17 号染色体上的 ss715628133 (7042685)各自上下游 10 kb 区域内分别鉴定到 7 和 13 个基因。令人遗憾的是,对于候选基因是否对大豆产量有影响,以及何种优势单倍型与大豆产量紧密相关未见报道,但也为育种工作者提供了借鉴。Yoosefzadeh-Najafabadi 等^[15]结合 250 个大豆品种的 40 712 个 SNPs,利用机器学习法中 SVR (Support Vector Regression)算法识别出 8 个与大豆籽重相关的 SNPs,分布于 7、15、18、19 和 20 号染色体,包含了 121 个候选基因,其中 S15_34958361 标记附近的 *Glyma. 15G214600* 和 *Glyma. 15G214700* 基因为大豆产量相关的候选基因,推测后续对其基因功能的揭示对应用分子标记辅助育种改良大豆粒重性状具有重要作用。

1.2.2 大豆株型 大豆株型主要由大豆分枝数、分枝角以及株高组成,影响着大豆生产力和产量。CLARK 等^[16]利用驯化大豆品种 *WILLIAMS 82* 和野生大豆 (*GLYCINE SOJA*)品种 PI 479752 构建了重组自交系群体,用来观测品系间分枝角的变化,为了鉴定大豆分枝角相关的 QTL,选择极宽分枝角 (Wide Branching Angle, WBA)的 RIL 品系与具有窄分枝角 (Narrow Branching Angle, NBA)的优良大豆品种 LD00-3309 进行杂交,以 F₂后代和 F₂衍生的 F₃后代组成的群体用于遗传分析和 QTL 作图,最后发现控制分支角度的主要 QTL 位点位于 19 号染色体,主要源自具有极宽分枝角 (WBA)的 NIL。Bhat 等^[17]将 3 个地点 211 份大豆品种表型和来自 SOYSP50K15 的 291 962 个 SNPs 相结合,鉴定出 3 个 (GM02: AX-93958260、GM17: AX-94154834 和 GM19: AX-93897200)与大豆株高相关性较强的 SNPs,并以 130 kb 作为其遗传变异区域,总计筛选出 124 个注释基因。相较前人所鉴别出 *QPH2* 位点

(15993654-41032570 KB)^[18], GM02: AX-93958260 附近 130 kb 区域做到了物理区间的缩小,进一步精确了与大豆株高相关的基因。Priyanatha 等^[13]利用 200 个品种表型和 32 K 高质量 SNPs 进行 GWAS 分析,鉴定出了 1 个与大豆株高相关的 SNP (GM05: S054,732,203),效应贡献为 - 4. 19,其区域内无所含候选基因被筛选,但也进一步揭示了大豆株高的遗传基础。Rani 等^[19]利用 96 个熟期不同的大豆品种和 100 个全基因组基因标记进行基因-环境相互作用 (GEI) 分析,结果揭示 9 个标记 (SATT194、SATT300、SATT316、SATT102、GMES6336、SATT154、GMES0902)与植株高有显著相关性,解释了 12% ~ 31% 的表型变异。其中 SATT300 标记与单株分枝数和单株粒重同样相关,因此,SATT300 标记附近区域寻找候选基因成为最优区间。

1.3 国际大豆品质性状研究

1.3.1 大豆蛋白 大豆蛋白质是优质的植物蛋白,富含人体所需的多种氨基酸,大豆籽粒蛋白是大豆改良育种常规目标。Cunicelli 等^[20]首先构建由 5601T 和 U99-310255 杂交后代所组成的 NIL 群体,再利用 6 种环境中生长的 138 个 NILs 的数据进行 QTL 检测分析,结果显示共有 12 个 QTLs 与大豆品质性状相关,与大豆蛋白性状相关的 QTL 有两个分别位于第 15 和 17 号染色体。CQPROT-003 是前人精确定位到的与大豆蛋白性状相关的 QTL 位点^[21],为了更好开发利用该位点,Marsh 等^[22]利用泛基因组数据集中的 985 个野生、地方品种和栽培品种大豆种质联合含有 544 997 个 SNPs 的 SNP 矩阵进行关联和连锁分析,结果鉴定出 1 个 173 kb 的连锁块,运用单倍体作图的方式对该区域的变体进行了表征,在高蛋白单倍型中鉴定了 34 个高置信度 SNPs、4 个插入、1 个缺失和 304 bp 的结构变体,其中存在可变长度的三核苷酸串联重复的基因 *GLYMA. 20G085100* 与高蛋白表型密切相关,推测为因果变异。Singer 等^[23]利用 311 个大豆种质和 35 570 个 SNPs 进行 GWAS 研究,鉴定出 23 个与大豆蛋氨酸含量相关的 SNPs,其中与位于 3 号染色体的 SS715586112、SS715586120、SS715586126、SS715586203 和 SS715586204、8 号染色体的 SS715599541 和 16 号染色体的 SS715599547 和 SS715625009 SNPs 位点强相关,为致力于提高蛋氨酸含量的大豆育种者的可行工具。

1.3.2 大豆油份 大豆作为世界范围内重要的油料作物,其产量约为世界油料作物总产量的 60%^[24]。大豆脂肪酸主要分两大类,第一类为不饱和脂肪酸:油酸、亚油酸、亚麻酸。第二类为饱和脂肪酸:棕榈酸和硬脂酸,而饱和脂肪酸和不饱和脂肪酸的含量及组成比例对大豆油的品质有直接的

影响。Xue 等^[25]利用东农 42 (棕榈酸含量较高) 和霍比特 (棕榈酸含量较低) 杂交后代构建 F_{2:6} 重组自交系, 以包含 9 980 个 SLAF 标记的高密度遗传图谱绘制出与棕榈酸含量相关的两个 QTL, 最高能够解释 3 个不同年份和环境 10.1% 的表型变异。Zuo 等^[26]采用 GWAS 方法中 3VmrMLM 模型对 286 个大豆品种的种子油含量及其脂肪酸性状表型与 106 013 个 SNPs 的性状表型进行关联分析, 定位到 7 个大豆油分相关的基因 (*GmSAPAD-A*、*GmSAPAD-B*、*Gmbzip123*、*Gm31-t39*、*GmFATB1A*、*Gmdgat2d* 和 *Gmdgat1B*), 其中 5 个基因在其他物种中被证实参与油份代谢, 佐证了实验结果的公信力。除此之外, 54 个 QEI 位点附近的 *GmFATB2B* 基因被重点关注, 拓展了影响大豆油份性状的遗传基础。为大豆分子标记辅助育种的应用提供理论依据。Kumar 等^[27]通过转录组和元 QTL 方法对大豆以往鉴定出的 1 000 个 QTL 区域进行分析, 确定了 14 个大豆油份相关元 QTL 和 94 个大豆油份相关基因, 在元 QTL 以及 0.1 Mb 侧翼区域 (METAQTL-OC_7.1、METAQTL-OC_13.2、METAQTL-OC_20.3) 所筛选到的 *Glyma. 07G034800*、*Glyma. 13G035200*、*Glyma. 20G111000* 存在多种单倍型, Hap1 组的油份均高于其他单倍型组。

2 国内相关研究动态

国内外分子标记在大豆上的应用范围大致相同, 如大豆产量和品质性状。抗病性方面的研究有所差异, 大豆疫霉根腐病作为常见病害, 应用分子标记对其分子遗传基础进行解析为目前最为常用的策略。除此之外, 2022 年国内大豆锈病研究的所占比重同样较高, 因此, 国内范围大豆抗病性的研究以大豆疫霉根腐病和大豆锈病为例进行阐述。

2.1 国内大豆抗病性状研究

2.1.1 大豆疫霉根腐病 Li 等^[28]利用黑龙江省范围内 242 份野生大豆种质对大豆的抗疫霉性状进行鉴定。使用符合 MAF ≥ 5% 且缺失率 ≤ 50% 规格的 999 800 个 SNPs 进行 GWAS 分析, 结果共检测到 9 个与大豆疫霉抗性重复相关的显著 SNPs, 分别位于 1、10、12、15、17、19 和 20 号染色体。筛选出的 8 个与大豆疫霉抗性相关的有利等位基因变异通过 t 检验得到验证, 其中 RKF3 (*Glyma. 19G051583* 和 *Glyma. 19G051582*) 和 RBK2 (*Glyma. 19G051581*) 被重新定位, 其余基因均位于显著 SNP 附近 5 kb 之内, 为 GWAS 结果的准确性提供了依据。(NBS-LRR) 结构域被鉴定为抗疫霉根腐病数量性状位点 (QTL) 相关, 判定为 R 等位基因的关键区域, *Glyma. 16G135500* 编码 NBS-LRR 型蛋白, 在 *GMA-MR1510* 基因的调控下通过 JA 和 SA 途径参与到抗

大豆疫霉根腐病的反应^[29]。随着大豆的种植年限延长, 大豆疫霉的致病性进化, 以往单一抗病性基因的策略难以为大豆作物提供有效的庇护, 有力地将多个抗病性基因导入同一植株, 保证抗病基因的多样化才能最大程度发挥其效应^[30]。

2.1.2 大豆花叶病 大豆花叶病 (Soybean Mosaic Disease, SMD) 是由多种病毒引起的, 大豆育种过程中大多数病毒常被忽略, 使大豆花叶病防治工作困难重重。为了有效缓解 SMD 所带来的不利影响, 培育具有广谱性的品种十分重要。Wang 等^[31]利用来自两种对病毒感染反应不同的大豆基因型构建了由 150 株 F_{7:9} 株系组成的 RIL 群体进行 QTL 分析, 结果显示, 1 个区间大小为 157 kb 的 QTL 位点被定位于第 13 号染色体上, 其区间包含的含有 LRR 结构域的 *Glyma. 13G190000*、*Glyma. 13G190300* 和 *Glyma. 13G190400* 的编码序列在抗性亲本和易感亲本之间存在显著差异。而 Jiang 等^[32]将垦丰 1 号 (抗性) × NN1138-2 组成的 427 个 RILs 构建的种群的抗性 or 易感性表型与 3 683 个连锁不平衡块共同用作形态标记, 以用来 QTL 鉴定, 与 W82. a1. v1 基因组相比, 所鉴定出的 462 kb R_{SC11K} 区域含有 429 个 SNPs, 142 个 InDels, 并鉴定了 34 个推测基因, 其中, 10 个基因含有对蛋白质具有高、中度功能影响的 SNPs/InDels 变体, 说明 R_{SC11K} 区域为大豆蛋白性状显著相关的 QTL 位点, 有利于大豆的改良育种。Zhang 等^[33]通过批量分离分析 (BSA) -seq 和精细定位相结合的方法, 确定了 1 个新的抗大豆花叶病毒的基因座 *RSMV-11*, 覆盖了 11 号染色体约 207 kb 区域, 包含了 25 个注释基因。对东农 50 和小青豆两种大豆品种的 *RSMV -11* 区域进行比对, 鉴定出 11 个具有非同义单核苷酸多态性 (SNP) 或插入缺失突变 (InDels) 的基因, 其中编码 MATE 转运蛋白的基因 *GmMATE68* (*Glyma. 11G028900*) 与植物抗病性相关, 为抗大豆花叶病的改良育种提供了潜在价值。

2.2 国内大豆产量性状研究

2.2.1 大豆百粒重 大豆籽粒大小是决定农作物产量的重要农艺性状之一。Shao 等^[34]通过深度重测序技术 (~20 ×) 对自然种群进行基因分型, 并在 6 种环境下对 6 种相关性状进行表型分析。最后结果表明共有 154 个 SNPs 与不同环境中的种子长度密切相关, 323, 483, 565, 394 和 2 038 个 SNPs 与种子宽度、种子直径、种子周长、多种环境下的种子面积和长宽比性状相关。此外, 同时通过 SNP 等位基因和转录组数据分析了 6 号和 10 号染色体上相关 SNP 两侧的 218 个基因的 DNA 突变和 RNA 表达, 发现候选基因 *Glyma. 10G035200* (特异性二酰基甘油脂肪酶 Sn1)、*Glyma. 10G035400* (转录因子) 和

Glyma. 10G058200(苯丙氨酸解氨酶)与种子大小和形状相关,推测它们在改良大豆产量过程中具有积极影响。同样,Duan等^[35]通过GWAS技术对1853个大豆品种的SNPs进行分析,在5号染色体上鉴定出多年稳定的180 kb 间隔段(43.58~43.76 Mb),注释了23个基因,其中*GmST05*(*SoyZH13_05G229200*)基因区域含有6个SNPs和2个Indels,外显子范围内存在7个间隔性SNPs,形成了两种Hap(HapI、HapII),*GmST05*^{HapI}为其优势单倍型,其OE和KO株系均能增加或减少大豆籽粒的重量。而Zuo等^[36]将驯化和改良分析、GWAS研究相结合,在286个大豆种质中确定了与种子大小相关性状的343个候选驯化基因(Candidate Domestication Gene, CDGs)和85个候选改良基因(Candidate Improved Gene, CIGs)。

2.2.2 大豆株型 株高是主要育种目标,与大豆的株型和产量密切相关。Wang等^[37]运用多位点关联分析的方法对具有63306个SNPs标记的455个种质群体进行分析,检测到62个与株高相关的QTNs位点,其中26个QTNs可在多个模型下被检测到。筛选出的候选基因*Glyma. 02G133000*和*Glyma. 05G240600*不仅进一步完善了大豆株高性状的调控网络,而且有助于大豆的分子标记辅助育种。Su等^[38]通过BSA技术对株高极端高和低各20个单株进行GWAS分析,结果显示1个峰值SNP存在于9号染色体的物理位置(39.78~44.85 Mb)范围内,包含了9个候选基因,其中*GLYMA. 09G189300*、*GLYMA. 09G193000*和*GLYMA. 09G199500*携带有非同义突变SNPs,*GMIAA27*(*GLYMA. 09G193000*)基因第二外显子上单核苷酸替代(C-T)导致*GmiAA27*的等位基因*GMIAA27*的效应为矮化植株。突变体的应用也是加速相关性状研究进程的策略之一。Song等^[39]对中品661品种突变体M₂进行全基因组测序分析,通过参照中品661的基因组序列,在矮化突变体中发现编码BEL-1蛋白的*GLYMA. 13G312900*基因CDS序列存在3个非同义突变(*SER107GLY*、*ALA111TR*和*TR112ALA*),推测其导致了矮化突变体表型建成。

2.3 国内大豆品质性状研究

2.3.1 大豆蛋白 Qin等^[40]基于284个大豆种质和180个重组自交系(RIL)的GWAS研究和连锁作图方法,对蛋白质含量进行了4年评估,利用混合线性模型(Mixed Linear Model, MLM)、一般线性模型(Generalized Linear Model, GLM)和贝叶斯区间作图方法共检测到22个与蛋白含量相关的SNPs和5个QTL位点,6号染色体上的25.77%贡献表型变异的QTL基因组区域(*Chr6_18844283-19315351*)受到重点关注,并且包含7个与蛋白含量性状相关

的候选基因,均在拟南芥中被注释。Feng等^[41]基于受限两阶段多位点全基因组关联研究(RTM-GWAS)对具有15501个SNPs连锁不平衡块标记的361个种质组成的中国东北大豆种质种群(NorthEast Chinese Soybean Germplasm Populations, NECSGP)的种子蛋白质含量(Seed Protein Content, SPC)进行评估,结果确定了包含273个等位基因的73个SPC数量性状位点(QTL),解释了71.70%表型变异,其中28个QTLs为新QTL。同时梳理了QTL-等位基因结构从老MGs到新MGs的进化变化关系,新MGs的等位基因近90%遗传自老MGs,其它可能来源于等位基因的重组。通过对H88×P73杂交后代群体进行连锁分析,定位到3个QTLs(*qPRO-20-1*、*qPRO-6-1*、*qPRO-6-2*),其大小和位置信息分别为2.46 Mb (28349696~30805913)、0.07 Mb (14779248~14846706)和32.38 Mb (15214173~47597755),贡献率大小为19.31%、2.65%和9.55%,并结合大豆200 K芯片将*qPRO-20-1*定位区间缩短至95.8 kb,筛选到与氮素合成相关的*Glyma. 20G081800*作为大豆蛋白质相关基因,为高蛋白品种大豆选育提供优异的种质资源和检测技术^[42]。

2.3.2 大豆油份 Zuo等^[36]利用GWAS分析、连锁分析和BSA方法确定了113个驯化和103个改良区域。位点size_loci_9(Gm:47318349-51550019)、size_loci_36(Gm06:620481-820481)、size_loci_88(Gm10:48991131-49360979)、size_loci_128(Gm14:3341523-3541523)和size_loci_155(Gm15:1016635-4585355)区域内鉴定到5个与大豆油份相关的基因,在UTR、基因上游和基因间隔区均存在大量变异,其SNP对精英单倍型频率的影响尚未可知。Goettel等^[43]通过GWAS技术对CR20上4 Mb区间(CR20:29,050,000-33,120,000)内候选基因进行关联分析,确定到在*GmPOWR1*(*Glyma. 20G085100*)的(31.5~32 Mb)范围内存在321 bp缺失,影响着大豆品种油份含量的高低。Li等^[44]通过GWAS分析和连锁分析,在多环境下检测到8号染色体影响大豆油份以及种子性状的SNP位点(*Seed thickness 1*, *ST1*),覆盖了约120 kb (Gm08:8350932~8477410)范围,在不同品种背景有1个TGTG序列的删除,导致*GLYMA. 08G109100*基因过早停止翻译,功能丧失。且在过表达植株实现了*ST1*位点作为大豆油份的遗传调控基础,表明其对分子标记辅助育种具有贡献。

3 大豆分子标记研究发展趋势与建议

通过总结2022年国内外大豆分子标记相关研究和应用,发现分子标记的相关研究主要集中在抗病、产量和品质性状方面,其他农艺性状涉及较少。

众多 QTL 位点、QTN 位点和候选基因被鉴定,性状相关的基因遗传机制被初步梳理,但分子标记的研究和应用过程也暴露出一些弊端,有待完善。首先,遗传位点的定位所需群体量大,部分研究单位所保存的种质资源难以满足首要要求;其次,遗传图谱的精密度不够,所定位到的遗传位点过于广泛,区间内遗传信息庞大,无法有效确定其与相关性状之间的关系;再者,遗传位点的定位对环境和研究方法的要求较高,容易造成定位结果精确度差的问题;最后,随着分子标记研究技术的应用,如连锁分析、关联分析和基因组选择分析,大量主效遗传位点被筛选,微效遗传位点被无意间剔除,造成部分基因组的缺失,难于追本溯源。

因此,为了能更好地定位到目标性状的位点,研究机构间进行合作、扩大群体样本量不失是一种有效的策略。随着研究的不断深入,分子标记的密度也在不断增多,从最初的数百标记到目前普遍小则几万,大则几十万、数百万,保证了定位的精确性。多年多环境下定位能够有效避免偶然性结果的出现,而新颖技术(3-Variance multi-locus random-SNP-effect Mixed Linear Model,3VmrMLM)的更新也保证了定位的精度。大多时候,筛选鉴定的基因组区域与目的性状相关性较强,区域内鉴定出优秀基因的概率更有保证,部分贡献较小表型差异的区域内也存在与目的性状相关的候选基因。为保证优秀基因不丢失,在关注主效基因的同时,应该结合多种评估策略对微效基因组区域内的候选基因进行深入分析,竭力最为准确地研究目的性状。相信随着生物技术的不断更新,分子标记辅助育种会使大豆育种工作更加便捷,为传统大田育种和分子育种提供有价值的信息。

参考文献

[1] 王军, 谢皓, 郭二虎, 等. DNA 分子标记及其在谷子遗传育种中的应用[J]. 北京农学院学报, 2005(1): 76-80. (WANG J, XIE H, GUO E H, et al. DNA molecular markers and their application in genetic breeding of millet[J]. Journal of Beijing University of Agriculture, 2005(1): 76-80.)

[2] LANDER E S. The new genomics: Global views of biology[J]. Science, 1996, 274(5287): 536-539.

[3] HASAN N, CHOUDHARY S, NAAZ N, et al. Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes[J]. Journal of Genetic Engineering and Biotechnology, 2021, 19(1): 128.

[4] WANG L, YANG Y, YANG Z, et al. *GmFtsH25* overexpression increases soybean seed yield by enhancing photosynthesis and photosynthates[J]. Journal of Integrative Plant Biology, 65(4): 1026-1040.

[5] SONG B, AN L, HAN Y, et al. Transcriptome profile of near-isogenic soybean lines for β -conglycinin α -subunit deficiency

Genetics, 2023, 13: 1090994.

[20] CUNICELLI M, OLUKOLU B A, SAMS C, et al. Mapping and identification of QTL in 5601T × U99-310255 RIL population using SNP genotyping: Soybean seed quality traits [J]. Molecular Biology Reports, 2022, 49(7): 6623-6632.

[21] DIERS B W, KEIM P, FEHR W R, et al. RFLP analysis of soybean seed protein and oil content [J]. Theoretical and Applied Genetics, 1992, 83:608-612.

[22] MARSH J I, HU H, PETEREIT J, et al. Haplotype mapping uncovers unexplored variation in wild and domesticated soybean at the major protein locus *cqProt-003* [J]. Theoretical and Applied Genetics, 2022, 135(4): 1443-1455.

[23] SINGER W M, SHEA Z, YU D, et al. Genome-wide association study and genomic selection for proteinogenic methionine in soybean seeds [J]. Frontiers in Plant Science, 2022, 13: 859109.

[24] KIM J M, LYU J I, KIM D G, et al. Genome wide association study to detect genetic regions related to isoflavone content in a mutant soybean population derived from radiation breeding [J]. Frontiers in Plant Science, 2022, 13: 968466.

[25] XUE Y, GAO H, LIU X, et al. QTL mapping of palmitic acid content using specific-locus amplified fragment sequencing (SLAF-Seq) genotyping in soybeans (*Glycine max* L.) [J]. International Journal of Molecular Sciences, 2022, 23(19): 11273.

[26] ZUO J F, CHEN Y, GE C, et al. Identification of QTN-by-environment interactions and their candidate genes for soybean seed oil-related traits using 3VmrMLM [J]. Frontiers in Plant Science, 2022, 13: 1096457.

[27] KUMAR V, GOYAL V, MANDLIK R, et al. Pinpointing genomic regions and candidate genes associated with seed oil and protein content in soybean through an integrative transcriptomic and QTL Meta-analysis [J]. Cells, 2022, 12(1): 97.

[28] LI W, LIU M, LAI Y C, et al. Genome-wide association study of partial resistance to *P. sojae* in wild soybeans from Heilongjiang province [J]. Current Issues in Molecular Biology, 2022, 44(7): 3194-3207.

[29] ZHOU L, DENG S H, XUAN H D, et al. A novel TIR-NBS-LRR gene regulates immune response to *Phytophthora* root rot in soybean [J]. Science, 2022, 10(6): 1644-1653.

[30] BELZILE F, JEAN M, TORKAMANEH D, et al. The SoyaGen Project: Putting genomics to work for soybean breeders [J]. Frontiers in Plant Science, 2022, 13: 887553.

[31] WANG D, CHEN S, HUANG Z, et al. Identification and mapping of genetic locus conferring resistance to multiple plant viruses in soybean [J]. Theoretical and Applied Genetics, 2022, 135(9): 3293-3305.

[32] JIANG H, JIA H, HAO X, et al. Mapping Locus RSC11K and predicting candidate gene resistant to Soybean Mosaic Virus strain SC11 through linkage analysis combined with genome resequencing of the parents in soybean [J]. Genomics, 2022, 114(4): 110387.

[33] ZHANG Y, SONG J, WANG L, et al. Identifying quantitative trait loci and candidate genes conferring resistance to Soybean Mosaic Virus SC7 by quantitative trait loci-sequencing in soybean [J]. Frontiers in Plant Science, 2022, 13:843633.

[34] SHAO Z, SHAO J, HUO X, et al. Identification of closely associated SNPs and candidate genes with seed size and shape *via* deep re-sequencing GWAS in soybean [J]. Theoretical and Applied Genetics, 2022, 135(7): 2341-2351.

[35] DUAN Z, ZHANG M, ZHANG Z, et al. Natural allelic variation of *GmST05* controlling seed size and quality in soybean [J]. Plant Biotechnology Journal, 2022, 20(9): 1807-1818.

[36] ZUO J F, IKRAM M, LIU J Y, et al. Domestication and improvement genes reveal the differences of seed size- and oil-related traits in soybean domestication and improvement [J]. Computational and Structural Biotechnology Journal, 2022, 20: 2951-2964.

[37] WANG J, HU B, JING Y, et al. Detecting QTL and candidate genes for plant height in soybean via linkage analysis and GWAS [J]. Frontiers in Plant Science, 2022, 12: 803820.

[38] SU B, WU H, GUO Y, et al. *GmIAA27* encodes an AUX/IAA protein involved in dwarfing and multi-branching in soybean [J]. International Journal of Molecular Sciences, 2022, 23(15): 8643.

[39] SONG J, WANG X, HUANG L, et al. Genetic dissection of the soybean dwarf mutant *dm* with integrated genomic, transcriptomic and methylomic analyses [J]. Frontiers in Plant Science, 2022, 13: 1017672.

[40] QIN J, WANG F, ZHAO Q, et al. Identification of candidate genes and genomic selection for seed protein in soybean breeding pipeline [J]. Frontiers in Plant Science, 2022, 13:882732.

[41] FENG W, FU L, FU M, et al. Transgressive potential prediction and optimal cross design of seed protein content in the northeast China soybean population based on full exploration of the QTL-allele system [J]. Frontiers in Plant Science, 2022, 13: 896549.

[42] 杨硕. 大豆蛋白质含量的 QTL/基因定位 [D]. 哈尔滨: 东北农业大学, 2022. (YANG S. QTL/gene mapping of soybean protein content [D]. Harbin: Northeast Agricultural University, 2022.)

[43] GOETTEL W, ZHANG H, LI Y, et al. POWR1 is a domestication gene pleiotropically regulating seed quality and yield in soybean [J]. Nature Communications, 2022, 13(1): 3051.

[44] LI J, ZHANG Y, MA R, et al. Identification of ST1 reveals a selection involving hitchhiking of seed morphology and oil content during soybean domestication [J]. Plant Biotechnology Journal, 2022, 20(6): 1110-1121.