



关博文,陈庆山,武小霞,等.大豆遗传育种中全基因组关联分析的应用进展[J].黑龙江农业科学,2022(4):94-99.

大豆遗传育种中全基因组关联分析的应用进展

关博文¹,陈庆山¹,武小霞¹,杨学²

(1.东北农业大学农学院,黑龙江哈尔滨 150030;2.黑龙江省农业科学院草业研究所,黑龙江哈尔滨 150086)

摘要:目前我国缺少高产优质的大豆品种,需加大研发力度。全基因组关联分析(GWAS)具有高通量、高精度和费时少等显著优点。为促进 GWAS 在大豆遗传和育种过程中的应用,本文在介绍 GWAS 方法的基础上,总结其在大豆各主要农艺特性的遗传改良和育种工作中的应用进展,并对 GWAS 的优缺点和其发展趋势等问题进行讨论,并提出解决方法。

关键词:大豆;全基因组关联分析;遗传育种;性状改良;应用进展

大豆 [*Glycine max* (L.) Merr.] 含有约 40% 的蛋白质和 20% 的脂肪,是一种重要的经济和油料作物。然而目前,我国大豆的产量远远不能满足人们的需求,大部分大豆供给仍依靠进口。我国曾是大豆主要的出口国,如今却成为大量进口大豆的国家,研发高产优质的大豆品种势在必行。通过传统的育种方法培育高产优质的大豆品种需要在多个环境中进行数年的评估,并且昂贵、耗时耗力。近年来兴起的全基因组关联分析(GWAS)为培育高产优质的大豆品种提供了高效的研究方法^[1]。

GWAS 已成为检测作物农艺性状的主要方法^[2]。GWAS 用于确定遗传位点和性状之间的关联,其中的高密度标记使人们能够预测或识别相应基因。本文总结 GWAS 在大豆各主要农艺性状遗传育种中的应用进展、存在的问题,并分析解决途径,为 GWAS 在大豆遗传育种中的充分应用提供相关依据。

1 全基因组关联分析

1.1 全基因组关联分析的概述

全基因组关联分析(GWAS)方法检测全基因组的单核苷酸多态性(Single Nucleotide Polymorphism, SNP),以其为分子遗传标记,通过与特定性状进行对照分析或相关性分析,精确地获

得与这些性状关联的候选基因^[3]。

连锁不平衡(Linkage Disequilibrium, LD),是指位于两个或多个不同基因座上的等位基因非随机组合^[4]。GWAS 利用了连锁不平衡的原理,在大量具有不同特征的作物个体中准确地识别出高密度 SNP 标记,然后从其中甄别得出与各作物性状变异密切相关的 SNP 标记^[5]。

1.2 全基因组关联分析的一般流程

1.2.1 种质资源的选择 样本需要多样化,并包含尽可能多的该物种的完整遗传信息。在选择种质资源时,需要注意种质材料要反映最大的遗传变异,主要包括基因型的差异及其对环境的反应。

1.2.2 基因分型 随着测序技术的快速发展,研究者能够更方便地获得高通量、低成本的测序服务。SNP 标记因其遗传稳定性高、数量丰富等特点,在 GWAS 研究中广泛应用。

1.2.3 群体结构分析 种群结构会导致一些等位基因频率在不同的亚群之间明显不同,从而导致结果的假阳性,影响关联分析的精确度。因此在进行 GWAS 分析之前要对群体结构和亲缘关系进行相应的评测。群体系统进化树构建、主成分分析、Structure 分析可用于研究样本间的群体结构和亲缘关系,这 3 种分析结果可以相互验证^[6]。

1.2.4 表型考察 目标性状的评估会对 GWAS 的结果产生很大影响。表型性状会受基因型和周围环境等多重因素的影响。所以制定良好的试验计划,从时间、空间上增加表型的重复性显得非常重要。

1.2.5 关联分析 关联分析可以利用广义线性模型(GLM)和混合线性模型(MLM)来进行,其

收稿日期:2021-10-10

基金项目:国家自然科学基金区域创新发展联合基金(U20A20271004394)。

第一作者:关博文(1997—),男,硕士研究生,从事大豆遗传育种及生物技术应用研究。E-mail:563134240@qq.com。

通信作者:杨学(1969—),男,硕士,研究员,从事大豆抗病育种研究。E-mail:yxflax@126.com。

中 GLM 以群体结构作为协变量,MLM 采用 Q(群体结构)+K(亲缘关系矩阵)模型,两种模型均可以使用,基于此两种模型的软件 Tassel 被广泛应用于关联分析^[7]。

1.3 全基因组关联分析的优点

(1)群体易获得,一般以自然群体为研究对象,不用构建专门的群体,缩短研究时间;(2)效率高、范围广,GWAS 可同时分析多个性状和多个等位基因;(3)精度高,自然种群进化过程中大量积累重组信息,位点间分辨率提高,可精确定位目标性状相关位点;(4)高通量、低成本,GWAS 研究多采用自然群体,自然群体在许多性状上都存在丰富的遗传变异,因此利用同一套种质材料和基因型信息就可对多个不同性状同时进行遗传学分析,提升研究通量,同时节约成本^[8-9]。

2 大豆遗传育种中 GWAS 分析的应用进展

2.1 在大豆品质相关性状研究上的应用

2.1.1 蛋白质 大豆含有丰富的蛋白质,非常受人们欢迎,因此找出大豆蛋白的遗传位点有着重要意义。Zhang 等^[10]对 211 种不同大豆种质进行 GWAS 分析,鉴定了 8 个与蛋白质含量相关的 QTL,分析表明 qPC-15-1 中的 Glyma.15G049200 基因影响蛋白质合成。Zhang 等^[11]在 144 个四向重组自交系中对蛋白进行了多位点 GWAS 分析,产生 109 676 个非冗余 SNP 的基因型数据,通过 5 种多位点 GWAS 方法共鉴定了 129 个重要的 QTN。Hwang 等^[12]应用 Illumina Infinium 和 GoldenGate 等多种方法对 55 159 个 SNPs 进行了基因分型,在 17 个区域发现 40 个与蛋白质含量显著相关的 SNP。Dias 等^[13]用 6 000 个 SNP 标记对 169 个巴西大豆品种进行了基因分型,采用近红外反射法测定蛋白质含量,在 6 条染色体(2 号、6 号、11 号、12 号、13 号和 16 号)上检测到 7 个与蛋白质含量相关的 QTL。以上研究成果均有助于发现控制蛋白质含量的大豆基因,为培育高蛋白品种打下基础。

2.1.2 油脂 大豆作为重要的油料作物之一,为人们提供丰富的植物油,油分是大豆重要的品质性状之一。Yan 等^[14]利用冀豆 12 和黑豆的重组自交系群体,确定了 6 个与含油量相关的 QTL,检测到 qOIL_8_1,通过 GWAS,对 qOIL_8_1 进行了分析,发现该地区单倍型频率存在显著差异,

表明该地区存在较强的驯化选择足迹。Liu 等^[15]获得 2 311 337 个 SNP 标记,GWAS 分析表明,与油酸含量相关的基因很多。在 2018 年和 2019 年两次分析中,均检测到位于第 4 染色体的 Glyma.04G102900.1 和位于第 11 染色体的 Glyma.11G229600.1。Liu 等^[16]通过全基因组关联分析和多组学分析发现 284 个与种子油分性状相关的候选基因,MCP 和 SCAD 分析发现 6 个种子油分相关性状。Li 等^[17]将一个由 144 个品系组成的四向重组自交系群体种植在 10 个环境中,并将其和一个由 455 份种质组成的种质小组在两个环境中种植,收集油分表型。通过 QTL 和 QTN 的比较表明,有 4 个 QTN 控制含油量,对这 4 个 QTN 衰减区的基因进行通径分析,筛选出两个与大豆油脂合成或代谢相关的候选基因。Fang 等^[18]收集了全球 809 份大豆品种,并在两年内在 3 个地点对 84 个农艺性状进行了表型分析,确定 14 个与油合成相关的基因负责大豆中脂肪酸的积累,并根据加性模型发挥作用。这些研究结果为今后高油大豆的培育打下基础。

2.2 在大豆产量相关性状研究上的应用

2.2.1 株高、主茎节数、分枝数等性状 株高、主茎节数、分枝节数等性状均是直接影响大豆生长和产量的主要因素,对这些性状进行研究和分析可更好地提高大豆的产量。Shim 等^[19]通过 GWAS 分析结合连锁分析确定了一个控制大豆分枝数的候选基因。发现在一个大豆核心种质中,5 个数量性状核苷酸(QTN)与分枝数相关。在这些 QTN 中存在一个跨越 20 个基因的连锁不平衡(LD)块 qtnBR6-1,它与先前发现的主数量性状基因座 qBR6-1 重叠,最终分析建议 BRC1 作为大豆分枝数的候选基因。Lü 等^[20]利用 192 份天然大豆种质和 1 536 个 SNP 标记鉴定了与株高相关的基因组区域,分析得出 BARC-040651-07807、BARC-030433-06867 和 BARC-042475-08274 这 3 个 SNP 与生育末期的株高显著相关,BARC-038795-07333 和 BARC-013749-01246 两个 SNP 与生育前期的株高显著相关。Borash 等^[21]对 63 个不同的大豆品种用 284 923 个 SNP 标记进行了基因分型,并在两个地点进行了表型分析。通过混合线性模型进行的 GWAS 确定了几种与植物高度和分枝数相关的定量性状核苷酸。此类研究有助于对产量相关候选基因的筛选,培育高产品种。

2.2.2 百粒重和单株荚数等性状 百粒重、单株荚数等性状对大豆产量有着重要的影响,同时百粒重对于大豆品质也有着重要意义。Zhao 等^[1]选择了由 185 个种质组成的样本,在 3 个测试环境中进行分析,通过 GWAS 确定了 34 个关联信号与百粒重相关。其中,19 个 QTN 是新发现的,另外 15 个 QTN 重叠或位于已知百粒重 QTL 的基因组区域附近。Ikram 等^[22]使用多基因座方法共检测到 218 个重要 QTN,而通过单基因座方法识别了 8 个 QTN,得出 Glyma07g07850、Glyma10g03440 和 Glyma10g36070g 为百粒重候选基因。Qi 等^[23]根据来自 20 个不同环境的表型数据和 144 个四向重组自交系的 109,676 个单核苷酸多态性的基因型数据,利用 5 个多位点 GWAS 方法,获得了 118 个控制百粒重的 QTN。Song 等^[24]通过连锁分析结合 GWAS 研究,鉴定 4 个大豆品种(垦丰 14×垦丰 15)×(黑农 48×垦丰 19)后代群体,共鉴定出 602 个 QTL,确定了 11 个可能涉及豆荚生长发育的潜在候选基因。目前来看,对大豆百粒重、单株荚数等性状的分析已取得初步成效,后续需要对其候选基因进行验证,助力培育高产大豆品种。

2.2.3 生育期性状 生育期性状是影响大豆产量的重要性状之一,深入了解生育期性状遗传基础,可为大豆品种种植的合理布局和分子育种提供相应的理论依据。Zuo 等^[25]利用关联分析和连锁作图方法确定了开花天数、开花至成熟天数和成熟期天数的数量性状基因座,共识别出 96 个关联,涉及 19 个 SSR,利用这两种作图方法,发现 Satt150、Satt489、Satt172 和 Sat_312 与生育期性状相关。Zhou 等^[26]对 286 份自然群体进行 RAD-seq 基因分型,鉴定出 48 个与驯化相关的基因座。其中 12 号染色体上的 4 个和 11、15 号染色体上的另外 2 个与开花时间有关。Zhang 等^[27]利用 152 个重组自交系个体进行连锁定位,在 6 号染色体上找到了一个影响开花时间的主效 QTL 位点 *qFT6*,通过分析表明,*qFT6* 区域的高分辨率图谱不仅可用于开花时间的标记辅助选择,还可用于目标基因的进一步定位克隆。Liu 等^[28]使用 4 032 个 SNP 标记检测主要来自中国东北的 146 个栽培品种。结果表明,种群中存在丰富的表型变异,基因型解释的变异、环境解释的变异和环境相互作用解释的基因型对每个生育期性状的影响都很显著。Li 等^[29]对 279 份大豆不

同地方品种进行 GWAS 研究,确定 16 个与开花时间和成熟时间相关的候选数量位点,并检测到含有 *GmPRR3b* 的基因组区域与开花和成熟时间的强关联信号。对生育期性状的深入挖掘,有助于培育出适宜不同地方环境的大豆品种,是培育优良品种过程中不可缺少的一环。

2.3 在大豆抗逆相关性状研究上的应用

干旱、盐碱、矿物质缺乏以及根腐病等一系列病害都会严重影响大豆产量,为此加大对此类性状的分析研究,有助于大豆的生长发育。Zhao 等^[30]用 225 个大豆品种和 109 个重组自交系与东农 L-28(抗大豆疫霉小种 1 号)和合丰 25(感病大豆小种 1 号)杂交,鉴定大豆疫霉小种 1 号的抗性。利用该特异位点扩增片段对 225 个大豆品种进行了测序,通过 GWAS 方法定位抗性基因座,获得 28 722 个 SNPs,8 个 QTN 与大豆疫霉小种 1 号抗性相关。Rolling 等^[31]以大豆疫霉分离物 OH.121 和 C2.S1 为材料,在全基因组关联分析中,分别对 478 个和 495 个植物引种中的 QDR 进行测定,以确定与 QDR 基因座(QDRL)连锁的遗传标记。总共有 48 个显著标记分布在 16 条染色体上,在连锁分析的基础上,共获得 44 个 QDRL。Lai 等^[32]从利用大豆耐涝相关的多维基因组特征收集的 36 705 个测试基因中,根据合适的分界点,筛选出 83 个耐涝基因,利用来自 SoyNet 的独立样本、全基因组关联研究、SoyBase、GO 数据库和转录组数据库进行验证,结果都表现出很好的一致性,表明这 83 个耐涝基因明显优于其他基因。Zhang 等^[33]从 235 份野生大豆种质中鉴定对 SCNHG2.5.7 型具有抗性的基因型材料,总共鉴定了 43 个抗性种质,其中 10 个 SNP 与大豆胞囊线虫的 SONHG 2.5.7 抗性显著相关。Zeng 等^[34]使用 33 009 个 SNP 进行大豆耐盐性与叶氯化物浓度和叶绿素浓度的关联分析。发现 45 个 SNP 与其显著相关,为培育耐盐大豆提供依据。Wu 等^[35]对 384 份大豆植株引种,连续两年进行田间耐涝鉴定,共获得 42 291 个 SNP 标记,确定了 14 个 SNP 与所有环境和模型的洪涝耐受性相关。Wen 等^[36]通过 GWAS 方法鉴定了控制白霉抗性的数量性状位点,发现在田间和温室中分别有 16 个和 11 个与抗性显著相关的位点,其中 8 个位点定位于之前定位的 QTL 区间,1 个位点在两种环境中均与抗性显著相关。

3 全基因组关联分析中存在的问题及解决方法分析

假阳性关联的识别以及已识别位点和家族错误率的再现性现在被认为是 GWAS 方法的主要局限性^[37]。通常,有两种策略用于验证使用 GWAS 发现的关联:一种是联合分析,另一种是基于复制的分析。联合分析使用同一群体中相同特征的所有可用 GWAS 数据,通过汇集多阶段基因分型数据或使用荟萃分析来寻找相关的 SNP。基于复制的分析将数据分成两部分,一部分用于发现,另一部分用于验证^[38]。

目前现有的 GWAS 方法大多数只考虑加性效应,进行 GWAS 时使用单位点分析会导致一些稀有突变无法被有效地检测出。今后,多位点分析、非加性效应和互作效应等因素均应发展成为一个重要研究方向^[39]。

目前对不同植物种类的各种性状进行了大量的 GWAS 研究,而且还在不断增加,但还没有一个公共的平台可以访问所有研究数据。这与在人类身上进行的研究形成了鲜明对比,在美国国家人类基因组研究所的目录中,几乎每项研究都有很好的编目。研究的编目很重要,因为使用这样的编目,人们可以获得关于不同特征的最重要的关联信息,并计划未来的研究。在植物方面,仍没有这样全面的目录,虽然已经针对有些植物开发出用户友好的平台,但远远不够,仍需要各位植物学家共同努力打造一个公共平台,促进对特定物种的 GWAS 研究。大豆研究者应该加大 GWAS 分析的力度,共同努力打造共享平台,实现资源共享,充分利用 GWAS 培育出优良的大豆品种。

4 展望

基于连锁不平衡的 GWAS 的结果在大豆改良方面还没有得到充分利用,目前主要集中在对大豆蛋白、油分、抗逆性以及产量等相关性状上的初步探索。相比之下,通过基于连锁区间作图确定的 QTL 已成功渗入到几种重要作物的栽培品种中,从而实现品种改良。这也表明需要在大豆品种改良上充分利用 GWAS 分析得出的结果。

目前,大部分研究还只是停留在验证显著位点效果的阶段,后 GWAS 时代已经来临,需要进行更多的生物学研究,构建出一张完整的大豆分子调节网络图谱。GWAS 在大豆遗传育种的过

程中的应用还有着巨大的潜力。虽然目前 GWAS 仍存在许多不足,需要不断探索改进,但将来其在大豆育种中必将发挥更大的作用。

参考文献:

- [1] ZHAO X, DONG H R, CHANG H, et al. Genome wide association mapping and candidate gene analysis for hundred seed weight in soybean *Glycine max* (L.) Merrill[J]. BMC Genomics, 2019, 20(1): 648.
- [2] ZHANG X L, DING W T, XUE D, et al. Genome-wide association studies of plant architecture-related traits and 100-seed weight in soybean landraces[J]. BMC Genomic Data, 2021, 22(1): 10.
- [3] Humberto A G, Benjamin W, Braulio S C, et al. Association mapping of seed quality traits in *Brassica napus* L. using GWAS and candidate QTL approaches [J]. Molecular Breeding, 2015, 35(6): 143.
- [4] Gaut B S, Long A D. The lowdown on linkage disequilibrium [J]. Plant Cell, 2003, 15(7): 1502-1506.
- [5] 刘红占, 王俊生, 胡利宗, 等. 全基因组关联分析在油菜遗传育种中的应用和研究进展[J]. 分子植物育种, 2018, 16(8): 2563-2570.
- [6] 谈静, 郭俊杰, 曾杰. 多倍体植物复杂性状全基因组关联分析研究进展[J]. 分子植物育种, 2020, 18(4): 1282-1289.
- [7] 李廷雨, 黎永力, 甘卓然, 等. 全基因组关联分析在大豆中的研究进展[J]. 大豆科学, 2020, 39(3): 479-484.
- [8] FLINT-GARCIA S A, Thuillet A C, YU J, et al. Maize association population: A high-resolution platform for quantitative trait locus dissection[J]. The Plant Journal, 2005, 44(6): 1054-1064.
- [9] 万何平, 陈禅友, 陈高, 等. 全基因组关联分析在大豆遗传学上的研究进展[J]. 江汉大学学报(自然科学版), 47(3): 197-203.
- [10] ZHANG S S, HAO D R, ZHANG S Y, et al. Genome-wide association mapping for protein, oil and water-soluble protein contents in soybean [J]. Molecular Genetics and Genomics, 2021, 296(1): 91-102.
- [11] ZHANG K X, LIU S L, LI W B, et al. Identification of QTNs controlling seed protein content in soybean using multi-locus genome-wide association studies[J/OL]. Frontiers in Plant Science, 2018, 9 [2021-09-10]. https://www.researchgate.net/publication/329091396_Identification_of_QTNs_Controlling_Seed_Protein_Content_in_Soybean_Using_Multi-Locus_Genome-Wide_Association_Studies. DOI: 10. 3389/FPLS. 2018. 01690.
- [12] HWANG E Y, SONG Q J, JIA G F, et al. A genome-wide association study of seed protein and oil content in soybean [J]. BMC Genomics, 2014, 15: 1-12.
- [13] DIAS D A, POLO L R T, LAZZARI F, et al. Genome-wide association for mapping QTLs linked to protein and oil contents in soybean[J]. Pesquisa Agropecuaria Brasileira,

- 2017,52(10):896-904.
- [14] YAN L,DI R,WU C J,et al. Haplotype analysis of a major and stable QTL underlying soybean(*Glycine max*) seed oil content reveals footprint of artificial selection[J]. Molecular Breeding,2019,39(4):57.
- [15] LIU X Y,QIN D,PIERSANTI A,et al. Genome-wide association study identifies candidate genes related to oleic acid content in soybean seeds[J]. BMC Plant Biology, 2020,20(1):399.
- [16] LIU J Y,LI P,ZHANG Y W,et al. Three-dimensional genetic networks among seed oil-related traits,metabolites and genes reveal the genetic foundations of oil synthesis in soybean[J]. Plant Journal,2020,103(3):1103-1124.
- [17] LI X Y,ZHANG K X,SUN X,et al. Detection of QTL and QTN and candidate genes for oil content in soybean using a combination of four-way-RIL and germplasm populations [J]. Crop Journal,2020,8(5):802-811.
- [18] FANG C,MA Y M,WU S W,et al. Genome-wide association studies dissect the genetic networks underlying agro-nomical traits in soybean [J]. Genome Biology, 2017, 18:161.
- [19] SHIM S,HA J,KIM M Y,et al. *GmBRC1* is a candidate gene for branching in soybean[*Glycine max* (L.) Merrill] [J]. International Journal of Molecular Sciences, 2019, 20(1):135.
- [20] LÜ H Y,LI H W,FAN R,et al. Genome-wide association study of dynamic developmental plant height in soybean [J]. Canadian Journal of Plant Science, 2017, 97 (2): 308-315.
- [21] BORASH J,SINGODE A,TALUKDAR A,et al. Genome-wide association studies (GWAS) reveal candidate genes for plant height and number of primary branches in soybean *Glycine max* (L.) Merrill [J]. Indian Journal of Genetics and Plant Breeding,2018,78(4):460-469.
- [22] IKRAM M,HAN X,ZUO J F,et al. Identification of QTNs and their candidate genes for 100-seed weight in soybean (*Glycine max* L.) Using Multi-Locus Genome-Wide Association studies[J]. Genes,2020,11(7):714.
- [23] QI Z Y,SONG J,ZHANG K X,et al. Identification of QTNs controlling 100-seed weight in soybean using multi-locus genome-wide association studies[J/OL]. Frontiers in Genetics, 2020, 11 [2021-09-10]. https://www.researchgate.net/publication/342985970_Identification_of_QTNs_controlling_100-Seed_Weight_in_Soybean_Using_Multi-locus_Genome-Wide_Association_Studies. DOI:10.3389/fgene.2020.00689.
- [24] SONG J,SUN X,ZHANG K X,et al. Identification of QTL and genes for pod number in soybean by linkage analysis and genome-wide association studies [J/OL]. Molecular Breeding, 2020, 40 (6) [2021-09-10]. https://www.researchgate.net/publication/343787918_Identification_of_QTL_and_genes_for_pod_number_in_soybean_by_linkage_analysis_and_genome-wide_association_studies. DOI:10.1007/s11032-020-01140-w.
- [25] ZUO Q M,HOU J F,B. Zhou,et al. Identification of QTLs for growth period traits in soybean using association analysis and linkage mapping[J]. Plant Breeding, 2013, 132 (3): 317-323.
- [26] ZHOU L,WANG S B,JIAN J B,et al. Identification of domestication-related loci associated with flowering time and seed size in soybean with the RAD-seq genotyping method [J]. Scientific Reports,2015,5:9350.
- [27] ZHANG D,CHENG H,HU Z B,et al. Fine mapping of a major flowering time QTL on soybean chromosome 6 combining linkage and association analysis [J]. Euphytica, 2013,191(1):23-33.
- [28] LIU Z X,LI H H,FAN X H,et al. Phenotypic characterization and genetic dissection of growth period traits in soybean(*Glycine max*) using association mapping[J/OL]. PLoS One,2016,11(7)[2021-09-10]. <https://pubmed.ncbi.nlm.nih.gov/27367048/>. DOI: 10.1371/JOURNAL.PONE.0158602.
- [29] LI C,LI Y H,LI Y F,et al. A domestication-associated gene GmPRR3b regulates the circadian clock and flowering time in soybean [J]. Molecular Plant, 2020, 13 (5): 745-759.
- [30] ZHAO X,BAO D F,WANG W,et al. Loci and candidate gene identification for soybean resistance to Phytophthora root rot race 1 in combination with association and linkage mapping[J]. Molecular Breeding,2020,40(10):100.
- [31] ROLLING W,LAKE R,DORRANCE A E,et al. Genome-wide association analyses of quantitative disease resistance in diverse sets of soybean *Glycine max* (L.) Merr. plant introductions[J]. Plos One,2020,15(3):E0227710.
- [32] LAI M C,LAI Z Y,JHAN L,et al. Prioritization and evaluation of flooding tolerance genes in soybean *Glycine max* (L.) Merr. [J/OL]. Frontiers in Genetics, 2021, 11 [2021-09-10]. https://www.researchgate.net/publication/348810439_Prioritization_and_Evaluation_of_Flooding_Tolerance_Genes_in_Soybean_Glycine_max_L_Merr. DOI:10.3389/FGENE.2020.612131.
- [33] ZHANG H Y,LI C Y,DAVIS E L,et al. Genome-wide association study of resistance to soybean cyst nematode (*Heterodera glycines*) HG Type 2. 5. 7 in wild soybean (*Glycine soja*) [J]. Frontiers in Plant Science, 2016, 7:1214.
- [34] ZENG A,CHEN P,KORTH K,et al. Genome-wide association study(GWAS) of salt tolerance in worldwide soybean germplasm lines [J/OL]. Molecular Breeding, 2017, 37 (3) [2021-09-10]. <https://www.researchgate>.

net/publication/314241315_Genome-wide_association_study_GWAS_of_salt_tolerance_in_worldwide_soybean_germplasm_lines. DOI:10.1007/S11032-017-0634-8.

[35] WU C J, MOZZONI L A, MOSELEY D, et al. Genome-wide association mapping of flooding tolerance in soybean [J/OL]. *Molecular Breeding*, 2019, 40 (1) [2021-09-10]. https://www.researchgate.net/profile/Evangelina-Ella/publication/272623796_Genome-wide_association_mapping_of_tolerance_of_flooding_during_germination_in_rice/links/54eac8220cf27a6de114d144/Genome-wide-association-mapping-of-tolerance-of-flooding-during-germination-in-rice.pdf. DOI: 10.13140/2.1.1663.3125.

[36] WEN Z X, TAN R J, ZHANG S C, et al. Integrating GWAS and gene expression data for functional characterization of resistance to white mould in soya bean[J]. *Plant Biotechnology Journal*, 2018, 16(11):1825-1835.

[37] GUPTA P K, KULWAL P L, JAISWAL V. Association mapping in plants in the post-GWAS genomics era[J]. *Advances in Genetics*, 2019, 104:75-154.

[38] CHAUDHARY J, Patil G B, Sonah H, et al. Expanding omics resources for improvement of soybean seed composition traits[J]. *Frontiers in Plant Science*, 2015, 6:1021.

[39] 郝兴杰, 胡林, 张淑君. 全基因组关联分析方法的研究进展[J]. *畜牧兽医学报*, 2016 47(2):213-217.

Application Progress of Genome-wide Association Analysis on Soybean Genetic Breeding

GUAN Bo-wen¹, CHEN Qing-shan¹, WU Xiao-xia¹, YANG Xue²

(1. School of Agriculture, Northeast Agricultural University, Harbin 150030, China; 2. Prataculture Research Institute, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China)

Abstract: At present, there is a lack of high-yield and high-quality soybean varieties in China, so we need to strengthen research and development. Genome wide association analysis (GWAS) has the advantages of high throughput, high precision and less time-consuming. In order to promote the application of GWAS in the process of soybean genetics and breeding, based on the introduction of the theory of GWAS method, this paper summarized its application progress in the genetics and breeding of main agronomic characteristics of soybean, discussed the advantages and disadvantages of GWAS and its development trend, and put forward solutions.

Keywords: soybean; genome-wide association analysis; genetic breeding; character improvement; application progress

(上接第 89 页)

Study on Grain Production Efficiency in Heilongjiang Province

JIANG Yu-bo^{1,2}, LI Shuang², YU Yang¹, GONG Xiu-jie¹, HAO Yu-bo¹, LI Liang¹, LYU Guo-yi¹, QIAN Chun-rong¹

(1. Institute of Crop Cultivation and Tillage, Heilongjiang Academy of Agricultural Sciences, Harbin 150028, China; 2. College of Economics and Management, Northeast Agricultural University, Harbin 150030, China)

Abstract: In order to clarify the technical efficiency level and influencing factors of grain production in Heilongjiang Province, the data envelopment analysis method and the random frontier analysis method were used comprehensively. Through the Three-stage DEA analysis model, the input variables of human force, machinery, land and other inputs in agricultural production were modified by using the relevant variables of economic environment, natural environment and social environment. From the macro perspective, modified the comprehensive technical efficiency of grain between Heilongjiang Province and other regions were compared and analyzed. The results showed that the environmental factors had a significant impact on the agricultural production investment in various regions, and the calculation results in most regions had changed after excluding the environmental factors. Although the comprehensive technical efficiency of grain production in Heilongjiang Province was at the forefront of China's agricultural production, the results of super efficiency analysis showed that Heilongjiang Province still had some room to improve.

Keywords: grain crops; Heilongjiang Province; production efficiency; Three-stage DEA analysis model