

全基因组重测序在大豆育种上的研究进展

赵海红¹, 郭泰¹, 王志新¹, 郑伟¹, 李灿东¹, 徐杰飞¹, 张振宇¹, 赵星棋¹, 王士强²

(¹黑龙江省农业科学院佳木斯分院, 黑龙江佳木斯 154007;

²黑龙江八一农垦大学农学院, 黑龙江大庆 163319)

摘要:大豆是世界上最主要的油料作物之一。随着基因测序技术的快速发展及广泛应用,大豆育种研究受到了深刻的影响。本文简述了基因测序的主要方法及全基因组重测序在大豆育种、遗传转化研究中的应用,综述了目前全基因组重测序从发现突变位点、揭示进化关系、挖掘功能基因等基因功能和结构的角度,来研究大豆育种及疾病发生过程中的分子机理,为今后基因测序技术在大豆育种中的研究与应用提供参考。

关键词:大豆育种;全基因组重测序;高通量测序技术

中图分类号:S565.1

文献标志码:A

论文编号:cjas19020011

Whole Genome Resequencing in Soybean Breeding: Research Progress

Zhao Haihong¹, Guo Tai¹, Wang Zhixin¹, Zheng Wei¹, Li Candong¹, Xu Jiefei¹,

Zhang Zhenyu¹, Zhao Xingqi¹, Wang Shiqiang²

(¹Jiamusi Branch of Heilongjiang Academy of Agricultural Sciences, Jiamusi 154007, Heilongjiang, China;

²College of Agronomy, Heilongjiang Bayi Agricultural University, Daqing 163319, Heilongjiang, China)

Abstract: Soybean is one of the most important oil crops in the world. With the rapid development and the wide application of gene sequencing technology, soybean breeding research has been deeply affected. This paper briefly described the main methods of gene sequencing and the whole genome resequencing application in soybean breeding and genetic transformation research, summarized the whole genome resequencing to study the molecular mechanisms of soybean breeding and disease development from the perspective of functions and structures of functional genes through discovering mutation sites, revealing evolutionary relationships, and mining functional genes. The study could provide certain reference for the research and application of the gene sequencing technology in the future.

Keywords: Soybean Breeding; Whole Genome Resequencing; High-throughput Sequencing

0 引言

大豆是目前全世界种植面积相对比较大的作物之一,大豆作为重要的食品、饲料和工业原料,在人类的生产和生活中发挥着重要的作用。2017年中国大豆播种面积为819.4万hm²,较上年增加13.8%,全国大豆

平均单产为1817 kg/hm²,较上年增加1.2%^[1],这些大豆主要用于国民食用。然而,据估计,中国国内饲料用粮需求在3.8亿~4.0亿t,而目前饲料用粮仅为2.7亿t,饲料粮缺口较大^[2],可见,中国对大豆的需求急需解决。2017年中国大豆进口量就约为9553万t^[3],而且多年

基金项目:北方早熟大豆优质高产广适新品种培育“国家重点研发项目”(2017YFD0101302-2);主要农作物种质资源创新与规模化制种技术研究“黑龙江省重点项目”(GA18B101);优良新品种培育和示范推广国家重点研发项目(2016YFD0101905);国家大豆产业技术体系综合试验站(CARS-04-CES05);大豆等经济作物诱变育种技术与新品种创制“主要农作物诱变育种项目”(2016YFD0102105)。

第一作者简介:赵海红,女,1981年出生,黑龙江五常人,助理研究员,硕士,主要从事大豆相关研究。通信地址:154007 黑龙江省佳木斯市安庆街531号, Tel:0454-8351161, E-mail: haihong51job@163.com。

通讯作者:郭泰,男,1964年出生,黑龙江佳木斯人,研究员,硕士,主要从事大豆育种研究。通信地址:信地址:154007 黑龙江省佳木斯市安庆街531号, Tel:0454-8351161, E-mail: guotaidadou@163.com。

收稿日期:2019-02-21, **修回日期:**2019-05-31。

来,中国大豆市场供给高度依赖进口,且进口渠道较为集中,美国作为中国大豆第二大进口国,2018年以来,却与中国的贸易摩擦不断升级^[4]。振兴中国大豆产业势在必行。因此,必须通过大豆品种改良、机械化生产、规模化种植等多种途径,持续提高中国大豆单产及品质,为改善中国大豆种植提供技术保障。那么,研究大豆在特定条件下的基因表达模式,对于开发和利用大豆种质资源也就十分必要,全基因组重测序工作对大豆育种及相关分子生物学研究意义重大。

2010年,大豆基因序列的揭示^[5],对于开展大豆靶向性分子育种、聚合育种,高效精准地培育大豆新品种及创制新种质具有极大的促进作用,基因测序为了解大豆基因组的结构和进化提供了宝贵资源^[6],有利于大豆的农艺性状的分析,利用基因测序研究成果可选育出高抗病、抗多种病害、抗除草剂、高品质、高产量的新品种^[7-8],极大丰富大豆品种类型,大大缩短育种周期的同时,为基于测序技术的大豆分子育种打好坚实基础,使农民增产增收增效的同时,增强中国大豆生产在国际市场上的地位。

1 基因测序技术的发展

基因测序技术在分子生物学和基础医学领域应用广泛,已经从第一代发展到第四代^[9]。

1.1 第一代测序技术

1977年Maxam和Gibert^[10]报道了通过化学降解测定DNA序列的方法,由于该方法程序复杂、操作繁琐,后来逐渐被Sanger法取代。随后,在Sanger技术的基础上,发展起来的荧光自动测序技术是应用最为广泛的第一代测序技术^[9]。该测序技术测序精确,但通量较低,成本高,后来逐渐被二代测序技术取代。

1.2 第二代测序技术

2004年,随着Roche公司首次推出以高通量低成本为主要特点的测序平台,生物学研究正式进入了以二代测序为核心测序手段的时代^[11]。第二代测序技术采用的策略是边合成边测序,依照Sanger测序原理,捕捉末端荧光标记来确定DNA序列^[12]。其最显著的特征是高通量、低成本、自动化程度高,能在很短的时间内完成上百亿个碱基对的测序,一次能对几十万到几百万条DNA分子进行序列测序,满足极短时间内对基因组进行高分辨率检测的要求^[13]。

1.3 第三代测序技术

鉴于一代和二代测序存在依赖于模板扩增以及产生的序列较短等缺点,在需要进行组装,比对以及注释等生物信息学分析时存在着很大的困难^[14],为了补充和进一步完善测序技术,近几年研发出第三代的测序

方法—单分子测序技术(Single-molecule Sequencing),主要包括生物科学公司的Heli Scope单分子测序仪^[15]、太平洋生物科学公司的单分子实时DNA测序技术(Single-molecule Real-time, SMRT)和Oxford Nanopore公司的单分子纳米孔测序技术(The Single-molecule Nanopore DNA Sequencing)^[16]等。现阶段大多数的研究还采用第二代和第三代测序数据的结合使用,这能够有效的解决由二代测序本身局限所导致的研究难题。

1.4 第四代测序技术

英国牛津纳米公司研发的MinION测序仪是基于纳米孔单分子测序技术的第四代测序技术,其体积比普通U盘稍大,携带十分方便,MinION默认运行时间为48 h,其标准识别速度为一秒钟识别250个碱基(250 bps)^[17],运行2 min产生的数据就可以用于相关分析,平均读长可以达到13~20 kb^[18],通过提高测序深度,MinION测序可以达到98%的准确率^[19]。但是生物纳米孔使用的脂基质,其机械稳定性较弱,反复洗涤会降低测序芯片的使用寿命^[20]。

2 全基因组重测序在大豆育种中的应用

近年来基因组重测序也越来越广泛地应用于物种进化研究和育种研究领域^[21,26-31]。2002年水稻全基因组计划率先宣布完成^[22]。在随后几年里,科学家们先后完成了包括黄瓜^[23]、玉米^[24]、高粱^[25]、大豆^[5]等物种的全基因组测序,而后,通过基因组重测序从更多的角度来研究作物育种学,比如发现突变位点^[26-27],揭示进化关系^[28]、挖掘功能基因^[29-30]、分子辅助选择育种^[31]等,这加快了优质新品种的培育进程。全基因组重测序技术逐渐成为重要的育种手段。

在大豆育种研究中,开发诱变群体、获得有感兴趣的变异表型的突变株之后,传统的正向遗传学方法往往是通过构建F₂分离群体、重组自交系等方法定位与表型相关的基因,要连续选育多代,工作量大而且周期较长。通过全基因组重测序等反向遗传学手段,可以快速发现与表型相关的基因位点,加快育种进程。

大豆基因序列的揭示^[5]后,2010年Lam等^[32]对31份大豆包括野生种和栽培种进行全基因组重测序,共获得205614个SNP位点,并发现野生大豆中等位基因多态性水平要明显高于栽培大豆。

2012年李泽锋^[33]利用二代测序技术,对来自中国江南的一个野生大豆基因组(‘兰溪1号’,Lanxi1)进行了深度测序(>50X),发现‘兰溪1号’相对其他北方野生大豆来说,与栽培大豆亲缘关系更远,这有可能暗示出大豆是从中国北方开始驯化的。

2015年Zhou等^[34]人通过对302株野生大豆、地方品种和驯化品种大豆进行大于11X深度的高通量测序,发现979万个SNPs,87.68万个Indels,还有1614个CNVs和6388个大片缺失。发现230个受选择区域和162个拷贝数变异。全基因组关联分析表明10个受选择区域和9个驯化性状相关联,发现13个被注释为与油脂、株高等农艺性状相关的位点。与之前QTL定位结果比较分析发现,230个受选择区域中96个与调控油脂的QTL相关,21个区间内包含脂肪合成关键基因。

2016年张彦威^[9]等对大豆品种‘齐黄34’进行全基因组重测序,共检测到1519494个SNP位点,357549个小片段InDel位点,4506个结构变异,为分子标记辅助选择提供了重要的标记资源。同年,33份栽培品种和68份野生大豆品种这两个基因的序列被分析^[35],显示野生型大豆中大豆孢囊线虫抗性位点基因缺失,大豆孢囊线虫抗性基因可能是在大豆驯化过程中人工选育出来的。

2017年周航^[36]对细胞分裂素信号转导具有反向调节作用的GmAP1.2基因在大豆中的功能进行了研究,探讨了其与GmARR10的差别,为大豆抗逆和开花基因的发掘应用以及大豆分子育种提供了理论依据。

2018年葛逢勇^[37]通过表型鉴定在大豆孢囊线虫表现广谱抗性的PI437654突变群体中,获得对大豆孢囊线虫4号生理小种侵染表型发生变化的序列初步分析,以期后续研究最终鉴定出大豆抗大豆孢囊线虫4号生理小种基因。同年,张峰阁^[38]研究表明,突变体ZK1917和‘黑农35’在Dt2和GmSOC1CDS区未发生变异,通过对Dt1/Dt2/GmSOC1表达量的监控,发现目的基因可能在V2期生长点通过影响GmSOC1与Dt1的结合作用,从而释放Dt1的表达,进而影响结荚习性。

此外,通过全基因组重测序技术,对大豆矮秆基因^[39]、大豆根中磷酸盐缺乏的响应基因^[40]、对黑豆抗胞囊线虫基因^[41]等进行了相关研究,为大豆分子育种提供了理论依据。

3 展望

随着第三代、第四代测序技术的不断完善和发展,将来测序成本会大幅度降低,测序的通量和准确性也会不断提高,高通量测序技术将成为一项常规的实验方法,通过这一方法,转基因技术在应用中遇到的一些问题将迎刃而解,比如大豆的许多经济性状是数量性状,受微效多基因控制,而且基因数量庞大,分离目的基因困难;外源DNA直接导入,一般导入的是总

DNA,缺乏标记基因,鉴定难度大;转化后表达基因的调控、标记基因的去除等,这些目前研究热点问题将会得到有效地解决。随着分子生物学的快速发展,全基因组重测序技术必将会在大豆育种中大显身手。最终人类将会突破了自然资源的限制,按照需要选择基因,更进一步的改造大豆的遗传物质,赋予其新的性状,大幅度提高了大豆的产量和品质,使大豆可能同时具有高产、优质、抗病虫、抗寒、抗旱、抗除草剂等多重优点。大豆育种已进入了新的时期。

参考文献

- [1] 殷瑞锋,徐雪高,张振. 2017年大豆市场形势分析与2018年展望[J]. 农业展望,2017(12):4-7.
- [2] 成升魁,徐增让,谢高地,等. 中国粮食安全百年变化历程[J]. 农学报,2018,8(1):186-192.
- [3] 李顺萍. 世界大豆生产布局及中国大豆对外依存度分析[J]. 世界农业,2018(11):108-112.
- [4] 原梓涵,邵娜. 中美贸易摩擦对大豆市场的影响及前景分析[J]. 农业展望,2018,10,89-93.
- [5] Schmutz J, Cannon S B, Schlueter J, et al. Genome sequence of the palaeopolyploid soybean[J]. Nature,2010,463:178-183.
- [6] 张彦威,李伟,张礼凤,等. 基于重测序的大豆新品种齐黄34的全基因组变异挖掘[J]. 中国油料作物学报,2016,38(2):150-158.
- [7] Park J, Lee Y, Kang B, Wi. Co-transformation using a negative selectable marker gene for the production of selectable marker gene-free transgenic plants[J]. Theoretical and Applied Genetics,2004,109:1562-1567.
- [8] Vidal J R, Kikkert J R, Wallace P G, Reich B I. High-efficiency biolistic co-transformation and regeneration of 'Chardonnay' (*Vitis vinifera* L.) containing npt-II and antimicrobial peptide genes[J]. Plant Cell Reports,2003,22(4):252-260.
- [9] 马潞林,宋一萌,葛力源. 基因测序技术的发展与临床应用概述[J]. 重庆医科大学学报,2018,43(4):477-479.
- [10] Maxam A M, Gilbert W. A New Method for Sequencing DNA[J]. Proc Natl Acad Sci USA,1977,74(2):560-564.
- [11] 陶申童. 基于重测序的杨树基因组重组事件的研究[D]. 南京:南京林业大学,2017.
- [12] Mardis E R. The impact of next-generation sequencing technology on genetics[J]. Trends in Genetics,2008,24(3):133-141.
- [13] Shendure J, Porreca G J, Reppas N B, et al. Accurate Multiplexed Polony Sequencing of an Evolved Bacterial Genome [J]. Science, 2005,309(5741):1728-1732.
- [14] Pop M, Salzberg S L. Bioinformatics challenges of new sequencing technology[J]. Trend in Genetics,2008,24(3):133-141.
- [15] Steinmann K E, Hart C E, Thompson J F, et al. Helicos Single-molecule Sequencing of Bacterial Genomes[J]. Methods Mol Biol. , 2011,733:3-24.
- [16] Ying Y L, Cao C, Long Y T. Single Molecule Analysis by Biological Nanopore Sensors[J]. Analyst,2014,139(16):3826-3835.
- [17] Miten J, Hugh E. Olsen, Benedict Paten, et al. The Oxford

- Nanopore Min ION: delivery of nanopore sequencing to the genomics community[J]. Genome Biology, 2016, 17(1): 239.
- [18] Tyson J R, O'Neil N J, Jain M, et al. Min ION-based long-read sequencing and assembly extends the *Caenorhabditis elegans* reference genome[J]. Genome Research, 2018, 28: 266-274.
- [19] Giordano F, Aigrain L, Quail M A., et al. De novo yeast genome assemblies from Min ION, Pac Bio and Mi Seq platforms[J]. Scientific Reports, 2017, 7(1): 3935.
- [20] Haque F, Li J, Wu H C, et al. Solid-State and Biological Nanopore for Real-Time Sensing of Single Chemical and Sequencing of DNA [J]. Nano Today, 2013, 8(1): 56-74.
- [21] 冉洪. 厚壁毛竹种质性状的重测序研究[D]. 北京: 中国林业科学研究院, 2016.
- [22] GOFF, Stephen A, RICKE, et al. A draft séquence of the rice genome (*Oryza sativa* L. ssp. japonica): The rice genome[J]. Science, 2002, 296(5565): 92-100.
- [23] Huang S, Li R, Zhang Z, et al. The genome of the cucumber, *Cucumis sativus* L[J]. Nature genetics, 2009, 41(12): 1275-81.
- [24] Schnable P S, Ware D, Fulton R S, et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics[J]. Science, 2009, 326(5956): 1112-5.
- [25] Paterson A H, Bowers J E, Bruggmann R, et al. The Sorghum bicolor genome and the diversification of grasses[J]. Nature, 2009, 457(7229): 551-6.
- [26] 陈璇, 郭蓉, 王璐, 等. 基于全基因组重测序的野生型大麻和栽培型大麻的多态性 SNP 分析[J]. 分子植物育种, 2018, 16(3): 893-897.
- [27] Barabaschi D, Tondelli A, Desiderio F, et al. Next generation breeding[J]. Plant Science, 2016, 242: 3-13.
- [28] 胡鸣, 姚圣黎, 程晓晖, 等. 基于高深度重测序的春性、半冬性和冬性甘蓝型油菜基因组遗传变异分析[J]. 中国油料作物学报, 2018, 40(4): 469-478.
- [29] 任民, 程立锐, 刘旦, 等. 基于 RAD 重测序技术开发烟草品种 SNP 位点[J]. 中国烟草科学, 2018, 39(3): 10-17.
- [30] Wang W S, Ramil M, Hu Z Q, et al. Genomic variation in 3010 diverse accessions of Asian cultivated rice[J]. Nature, 2018, 557: 43-49.
- [31] 赵庆英, 张瑞娟, 王瑞良, 等. 基于名优谷子品种晋谷 21 全基因组重测序的分子标记开发[J]. 作物学报, 2018, 44(5): 686-696.
- [32] Lam H M, Xu X, et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection [J]. Nat Genet, 2010, 42: 1053-1059.
- [33] 李泽锋. 一个中国南方野生大豆基因组深度测序及其分析[D]. 杭州: 浙江大学, 2012.
- [34] Zhou L, Wang S B, Jian J, 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.
- [35] Wu X, Zhou G, Chen Y, et al. Soybean cyst nematode resistance emerged via artificial selection of duplicated serine hydroxymethyltransferase genes[J]. Frontiers in Plant Science, 2016, 7(17): 998.
- [36] 周航. 利用 Gm ARR10 基因创制大豆新种质[D]. 贵阳: 贵州大学, 2017.
- [37] 葛逢勇. 抗大豆孢囊线虫 4 号生理小种的 P1437654 突变体的筛选与全基因组测序[D]. 北京: 中国农业科学院, 2018.
- [38] 张峰阁. 大豆结荚习性控制基因的遗传定位[D]. 北京: 中国科学院大学, 2018.
- [39] Zhong-feng Li, Yong Guo, Lin Ou, et al. Identification of the dwarf gene GmDW1 in soybean (*Glycine max* L.) by combining mapping-by-sequencing and linkage analysis[J]. Theoretical and Applied Genetics, 2018, 131: 1001-1016.
- [40] Zeng H Q, Wang G P, Zhang Y Q, et al. Genome-wide identification of phosphate-deficiency-responsive genes in soybean roots by high-throughput sequencing[J]. Plant Soil, 2016, 398: 207-227.
- [41] 朱晓岚. 黑豆抗胞囊线虫基因序列多样性及功能分析[D]. 沈阳: 沈阳农业大学, 2017.