

大豆 3 个核心亲本及其衍生品系基于 PAV 分子标记的亲缘关系研究

王自力,张吉顺,郭呈宇,孙峰峦,李忠洋,孔杰杰,盖钧镒,赵团结

(南京农业大学 大豆研究所/国家大豆改良中心/农业部大豆生物学与遗传育种重点实验室(综合)/作物遗传与种质创新国家重点实验室,江苏 南京 210095)

摘要:基因组大片段序列存在/缺失变异(presence/absence variation, PAV)作为一种基于 PCR 技术、方便快捷的新标记受到人们的关注,但在大豆遗传育种工作中的应用较少。大豆育种计划中一般会有一批核心亲本,揭示其作用特点有助于亲本选配。应用 PAV 标记对国家大豆改良中心种质创新计划中 3 个核心亲本(南农 86-4、南农 88-48 和诱处 4 号)与国内外材料杂交所获得 34 个组合衍生品系及亲本共 154 份材料所构建的样本进行核心亲本对其衍生品系遗传贡献分析。结果表明:221 个 PAV 标记的平均等位变异数为 2.1,多态性信息含量指数(PIC)平均为 0.239,位于基因内和基因间标记的丰富度和 PIC 平均值相当。基于 PAV 标记信息可将 34 个供试组合聚为 6 类,其中 3 个大类可分别与 3 个核心亲本的组合相对应,核心亲本与其衍生品系的遗传距离最小。154 份材料可聚为 9 类,来自同一核心亲本、同一组合的材料多聚在一起,但也存在交叉现象。对 23 个单交和 6 个三交组合的亲本遗传贡献率分析表明,共有 22 个组合(占 75.90%)的核心亲本对衍生品系的平均贡献率高于基于系谱的理论值,其中来自本地地区的南农 86-4、南农 88-48 对衍生品系的平均遗传贡献值总体上高于其它杂交亲本,而异生态区的诱处 4 号的贡献率相对较低。PAV 标记能在系谱信息基础上进一步反映大豆亲本及衍生品系的亲缘关系。

关键词:大豆;PAV 标记;育成系;遗传多样性;核心亲本

中图分类号:S565.1 **文献标识码:**A **DOI:**10.11861/j.issn.1000-9841.2016.01.0001

Genetic Relationship of Three Soybean Core Parents and Their Derived Breeding Lines Detected by Presence/Absence Variation Markers

WANG Zi-li, ZHANG Ji-shun, GUO Cheng-yu, SUN Feng-luan, LI Zhong-yang, KONG Jie-jie, GAI Jun-yi, ZHAO Tuan-jie

(Soybean Research Institute/National Center for Soybean Improvement/Key Laboratory for Biology and Genetic Improvement of Soybean (General), Ministry of Agriculture/National Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: A new molecular marker system, presence/absence variation (PAV) of large fragment genomic sequences shows potential because of its convenient PCR-based technique and distinct band distribution on. However, its application in soybean breeding and genetic research remains to be carried. There are some core parents in soybean breeding program, which are the key to success in breeding, revealing their role is helpful for parent selection in hybrid breeding work. In this study, total 154 soybean materials including 28 parents and 126 breeding lines were used to reveal their genetic relationship. These lines were derived from the crosses between three core parents (Nannong 86-4, Nannong 88-48 and Youchu 4) and a set of domestic and exotic lines in a breeding program in National Center of Soybean Improvement. The average number of allele per PAV was 2.1. The average value of polymorphic information content (PIC) per PAV was 0.239. The average allele number of the gene-based and non-coding region PAV marker and the average values of polymorphic information content (PIC) of the two PAV types were almost the same. All 34 crosses were clustered into 6 groups based on PAV markers data, among them, three big groups correspond to the three core parents could be identified. All 154 materials were clustered into 9 groups. Most lines from the same core parent or same cross were classified as a group. According to the genetic contribution rate of PAV allele of the core parents, 22 crosses had higher average contribution rate than the expected ratio in the 29 crosses. The observed contribution values of core parent Nannong 86-4, Nannong 88-48 were higher than the expected values based on pedigree analysis, while Youchu 4 from other region had less contribution to the derived lines, indicating the rationalities of using local elite line as core parent in breeding program. In comparison with pedigree analysis, PAV marker can reveal detail genetic relationship among the parents and their derived lines.

Keywords: Soybean; PAV marker; Breeding line; Genetic diversity; Core parent

收稿日期:2015-03-23
基金项目:国家高技术研究发展计划“863 计划”(2012AA101106);国家自然科学基金(31271750);国家公益性行业(农业)科研专项经费(201203026-4);长江学者和创新团队发展计划(PCSIRT13073);江苏省优势学科建设工程专项(PAPD);江苏省现代作物生产协同创新中心项目(JCIC-MCP)。
第一作者简介:王自力(1990-),男,硕士,主要从事大豆分子育种研究。E-mail: njauwang@126.com。
通讯作者:盖钧镒(1936-),男,教授,博导,主要从事大豆遗传育种和数量遗传研究。E-mail: sri@njau.edu.cn;
赵团结(1969-),男,教授,博导,主要从事大豆遗传育种研究。E-mail: tjzhao@njau.edu.cn。

与传统的表型标记相比较,利用分子标记对大豆材料进行遗传分析,不受基因表达的时空限制和环境条件的影响,能够更快捷、准确地了解育种材料的遗传变异信息,针对性地选择亲本,从而培育出优良的品种组合^[1]。目前,已有多种 DNA 分子标记(RFLP、AFLP、RAPD、SSR、SNP 等)被用于多种农作物的分子连锁图的构建、遗传多样性分析、品种鉴定、分子标记辅助育种等方面的研究^[2]。作物全基因组序列信息使新型标记的开发与应用成为研究热点^[3]。基因组大片段存在/缺失(Presence/Absence variation, PAV)标记即结构变异标记是指序列在一个基因组中存在,但在另一个基因组中完全缺失^[4-5]。PAV 标记是新型的分子标记类型,普遍存在于各种生物的基因组中,可根据插入缺失位点两端的序列设计引物。在过去几年中,已经普遍在人类^[6]、果蝇^[7]、细菌^[8]以及植物^[9-13]研究中观察到。研究表明,PAV 标记可结合 SSR、SNP 等标记满足精细定位目的基因的需要,且可以较容易地通过生物信息学的方法获得,特异性高、经济实用、稳定性好。PAV 标记已用于植物基因组及供试群体遗传多样性研究^[9-13]。近年来,PAV 标记已应用于大豆多样性及遗传变异的研究^[14]。

在农作物新品种选育过程中,一些亲本具有配合力高、优良性状的遗传力强等优点,由其衍生许多新品种,这类亲本被称为核心亲本^[15]。大豆育种中也有类似情况。21 个祖先核心亲本对美国 1947-1988 年 258 个公共单位育成的大豆品种的贡献达 84.60%;对 1999-2001 年公共和私营单位育成的 312 个品种的贡献占 87.30%^[16]。熊冬金等^[17]通过系谱分析,归纳出了 1986-2005 年间中国 941 个大豆育成品种最重要的 54 个祖先亲本和 37 个直接亲本。在育种计划中,一些核心亲本被经常使用,育成大量新品系,这些品系具有的育种和生产利用潜力值得重视。本研究利用 PAV 标记分析 3 个核心亲本(南农 86-4、南农 88-48 和诱处 4 号)与国内外优良品种(品系)及中间材料的衍生品系的遗传关系,并通过比较其遗传背景、种间亲缘关系、系谱和标记的聚类分析结果,揭示其遗传关系,分析核心亲本形成的基本条件,同时解释核心亲本对衍生品系的遗传物质的传递及分布,旨在为今后的大豆设计育种和核心亲本预测奠定基础,也为高效利用这些种质提供理论依据。

1 材料与方法

1.1 材料

所用材料均来自国家大豆改良中心种质创新计划,主要面向江淮地区选育熟期适宜、优质、高产新种质。供试材料共 154 份,包括以南农 86-4、南农 88-48 和诱处 4 号 3 个核心亲本与不同来源种质杂交共 34 个组合衍生的新品系 126 份及其亲本 28 个(表 1)。3 个核心亲本中,南农 86-4 是从骨干亲本南农 1138-2 通过自然变异选择育种获得,关荣霞等^[18]利用 48 对 SSR 标记分析发现南农 86-4 和南农 1138-2 的相似性达 77.08%,具有高产高油等优良特性,其系谱可追溯到南方核心祖先亲本奉贤穗稻黄^[19]。南农 88-48 选自南农 73-935 × SRF400 组合,南农 73-935 则为骨干祖先亲本奉贤穗稻黄和 51-83(南农 493-1 亲本)杂交后代,具有早熟高产等优点^[20]。诱处 4 号是早熟 3 号 × 蒙城大青豆的辐射诱变选系,系谱可追溯到 58-161、徐豆 1 号、金元、铁荚四粒黄、嘟噜豆等黄淮和东北的骨干亲本,其具有高产高光效特点,曾获得 $4\ 500\ \text{kg} \cdot \text{hm}^{-2}$ 的高产纪录^[21]。育成品系采用混合法,各组合 $F_2 \sim F_4$ 代摘荚处理, $F_5 \sim F_6$ 代选丰产性、株型好的单株种成株行,再从中选择丰产性好、性状稳定的品系用于本研究。每个组合有 2~7 个品系,世代已到 $F_8 \sim F_{14}$ 代。为便于叙述,南农 86-4、南农 88-48 和诱处 4 号的衍生品系分别以 N、E 和 Y 表示,部分同时含两个核心亲本的以先出现的亲本表示,亲本则以 P 表示,第一个数字表示组合号,第二个数字表示衍生品系号。

1.2 PAV 标记基因型分析

采用 CTAB 法^[22]提取样品 DNA。利用已公布的大豆 PAV 引物开发和多态性分析数据^[14],选取在供试样本中存在多态性的 221 对 PAV 引物(表 2),在染色体上的分布密度是 8~22 个,标记间平均遗传距离约为 10 cM。基因型检测程序按文献 Wang 等^[14]所述,基本过程简述如下:PCR 反应体系为 $20\ \mu\text{L}$, $10 \times \text{PCR buffer}$ $2\ \mu\text{L}$, $2.5\ \mu\text{mol} \cdot \text{L}^{-1}$ dNTP $1.6\ \mu\text{L}$, $25\ \text{mmol} \cdot \text{L}^{-1}$ MgCl_2 $1.2\ \mu\text{L}$ 、上下游引物 ($10\ \mu\text{mol} \cdot \text{L}^{-1}$) 各 $0.1\ \mu\text{L}$ 、模板 ($40\ \text{ng} \cdot \mu\text{L}^{-1}$) $0.5\ \mu\text{L}$ 和 $0.2\ \mu\text{L}$ *Taq* 酶 ($5\ \text{U} \cdot \mu\text{L}^{-1}$)。PCR 反应过程如下: 94°C 预变性 5 min, 95°C 40 s, 55°C 退火 1 min, 72°C 延伸 2 min,共 35 个循环,最后 72°C 延伸

10 min。PCR 反应均在 GeneAmp PCR 9700(Applied Biosystems, Foster City, CA, USA) 中进行。利用 1.5% 琼脂糖凝胶检测产物片段大小。

1.3 基于 PAV 标记的遗传多样性及亲本遗传贡献分析

使用 PowerMarker V 3.25 软件^[23] 对 154 份材料进行遗传多样性分析,计算每个标记的等位变异数和多态性信息含量指数^[24] (polymorphism information content, PIC)。其计算公式为:

$$PIC = 1 - \sum_{i=1}^m p_i^2 - \sum_{i=1}^{m-1} \sum_{j=i+1}^m 2p_i^2 p_j^2$$

式中, p_i 是第 i 个等位基因在群体中的频率, p_j

为第 j 个等位基因在群体中的频率, m 为等位基因数。依照基因型数据运用 Nei^[25] 求算遗传距离,根据所计算的遗传距离,按非加权类平均法(unweighted pair group method with arithmetic mean, UPGMA) 进行聚类分析。利用软件 FigTree V 1.4.0 绘制亲缘关系树状图。进一步比较 34 个组合材料基于系谱和 PAV 标记的供试材料亲缘关系。

核心亲本对后代的遗传贡献分析:先统计单交、三交组合在亲本间表现多态的 PAV 标记。统计后代品系中表现为核心亲本等位变异类型的标记所占比例,作为核心亲本遗传贡献的估计值。

表 1 供试材料及其基于 PAV 标记信息的聚类结果

Table 1 List of the materials and their classification based on the data of PAV markers								
编号 Code	组合/亲本名称 Cross/Parent name	聚类结果 Cluster	编号 Code	组合/亲本名称 Cross/Parent name	聚类结果 Cluster	编号 Code	组合/亲本名称 Cross/Parent name	聚类结果 Cluster
N1-1	N7241/NN86-4	IX	E1-3	22369/NN88-48	IX	Y5-3	Sp//NN493-1/YC4	IX
N1-2	N7241/NN86-4	IX	E1-4	22369/NN88-48	IX	Y6-1	Chuxiu//NXHD/YC4	IX
N1-3	N7241/NN86-4	IX	E2-1	NN88-48/NG4690	VIII	Y6-2	Chuxiu//NXHD/YC4	IX
N1-4	N7241/NN86-4	IX	E2-2	NN88-48/NG4690	VIII	Y6-3	Chuxiu//NXHD/YC4	IX
N1-5	N7241/NN86-4	IX	E2-3	NN88-48/NG4690	VIII	Y7-1	HCBYB/YC4	VII
N2-1	NN86-4/NH5	IX	E3-1	NN88-48/NN87-17	VIII	Y7-2	HCBYB/YC4	VII
N2-2	NN86-4/NH5	VIII	E3-2	NN88-48/NN87-17	VIII	Y7-3	HCBYB/YC4	VII
N2-3	NN86-4/NH5	IX	E3-3	NN88-48/NN87-17	VIII	Y7-4	HCBYB/YC4	VII
N2-4	NN86-4/NH5	IX	E3-4	NN88-48/NN87-17	VIII	Y8-1	XSD/YC4	IX
N3-1	NN86-4/NT-1	V	E4-1	NN88-48/AGH	VIII	Y8-2	XSD/YC4	IX
N3-2	NN86-4/NT-1	V	E4-2	NN88-48/AGH	VIII	Y9-1	YC4/NG5545	VI
N3-3	NN86-4/NT-1	VIII	E4-3	NN88-48/AGH	VIII	Y9-2	YC4/NG5545	VI
N4-1	NN86-4/Sp	IX	E4-4	NN88-48/AGH	VIII	Y9-3	YC4/NG5545	VI
N4-2	NN86-4/Sp	IX	E4-5	NN88-48/AGH	IX	Y10-1	YC4/Graham	VI
N4-3	NN86-4/Sp	IX	E4-6	NN88-48/AGH	VIII	Y10-2	YC4/Graham	VI
N4-4	NN86-4/Sp	IX	E5-1	NN88-48/NT-1	VIII	Y11-1	YC4/Chudou 1	IX
N4-5	NN86-4/Sp	IX	E5-2	NN88-48/NT-1	VIII	Y11-2	YC4/Chudou 1	IX
N5-1	NN86-4/Vance	IX	E5-3	NN88-48/NT-1	VIII	Y11-3	YC4/Chudou 1	IX
N5-2	NN86-4/Vance	IX	E6-1	NN88-48/Sp	VIII	Y11-4	YC4/Chudou 1	VIII
N6-1	NN86-4/XSD	IX	E6-2	NN88-48/Sp	IX	Y12-1	ZGDD/YC4	IX
N6-2	NN86-4/XSD	IX	E6-3	NN88-48/Sp	VIII	Y12-2	ZGDD/YC4	IX
N7-1	NN86-4/XSD//NN87-23	V	E6-4	NN88-48/Sp	VIII	Y12-3	ZGDD/YC4	IX
N7-2	NN86-4/XSD//NN87-23	IX	E7-1	NN88-48/T173	VIII	P1	NG4690	VI
N7-3	NN86-4/XSD//NN87-23	IX	E7-2	NN88-48/T173	VIII	P2	NN73-935	IX
N7-4	NN86-4/XSD//NN87-23	IX	E7-3	NN88-48/T173	VIII	P3	NN86-4	IX

续表 1

编号 Code	组合/亲本名称 Cross/Parent name	聚类结果 Cluster	编号 Code	组合/亲本名称 Cross/Parent name	聚类结果 Cluster	编号 Code	组合/亲本名称 Cross/Parent name	聚类结果 Cluster
N7-5	NN86-4/XSD//NN87-23	IX	E7-4	NN88-48/T173	VIII	P4	NN88-48	VIII
N7-6	NN86-4/XSD//NN87-23	IX	E8-1	NN88-48/NXHD//sb	V	P5	自贡冬豆 ZGDD	IX
N7-7	NN86-4/XSD//NN87-23	IX	E8-2	NN88-48/NXHD//sb	V	P6	T173	II
N8-1	NN86-4/YC4	IX	E8-3	NN88-48/NXHD//sb	V	P7	NT-1	V
N8-2	NN86-4/YC4	IX	E8-4	NN88-48/NXHD//sb	V	P8	南雄黄豆 NXHD	V
N8-3	NN86-4/YC4	IX	E8-5	NN88-48/NXHD//sb	V	P9	汉川八月爆 HCBYB	VII
N8-4	NN86-4/YC4	IX	E9-1	NN88-48/YC4	IX	P10	滁豆 1 号 Chudou 1	IX
N8-5	NN86-4/YC4	IX	E9-2	NN88-48/YC4	VII	P11	NN493-1	IX
N8-6	NN86-4/YC4	IX	E10-1	N//XSD/NN88-48	VIII	P12	NN88-31	IX
N9-1	NN87-23/NN86-4	IX	E10-2	N//XSD/NN88-48	VIII	P13	NN87-23	IX
N9-2	NN87-23/NN86-4	IX	E10-3	N//XSD/NN88-48	IX	P14	YC4	IX
N9-3	NN87-23/NN86-4	IX	Y1-1	NN493-1/YC4	IX	P15	N7241	III
N9-4	NN87-23/NN86-4	IX	Y1-2	NN493-1/YC4	IX	P16	Vance	VI
N9-5	NN87-23/NN86-4	IX	Y1-3	NN493-1/YC4	IX	P17	NN1138-2	IX
N9-6	NN87-23/NN86-4	IX	Y1-4	NN493-1/YC4	IX	P18	油 94-412 You94-412	IX
N9-7	NN87-32/NN86-4	IX	Y1-5	NN493-1/YC4	IX	P19	Graham	VI
N10-1	Sp//NN86-4/YC4	IX	Y1-6	NN493-1/YC4	IX	P20	香水豆 XSD	IX
N10-2	Sp//NN86-4/YC4	IX	Y2-1	NN88-31/YC4//NN73-935	IX	P21	油 96-3 You 96-3	VII
N10-3	Sp//NN86-4/YC4	IX	Y2-2	NN88-31/YC4//NN73-935	IX	P22	NN87-17	IX
N11-1	You 94-412/NN86-4	IX	Y3-1	You 96-3//YC4/NN88-31	IX	P23	油春 96-4 Youchun 96-4	IX
N11-2	You 94-412/NN86-4	IX	Y3-2	You 96-3//YC4/NN88-31	IX	P24	NH5	VI
N11-3	You 94-412/NN86-4	VIII	Y4-1	NIL91-1N/YC4	IV	P25	NG5545	VI
N12-1	Youchun 96-4/NN86-4	IX	Y4-2	NIL91-1N/YC4	IV	P26	Sp	VI
N12-2	Youchun 96-4/NN86-4	IX	Y4-3	NIL91-1N/YC4	IV	P27	AGH	I
N12-3	Youchun 96-4/NN86-4	IX	Y4-4	NIL91-1N/YC4	IV	P28	楚秀 Chuxiu	IX
E1-1	22369/NN88-48	IX	Y5-1	Sp//NN493-1/YC4	VIII			
E1-2	22369/NN88-48	IX	Y5-2	Sp//NN493-1/YC4	IX			

材料编号命名方式为组合号-品系号;N、E、Y 分别代表含南农 86-4、南农 88-48 和诱处 4 号核心亲本的组合。NN86-4 为南农 86-4,NN88-48 为南农 88-48,YC4 为诱处 4 号。

The material code is named as crossidentifier-line code. N, E, Y represent the crosses using the core parent Nannong 86-4, Nannong 88-48, Youchu 4 respectively. NN86-4 indicates Nannong 86-4, NN88-48 indicates Nannong 88-48, YC4 indicates Youbian 4.

1.4 基于系谱的组合品系亲缘关系分析

根据系谱,供试品系上溯至 51 个原始亲本,参考崔章林等^[19]方法计算原始亲本在每个品系中所占的比例。使用 NTSYSpC Version 2.10e 软件^[26],计算两两品系余弦距离,然后按 UPGMA 进行聚类分析。

2 结果与分析

2.1 供试材料 PAV 标记的多样性特点

221 个 PAV 标记在 154 份材料中共有 463 个等位变异,平均每个标记为 2.1 个(表 2)。其中 202 个标记有 2 个等位变异,占总标记数 91.40%,17 个标记等位变异数为 3,占总标记数 7.69%,2 个标记

等位变异数为 4, 占总标记数 0.90%。总的看, PAV 较大变异。丰富度和 PIC 值最高的标记是位于 6 号
标记等位变异数目较少。标记多态性信息含量指染色体的 06P0352。
数 PIC 值平均为 0.239, 变幅为 0.007 ~ 0.611, 存在

表 2 221 个 PAV 标记在供试 154 份材料中的多样性指数
Table 2 Diversity parameters of 221 PAV markers in the tested 154 lines

标记名称	等位变	多态信息	标记名称	等位变	多态信息	标记名称	等位变	多态信息	标记名称	等位变	多态信息
Marker	异数	含量	Marker	异数	含量	Marker	异数	含量	Marker	异数	含量
name	AN	PIC	name	AN	PIC	name	AN	PIC	name	AN	PIC
01P0206	2	0.371	05P0522	2	0.019	09P1243	2	0.329	15P0687	2	0.067
01P0361	2	0.068	05P0726	2	0.099	09P1340	2	0.169	15P1300	2	0.267
01P0624	2	0.109	05P0124	2	0.298	09P1364	2	0.184	15P1409	2	0.031
01P0849	2	0.209	05P0496	2	0.099	09P1428	2	0.025	15P1450	2	0.201
01P1094	2	0.353	05P0686	2	0.027	09P1523	2	0.234	15P1451	2	0.095
01P0504	2	0.332	05P0655	2	0.253	10P0293	2	0.235	15P1487	2	0.328
01P0998	2	0.289	05P0868	3	0.545	10P0368	2	0.110	15P1562	2	0.212
01P1352	2	0.321	06P0596	2	0.348	10P0503	2	0.078	16P0403	2	0.169
02P0124	3	0.400	06P1249	2	0.216	10P0576	2	0.375	16P0436	2	0.132
02P0820	2	0.364	06P1288	2	0.204	10P0683	2	0.031	16P0595	2	0.153
02P0946	2	0.195	06P1301	2	0.183	10P0919	2	0.014	16P0781	2	0.330
02P1379	2	0.315	06P1360	2	0.242	10P1217	3	0.104	16P1098	2	0.315
02P0564	2	0.341	06P1362	2	0.265	10P1243	2	0.050	16P0272	2	0.368
02P0784	2	0.374	06P1365	2	0.283	11P0009	2	0.126	16P0616	2	0.359
02P0169	3	0.394	06P1371	2	0.223	11P0027	2	0.329	16P0626	2	0.354
02P0368	2	0.013	06P0067	2	0.373	11P0034	2	0.119	16P0856	2	0.355
02P0480	3	0.541	06P0568	2	0.351	11P0042	2	0.119	16P1017	2	0.188
02P1028	2	0.309	06P1284	2	0.352	11P0058	2	0.112	17P0018	2	0.374
03P1236	2	0.177	06P1366	2	0.362	11P0269	2	0.367	17P0117	3	0.093
03P1072	2	0.193	06P1581	3	0.297	11P0400	2	0.299	17P0136	2	0.019
03P0995	2	0.360	06P0222	2	0.067	11P1085	2	0.375	17P0253	3	0.508
03P0504	2	0.271	06P0352	4	0.611	12P0437	2	0.306	17P1020	2	0.254
03P0416	2	0.372	06P1392	2	0.348	12P0467	2	0.095	17P1033	2	0.242
03P0069	2	0.067	06P1411	2	0.355	12P0669	2	0.278	17P1109	2	0.295
03P0273	3	0.353	06P1426	2	0.344	12P1027	3	0.390	17P1187	2	0.342
03P0364	2	0.320	07P0051	2	0.144	12P1060	3	0.382	17P0788	2	0.309
03P0390	2	0.331	07P0389	2	0.374	12P1230	2	0.129	17P1384	2	0.323
03P0447	2	0.294	07P0595	2	0.370	12P1234	2	0.178	18P0141	2	0.185
03P0457	2	0.293	07P0765	2	0.303	12P0081	2	0.362	18P0247	2	0.307
03P0490	2	0.297	07P0833	2	0.159	12P0195	2	0.346	18P0397	2	0.280
03P0503	2	0.267	07P1368	2	0.283	12P0390	2	0.372	18P0599	2	0.322
03P0606	2	0.345	07P0167	2	0.183	13P0424	2	0.372	18P1108	2	0.104
03P0633	2	0.133	07P0793	2	0.191	13P1374	2	0.088	18P1146	2	0.365
03P0642	2	0.347	07P0917	2	0.038	13P1655	2	0.095	18P1186	2	0.349
03P0689	2	0.371	07P1175	2	0.273	13P1810	2	0.090	18P1405	2	0.152
03P0694	2	0.371	08P0549	2	0.356	13P1852	2	0.055	18P1783	2	0.360
03P0926	2	0.350	08P0671	2	0.129	13P1902	2	0.043	18P1786	2	0.117
03P0926	2	0.374	08P0784	2	0.370	13P0179	2	0.363	19P0128	3	0.137
03P1147	2	0.281	08P1233	2	0.171	13P0530	2	0.218	19P0363	2	0.160
03P1356	2	0.274	08P0381	2	0.369	13P0627	2	0.299	19P0676	2	0.201
04P0604	2	0.019	08P0941	2	0.251	13P1078	2	0.364	19P0866	2	0.050
04P0619	2	0.145	08P0994	2	0.375	14P0042	2	0.296	19P0900	2	0.333
04P0720	2	0.202	08P1206	2	0.072	14P0077	2	0.019	19P1169	3	0.482
04P0736	2	0.109	09P0140	2	0.050	14P0417	2	0.372	19P1286	3	0.399

续表 2

标记名称	等位变	多态信息	标记名称	等位变	多态信息	标记名称	等位变	多态信息	标记名称	等位变	多态信息
Marker	异数	含量	Marker	异数	含量	Marker	异数	含量	Marker	异数	含量
name	AN	PIC	name	AN	PIC	name	AN	PIC	name	AN	PIC
04P0833	2	0.038	09P0222	2	0.038	14P0598	2	0.109	19P1295	2	0.038
04P0851	2	0.084	09P0314	2	0.026	14P0641	2	0.236	20P0160	2	0.323
04P0061	2	0.212	09P0516	2	0.176	14P0979	2	0.370	20P0796	2	0.038
04P0084	2	0.007	09P0816	2	0.373	14P1166	2	0.359	20P0850	2	0.177
04P0104	4	0.190	09P0864	2	0.326	14P1209	2	0.374	20P0874	2	0.099
04P0348	2	0.166	09P0916	3	0.428	14P0760	2	0.292	20P0978	3	0.412
04P0461	2	0.375	09P1016	2	0.235	14P0811	2	0.296	20P1079	2	0.132
04P0678	2	0.026	09P1033	2	0.146	15P0363	2	0.067	20P1095	2	0.165
04P0687	2	0.013	09P1183	2	0.110	15P0365	2	0.118	20P1139	2	0.232
04P0979	2	0.319	09P1202	2	0.273	15P0551	2	0.370	Min	2	0.007
04P1037	2	0.056	09P1208	2	0.273	15P0554	2	0.038	Max	4	0.611
05P0332	2	0.286	09P1211	3	0.533	15P0642	2	0.147	Mean	2.1	0.239

标记命名方式为染色体序号 + P + 标记编号,与 Wang 等^[14]对应。AN 表示等位变异数;PIC 表示多态性信息含量指数。

The marker name is showed as chromosome number + P + marker code followed Wang^[14]. AN: Allele number; PIC: Polymorphic information content.

221 个标记中,74 个位于注释基因序列区段内,在供试材料中共发现 156 个等位变异,平均每个标记有 2.11 个,标记多态性信息含量指数平均为 0.250,变幅为 0.007~0.533(表 3)。147 个位于非基因区段的标记,共发现 307 个等位变异,平均每个标记有 2.09 个,标记 PIC 值平均为 0.234,变幅为

0.013~0.611。总的看,基因内的标记数比基因间少。基因内 PAV 标记的平均等位变异数和多态性信息含量指数与非基因区相当,丰富度和 PIC 值最高的标记 06P0352 位于非基因区段,两类标记均存在一批多态性高的标记可供选用(表 2)。

表 3 大豆基因内与基因间 PAV 标记多态性信息

Table 3 Polymorphism of the PAV markers located in gene or non-coding sequence regions

标记类型	标记数	等位变异数 Allele number		多态性信息含量指数 PIC value	
		平均 Average	变幅 Range	平均 Average	变幅 Variation
基因内 Gene-based PAV	74	2.11	2~4	0.250	0.007~0.533
基因间 Non-coding region PAV	147	2.09	2~4	0.234	0.013~0.611

2.2 供试杂交组合及材料间的遗传关系

2.2.1 杂交组合基于系谱信息的聚类分析 从图 1a 可见,34 个组合在 0.158 处被分为 5 类,第 I 类包括 24 个组合,可分 2 个亚类,第 1 亚类包括 12 个含南农 86-4(P3)核心亲本的组合以及含南农 88-48(P4)和诱处 4 号(P14)的组合各 1 个(E6、Y5),即该类以南农 86-4 血缘为主。其中 Y5 与 N3、N11 和 N8 聚为一小类,原因是该三交组合(Sp//NN493-1/YC4)的亲本中 Sp 是以南农 86-4 为回交亲本与国外引种材料 D76-1609 回交 2 次育成的 BC₂F₄选系,南农 86-4 的血缘理论上占 0.375。E6(NN88-48/Sp)与 N4(NN86-4/Sp)两个则含有共同父本 Sp。第 2 亚类 10 个组合中有 7 个含有核心亲本南农 88-48,另 3 个含诱处 4 号的组合 Y1(NN493-1/YC4)、Y2(NN88-31/YC4//NN73-935)和 Y4(NIL91-1N/YC4)归在一起,系谱分析发现各组合中南农 493-1 的血缘分别为占 0.5、0.22 和 0.19,均占较高的比

例,故与南农 88-48(南农 493-1 的血缘占 0.25)为亲本的组合聚在一起。三交组合 Y3(You 96-3//YC4/NN88-31)被单独归为第 II 类。第 III 类的 2 个组合 E10(N//XSD/NN88-48)和 Y8(XSD/YC4)都有亲本香水豆。第 IV 类的 2 个组合 E8(NN88-48/NXHD//sb)和 Y6(Chuxiu//NXHD/YC4)都有亲本南雄黄豆。第 V 类的 5 个组合都含核心亲本诱处 4 号。总的看,34 个组合根据系谱关系大致可归为与 3 个核心亲本对应的类中。南农 86-4 与南农 88-48 因含有共同祖先亲本奉贤穗稻黄,二者关系较近。符合程度与 3 个核心亲本的遗传背景有关,含南农 86-4 的组合基本集中在一类,含诱处 4 号的组合较为分散。

2.2.2 杂交组合基于 PAV 标记的聚类分析 使用 PowerMarker V 3.25 软件计算 34 个组合间的遗传距离,变幅为 0.035~0.375,平均为 0.184。从图 1b 可见,在 0.052 4 处被分为 6 类,第 I~III 类分别只

有 1 个组合(Y4、E1 和 Y7),这 3 个组合除核心亲本外的另一个亲本的系谱都无法追溯。第Ⅳ~Ⅵ类分别与南农 88-48、南农 86-4 和诱处 4 号为亲本的组合对应,3 个核心亲本也分别位于相应类中。第Ⅳ类 12 个组合中,8 个含有核心亲本南农 88-48,其祖先亲本包括奉贤穗稻黄、南农 493-1、SRF400。该类中 N2、N3、N11、Y5 四个组合中(表 1),N 的 3 个组合的核心亲本南农 86-4 的原始亲本为奉贤穗稻黄,而且 N2 的亲本 NH5 为美国品种,N3 的亲本 NT-1 也含有国外引种材料 PI227224 血缘,Y5 的亲

本则包括南农 493-1。即这 4 个组合与南农 88-48 拥有个别相同的祖先亲本,衍生品系的血缘;第Ⅴ类 8 个都为 N 编号组合,即亲本都有核心亲本南农 86-4;第Ⅵ类 11 个组合的亲本中都有诱处 4 号。

与基于系谱信息聚类结果相比,利用 PAV 标记可将核心亲本南农 86-4、南农 88-48 和诱处 4 号分成相对独立的 3 大类。南农 86-4 和南农 88-48 的祖先亲本中都有奉贤穗稻黄,含这两个亲本的组合根据系谱关系被聚为一类(图 1a),而利用 PAV 标记则能区分出二者的差异。

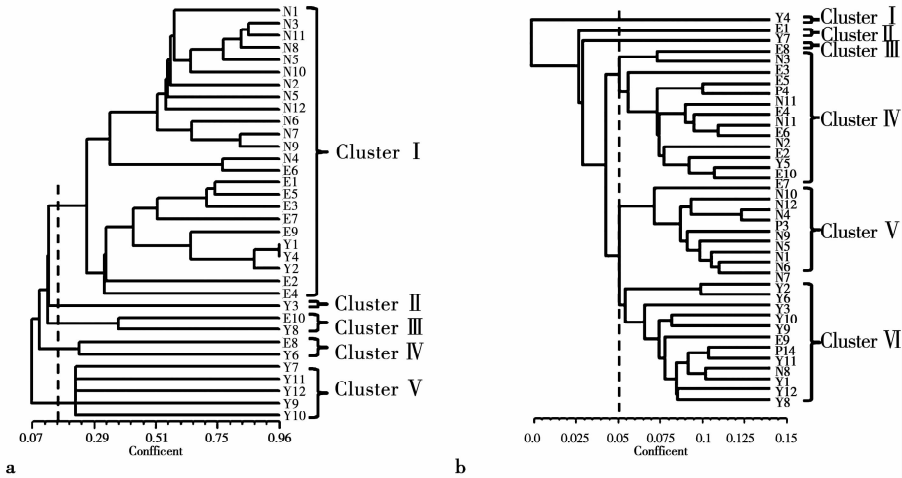


图 1 34 个组合按系谱(a)和 PAV 标记数据(b)所得的聚类树

Fig. 1 The cluster results of the 34 crosses based on the data of pedigree (a) and PAV marker (b)

2.2.3 亲本及品系间基于 PAV 标记的亲缘关系

供试 154 份大豆材料间的 Nei 遗传距离在 0~0.442 之间,平均为 0.264。所有材料可分为 9 类(表 1),第Ⅰ、Ⅱ、Ⅲ、Ⅳ和Ⅶ类材料较少,第Ⅴ类有 10 个材料,第Ⅵ类有 11 个材料,第Ⅷ类有 30 个材料,第Ⅸ类有 89 个材料。总的看,同一组合的材料多聚在一起。第Ⅷ类包括 1 个亲本和 29 个品系,E 编号的品系有 24 个。第Ⅸ类包括 14 个亲本和 75 个品系,N 编号的品系有 44 个,Y 编号的有 23 个,可大致分为 2 个亚群。表明由核心亲本南农 88-48、南农 86-4、诱处 4 号衍生的大部分品系可按各自核心亲本归在一起,后两个亲本关系较近。此外,不同类群间也出现系谱相同或相似的材料并没有聚为一组的现象,如 N3 组合、E6 组合等,这可能与遗传背景有关,反映了杂交重组后的不同品系间遗传差异。

遗传距离值的大小可大致反映材料间亲缘关系的远近。28 个亲本中,16 个南方亲本与 126 个品系间的平均遗传距离为 0.266,而黄淮和国外引种的 5 和 7 个亲本与所有品系间的遗传距离平均值分别为 0.285 和 0.310,反映了不同地理来源亲本对后代品系的贡献存在差异,本地亲本作用更大。3

个核心亲本南农 86-4、南农 88-48 和诱处 4 号在含有各自血缘的亚群中的平均遗传距离分别为 0.15、0.19 和 0.18,均小于其他亲本相应值。所有品系平均遗传距离最小的 5 个亲本为南农 493-1(0.21)、南农 86-4(0.22)、南农 1138-2、香水豆、南农 73-935(均为 0.23);最大的 5 个是南雄黄豆、汉川八月爆、T173、AGH、N5545(均为 0.33),主要是异地的地方品种。

2.3 育种核心亲本对衍生品系的遗传贡献

进一步计算核心亲本中 PAV 标记特有等位变异在单交组合后代品系中的遗传贡献率。从表 4 可见,核心亲本特有等位变异的平均遗传贡献率在 16 个组合均高于 50%,占 69.6%。9 个 N 类型组合中,南农 86-4 特有等位变异的平均遗传贡献率高于 50% 的有 7 个,最高的 N9 组合中,核心亲本南农 86-4 的贡献率达 76%。低于 50% 的 N3 和 N6 分别为 41.29%、48.91%,但其中也有品系高于 50%。7 个 E 类型组合中,南农 88-48 对后代品系的平均遗传贡献率均高于 50%,最大值为 76.54%(E4),最小值为 56.01%(E9)。在所有组合中,只有 E6 中有 1 个品系(44.64%)低于 50%。7 个 Y 类型组合

中,核心亲本诱处 4 号只在 Y1 (70. 95%) 和 Y12 (69. 76%) 的平均贡献率高于 50% ,且都作为父本。全部三交组合核心亲本 PAV 标记等位变异的平均贡献率都超过了理论值(表 4),只有 N10 中有 1 个品系(39. 47%) 低于理论值 43. 75% 。总的看,在 29 个组合中,有 22 个组合(占 75. 9%) 的核心亲本平均贡献率高于理论值。

表 4 组合中核心亲本 PAV 标记等位变异对衍生品系的遗传贡献
Table 4 Contribution rate of core parents to the derived lines based on PAV markers

组合 Cross	家系数 Number of line	多态标记数 Marker number	核心亲本标记等位变异贡献率 Contribution rate of core parent / %			
			理论值 Expected value	实际值 Observed values		平均数 Average
N1 [#]	5	64	50. 00	43. 84, 58. 67, 70. 15, 71. 11, 74. 19		63. 59 ± 12. 52
N2	4	83	50. 00	42. 86, 52. 94, 54. 43, 62. 67		53. 22 ± 8. 13
N3	3	75	50. 00	32. 88, 35. 14, 55. 84		41. 29 ± 12. 66
N5	2	85	50. 00	52. 27, 64. 20		58. 24 ± 8. 43
N6	2	46	50. 00	43. 48, 54. 35		48. 91 ± 7. 69
N8	6	72	50. 00	37. 14, 51. 56, 53. 33, 56. 58, 59. 46, 63. 51		53. 60 ± 9. 13
N9 [#]	7	70	50. 00	67. 65, 68. 12, 68. 57, 81. 43, 81. 69, 81. 69, 82. 86		76. 00 ± 7. 40
N11 [#]	3	65	50. 00	50. 00, 58. 67, 66. 67		58. 44 ± 8. 34
N12 [#]	3	57	50. 00	56. 14, 59. 65, 73. 21		63. 00 ± 9. 02
E2	3	66	50. 00	51. 47, 56. 72, 60. 94		56. 37 ± 4. 74
E3	4	71	50. 00	57. 14, 57. 97, 59. 15, 61. 11		58. 84 ± 1. 72
E4	6	64	50. 00	60. 32, 71. 93, 81. 54, 81. 82, 81. 82, 81. 82		76. 54 ± 8. 87
E5	3	70	50. 00	61. 11, 62. 50, 63. 08		62. 23 ± 1. 01
E6	4	57	50. 00	44. 64, 80. 85, 80. 95, 80. 95		71. 85 ± 18. 14
E7	4	85	50. 00	55. 95, 67. 47, 68. 24, 68. 60		65. 07 ± 6. 09
E9	2	74	50. 00	54. 22, 57. 81		56. 01 ± 2. 54
Y1 [#]	6	77	50. 00	60. 26, 61. 84, 72. 73, 76. 62, 76. 62, 77. 63		70. 95 ± 7. 87
Y7 [#]	4	82	50. 00	45. 78, 48. 75, 48. 78, 50. 62		48. 48 ± 2. 00
Y8 [#]	2	44	50. 00	30. 43, 38. 10		34. 27 ± 5. 42
Y9	3	77	50. 00	42. 47, 50. 00, 53. 85		48. 77 ± 5. 79
Y10	2	68	50. 00	35. 82, 47. 06		41. 44 ± 7. 95
Y11	4	42	50. 00	41. 67, 43. 48, 56. 10, 58. 14		49. 85 ± 8. 47
Y12 [#]	4	65	50. 00	65. 63, 70. 69, 72. 97		69. 76 ± 3. 76
N7 [§]	7	28	25. 00	50. 00, 50. 00, 56. 25, 58. 33, 59. 26, 65. 52, 70. 37		58. 53 ± 7. 53
N10 [§]	3	43	43. 75	39. 47, 45. 65, 50. 00		45. 04 ± 5. 29
Y2 [§]	2	29	25. 00	25. 64, 26. 32		25. 98 ± 0. 48
Y3 [§]	2	27	25. 00	60. 71, 65. 38		63. 05 ± 3. 30
Y5 [§]	3	42	25. 00	33. 33, 55. 00, 58. 14		48. 82 ± 13. 51
Y6 [§]	3	31	25. 00	46. 88, 46. 88, 51. 72		48. 49 ± 2. 80

[#]表示该组合中核心亲本作为父本; [§]指三交组合。
The cross with the [#] represents the core parent was used as male parent; [§] indicates a three-way cross.

单交组合,核心亲本标记等位变异平均贡献率高于 50% 的组合中,只有 N1、N2、N8 和 E6 组合中各存在 1 个品系低于 50%; 低于 50% 的组合中(N3、N6、Y7、Y8、Y9、Y10 和 Y11), N3、N6、Y7、Y9 和 Y11 各自有品系高于 50%。含核心亲本南农 86-4(N 类型)、南农 88-48(E 类型) 衍生的 61 个品系中,有 54 个(占 88. 5%) 品系基于 PAV 标记的核心亲本血缘均超过 50% 的理论值。衍生品系的遗传组成与核心亲本有更高的相似性,而不是双亲的平均值。在三交组合中,情况同样如此。

3 结论与讨论

大量的全基因组测序研究使人们对植物基因组结构的了解不断深入,分子标记开发工作有了较大进展。基因组段结构变异(structure variation) 作为一种特殊的变异类型受到关注。较长 DNA 片段序列变异(几百到几千 bp) 包括 DNA 片段缺失、重复、拷贝数变异(copy number variations, CNVs) 以及插入、倒位。与 SNP 比较,结构变异虽然发生的频率较低,在基因组内也是广泛存在的^[27]。大豆中有

大量 PAV 结构变异标记^[28]。Wang 等^[14]在 33 份野生和栽培大豆中检测覆盖全基因组的 33 127 个 PAV 标记,为进一步利用奠定基础。Li 等^[29]利用 99 个 SSR 标记和 554 个 SNP 分析中国 303 个野生栽培大豆,发现 SSR 标记的平均等位变异数为 21.5 个,Nei's 基因多样性指数为 0.77,而 SNP 的遗传多样性指数只有 0.35。本研究中 154 个育成品系及其亲本基于 221 个 PAV 标记的遗传多样性指数为 0.29。虽然该标记的多态性较低^[30],但该标记带型清晰、容易分辨。本文利用 PAV 分子标记能较好地揭示了这些亲本之间的相互关系,进一步证实该标记可用于种质鉴定,其中也有一批多样性指数高的标记,如 PIC>0.35 的有 56 个标记,这些可作为遗传多样性等研究的优先选用的标记(表 2)。

大豆为严格自花授粉作物,育成品种通常有明确的系谱信息。基于系谱来研究大豆品种的遗传关系已被育种家普遍采用^[19,31-32]。但系谱分析只能将来自相同组合的品系聚为一类,PAV 标记可进一步检测组合内不同品系间的遗传变异。从 34 个组合聚类分析结果来看,按 PAV 标记数据的聚类结果和系谱聚类结果有很大部分是相同的。另一方面,由于育种过程中的差异选择,同一组合的不同品系间会存在差异,有部分材料在两种方法的聚类中不一致。综合分子标记和系谱信息能更准确研究品系间的遗传差异^[33],这些信息可用于育种实践。本研究中南农 86-4、南农 88-48 为南方夏大豆最主要核心亲本南农 1138-2 和南农 493-1 衍生品种,是本地品种最主要血缘,而诱处 4 号则可追溯到个黄淮和东北的核心祖先亲本。以诱处 4 号作为外来血缘组配的 Y5 组合 (Sp//NN493-1/YC4) 后代表现突出,从中选育出的南农 34 已通过江苏省和江西省审定。

在当今作物育种过程中,育种家不断地对作物目标性状进行选择,致使后代传承不同亲本血缘的份额出现差异。核心亲本可能携带更多控制重要农艺性状的基因,在育种过程中经历了较强人为选择压力,从而造成更多选择牵连效应,保留传递给后代,导致了这种现象的发生,其中许多优异等位变异有待于进一步发掘。Shoemaker 等^[34]认为植物育种实际上反映了育种家对农艺性状表型的“操纵”,高产、抗病和优质等性状一直受到强烈选择。Christopher 等^[35]采用 SSR 分子标记分析了澳大利亚小麦品种“Cook”和大麦品种“Triumph”分别在衍生品种的遗传,发现一些高频率选择位点与重要农艺性状相关。找到发挥重要作用的关键基因组区域将有助于阐明核心亲本起作用的遗传本质^[36]。

本研究也发现一些核心亲本的高频率传递标记所在染色体区间存在许多农艺品质性状的 QTL,如 06P1426 两侧各 10 cM 区间内,与产量及百粒重、生育期、株高和倒伏性、蛋白质和优质含量等相关的 QTL 在 Soybase 的 QTL 数据库已报道有 65 个(次);同样,标记 07P0389、09P0916 两侧分别报道有 22 和 24 个;15P0551 和 18P0397 两侧各有 9 个 QTL。

221 个 PAV 标记在供试 154 份大豆材料的平均等位变异数为 2.1 个,多态性信息含量指数 (PIC) 平均为 0.239。34 个组合按 PAV 标记信息的聚类结果与基于系谱的归类相对一致,前者可进一步反映同一组合中不同品系间的遗传差异。对单交和三交组合 PAV 多态标记分析表明核心亲本对衍生品系的遗传贡献高于其他亲本,来自本生态区的亲本与所有品系的遗传距离最小。PAV 标记可以有效揭示供试材料的遗传多样性,相较于系谱信息,其能更好地揭示核心亲本血缘在后代中的传递特点。

参考文献

[1] Collard B C Y, Jahufer M Z Z, Brouwer J B, et al. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts [J]. *Euphytica*, 2005, 142(1-2): 169-196.

[2] Jones N, Ougham H, Thomas H, et al. Markers and mapping revisited: finding your gene[J]. *New Phytologist*, 2009, 183(4): 935-966.

[3] Xu Y, Lu Y, Xie C, et al. Whole-genome strategies for marker-assisted plant breeding[J]. *Molecular Breeding*, 2012, 29(4): 833-854.

[4] Springer N M, Ying K, Fu Y, et al. Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content [J]. *PLoS Genetics*, 2009, 5(11): e1000734.

[5] Gao Q, Yue G, Li W, et al. Recent progress using high-throughput sequencing technologies in plant molecular breeding [J]. *Journal of Integrative Plant Biology*, 2012, 54(4): 215-227.

[6] Iafrate A J, Feuk L, Rivera M N, et al. Detection of large-scale variation in the human genome [J]. *Nature Genetics*, 2004, 36(9): 949-951.

[7] Kern A D, Begun D J. Recurrent deletion and gene presence/absence polymorphism: telomere dynamics dominate evolution at the tip of 3L in *Drosophila melanogaster* and *D. simulans* [J]. *Genetics*, 2008, 179(2): 1021-1027.

[8] Arrach N, Porwollik S, Cheng P, et al. Salmonella serovar identification using PCR-based detection of gene presence and absence [J]. *Journal of Clinical Microbiology*, 2008, 46(8): 2581-2589.

[9] Batley J, Barker G, O'Sullivan H, et al. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data [J]. *Plant Physiology*, 2003, 132(1): 84-91.

- [10] García-Lor A, Luro F, Navarro L, et al. Comparative use of InDel and SSR markers in deciphering the interspecific structure of cultivated citrus genetic diversity: A perspective for genetic association studies[J]. *Molecular Genetics and Genomics*, 2012, 287(1): 77-94.
- [11] Ollitrault F, Terol J, Martin A A, et al. Development of InDel markers from *Citrus clementina* (Rutaceae) BAC-end sequences and interspecific transferability in *Citrus*[J]. *American Journal of Botany*, 2012, 99(7): e268-e273.
- [12] Liu B, Wang Y, Zhai W, et al. Development of In Del markers for *Brassica rapa* based on whole-genome re-sequencing[J]. *Theoretical and Applied Genetics*, 2013, 126(1): 231-239.
- [13] Zhang L M, Luo H, Liu Z Q, et al. Genome-wide patterns of large-size presence/absence variants in sorghum[J]. *Journal of Integrative Plant Biology*, 2014, 56(1): 24-37.
- [14] Wang Y, Lu J, Chen S, et al. Exploration of presence/absence variation and corresponding polymorphic markers in soybean genome[J]. *Journal of Integrative Plant Biology*, 2014, 56(10): 1009-1019.
- [15] 庄巧生. 中国小麦品种改良及系谱分析[M]. 北京: 中国农业出版社, 2003. (Zhuang Q S. Chinese wheat improvement and pedigree analysis[M]. Beijing: China Agriculture Press, 2003.)
- [16] Carter T E Jr, Nelson R L, Sneller C H, et al. Genetic diversity in soybean[M]// Boerma H R, Specht J E. Soybean: Improvement, production and uses. 3rd ed. Agronomy Monograph. Madison: ASA and CSSA, 2004: 303-416.
- [17] 熊冬金, 赵团结, 盖钧镒. 中国大豆育成品种亲本分析[J]. *中国农业科学*, 2008, 41(9): 2589-2598. (Xiong D J, Zhao T J, Gai J Y. Parental analysis of soybean cultivars released in China[J]. *Scientia Agricultura Sinica*, 2008, 41(9): 2589-2598.)
- [18] 关荣霞, 张磊, 刘章雄, 等. DNA 导入和系选大豆品种及其亲本遗传关系的 SSR 标记分析[J]. *植物学报*, 2010, 45(1): 36-43. (Guan R X, Zhang L, Liu Z X, et al. Relationship analysis of soybean cultivars bred by intracultivar selection or pollen tube pathway and parents by simple sequence repeat markers[J]. *Chinese Bulletin of Botany*, 2010, 45(1): 36-43.)
- [19] 崔章林, 盖钧镒, Carter T E J, 等. 中国大豆育成品种及其系谱分析(1923-1995)[M]. 北京: 中国农业出版社, 1998. (Cui Z L, Gai J Y, Carter T E Jr, et al. The released Chinese soybean cultivars and their pedigree analysis (1923-1995)[M]. Beijing: China Agriculture Press, 1998.)
- [20] 刘佑斌, 盖钧镒, 游明安. 高产早熟夏大豆新品种南农 88-48 的选育研究[J]. *南京农业大学学报*, 1995, 18(3): 1-6. (Liu Y B, Gai J Y, You M A. Development of soybean cultivar with high yield and early maturity: Nannong 88-48[J]. *Journal of Nanjing Agricultural University*, 1995, 18(3): 1-6.)
- [21] 张性坦, 赵存, 柏惠侠, 等. 夏大豆诱处 4 号公顷产 4500 kg 生理指标研究[J]. *中国农业科学*, 1996, 29(6): 46-54. (Zhang X T, Zhao C, Bai H X, et al. Study on physiological indexes of summer soybean variety Youchu 4 yielding 4500 kg·ha⁻¹ [J]. *Scientia Agricultura Sinica*, 1996, 29(6): 46-54.)
- [22] Doyle J J. Isolation of plant DNA from fresh tissue[J]. *Focus*, 1990, 12: 13-15.
- [23] Liu K J, Muse S V. PowerMarker: An integrated analysis environment for genetic marker analysis[J]. *Bioinformatics*, 2005, 21(9): 2128-2129.
- [24] Botstein D, White R L, Skolnick M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms[J]. *American Journal of Human Genetics*, 1980, 32(3): 314.
- [25] Nei M, Tajima F, Tateno Y. Accuracy of estimated phylogenetic trees from molecular data[J]. *Journal of Molecular Evolution*, 1983, 19(2): 153-170.
- [26] Rohlf F J. NTSYS-PC: Numerical taxonomy and multivariate analysis system[G]. New York: Exeter Software, 1992.
- [27] Saxena R K, Edwards D, Varshney R K. Structural variations in plant genomes[J]. *Briefings in Functional Genomics*, 2014, 13(4): 296-307.
- [28] Lam H M, Xu X, Liu X, et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection[J]. *Nature Genetics*, 2010, 42(12): 1053-1059.
- [29] Li Y H, Li W, Zhang C, et al. Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci [J]. *New Phytologist*, 2010, 188(1): 242-253.
- [30] 冯芳君, 罗利军, 李英, 等. 水稻 InDel 和 SSR 标记多态性的比较研究[J]. *分子植物育种*, 2005, 3(5): 725-730. (Feng F J, Luo L J, Li Y, et al. Comparative analysis of polymorphism of InDel and SSR markers in rice[J]. *Molecular Plant Breeding*, 2005, 3(5): 725-730.)
- [31] Bernard R L, Juvik G A, Hartwig E E, et al. Origins and pedigrees of public soybean varieties in the United States and Canada [R]. US Department of Agriculture, Technical Bulletin. 1988, No. 1764.
- [32] Carter T E J, Gizlice Z, Burton J W. Coefficient-of-parentage and genetic similarity estimates for 258 north America soybean cultivars by public agencies during 1945-1988[R]. US Department of Agriculture, Technical Bulletin, 1993, No. 1814.
- [33] 赵战胜, 喻树迅, 范术丽, 等. 北疆早熟陆地棉品种的遗传多样性分析[J]. *棉花学报*, 2012, 24(6): 473-480. (Zhao Z S, Yu S X, Fan S L, et al. Analysis of genetic diversity of early maturing upland cotton varieties in northern Xinjiang[J]. *Cotton Science*, 2012, 24(6): 473-480.)
- [34] Shoemaker R C, Guffy R D, Lorenzen L L, et al. Molecular genetic mapping of soybean: Map utilization[J]. *Crop Science*, 1992, 32(5): 1091-1098.
- [35] Christopher M, Mace E, Jordan D, et al. Applications of pedigree-based genome mapping in wheat and barley breeding programs [J]. *Euphytica*, 2007, 154(3): 307-316.
- [36] 李小军, 徐鑫, 刘伟华, 等. 利用 SSR 标记探讨骨干亲本欧柔在衍生品种的遗传[J]. *中国农业科学*, 2009, 42(10): 3397-3404. (Li X J, Xu X, Liu W H, et al. Genetic diversity of the founder parent Orofen and its progenies revealed by SSR markers [J]. *Scientia Agricultura Sinica*, 2009, 42(10): 3397-3404.)