

## 华南野生大豆 *GmPAP1* 基因多样性分析

程春明<sup>1,2</sup>, 杨存义<sup>1</sup>, 马启彬<sup>1</sup>, 年海<sup>1</sup>

(1. 华南农业大学 农学院, 广东 广州 510642; 2. 江西省农业科学院 作物研究所, 江西 南昌 330200)

**摘要:** *GmPAP1* 属低磷诱导基因, 通过同源克隆测序, 分析 68 份来源于华南亚热带地区江西、湖南、福建、广西等地野生大豆 *GmPAP1* 基因的多样性。与原基因序列 *GmPAP1* (AF236108) 比较, 68 份种质资源共有 22 个 SNPs 和 6 个 Indels, 22 个 SNPs 中有 13 处发生转换, 9 处发生颠换, 转换所占比例接近 2/3, 有 14 个发生了非同义突变, 表明该基因在不同地区的大豆中具有较高的多样性。

**关键词:** 野生大豆; *GmPAP1*; 多样性

中图分类号: S565.1

文献标识码: A

文章编号: 1000-9841(2010)06-0920-05

## Diversity Analysis of *GmPAP1* of Wild Soybean in South China

CHENG Chun-ming<sup>1,2</sup>, YANG Cun-yi<sup>1</sup>, MA Qi-bin<sup>1</sup>, NIAN Hai<sup>1</sup>

(1. College of Agriculture, South China Agricultural University, Guangzhou 510642, Guangdong; 2. Crop Institute of Jiangxi Academy of Agricultural Sciences, Nanchang 330200, Jiangxi, China)

**Abstract:** *GmPAP1* is a gene induced by low-phosphorus condition, the sequence polymorphisms of *GmPAP1* in comparison with the control gene sequence (AF236108) from Genbank were analyzed by cloning in 68 wild soybean genotypes mainly from the subtropical zone: Jiangxi, Hunan, Fujian province and Guangxi Autonomous Region in South China. The results showed that, 22 SNPs and 6 Indels were found. Among these polymorphic sites, 13 transition, 9 transversion and 14 non-synonymous mutations were detected, indicating the abundance of gene diversity among these genotypes.

**Key words:** Wild soybean; *GmPAP1*; Diversity

SNPs (Single Nucleotide Polymorphisms, 单核苷酸多态性) 是指基因组 DNA 序列中由于单个核苷酸 (A, T, C, G) 的替换 (转换和颠换) 而引起的多态性。这种新型分子标记是继限制性片段长度多态性 (RFLP) 和微卫星标记 (SSR) 之后最有前途的第三代分子标记。SNPs 是基因特异性标记, 对基因组或功能基因组的测序分析结果可以提供精确的等位基因突变信息, 如果突变发生在编码区并且改变了氨基酸序列, 那么这种 SNPs 就可能同表型相关<sup>[1]</sup>。从高度稳定的 SNPs 中易于发掘出临近目标位点的高密度标记<sup>[2]</sup>。发掘 SNPs 最简捷、最彻底的方法是 DNA 直接测序法<sup>[3]</sup>。

缺磷诱导根系分泌酸性磷酸酶 (APA) 是高等植物适应低磷胁迫的一个普遍机制。低磷条件能增加细胞外磷酸酶的分泌和细胞内磷酸酶的合成, 因为细胞外分泌磷酸酶能促进有机磷的活化, 提高土壤磷素的有效性, 细胞内合成磷酸酶有利于体内

有机磷与无机磷的相互转化<sup>[4]</sup>。紫色酸性磷酸酶 (PAP) 是酸性磷酸酶的一种, 是一类金属水解酶, 植物受低磷胁迫诱导时一些 PAP 会大量表达, 并分泌到根际土壤中水解根际土壤中的磷脂化合物释放出 Pi 供植物利用<sup>[5]</sup>。Schenk 等<sup>[6]</sup> 基于 EST 序列, 已经从大豆中克隆出了一些重要的 *GmPAP1* 基因的 cDNA 序列, 通过验证, 属于紫色酸性磷酸酶基因家族。Liao 等<sup>[7]</sup> 通过低磷诱导证实了上述基因与低磷相关, 在发芽 7 d 后进行低磷诱导, 12 d 后提取 RNA, 经 Northern blot 杂交分析认为 *GmPAP1* 属低磷诱导基因, 在磷缺乏时主要在叶部和根部表达。

研究通过对 68 份来自华南不同省份的野生大豆的 *GmPAP1* 基因的同源克隆, 通过测序, 比较分析并明确出该基因多样性 (SNPs) 变化情况, 以期开发与磷效率性状有关的 SNPs 标记, 应用于大豆磷效率分子标记辅助育种。

收稿日期: 2010-07-23

基金项目: 国家自然科学基金资助项目 (30771364)。

第一作者简介: 程春明 (1972-), 男, 副研究员, 博士, 从事大豆遗传育种研究。E-mail: ccmccmcc@126.com。

通讯作者: 年海, 教授, 博士生导师。E-mail: hnian@scau.edu.cn。

## 1 材料与方 法

### 1.1 供试材料

选择来源于华南 68 份种质进行基因同源克隆。其中江西 20 份,湖南 18 份,福建 23 份,广西 7 份(表 1)。

### 1.2 试验方法

1.2.1 引物设计 以 NCBI 中的 *GmPAP1* (AF236108)序列为模板,用 Primer 5.0 在候选基因两端设计引物。引物序列为:

Forward Primer: 5'-ATGGGCACTCAGAGAAGCAA-3'

Reverse Primer: 5'-TTACATGTCACCTGGAGTCAA-3'

1.2.2 低磷诱导 播种:选取较为饱满的野生大豆种子,每个品种 10 粒左右,用 10% 的双氧水消毒,随后用蒸馏水冲洗干净,然后用小刀刮破种皮。分批种在盛有石英砂的盆中,用 1/2 浓度的 Hoagland 营养液把砂均匀湿润,每个品种种 1 行,放置人工培养箱中,培养箱温度控制在 28℃ 左右。生长过程中,同时喷洒同样的营养液保持沙面湿润至移苗。移苗诱导:待真叶完全展开后,移至盛有低磷的 1/2 浓度的 Hoagland 营养液 ( $0.2 \mu\text{mol} \cdot \text{L}^{-1} \text{KH}_2\text{PO}_4$ ) 的培养箱中,培养 7 d 后,采用 Trizol 法提取 RNA。

表 1 试验材料编号和来源

Table 1 Accession number and origin of experiment materials

编号 Accession number	来源 Origin						
JW9	江西九江	JW189	江西会昌	W363	湖南衡山	BW53	福建沙县
JW34	江西都昌	JW197	江西南昌	W364	湖南衡阳	BW55	福建明溪
JW48	江西武宁	JW200	江西进贤	W367	湖南绥宁	BW60	福建尤溪
JW53	江西武宁	W231	湖南新田	W375	湖南郴州	BW61	福建清流
JW66	江西铜鼓	W303	湖南湘潭	BW17	福建浦城	BW64	福建永安
JW73	江西宜丰	W317	湖南新田	BW20	福建崇安	BW66	福建永安
JW80	江西万载	W334	湖南华阴	BW23	福建光泽	BW69	福建连城
JW101	江西安义	W342	湖南常德	BW26	福建松溪	BW72	福建上杭
JW108	江西新建	W343	湖南常德	BW28	福建建阳	BW73	福建霞浦
JW120	江西乐平	W345	湖南石门	BW33	福建邵武	BW74	福建建瓯
JW135	江西婺源	W346	湖南溪口	BW35	福建建瓯	BW81	广西兴安
JW141	江西德兴	W349	湖南龙山	BW38	福建泰宁	BW84	广西灌阳
JW159	江西鹰潭	W350	湖南湘阴	BW42	福建建宁	BW94	广西贺县
JW168	江西崇仁	W355	湖南凤皇	BW45	福建顺昌	BW99	广西融安
JW174	江西宜黄	W359	湖南浏阳	BW47	福建将乐	BW101	广西富川
JW181	江西南丰	W361	湖南黔阳	BW48	福建将乐	BW105	广西昭平
JW186	江西南都	W362	湖南黔阳	BW50	福建龙岩	BW106	广西象州

1.2.3 RT-PCR RT:1  $\mu\text{L}$  RNA,1  $\mu\text{L}$   $10 \mu\text{mol} \cdot \text{L}^{-1}$  引物 Olig(dT),3  $\mu\text{L}$  无 RNase 的水,65℃ 变性 3 min 后冰浴 3 min,依次加入 1  $\mu\text{L}$   $0.01 \text{mmol} \cdot \text{L}^{-1}$  的 DDT,2  $\mu\text{L}$   $1 \times$  第一链合成缓冲液,1  $\mu\text{L}$   $10 \mu\text{mol} \cdot \text{L}^{-1}$  dNTP,1  $\mu\text{L}$  200 U 的反转录酶,0.5  $\mu\text{L}$  Ribonu-

lease Inhibitor,置于 42℃ 水浴反应 45 ~ 75 min,70℃ 放置 3 min,终止反应。

PCR 反应体系:5  $\mu\text{L}$   $10 \times$  pyrobest PCR Buffer ( $\text{Mg}^{2+}$  Plus),0.5  $\mu\text{L}$  TakaRa pyrobest Taq (5 U  $\cdot \mu\text{L}^{-1}$ ),8  $\mu\text{L}$   $2.5 \text{mmol} \cdot \text{L}^{-1}$  dNTP Mixture,2  $\mu\text{L}$  cD-

NA, 2  $\mu\text{L}$  10  $\mu\text{mol} \cdot \text{L}^{-1}$  引物, ddH<sub>2</sub>O 补足。PCR 反应程序为: 95 $^{\circ}\text{C}$  4 min, 94 $^{\circ}\text{C}$  30 s, 55 $^{\circ}\text{C}$  30 s, 72 $^{\circ}\text{C}$  1 min, 35 个循环, 72 $^{\circ}\text{C}$  10 min, 4 $^{\circ}\text{C}$  保存。RT-PCR 反应产物直接交给北京奥科生物公司双向测序。

1.2.4 基因序列分析 利用 DNASTAR 5.0 软件中的 EditSeq 和 SeqMan 功能模块对所有材料的测序产物进行编辑和拼接, 采用 Clustal xl 软件进行序列比对分析。利用 DNASTAR5.0 中的 EditSeq 和 MegAlign 功能模块进行翻译和比较。用 TASSEL 2.01 软件中相关功能模块完成基因序列多样性和聚类分析。

### 3 结果与分析

#### 3.1 基因序列多样性分析

基因序列拼接成功后, 将 68 份种质与对照序列进行比较, 经统计(表 2), 在 68 个序列中, 共统计到多态性位点 28 个, 平均每 35 bp 检测到 1 个多态性位点。其中有 22 个 SNPs, 平均每 45 bp 检测到 1 个 SNP, 有 6 个 Indel, 平均每 167.5 bp 检测到 1 个。检测到的 SNP 数量远高于 Indel。在 22 个 SNPs 中, 有 13 处发生转换, 9 处发生颠换, 转换所占比例接近 2/3。另外, 比对结果表明, 所检测到的 SNP 和 Indel 均为单一碱基变化。表明 68 份种质的 *GmPAP1* 基因由于自然突变产生了一定数量碱基转换、颠换及缺失, 基因序列具有较为丰富的多样性。

表 2 *GmPAP1* 基因序列多样性分布  
Table 2 Polymorphisms of gene *GmPAP1* sequence of the 68 wild soybean genotypes

	序列总长度 No. of bp	变异位点数 No. of polymorphic sites	备注 Comments
	999 bp	—	—
SNP	—	22	1 个/45 bp
Indel	—	6	1 个/167.5 bp
total	—	28	1 个/35 bp

另外, 在所检测到的 22 个 SNPs 中, 有 5 个 SNPs 仅仅在参照序列中存在, 在其它材料中变化一致, 均不同于参照序列(表 3)。由于 *GmPAP1* 原基因的克隆来源于大豆, 而所有参试材料均是野生大豆, 大豆是由野生大豆遗传进化而来, 所以在这几个突变位点中, 有可能其原始基因比较一致, 只是由于进化过程中这些位点发生了碱基的改变。

表 3 种质资源与对照序列相比较

Table 3 Gene *GmPAP1* sequence polymorphisms among genotypes and AF236108

	401bp	497bp	572bp	659bp	741bp
AF236108	A	A	C	A	G
68 genotypes	C	T	T	G	A

SNPs 存在于基因组中的各个区域, 根据它们在基因组中分布的位置可分为基因编码区 SNPs (coding-region SNPs, cSNPs) 和非编码区 SNPs (noncoding-region SNPs, ncSNPs) 2 类。由于选择压力的原因, cSNPs 比 ncSNPs 要少。另外根据 cSNPs 对编码蛋白质的影响可以将它分为 2 种: 一种是同义 cSNPs (synonymous cSNPs), 即 SNPs 所致编码序列的改变并不影响其所翻译的氨基酸序列, 突变碱基与未突变碱基具有相同的“含义”, 是由于氨基酸密码子的兼并性而产生的; 另一种是非同义 cSNPs (non-synonymous cSNPs), 即碱基序列的改变将导致编码氨基酸的改变, 从而产生蛋白质序列的改变, 最终可能影响到蛋白质的功能<sup>[8]</sup>。该研究针对的是 cDNA 序列, 全部为 cSNPs, 并通过对所有基因序列翻译成氨基酸序列后进行了比较, 发现在 22 个 SNPs 的变异中, 只有 8 个 SNPs 发生了同义突变, 14 个 SNPs 发生了非同义突变, 非同义突变较同义突变频率高出近一倍。由于大豆是由野生大豆进化而来, 所以应该说由野生大豆进化到大豆过程中, 该基因发生了系列碱基的突变, 其基因功能可能随之有所改变。

#### 3.2 基因序列聚类结果分析

以各材料 *GmPAP1* 的 SNP 和 Indel 变化差异为依据, 利用 TASSEL 2.01 软件构建了 69 个不同基因型的 *GmPAP1* 的聚类图(包括对照)(图 1)。从聚类图中可以看出, 基于基因序列差异的基础上, 将这些材料分成三大类和 9 个小类。三大类中, 第一大类只有 1 份材料, 即 BW69 (福建连城), 从该基因变异情况来看, 该材料与其它材料的变异相对较大; 第三大类也只有 3 份材料, 即 W334 (湖南华阴)、JW200 (江西进贤) 及 W317 (湖南新田) 等; 其余材料基本上集中聚类到第二大类材料中, 第二大类又可细分为 7 个小类, 7 个小类当中, 每个小类所包含的种质材料基本来自不同地区。这与利用 SSR 分子标记所得出的所有种质材料的聚类情况有所差异, 表明该基因受到环境的影响更大。与对照序列进行比较分析来看, 来自广西的野生大豆 BW81 的基因序列与对照基因序列最为接近, 只在 41 bp 及 401 bp 两处发生了碱基的变化。同时与对照序

列聚到一小类的基本为湖南、广西及江西材料,没有福建的材料。



图 1 基于 *GmPAP1* 基因序列的聚类图

Fig. 1 The cladogram tree by gene *GmPAP1* sequence of the 68 wild soybean genotypes

### 3 结论与讨论

单核苷酸多态性 (SNPs, Single Nucleotide Poly-

morphisms) 在基因组中数量多、分布广泛,属二等位基因分子标记,可以稳定遗传的点突变。所以自 1996 年 Lander<sup>[9]</sup> 首次正式提出 SNPs 为新一代分子

标记后,动植物 SNPs 研究迅猛发展,在作物上,SNPs 已广泛应用于高密度遗传图谱构建、性状作图和基因的精确定位、群体遗传结构分析以及系统发育分析等方面的研究<sup>[10]</sup>。

*GmPAP1* 基因在栽培大豆中克隆成功<sup>[3]</sup>,并被证明在磷缺乏时表达量增加<sup>[4]</sup>。68 份野生大豆同源克隆测序分析后,与对照序列相比较,共检测出 28 个 SNPs 和 Indel,28 个 SNPs 全部是单个碱基突变。同时这些点突变与对照比较,有些位点只在对照中才有,也有少数位点在少数野生大豆中才有,这些都表明了 SNPs 具有数量多、单拷贝等优势。

另外,*GmPAP1* 基因还表现出以下特点:SNPs 中有近 2/3 碱基变化属于转换类型,与生物体碱基变化类型相似。从氨基酸序列分析来看,22 个 SNPs 中有近 2/3 发生了非同义突变,与邹洪锋<sup>[8]</sup>的研究结果差别较大,其在野生大豆、半野生大豆及栽培大豆编码区发现的 14 个 SNPs 中,引起同义突变的 SNPs 为 12 个,引起非同义突变的仅为 2 个。并且 *GmPAP1* 基因非同义突变的频率也远高于人类的 1: 3.4<sup>[11]</sup> 和玉米的 1: 4.4<sup>[12]</sup>。表明野生大豆 *GmPAP1* 基因表现出较高的多态性和多样性,种质资源之间基因突变较高,可能受自然环境的作用,逐渐形成了具有调节不同环境的功能基因。

在长期的自然选择作用下,相同环境条件下野生大豆生物学性状具有趋同性。野生大豆遗传差异与不同环境、不同地理来源有紧密联系,不同生态区野生大豆生物学性状具有较大遗传差异<sup>[13-16]</sup>。但基于 *GmPAP1* 基因的 SNPs 变化将 68 份野生大豆分成了三大类 9 个小类,各小类聚类的资源组成与地理来源没有表现出一致的现象。认为基于 *GmPAP1* 基因 SNPs 虽然具有较高的稳定性,但只局限于基因组的某一小区段,并且该基因可能容易受到自然地理环境的影响,突变率较高,所以在该基因序列的变化差异上未能很好地明确野生大豆种质资源之间的亲缘关系及差异。

## 参考文献

- [1] Kim M Y, Ha B K, Jun T H, et al. Single nucleotide polymorphism discovery and Linkage mapping of lipoxygenase-2 gene (Lx2) in soybean[J]. Euphytica, 2004, 135: 169-177.
- [2] Batley J, Barker G, Sullivan H O, et al. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data [J]. Plant Physiology, 2003, 132: 84-91.
- [3] 李英慧. 中国大豆地方品种遗传多样性分析及大豆抗胞囊线虫 SNAP 标记发掘[D]. 北京:中国农业科学院,2005:56-58. (Li Y H. Genetic structure and diversity for chinese soybean (*Glycine max* L.) landraces in China and SNPs marker development at *rhg1* locus underlying resistance to Soybean Cyst Nematode[D]. Beijing: Chinese Academy of Agricultural Sciences, 2005: 56-58.)
- [4] 严小龙,廖红,年海,等. 根系生物学:原理与应用[M]. 北京:科学出版社,2007: 138-140. (Yan X L, Liao H, Nian H, et al. Root biology: Principles and Applications[M]. Beijing: Science Press, 2007: 138-140.)
- [5] 李东屏,王道文. 拟南芥紫色酸性磷酸酶基因(AtPAP)对磷饥饿的响应[J]. 生命科学研究,2003,7(1):65-69. (Li D P, Wang D W. Responses of putative purple acid phosphatase genes in *Arabidopsis thaliana* (AtPAP) to phosphorus starvation[J]. Life Science Research, 2003, 7(1):65-69.)
- [6] Schenk G, Guddat L W, Ge Y, et al. Identification of mammalian-like purple acid phosphatases in a wide range of plants[J]. Gene, 2000, 250: 117-125.
- [7] Liao H, Wonga F L, Phang T H, et al. *GmPAP3*, a novel purple acid phosphatase-like gene in soybean induced by NaCl stress but not phosphorus deficiency [J]. Gene, 2003, 318: 103-111.
- [8] 邹洪锋. 野生、半野生和栽培大豆的单核苷酸多态性[D]. 南昌:南昌大学,2005. (Zhou H F. Single nucleotide polymorphisms in *G. soja*, *G. gracilis* and *G. max*[D]. Nanchang: Nanchang University, 2005.)
- [9] Lander E S. The new genomics: global views of biology [J]. Science, 1996, 274: 536-539.
- [10] 邹喻莘,葛颂. 新一代分子标记—SNPs 及其应用生物多样性[J]. 生物多样性,2003,11(5):370-382. (Zhou Y P, Ge S. A novel molecular marker—SNPs and its application [J]. Biodiversity Science, 2003, 11(5):370-382.)
- [11] Halushka M K, Fan J B, Bentley K, et al. Patterns of single nucleotide polymorphisms in candidate genes for blood-pressure homeostasis[J]. Nature Genetics, 1999, 22(3):239-247.
- [12] Tenaillon M I, Sawkins M C, Long A D, et al. Patterns of DNA sequence polymorphisms along chromosome 1 of maize (*Zea mays* L.) [J]. Proceedings of the National Academy of Sciences, 2001, 98(16):9161-9166.
- [13] 徐豹,徐航,庄炳昌,等. 中国野生大豆(*G. soja*)籽粒性状的遗传多样性及其地理分布[J]. 作物学报,1995,21(6):733-739. (Xu B, Xu H, Zhuang B C, et al. Polymorphism and geographical distribution of seed characters of wild soybean (*G. soja*) in China[J]. Acta Agronomica Sinica, 1995, 21(6):733-739.)
- [14] 董英山,庄炳昌,赵丽梅,等. 中国野生大豆遗传多样性中心[J]. 作物学报,2000,26(5):521-527. (Dong Y S, Zhuang B C, Zhao L M, et al. The Genetic diversity centers of annual wild soybean in China [J]. Acta Agronomica Sinica, 2000, 26(5): 521-527.)
- [15] 丁艳来,赵团结,盖钧镒. 中国野生大豆的遗传多样性和生态特异性分析[J]. 生物多样性,2008,16(2):133-142. (Ding Y L, Zhao T J, Gai J Y. Genetic diversity and ecological differentiation of chinese annual wild soybean (*G. soja*) [J]. Biodiversity Science 2008, 16(2):133-142.)
- [16] Wang K J, Li X H, Li F S. Phenotypic diversity of the big seed type subcollection of wild soybean (*Glycine soja* Sieb. et Zucc.) in China [J]. Genetic Resources and Crop Evolution, 2008, 55: 1335-1346.