

水稻广亲和基因(WCG)近等基因系的随机引物 PCR 分析

戴欣 许莉萍 刘钿连 冯亮 刘良式

(中山大学生命科学院, 广州 510275)

摘要 应用随机引物 PCR (Random Primer PCR) 技术分别在水稻广亲和基因 (WCG) 的近等基因系和色素原基因 (C) 的分离群体库中寻找与 WCG 和 C 基因连锁的分子标记。对于 WCG 近等基因系, 在 226 个随机引物中初筛到 22 个显示多态性片段的引物。根据理论值计算, 在 22 个多态性片段中预期有 20 个与 WCG 连锁, 在这些连锁标记中距 WCG 最近的可达 0.5cM。同样在分离群体库的筛选中有 10 个扩增产物与 C 基因连锁。

关键词 随机引物; 聚合酶链反应; 广亲和基因; 近等基因系; 水稻

RANDOM AMPLIFIED POLYMORPHIC DNA ANALYSIS OF NEAR-ISOGENIC LINES OF WIDE COMPATIBILITY GENE IN RICE

Dai Xin Xu Liping Liu Dianlian Feng Liang Liu Liangshi

(Zhongshan University, School of Life Sciences, Guangzhou 510275)

Abstract There are growing evidences that the fertility of hybrid between indica and japonica varieties of rice (*Oryza sativa* L.) is controlled mainly by a wide compatibility gene (WCG), which is closely linked to the chromogen gene (C gene). Screening new molecular markers nearby WCG and C gene by RAPD will be an effective approach to enrich the present molecular map. Some results obtained from RAPD analysis and bulked segregants of C gene for Near Isogenic Lines (NILs) of WCG are shown. By method of polymerase chain reaction (PCR), 22 polymorphic primers in 226 random primers are found for NILs of WCG, among which 20 polymorphic DNA fragments should be linked with WCG under theoretical calculation, and the expected minimum map distance is 0.5 cent morgan. By using bulked segregants of C gene, 10 polymorphic DNA fragments in 200 random primers are found out.

Key words Random primer; Polymerase chain reaction; Near-isogenic line; Wide compatibility gene; Rice

水稻遗传图的研究经过了一个漫长的历程, 自从四十年代 Jodon, N. E⁽¹⁾ 研究水稻叶绿体突变的连锁关系, Hara, S.⁽²⁾ 研究水稻不育性和花青素因子的连锁关系和五十年代初 Oka, H. J⁽³⁾ 研究水稻品种间杂交不育机理及不育因子和色素基因的连锁关系以来, 基于形态标记和同工酶标记建立了水稻染色体连锁图⁽⁴⁾。迄今为止, 通过经典遗传学分析, 在水稻中发现的基因共 313 个, 在近五十年的时间里, 定位基因 168 个, 已知其所属连锁群但尚未具体定位的基

因有 190 个, 覆盖了 1398 厘摩 (cM) 的水稻基因组^[5]。由此可见, 遗传图的绘制是一项很费时的研究。1988 年, Tanksley^[6]将 RFLP (Restriction fragment length polymorphism) 分析应用于构建水稻基因组图谱, 在最近短短四、五年时间内, 得到了一幅具有 568 个分子标记、覆盖 1380cM 水稻基因组的分子图谱。日本科学家亦发表了密度更高(678 个标记, 1600cM)的 RFLP 图谱^[7], 如果将经典遗传图的基因座位与 RFLP 分子图的分子标记重合, 那就会为我们克隆在分子标记附近的基因带来极大的方便。

近年来, 以 PCR (Polymerase chain reaction) 技术和新的电泳技术为基础的各种新型分子标记进一步促进了高密度遗传连锁图谱的发展。1990 年 Williams 等^[8]发展起来的一种新型分子标记 RAPD (Random amplified polymorphic DNA), 在基因组作图、基因定位, 以及生物进化和分类的研究中取得了令人瞩目的进展^[8,9,10]。近等基因系可用来研究一些遗传标记和未知分子标记的连锁关系。1991 年 Martin 等人^[11]通过比较番茄近等基因系 (B6S1) 的 RAPD 多态性, 分离到 3 个与抗 *Pseudomonas* 基因紧密连锁的 DNA 标记。同年, Kesseili 等人^[12]提出了利用杂交后代 F₂ 分离群体库 (Bulk segregant) 进行基因定位的方法, 并在莠苣 Dm5/8 (棉疫病) 抗性基因的分离群体库中找到了三个与 Dm5/8 抗性基因连锁的 RAPD 标记。本文利用这一技术对水稻广亲和基因 (WCG) 和色素原基因 (C) 进行研究, 以期找到分别与它们紧密连锁的分子标记。

水稻有籼稻、粳稻之分, 两者各具一些优良性状, 通过杂交实现二者之间有利基因的重组或利用子一代的杂种优势一直是水稻育种研究中的重要课题。目前认为, 籼粳杂种的育性涉及 17 个 S 基因, 除 6 个位于第 3、4、7、11 连锁群上之外, 有 11 个位于第 6 连锁群上, 但主要由 S₅ 座位控制, S₅ (广亲和基因)^[13]的存在可部分克服籼粳杂交的半不育现象, 所以人们对广亲和基因的研究有着浓厚的兴趣; 又由于水稻色素原基因 (C gene) 与之紧密连锁^[14], 因此本研究除了涉及 WCG 之外, 也寻找与 C 基因连锁的分子标记, 以丰富水稻现有的遗传连锁图谱, 这将有利于精确定位和分离广亲和基因。

1 材料与方法

1.1 实验材料

1.1.1 植物材料

(1) 近等基因系为台北 8 和 Cps1017 (WCG) B8S1。Cps1017 (WCG) B8S1 为含广亲和基因的品种 Cps1017 (供体) 与不含广亲和基因的粳稻品种台北 8 的杂种植株 (F₁) 经 8 次回交、1 次自交获得, 并经测交证实 WCG 的存在。

(2) F₂ 分离群体库是利用色素原基因纯合显性 (CC) 的 Cps1017 与该基因纯合隐性 (cc) 的品种巴利拉 (Balila) 杂交 F₂ 代的个体经测交证实其基因型后, 各取 20 株 CC 和 cc 分别构成。

1.1.2 引物

Operon 公司出品的 26 组随机引物试剂盒, 分组标号为 A 至 Z, 每组含 20 种不同序列的 10mer 的随机引物共 520 种。本实验从中随机挑选了来自 A、B、D、E、F、J、K、L、M、N、O、P、Q、R、S、T、U、V、W、X、Y 组的引物共 420 个。

四种三磷酸脱氧核糖核苷酸 (dNTPs), 超纯水和明胶为 Sigma 公司产品。Taq DNA 聚合

酶采用 Promega 公司产品和华美公司的试剂盒 (10×PCR Buffer, 25 mM MgCl₂)。PCR 扩增使用 Perkin—Elmer 公司的 480 型 DNA 扩增仪。

1.2 实验方法

1.2.1 DNA 的提取纯化

水稻 DNA 的提取和纯化按照文献^[15]进行。水稻叶片 15—20g 在液氮中磨碎后置于 25ml (60℃) 抽提缓冲液 (35mmol/L NaCl, 50mmol/L Tris-Cl pH8.0, 0.2mmol/L EDTA, 7mol/L 脲, 2% Sarcosyl 及 5% TE 饱和酚) 中混匀, 加 1.5 ml 10% SDS 于 60℃ 保温 10min, 加 15ml 氯仿/异戊醇 (v/v, 24:1) 于室温下抽提 15min, 4000×g 离心 15min, 上清液经纱布过滤后, 加 0.6 体积异丙醇置室温下 15—30min, 此时水稻 DNA 呈纤维状析出。将 DNA 挑出后用 70% 乙醇洗涤 2 次, 真空干燥 10min。加 4ml TE 缓冲液 (pH8.0), 在 60℃ 水浴中溶解 60min, 经 10000×g 离心 10min 去除不溶物后加 RNA 酶至终浓度为 10μg/ml, 37℃ 下保温 30min, 加 1/10 体积 5mol/L LiCl 及 2 倍体积预冷无水乙醇, -20℃ 沉淀 20min, 4℃、6000×g 离心 10min, 70% 乙醇洗涤沉淀 2 次, 最后将 DNA 溶于适量的超纯水中。

1.2.2 PCR 扩增

反应总体积为 25μl, 其中包括: 10mmol/L Tris-HCl (pH9.0), 50mmol/L KCl, 2mmol/L MgCl₂, 0.001% 明胶, 0.2mmol/L dNTPs/每种, 0.2μmol/L Primer, 20—25ng 水稻 DNA, 5 U Taq DNA polymerase。反应程序: 前 5 个循环为 94℃ 变性 1 min, 35℃ 退火 1 min, 72℃ 延伸 2 min; 后 35 个循环为 94℃ 变性 10sec, 36℃ 退火 30sec, 72℃ 延伸 1.5 min。共 40 个循环, 最后在 72℃ 下保温 10min。

1.2.3 电泳

扩增产物在琼脂糖凝胶中电泳, 溴化乙锭染色, 紫外灯下观察结果并照相。

2 结果与讨论

在本实验中所用的台北 8 和 Cps1017 (WCG) B8S1 是一对近等基因系, 随着回交代数的增加, 作为轮回亲本台北 8 的基因组在近等基因系 Cps1017 (WCG) B8 中所占比例预期为 $1 - (1/2)^t$, 其中 $t = \text{回交代数} + \text{自交代数}$ 。由此可知, 只有当回交代数极高以至无限时, 才可能消除供体亲本的基因组成份, 一般回交 5—10 代后, 仍有若干座位的供体基因, 因而只能称之为近等位基因系 (Near Isogenic Lines, NILs)。近等基因系可用来鉴定一些遗传学标记与一个或多个分子标记的连锁关系, 以及鉴定未知分子标记与各种业已导入到近等基因系中的遗传学标记的连锁关系, 故它是基因组制图技术中一种十分有用的材料^[16]。以近等基因系台北 8 和 Cps1017 (WCG) B8S1 的 DNA 为模板, 用不同引物经 PCR 扩增后, 将其产物成对点样进行琼脂糖凝胶电泳可见, 绝大多数电泳谱带是完全相同的, 仅有少数谱带呈现多态性 (图 1); 以色素原基因的 F₂ 分离群体库的 DNA 为模板, 用不同引物经 PCR 扩增后也得到相似结果 (图 2)。这和以 10 个碱基长的核苷酸片段为检出基因组多态性位点的探针具有相似的原理。在所述 PCR 反应条件下, 只有那些能在基因组中找到完全同源的序列并为之退火的引物, 才能得到在电泳图谱中可见的扩增产物。每个反应得到的扩增带从 0—15 条不等, 片段大小在 300—4000bp 之间。经过对 226 个引物在近等基因系台北 8 号-Cps1017 (WCG) B8S1 DNA 的 PCR 扫描和 200 个引物在 F₂ CC-cc 分离群体库 DNA 的 PCR 扫描分别得到表 1 所列结果 (注: 表中所列结果均

为三次以上重复实验的结果,只统计了出现多态性带的次数,详见讨论部分)。

Primer: E-18 E-19 E-20 F-01 F-02 F-03

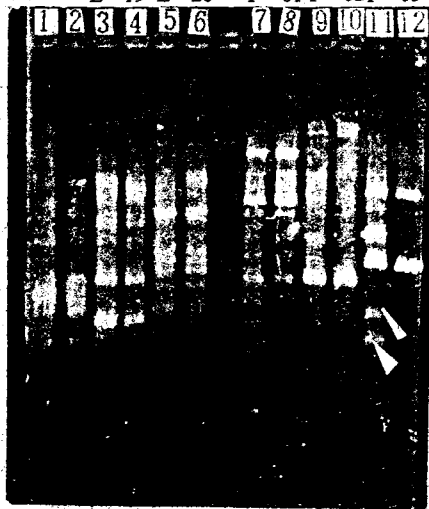


图1 以近等基因系台北8(单号)和Cps1017(WCG)B8S1(双号)的DNA为模板的RAPD产物电泳图谱。显示引物E-18, E-19, E-20, F-01, F-02的PCR产物谱带完全相同, F-03谱带呈多态性(箭头)。

Fig. 1 Electrophoresis pattern of RAPD products with the template DNA of NILs of WCG [left; Taipei8, right; Cps-1017 (WCG) B8S1]. Arrows show the polymorphic bands of F-03

Primer: K-06 K-09 K-08 M-04 M-20 λMarker

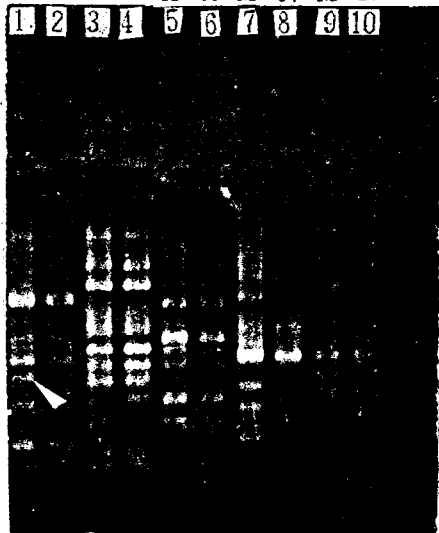


图2 以色素原基因的F₂分离群体库(cc;单号,CC;双号)的DNA为模板的RAPD产物电泳图谱。显示引物K-06的PCR产物谱带呈多态性(箭头处), K-09, K-08, M-04, M-20的谱带完全相同(模糊带型不计)。

Fig. 2 Electrophoresis pattern of RAPD products with the template DNA of bulked segregants of C gene (left; cc, right; CC). Arrow shows the polymorphic band of K-06

表1 RAPD多态性产物检测结果

Table 1 Results of the polymorphic product of RAPD

模板DNA和检测的引物	产生多态性的引物编号	引物序列	多态性带出现次数
WCG等基因系DNA, A B D J E F P Q R S T等11组 共226个引物	A-03	AGTCAGCCAG	4
	A-05	AGGGGTCTTG	4
	A-08	GTGACGTAGG	3
	A-10	GTGATCGCAG	1
	A-13	CAGCACCCAC	1
	B-03	CATCCCCCTG	4
	B-13	TTCCCCGCT	3
	B-14	TCCGCTCTGG	1
	B-16	TTTGCCCCGA	3
	B-20	GGACCCITAC	1
	J-01	CCCGGCATAA	1
	J-03	TCTCCGCTTG	1
	J-14	CACCCGGATG	2
	J-17	ACGCCAGTTC	2
	E-08	TCACCACGGT	1

续表 1

模板 DNA 和检测的引物	产生多态性的引物编号	引物序列	多态性带出现次数
	E-18	GGACTGCAGA	1
	F-03	CCTGATCACC	2
	Q-09	GGCTAACC GA	1
	P-08	ACATCGCCCA	1
	P-13	GGAGTGCC TC	1
	R-11	GTAGCCGTCT	1
	T-11	TTCCCCGCGA	1
	K-06	CACCTTTCCC	3
	K-19	CACAGGCGGA	3
	W-06	AGGCCCGATG	3
CC-cc Bulk DNA,	L-02	TGGGCGTCAA	3
K L M N U V W	O-02	ACGTAGCGTC	3
X Y O 等十组共	V-17	ACC GGCTTGT	2
200 个引物	V-03	CTCCCTGCAA	2
	W-10	TCGCATCCCT	2
	N-10	ACAACTGGGG	1
	N-08	ACCTCAGCTC	1

在本实验所用 PCR 条件下,每个反应平均产生 6.2 个带位,在近等基因系 DNA 上扫描了 226 个引物,其带位总数约为 1400 个,表列结果表明有 22 个多态性带位,即显示频率为 1.5%。在近等基因系台北 8-Cps1017 (WCG) B8S1 中,供体 DNA 占 WCG 基因所在染色体的比例按下列公式计算^[16];

$$U_M = (2/L) [(1/t)(1 - e^{-tL/2})]$$

其中 L 为染色体长度, $t = b + s$ (b 为回交代数; s 为自交代数,本实验 $s = 1$); 相应的偏差为:

$$V_M = (2/L^2) \{ [1/t^2] (2 - (tL + 2)e^{-tL/2}) - [(1/t)(1 - e^{-tL/2})]^2 \}$$

按水稻染色体平均长度 112.5cM (1380 cM/12) 计算,回交 8 代后标记染色体上供体亲本来源的 DNA 所占比例为 19.628% (偏差为 ± 0.01826), 即该染色体中约 22.0815cM ($\pm 2.054cM$) 来源于供体 DNA。对于该近等基因系,其它与 WCG 无关的染色体中,供体 DNA 所占比例为:

$$U_N = (1/2)^t$$

相应的偏差为:

$$V_N = \{ (1/2)^t (2/L) [(1/t)(1 - e^{-tL/2})] \} - [(1/2)^t]^2$$

计算结果 U_M 为 0.001953, 偏差为 ± 0.0003795 , 即平均每条染色体中供体与 WCG 无关的 DNA 所占的长度为 0.2179 cM (± 0.04270 cM), 因而供体来源的与 WCG 无关的 DNA 片段大小总共约为 2.4168 cM。综合考虑标记染色体和非标记染色体,则来自供体的基因组占 B8 回交个

体中总基因组的平均比例为:

$$U_{Ti} = [(L)(U_{Mi}) + (n-1)(L)(U_{Ni})] / (nL), \text{其中 } n \text{ 为染色体总数.}$$

相应的偏差为:

$$V_{Ti} = [(L)(V_{Mi}) + (n-1)(L)(V_{Ni})] / [(nL)^2]$$

将 $b=8, t=b+1, n=12, L=112.5\text{cM}$ 代入上式可知, 总基因组中 1.815% 来源于供体 DNA, 偏差为 ± 0.003012024 。以水稻总基因组大小约为 1380 cM 计算, 则有 24.498 cM (± 4.0662 cM) 的片段为供体 DNA 提供。我们在近等基因系扫描得到的 1400 个带位 (即 1400 个位点) 中应有 1.815% 即 25.405 个带位应来源于供体 (即呈现出多态性), 这与我们的实验结果 22 个多态性片段相接近。按以上计算结果, 24.498 cM 的供体片段中 22.0815 cM 来源于目标基因染色体, 2.4168 cM 为非目标基因染色体 DNA, 即将近 10 个多态性位点中应有 1 个位点不与目标基因连锁, 故可推算在我们找到的 22 个多态性扩增带中应有约 20 个带与 WCG 连锁。

G. J. Muchlbauer 等⁽¹⁶⁾曾经设定一个模式近等基因系基因组, 平均每条染色体长 $L=50$ cM, 染色体总数为 20, 回交代数 $b=5, t=b+1=6$, 得到的结果如下: 供体来源染色体总长为 40.7 cM, 约占基因组的 4%, 其中 25.9 cM (占 64%) 是与标记基因连锁的, 14.80 cM (占 36%) 不与标记基因连锁, 因此在回交个体基因组中得到的供体染色体约 2/3 在标记基因附近。若随机选择 100 个分子标记 (本文为 RAPD 扩增产物), 则在 B5S1 的近等基因系中预期将有 4 个为供体来源的标记, 而这 4 个标记中应有 2—3 个标记与标记基因连锁。本实验的近等基因系为 B8S1, 显然可以预期其与标记基因的连锁程度将提高, 因为用随机引物 PCR 可检测到的距目标基因的最近距离理论值⁽¹²⁾为:

$$D = C/2 (nx + 1)$$

其中 C 为基因组大小, n 为所检测的引物数, x 为每个引物扩增的平均带数。据此公式可知,

对 WCG 近等基因系扫描的 226 个引物得到的多态性片段与 WCG 的最近距离可达 0.5 cM。同样对 C 基因的 F_2 分离群体库的 200 个引物得到的多态性片段中与 C 基因连锁的最近距离也可达 0.56 cM。有实验表明, 用随机引物 PCR 在 F_2 分离群体库中扫描的多态性座位与目标基因距离不超过 25 cM, 且在分离群体库中显示的多态性带与目标基因不连锁的机率为 $2 [1 - (1/4)^n] (1/4)^n$, 其中 n 为组成群体库的个体数。我们的实验是分别采用 20 株 CC-cc 子代个体构成的, 因此所得到的 10 个多态性带中与 C 基因不连锁的可能性为 1.8×10^{-12} 。我们还在 F_2 群体上证实这种多态性谱带的分离, 0-02 是一个显示多态性的引物, 在 42 株 F_2 个体中可以明显看到其扩增产物多态性带位的分离 (图 3), 因此可认为实验得到的多态性扩增带是很有价值的。

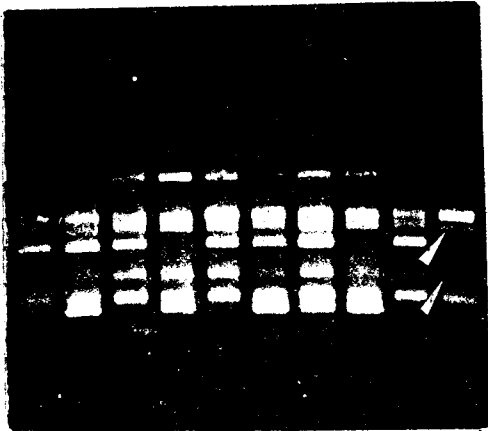


图 3 0-02 多态性带在色素原基因分离的 F_2 群体中的分离, 箭头处为多态性带位。

Fig. 3 The segregation of the 0-02 polymorphic bands in the bulked segregants of F_2 generation of C gene. Arrows show the polymorphic bands.

因此可认为实验得到的多态性扩增带是很有价值的。

在进行 PCR 产物多态性谱带分析时, 我们注意到多态性显示的稳定性, 即引物在模板

DNA 的一个以上的位点上退火的重复性。大部分多态性带位的分布是稳定的,但也有少数多态性不稳定,例如引物A-10在重复进行的三次PCR分析中,一次显示多态性谱带,两次无多态性谱带,即在这两次实验中,引物未曾找到相应的结合位点,而该多态性位点是客观存在的,故我们认为只要实验结果清晰,即使一次实验的多态性结果也应进行统计,并应及时将该多态性扩增产物回收克隆,以进行进一步鉴别。

在S. D. Tanksley⁽¹⁷⁾和Yoshiaki Nagamura⁽⁷⁾发表的水稻RFLP连锁图中可见,水稻第六染色体上标记位点RZ398至RZ612(RG264)之间和C226A, C191B至C1478之间约25-60 cM处得到的RFLP分子标记较少,而已有的证据表明广亲和基因的座位恰好落在此区,因此,丰富该区域的分子标记将有利于该基因的精确定位工作。目前我们正在克隆这些多态性带,以便最终鉴别出这些位点中与WCG-C基因连锁的标记,并计算其图距。预期可以在水稻第六染色体上增加一批新的图位和标记。

参考文献

- 1 Jodon N E. Inheritance and linkage relationships of a chlorophyll mutation in rice. *J Amer Soc Agron*, 1940, 32 (5): 342-346
- 2 Hara S. Linkage between factors for sterility and anthocyanin pigmentation in rice plant. *Jpn J Genet*, 1946, 21 (2): 32
- 3 Oka H I. The mechanism of sterility in the intervarietal hybrid. *Jpn J Breed*, 1953, 2 (4): 214-224
- 4 Ranjhan S J, Glaszmann C, Ramirez D A et al. Chromosomal location of four isozyme loci by trisomic analysis in rice (*Oryza sativa* L.). *Theor Appl Genet*, 1988, 75: 541-545
- 5 Kinoshita T. Report of the committee on gene symbolization, nomenclature and linkage groups. *Rice Genetics Newsletter*, 1993, 10: 7-39
- 6 McCouch S R, Tanksley S D et al. Molecular mapping of rice chromosomes. *Theor Appl Genet*, 1988, 76: 815-829
- 7 Yoshiaki Nagamura. A high density STS and EST linkage map of rice. *Rice Genome*, 1993, 2 (1): 11-13
- 8 Williams J G K, Kubelik A R, Livak K J et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl Acids Res*, 1990, 18: 6531-6535
- 9 Halward T, Stalker T, LaRue E et al. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Molecular Biology*, 1992, 18 (2): 315-325
- 10 Welsh J, Petersen C, McClelland M. Polymorphisms generated by arbitrarily primed PCR in the mouse; application to strain identification and genetic mapping. *Nucl Acids Res*, 1991, 19: 303-306
- 11 Martin G B et al. Rapid identification of markers near a *Pseudomonas* resistance gene in tomato using random primers and near-isogenic lines. *Proc Natl Acad Sci USA*, 1991, 88: 2336-2340
- 12 Michelmore R W, Paran I, Kesseli R V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating population. *Proc Natl Acad Sci USA*, 1991, 88: 9828-9832
- 13 Oka H I. Analysis of genes controlling FI sterility in rice by the use of isogenic lines. *Genetics*, 1974, 77: 521-534
- 14 Ikehashi H, Araki H. Genetics of FI sterility in remote crosses of rice. *Rice Genetics*, IRRI, 1986, 119-130
- 15 刘细连,王春新,刘良式.一种能表现水稻基因组DNA分化特征的重复DNA顺序. *植物学报*, 1994, 36 (6): 452-458
- 16 Muehlbauer G J, Speche J E, Thomas-Compton M A et al. Near-isogenic lines—a potential resource in the integration of conventional and molecular marker linkage maps. *Crop Science*, 1988, 28: 729-735
- 17 Tanksley S D et al. A high density molecular map of the rice genome. *Rice Genetics Newsletter*, 1992, 9: 10-13