

雷蒙德氏棉叶绿体基因组 Fosmid 文库构建

李朋波^{1,2}, 薛龙飞¹, 王彦霞¹, 张曦¹, 李召虎¹, 华金平^{1*}

(1. 中国农业大学农学与生物技术学院 / 作物遗传改良北京市重点实验室 / 作物基因组学与遗传改良农业部重点开放实验室 / 杂种优势研究与利用教育部重点实验室, 北京 100193; 2. 山西省农业科学院棉花研究所, 运城 044000)

摘要: 采用高盐、低 pH 值法提取雷蒙德氏棉叶绿体 DNA; 通过物理剪切法获得随机断裂的 DNA 片段; 剪切片段末端、补平修饰后与 pCC1FOS 载体连接; 用噬菌体包装蛋白包装重组 DNA, 侵染大肠杆菌 EPI300, 构建了雷蒙德氏棉叶绿体基因组文库。对于叶绿体 DNA 剪切, 以 1 mL 注射器中等速度吸打 18 次为最佳参数。叶绿体基因组 Fosmid 文库滴度为 1×10^4 cfu·mL⁻¹, 插入片段大小平均为 38 kb, 最终筛选出 39 个克隆用于后续研究, 覆盖叶绿体基因组 9.2 倍。以叶绿体特异标记筛选出能够覆盖雷蒙德氏棉叶绿体全基因组的 6 个克隆: F66, F46, F28, F8, F55 和 F3, 为基因组结构和功能基因分析提供了良好的基础。

关键词: 雷蒙德氏棉; 叶绿体 DNA; Fosmid 文库

中图分类号: Q812:S562 **文献标识码:** A

文章编号: 1002-7807(2011)01-0010-05

Construction of a Fosmid Library of Chloroplast Genome in *Gossypium raimondii*

LI Peng-bo^{1,2}, XUE Long-fei¹, WANG Yan-xia¹, ZHANG Xi¹, LI Zhao-hu¹, HUA Jin-ping^{1*}

(1. College of Agronomy & Biotechnology, China Agricultural University/Beijing Key Laboratory of Crop Genetic Improvement/Key Laboratory of Crop Genetic Improvement and Genome of Ministry of Agriculture/Key Laboratory of Crop Heterosis and Utilization of Ministry of Education, Beijing 100193, China; 2. Cotton Research Institute, Shanxi Academy of Agricultural Sciences, Yuncheng, Shanxi 044000, China)

Abstract: In this paper, a Fosmid library of *G. raimondii* chloroplast genome was constructed. The chloroplast DNA was isolated by high ionic strength and low pH buffer method. The DNA was randomly sheared and cloned into pCC1FOS vector. Recombinant DNA was packaged with the Lambda Packaging Extracts, then transfected into *E. coli* strain EPI300. The best sheared parameter employed in the study was 18 times with middle speed using a 1 mL injector. The library of chloroplast genome (titer: 1×10^4 cfu·mL⁻¹) was obtained in which the average inserted DNA fragment was 38 kb. Thirty-nine clones covering 9.2 fold the chloroplast genome were selected by selection marker to be further analyzed. Six clones, F66, F46, F28, F8, F55, and F3, which could span *G. raimondii* complete genome, were screened out by cotton chloroplast markers. The library would be a valuable resource for study on genome structure and functional genes investigation in cotton.

Key words: *Gossypium raimondii*; cpDNA; Fosmid library

雷蒙德氏棉 (*Gossypium raimondii*) 起源于秘鲁中部和西部, 细胞学上属于棉属 D5 基因组棉种, 多数研究认为它与异源四倍体棉种 D 亚基因组供体亲缘关系最近^[1-2], 可能为棉花异源四倍体提供了父本来源。叶绿体是植物细胞质中的重要细胞器, 其基因组与核基因组一起经历了协同进

化过程。因此, 尽管在异源四倍体棉种形成过程中雷蒙德氏棉没有提供细胞质, 但其细胞质遗传信息能够为棉种的演化提供重要信息。对棉花叶绿体基因组进行研究, 可从新的角度认识棉花起源、进化和传播。

基因组文库能够覆盖全部基因组, 是进行分

收稿日期: 2010-08-31

作者简介: 李朋波 (1976-), 男, 博士, 助研, lpbmhs@yahoo.com.cn; * 通讯作者: 华金平, 博士, 教授, jinping_hua@cau.edu.cn

基金项目: 教育部新世纪优秀人才支持计划 (NCET-06-0106); 国家 863 计划项目 (2006AA10A108, 2007AA10Z125); 教育部长江学者和创新团队发展计划; 高等学校学科创新引智计划 (111-2-03)

子克隆和基因组结构与功能研究的基础。BAC 文库和 Fosmid 文库是两类较常用的基因组文库^[3-4]。Fosmid 克隆载体以大肠杆菌 F- 因子为基础,以氯霉素抗性基因为选择标记,同时含有 λ DNA 的 cos 序列,可以包装外源 DNA 片段。Fosmid 文库插入片段长度一般为 30~40 kb,通常一个克隆即可包括基因全长,因此对于基因结构和功能研究,建立 Fosmid 文库是一种很好的方法。近年来,Fosmid 已被广泛应用于动植物基因组文库的构建,例如人^[5]、黑猩猩^[6]、扇贝^[7]、稻瘟病菌^[8]、线虫^[9]、摩擦禾^[10]、水稻^[11]等。对于叶绿体等小基因组文库构建,Fosmid 也具有显著的优势,如采用基因组文库测序在叶绿体基因组后拼接上更为方便准确,在功能基因研究中也具有更广泛的用途。据报道,利用 Fosmid 克隆的方法,已构建了绿色植物番薯 (*Ipomoea*)、丝兰 (*Yucca*)、钟萼草 (*Lindenbergia*)、地钱 (*Aneura*)及寄生植物安第斯菰属 (*Corynaca*)、菟丝子 (*Cuscuta*)等叶绿体基因组文库^[4,12],并应用于叶绿体基因组研究。因此,本研究采用 Fosmid 载体构建雷蒙德氏棉的叶绿体 DNA (chloroplast DNA, cpDNA) 文库,以期为其基因组测序及基因功能分析奠定基础。

1 材料和方法

1.1 实验材料

1.1.1 棉花材料。雷蒙德氏棉种子由中国农业科学院棉花研究所海南野生棉种植园(三亚)提供。将种子用刀片破壳以利吸水,催芽后种植于中国农业大学科学园温室,取幼嫩叶片备用。

1.1.2 主要试剂。CopyControl™ Fosmid 文库构建试剂盒 (Cat. No. CCFOS110) 购自 EPICENTRE

公司;质粒提取使用北京 TIANGEN 公司 TIANprep Mini 质粒小提试剂盒;内切酶 *EcoR* I 和 *Not* I 为 TaKaRa 公司产品。

1.2 实验方法

1.2.1 Fosmid 文库的构建。根据 Gong 等^[13]和刘少林等^[14]的高盐、低 pH 值法提取雷蒙德氏棉 cpDNA,用 *EcoR* I 酶切检测 cpDNA。将 cpDNA 用蒸馏水调至终浓度 $0.5 \mu\text{g} \cdot \mu\text{L}^{-1}$,作为文库构建的模板 DNA。取 $80 \mu\text{L}$ 稀释好的 DNA,用 1 mL 注射器中等速度吸打样品,一次吸打过程为 1 个处理,设置吸打次数为 2、4、6、8、10、12、14、16、18、20,共 10 个处理,用 1% 琼脂糖凝胶电泳检测,凝胶长度 20 cm,电压 30 V,电泳 20 h 检测。构建文库的其它步骤包括 cpDNA 的末端修饰与回收、pCC1FOS 载体连接与包装转染等参照试剂盒说明书进行。采用以下公式计算文库滴度与覆盖倍数:滴度 ($\text{cfu} \cdot \text{mL}^{-1}$) = (克隆数 \times 稀释倍数 \times $1000 \mu\text{L} \cdot \text{mL}^{-1}$) / 所用未稀释包装文库体积 (μL),其中,cfu (colony forming units) 表示单位体积的噬菌体稀释颗粒转染大肠杆菌后形成的单个菌落数;覆盖倍数 = (插入片段大小 \times 文库克隆数) / 基因组大小。

1.2.2 叶绿体 Fosmid 克隆的筛选。根据陆地棉叶绿体基因组序列,设计叶绿体特异标记筛选 Fosmid 克隆,引物序列见表 1。每个 PCR 反应体系为 $15 \mu\text{L}$,包括 20 ng 质粒 DNA, $1.5 \mu\text{L}$ $10 \times$ Taq 聚合酶反应缓冲液, $2.0 \mu\text{L}$ MgCl_2 ($15 \text{ mmol} \cdot \text{L}^{-1}$), $0.3 \mu\text{L}$ dNTP ($10 \text{ mmol} \cdot \text{L}^{-1}$),正向和反向引物各 $1 \mu\text{L}$ (引物浓度为 $2 \mu\text{mol} \cdot \text{L}^{-1}$) 以及 0.5 U Taq DNA 聚合酶。PCR 反应条件为 94°C 预变性 3 min; 94°C 变性 30 s, $54 \sim 55^\circ\text{C}$ 退火 30 s, 72°C 延伸 1

表 1 筛选 Fosmid 克隆的叶绿体特异标记

Table 1 Chloroplast-specific markers for screening Fosmid clones

标记 Markers	正向引物 Forward primers (5'-3')	反向引物 Reverse primer (5'-3')	基因组位置* Position in genome / bp	Tm/°C
GCID2	CTTGCAATTTTCATTGCAC	AATGCTCTGGTCAATCATT	1923-2178	55
GCID14	ACATGGCCATATGAGTTGAT	ATGCAAGGTAGAGGGATTTT	37412-37653	55
GCS22	GCTCCTTCGTCTCAAAATC	GTCCTTAGCCCTTGAATCTA	64849-65078	55
GCID28	ACATTGCTCTTTATGAGATGC	ATCTGGATCCAAAGAATCAG	86780-87022	54
GCID34	CACGACCTTGAACAGACAC	TATGACCATCGAGGAACCTT	115977-116209	55
GCS78	GGTTAGTTTCGACAATCCAG	GGATTCTTATTTTCCCCATC	130626-130828	54

注:* 表示在陆地棉叶绿体基因组上的位置。

Note: * Position in chloroplast genome of *G. hirsutum*.

min, 30次循环; 72°C延伸 10 min。PCR产物用 2% 琼脂糖凝胶电泳检测。

2 结果与分析

2.1 叶绿体 DNA 的检测

从 85 g 幼嫩的雷蒙德氏棉叶片中共获得 94.6 μg cpDNA, DNA 得率为 1.11 $\mu\text{g}\cdot\text{g}^{-1}$ 。酶切电泳结果见图 1。由图 1 可以看出, cpDNA 质量较好, 在普通琼脂糖凝胶上与 21 kb 的分子量 Marker 位置相当, 用 *EcoR* I 酶切 cpDNA, 能够产生酶切主带。因此, 提取的 cpDNA 质量可以满足 Fosmid 文库的构建。

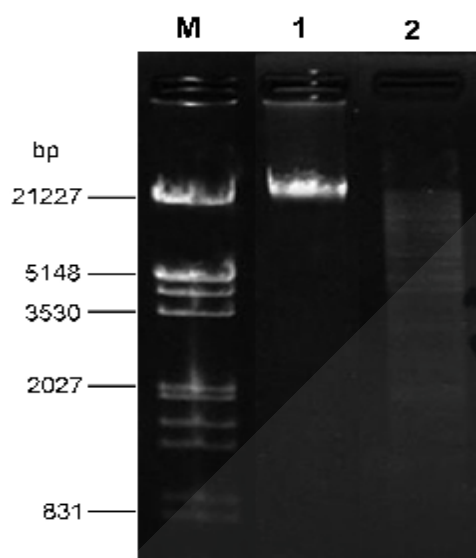


图 1 雷蒙德氏棉 cpDNA *EcoR* I 酶切图谱
Fig. 1 CpDNA of *G. raimondii* digested by *EcoR* I

M: λ DNA/*EcoR* I+*Hind* III, 1: cpDNA, 2: Digested band of cpDNA.

图 1 雷蒙德氏棉 cpDNA *EcoR* I 酶切图谱
Fig. 1 CpDNA of *G. raimondii* digested by *EcoR* I

2.2 cpDNA 克隆片段的获得

由于只有 25~40 kb 范围内的插入片段能够被噬菌体正常包装, 因此需要对剪切片段进行正确处理。用注射器中速吸打 12~20 次, 产生出 10% 以上的剪切 DNA 与对照 DNA 具有相同的迁移率(图 2), 吸打 18 次时产生的 36 kb 左右的片段最多, 因此以 1 mL 注射器中速吸打 18 次为最佳剪切次数。将剪切后的 DNA 电泳后, 小心切下含目标长度 cpDNA 的低熔点琼脂糖薄片回收, 电泳检测浓度, 回收片段浓度约为 50 ng·

μL^{-1} 。回收后共获得插入片段约 0.3 μg , 可满足后续连接的需要。

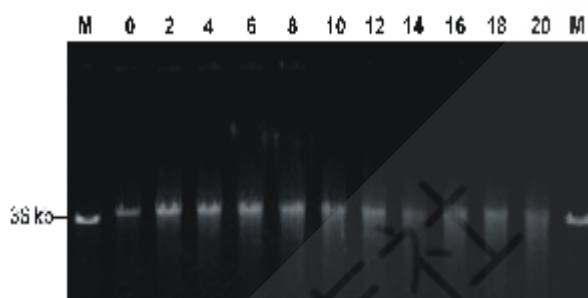


图 2 雷蒙德氏棉 cpDNA 随机剪切结果
Fig. 2 Random shearing of *G. raimondii* cpDNA

图 2 雷蒙德氏棉 cpDNA 随机剪切结果
Fig. 2 Random shearing of *G. raimondii* cpDNA

2.3 噬菌体颗粒的滴度计算

在文库铺板前, 计算包装 Fosmid 克隆的滴度, 将有助于确定铺板的数量和稀释倍数, 获得满足要求的文库。在本实验中, 由于 10^2 、 10^4 的稀释倍数过大而没有培养出克隆, 采用未稀释的噬菌体颗粒能够获得菌落密度合适的克隆, 每个平板可获得 100 个克隆。在本研究中, 未稀释的克隆滴度为 10^4 cfu·mL⁻¹, 即本次包装共可获得 1×10^4 个克隆。

2.4 文库的覆盖度

随机挑选 10 个克隆提取质粒, 用 *Not* I 酶切, 电泳检测片段的大小, 结果如图 3 所示, 每个克隆的酶切都产生了 35~40 kb 的插入片段带和 7.5 kb 的载体带, 插入片段平均约为 38 kb。随机挑取的克隆中, 没有发现空载现象。

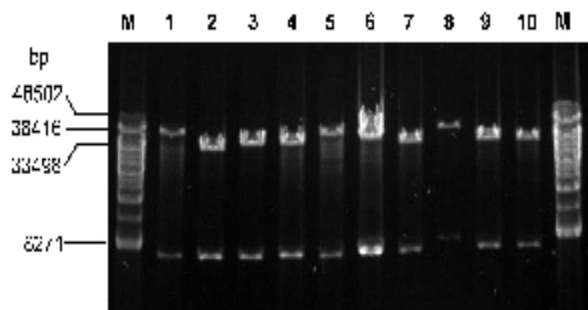


图 3 Fosmid 质粒 DNA *Not* I 酶切
Fig. 3 Fosmid plasmid DNA digested by *Not* I

图 3 Fosmid 质粒 DNA *Not* I 酶切
Fig. 3 Fosmid plasmid DNA digested by *Not* I

根据文库滴度的计算,用未稀释的噬菌体颗粒包装后涂布了 4 块 LB 抗性平板,培养后获得 386 个克隆。随机挑选了 118 个克隆提取质粒 DNA,采用叶绿体特异标记筛选,标记位置如图 4 所示。这些标记将叶绿体基因组分为六个区,各区跨度分别为 35.5 kb,27.2 kb,22.0 kb,29.2 kb,14.5 kb 及 31.7 kb,Fosmid 文库插入片段的大小为 38 kb 左右,因此,在叶绿体的阳性克隆上至少应该出现一个标记。结果表明,共有 71 个阳性克隆具有叶绿体特异标记,阳性率为 60.1%。在阳性克隆中,GCID2,GCID14,GCS22,GCID28,GCID34 和 GCS78 出现的次数分别为 10,13,29,31,6,5 次,在 GCS22 和 GCID28 两个标记处具有较多的冗余克隆。根据标记剔除部分冗余克隆,使每个标记在文库中出现的次数在 5~10 之间,从中筛选出 39 个克隆作为雷蒙德氏棉叶绿体基因组 Fosmid 文库,共覆盖雷蒙德氏棉叶绿体基因组 9.2 倍。

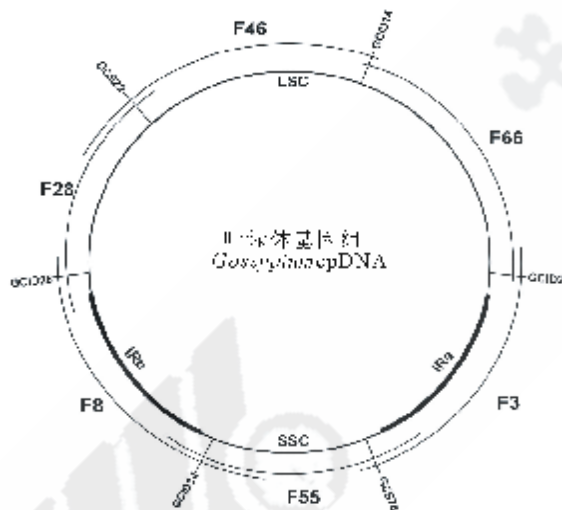


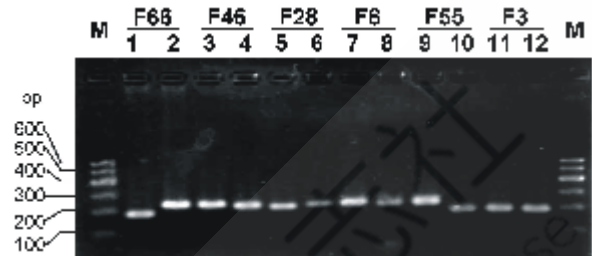
图 4 筛选的克隆在叶绿体基因组的位置

Fig. 4 The position of the screened clones

2.5 覆盖叶绿体全基因组克隆的筛选

对于基因组测序,应选取能够完全覆盖基因组的最少克隆。由于叶绿体基因组中两个 IR 区为反向重复序列,在后续序列拼接时容易引起错误拼接,因此,克隆的筛选原则是使两个 IR 区分别位于两个克隆中,其它克隆序列覆盖全基因组。根据图 4 所示的标记,能够同时出现相邻两个标记的克隆即为候选克隆,从保存的 39 个叶绿体克隆中共获得 9 个候选克隆,挑选片段最小

的克隆作为测序克隆,最终获得 6 个首尾重叠的克隆:F66,F46,F28,F8,F55 和 F3(图 4,图 5),总长约 228 kb,能够完全覆盖雷蒙德氏棉的叶绿体基因组。



1~12 分别为图 4 中叶绿体标记在各克隆的扩增带型。

1~12: The amplified bands of each clone by chloroplast markers in Fig. 4.

图 5 筛选克隆用两侧分子标记鉴定的结果

Fig. 5 Screened clones identified by flanking molecular markers

3 讨论

基因组文库的基本要求是克隆的稳定性和随机性。与高拷贝数的 Cosmid 文库不同,Fosmid 文库在宿主菌中以单拷贝形式存在,其克隆继代繁殖 160 代也不发生丢失与突变,具有很高的遗传稳定性^[3],是目前用于替代 Cosmid 的新粘粒文库。与 BAC 文库相比,Fosmid 文库克隆的 DNA 片段的随机性更高,机械剪切基因组保证了每段 DNA 在文库中出现的频率均等。同时,Fosmid 文库通过诱导培养可达到高拷贝数(10~50 个),方便获得足量的质粒 DNA 用于后续研究。并且由于插入片段较小,也降低了模板 DNA 制备的难度。

棉花叶绿体基因组大小为 160 kb^[15-16],参照试剂盒说明及草鱼核基因组的剪切参数^[17],用 1 mL 注射器中速吸打 20~100 次,结果造成 cpDNA 的过度剪切(数据未列);使用本研究的参数,可获得较理想的剪切效果。在文库构建中出现的部分假阳性克隆,可能是叶绿体膜粘附核及线粒体 DNA 的污染造成的。本研究最初为获得较纯的 cpDNA,使用 DNAase I 消化棉花叶绿体外 DNA,但引起 cpDNA 得率很低。在杨树 cpDNA 提取过程中也出现类似的结果^[18],因此,放弃酶消化操作。为消除核及线粒体 DNA 的污染,实验中采用棉花叶绿体基因组标记筛选阳性

克隆,获得了雷蒙德氏棉 cpDNA 特异克隆。本研究首次构建了棉花叶绿体的 Fosmid 文库,采用随机断裂 DNA 的方法能够较容易获得覆盖全基因组的克隆,并采用叶绿体基因组特异标记筛选克隆,构建的文库在每段序列上都具有相似的覆盖度,从而保证了雷蒙德氏棉叶绿体基因组 Fosmid 文库的均一性和可靠性。

基因组文库的构建在基因组物理图谱构建、基因的图位克隆、基因组测序、比较基因组研究、基因表达调控以及人类和动植物的基因组工程等方面都有极其重要的应用。本研究构建的雷蒙德氏棉叶绿体基因组 Fosmid 文库可用于叶绿体基因组的测序、基因功能分析和分离特定的基因,为研究棉属进化和叶绿体基因功能提供了重要基础。

参考文献:

- [1] WENDEL J F, Cronn R C. Polyploidy and the evolutionary history of cotton[J]. *Advances in Agronomy*, 2003, 78: 139-186.
- [2] 刘三宏,王坤波,宋国立. 四倍体棉种的起源与演化[J]. *棉花学报*, 2003, 15(5): 304-308.
LIU San-hong, Wang Kun-bo, Song Guo-li. The origin and evolution of tetraploid species of *Gossypium*[J]. *Cotton Science*, 2003, 15 (5): 304-308.
- [3] 李海权,刁现民. 基因组细菌人工染色体文库(BAC)的构建及应用[J]. *生物技术通报*, 2005(1): 6-11.
LI Hai-quan, Diao Xian-min. Construction and application of genomics bacterial artificial chromosome(BAC) library[J]. *Biotechnology Bulletin*, 2005 (1): 6-11.
- [4] JANSEN R K, Raubeson L A, Boore J L, et al. Methods for obtaining and analyzing whole chloroplast genome sequences [J]. *Methods in Enzymology*, 2005, 395: 348-384.
- [5] BIRREN B W, Tachi-iri Y, Kim U J, et al. A human chromosome 22 Fosmid resource: Mapping and analysis of 96 clones[J]. *Genomics*, 1996, 34: 97-106.
- [6] NEWMAN T L, Tuzun E, Morrison V A, et al. A genome-wide survey of structural variation between human and chimpanzee [J]. *Genome Research*, 2005, 15: 1344-1356.
- [7] 程洁,张玲玲,黄晓婷,等. 栉孔扇贝 Fosmid 文库的构建及基因组结构特征分析[J]. *中国海洋大学学报*, 2008, 38(1): 78-88.
CHENG Jie, Zhang Ling-ling, Huang Xiao-ting, et al. Fosmid library construction and genomic structure analysis in Zhikong Scallop (*Chlamys farreri*)[J]. *Periodical of Ocean University of China*, 2008, 38 (1): 78-88.
- [8] REHMEYER C, Li W, Kusaba M, et al. Organization of chromosome ends in the rice blast fungus, *Magnaporthe oryzae* [J]. *Nucleic Acids Research*, 2006, 34: 4685-4701.
- [9] DOLPHIN C T, Hope I A. Caenorhabditis elegans reporter fusion genes generated by seamless modification of large genomic DNA clones[J]. *Nucleic Acids Research*, 2006, 34(9): e72.
- [10] LAMB J C, Birchler J A. Retroelement genome painting: cytological visualization of retroelement expansions in the genera *Zea* and *Tripsacum*[J]. *Genetics*, 2006, 173: 1007-1021.
- [11] MIZUNO H, Wu J, Katayose Y, et al. Chromosome-specific distribution of nucleotide substitutions in telomeric repeats of rice (*Oryza sativa* L.) [J]. *Molecular Biology and Evolution*, 2008, 25: 62-68.
- [12] WICKETT N J, Zhang Y, Hansen S K, et al. Functional gene losses occur with minimal size reduction in the plastid genome of the parasitic liverwort *Aneura mirabilis*[J]. *Molecular Biology and Evolution*, 2008, 25: 393-401.
- [13] GONG X S, Yan L F. Improvement of the purification of chloroplast DNA from higher plants[J]. *Chinese Science Bulletin*, 1991, 36: 1633-1635.
- [14] 刘少林,靖深蓉,郭立平,等. 用改进的高盐低 pH 法分离和纯化棉花叶绿体及叶绿体 DNA[J]. *棉花学报*, 1997, 9(5): 239-241.
LIU Shao-lin, Jing Shen-rong, Guo Li-ping, et al. Purification of cotton chloroplast and chloroplast DNA using the improved method of high salt concentration and low pH[J]. *Cotton Science*, 1997, 9 (5): 239-241.
- [15] LEE S B, Kaittanis C, Jansen R K, et al. The complete chloroplast genome sequence of *Gossypium hirsutum*: Organization and phylogenetic relationships to other angiosperms[J]. *BMC Genomics*, 2006, 7: 61.
- [16] IBRAHIM R I H, Azuma J I, Sakamoto M. Complete nucleotide sequence of cotton (*Gossypium barbadense* L.) chloroplast genome with a comparative analysis of sequence among 9 dicot plants [J]. *Genes & Genetic Systems*, 2006, 81: 311-321.
- [17] 贾震虎. 草鱼基因组文库构建筛选及两个核心 *MHC I* 基因座序列图谱研究 [D]. 北京: 中国农业大学, 2007: 5-15.
JIA Zhen-hu. Construction and screening of a grass carp genome library and mapping of two core *MHC I* region [D]. Beijing: China Agricultural University, 2007: 5-15.
- [18] 崔彬彬,李云,冯大领,等. 杨树叶绿体分离及叶绿体 DNA 提取方法的研究[J]. *保定师范专科学校学报*, 2006, 19(2): 25-27.
CUI Bin-bin, Li Yun, Feng Da-ling, et al. A study for isolation of *Populus* chloroplast and cpDNA[J]. *Journal Baoding Teachers College*, 2006, 19 (2): 25-27. ●