

# 雷蒙德氏棉 DNase I 高敏感位点(DHSs)文库构建方法

王惠敏<sup>1,2</sup>, 刘玉玲<sup>2</sup>, 韦洋洋<sup>2</sup>, 张树林<sup>2</sup>, 刘震<sup>2</sup>, 卢全伟<sup>2</sup>, 杨太有<sup>1</sup>, 彭仁海<sup>1,2</sup>

(1.河南师范大学 生命科学学院,河南 新乡 453007; 2.安阳工学院 生物与食品工程学院,河南 安阳 455000)

**摘要:**染色质 DNA 对 DNase I 表现出高敏感性的位点(DNase I hypersensitive sites, DHSs)多是顺式作用元件所在的位置,包括启动子、增强子、调节子和衰减子等.研究 DHSs 位点的数目及其动态变化是探明顺式作用元件功能、揭示基因表达调控机制,乃至进行基因编辑和创新遗传材料的重要辅助手段.本文借鉴水稻和拟南芥 DHSs 文库构建方法,通过优化设计,初步建立了二倍体雷蒙德氏棉(*Gossypium raimondii* D<sub>5</sub>)的 DHSs 文库,为深入进行棉花功能基因组研究奠定基础.

**关键词:**雷蒙德氏棉; DNase I 高敏感位点(DHSs); 顺式作用元件; 文库构建

**中图分类号:**Q756

**文献标志码:**A

染色质 DNA 对 DNase I 表现出高敏感性的位点(DNase I hypersensitive sites, DHSs)是核小体缺乏的区域,容易被 DNase I 识别、酶切而引起 DNA 的降解<sup>[1-2]</sup>,DHSs 通常是启动子、增强子、调节子和衰减子等顺式作用元件所在的位置<sup>[3-4]</sup>.DHSs 研究开始于 20 世纪 70 年代<sup>[5-7]</sup>,首先在果蝇中检测到 DHSs<sup>[8]</sup>,随后在酵母<sup>[9]</sup>、线虫<sup>[10]</sup>和人类<sup>[11]</sup>等生物中进行了相关研究.植物因为具有细胞壁、细胞质富含次生代谢物质等原因,DHSs 研究较为滞后,Zhang 等首先在水稻<sup>[12]</sup>中开展 DHSs 研究,建立了较为成熟的 DHSs 研究技术,并在拟南芥<sup>[13]</sup>和番茄<sup>[14]</sup>等植物中应用.

棉花是重要的天然纤维作物,棉花细胞富含多糖、单宁和棉酚等代谢物质,其基因组研究远滞后于水稻和拟南芥等模式植物.棉花 DHSs 研究仅见陆地棉不同品系<sup>[15]</sup>,而作为异源四倍体 D 亚组的供体种,雷蒙德氏棉的 DHSs 研究尚未见报道.有鉴于此,本文针对雷蒙德氏棉细胞特殊性,开展雷蒙德氏棉 DHSs 文库构建方法优化,在材料选择、酶切浓度确定、载体连接等方面形成一套完善的技术体系,获得质量较高的 DHSs 文库.

## 1 材料与方法

### 1.1 实验材料

二倍体 D 基因组雷蒙德氏棉(*Gossypium raimondii* D<sub>5</sub>)幼叶取自中国农业科学院棉花研究所.鉴于雷蒙德氏棉不易栽培且成活率低、棉花细胞富含次生代谢物质的特性,本实验选择老株上的嫩叶进行细胞核的提取.

### 1.2 细胞核的提取、酶切和包埋

本研究参照 Zhang 等的方法<sup>[16]</sup>,用包埋法固定高分子量 DNA(High molecular weight DNA, HMW-DNA),以减少 HMW-DNA 的损伤,细胞核提取液中针对雷蒙德氏棉细胞特殊性,添加 2% (m/v) PVP40 (Sigma)、体积分数为 0.4% Triton X-100(索莱宝)和蔗糖.利用 DNase I(愚公生物)对获得的细胞核进行酶

收稿日期:2017-11-20;修回日期:2018-03-20.

基金项目:国家自然科学基金(31471548);河南省科技创新杰出人才(2018JR0009);棉花生物学国家重点实验室开放课题(CB2017A06);河南省高等学校重点科研项目(18A180010).

作者简介:王惠敏(1988-),女,河南濮阳人,河南师范大学硕士研究生,研究方向为分子细胞生物学,E-mail:780250660@qq.com.

通信作者:杨太有,教授,E-mail:13803731625@163.com;彭仁海,教授,E-mail:aydxprh@163.com.

切,固定温度 37 °C 和时间 10 min.在最佳酶切浓度的选定中设计了预实验和实验两个部分,在预实验中确定最佳酶切浓度的范围,实验确定最佳酶切浓度.酶切后加入反应终止液和 2% (m/v) 的低熔点琼脂糖 (Sigma),充分混合后加入 DNA 胶块模具中制成 80  $\mu$ L 的小胶块.最后将制好的胶块在 10 mL 十二烷基磺酸锂缓冲液 (Sigma) 中 37 °C 水浴过夜,以去除大分子蛋白质.

### 1.3 检测 DNase I 的酶切质量

使用 50 mM pH 为 8.0 的 EDTA 缓冲液清洗胶块 5 次,每次 1 h,去除十二烷基磺酸锂缓冲液的残留,然后,取 1/2 胶块放入点样孔中,脉冲场电泳 (Bio-Rad) 检测最佳酶切浓度和酶切质量.

### 1.4 Adaptor 1 和 Adaptor 2 的连接

选取酶切浓度较为合适的胶块提取高分子量 DNA,加入 T4 DNA 连接酶 (Promega) 和 Adaptor 1 (Takara) 后 20 °C 过夜连接,连接时设计不同的连接比例 (Adaptor 1 与大片段 DNA 的摩尔比例 1 : 3、1 : 8 和 1 : 15),并使用低熔点琼脂糖凝胶电泳分离出连接 Adaptor 1 的高分子量 DNA,经检测达到预期浓度后,用 *Mme* I (NEB) 酶切,获得已连接 Adaptor 1 的 20 bp 左右的 DNA 片段,磁珠 (Invitrogen) 回收后,加入 Adaptor 2 (Takara),室温连接 4 h.

### 1.5 产物的扩增和提纯

连接了 Adaptor 1 和 2 并吸附于磁珠上的产物,经 98 °C 30 S;12 个循环的 98 °C 10 S,60 °C 30 S,72 °C 15 S;72 °C 7 min 的程序进行 PCR 扩增后,用 (m/v) 12 % 的聚丙烯酰胺凝胶电泳回收纯化,再用 13  $\mu$ L 的 10 mM Tris-HCl pH 8.0 溶解 DNA 沉淀,检测浓度合格,即得 DHSs 文库.

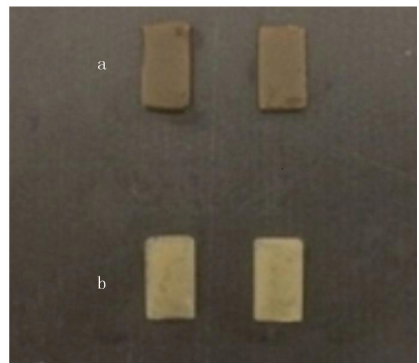
### 1.6 文库的检测

将制得的 DHSs 文库制备单克隆,随机挑取 10 个左右的单克隆送上海生工有限公司进行测序,并使用 DNAMAN 软件进行序列比对分析.

## 2 结果与讨论

### 2.1 细胞核提取

高质量的细胞核提取是获得 DHSs 可靠数据的前提.针对棉花细胞的特性,首先,为了彻底破坏细胞壁和其他纤维结构,释放细胞核,将样品用液氮充分研磨 1 h 以上,在研磨的过程中保证液氮不干,研磨充分后连同液氮一起转入冰冻的离心管中,-80 °C 保存备用.其次,细胞核提取液中不添加或添加 PVP40 进行比较,不加 PVP40 所制得的胶块颜色呈现黄褐色,是因为被棉花细胞中次生代谢物氧化的结果,细胞核提取质量不高 (图 1a),而添加 2% (m/v) 浓度的 PVP40 后所制得的胶块,细胞核提取质量高 (图 1b),因此实验确定在细胞提取液中加 2% (m/v) 浓度的 PVP40.再者,洗涤液中 Triton X-100 的用量非常关键,高浓度的 Triton X-100 损害外核膜使细胞核更脆弱,从而导致细胞核降解,低浓度的 Triton X-100 又不能很好地去除叶绿体等杂质,因此, Triton X-100 的浓度应根据不同的组织和物种的特性以及洗涤次数进行适当调整,我们设计了不同的浓度,结果表明在雷蒙德氏棉幼叶中确定的最佳用量是 0.4%.经过以上处理,所提取的细胞核质量较高,满足后续实验需要.



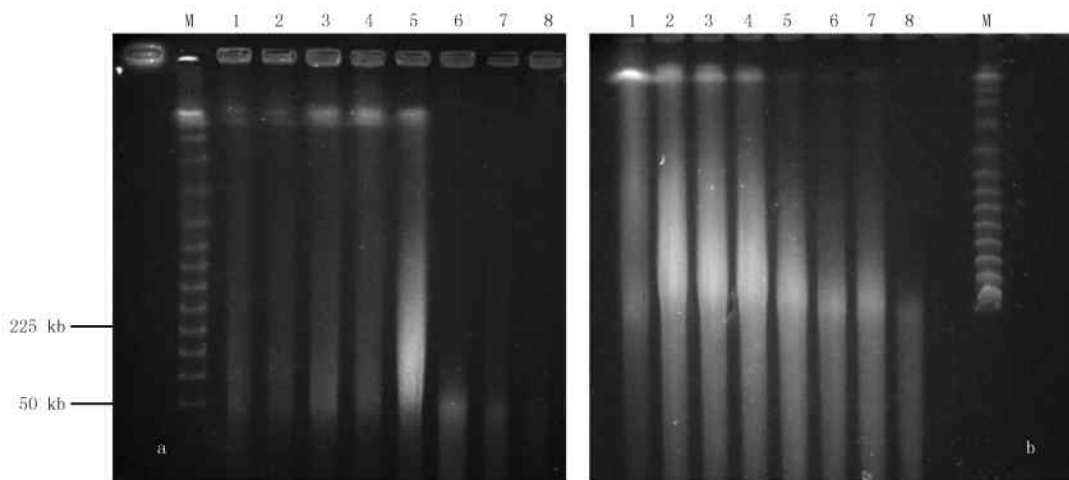
(a) 为不加 PVP40 所做的胶块; (b) 为加 PVP40 所做的胶块.

图 1 细胞核提取液中 PVP40 对胶块质量的影响

### 2.2 最佳 DNase I 用量的选择

最佳酶切浓度是 DHSs 文库构建中较为关键的步骤之一,因为 DHSs 文库构建不仅要有较高的 DNA 起始浓度,而且要保证高质量的完整酶切位点.本研究采用固定酶切时间和温度,调整酶切浓度以达到最佳酶切效果.首先是预实验,取 5 g 叶子所提取的细胞核悬浮液,分成 8 等份,分别加入  $10^{-5}$  U,  $10^{-4}$  U,  $10^{-3}$

U,  $10^{-2}$  U,  $10^{-1}$  U, 1.0 U, 4.0 U, 10 U 的 DNase I, 37 °C 孵育 10 min, 每 3 min 轻弹混匀, 酶切结束后立即置于冰上, 并快速加入酶切终止缓冲液, 初步确定最佳酶切浓度在 0.1 U 至 1 U 之间, 脉冲场电泳检测(图 2a). 取与预实验等体积等浓度的细胞核悬浮液, 分成 8 等份, 在 0 U 至 1 U 之间设定 0 U, 0.08 U, 0.12 U, 0.2 U, 0.4 U, 0.6 U, 0.8 U, 1.0 U 8 个不同的酶切浓度. 然后进行脉冲场电泳检测(图 2b). 由电泳结果图可以看出酶切浓度在 0.4 U 时细胞核已充分酶切, 且大片段主要集中在 50~225 kb 之间, 所以确定酶切浓度在 0.4 U 时为最佳酶切浓度. 需要注意的是, 为了使实验具有可参考性, 两次实验所用的细胞核的浓度和体积要相同.



(a) 为预实验组, 1-8 分别表示  $10^{-5}$  U,  $10^{-4}$  U,  $10^{-3}$  U,  $10^{-2}$  U,  $10^{-1}$  U, 1.0 U, 4.0 U, 10 U; (b) 为实验组, 1-8 分别表示 0 U, 0.08 U, 0.12 U, 0.2 U, 0.4 U, 0.6 U, 0.8 U, 1.0 U; M 表示 Marker.

图2 DNase I 酶切浓度的选择

### 2.3 高分子量 DNA 的回收和 Adaptor 1 的连接

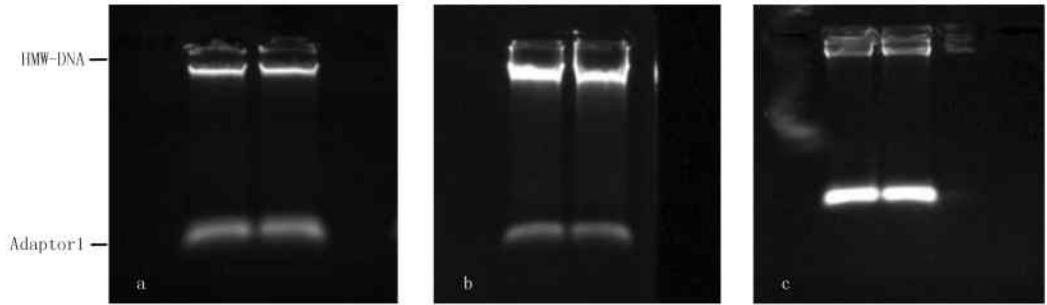
为了减少对大片段 DNA 造成的机械损伤, 本研究使用低熔点琼脂糖回收 Adaptor 1 连接之后的 HMW-DNA, 低熔点琼脂糖回收法<sup>[16]</sup>可以回收更加完整的高分子量 DNA, 重要的是对 DNA 分子平末端的损害较小. 电泳过程中不能使用过高的电压, 以免胶块受热融化, 为防止交叉污染所用的 TBE 缓冲液必须是新配置, 在紫外灯下切取目的片段时要准确快捷, 准确确定目的片段, 避免切割多余的琼脂糖, 影响回收效果. 苯酚-氯仿抽提法回收时, 要涡旋 2 min, 使溶胶液与目的片段充分混合, 但长时间的涡旋可能导致 DNA 大片段的断裂, 因此涡旋时间不宜过长.

本研究选用的 T4 DNA 连接酶在 20 °C 连接过夜, 这一步的关键是确定 Adaptor 1 和目的片段的连接比例, 我们得到的目的片段浓度约在  $500\sim 800\text{ ng}\cdot\mu\text{L}^{-1}$  之间, 设计了 Adaptor 1 与大片段 DNA 的摩尔比例进行了 1:3、1:8 和 1:15 的比例. 结果发现 1:8 的连接比例最佳, 连接率最高(图 3a). 当 Adaptor 1 与大片段 DNA 的摩尔比 1:15 时体系中接头的浓度过低, 影响 Adaptor 1 与大片段 DNA 的碰撞进而导致连接效率过低(图 3b); 1:3 的比例又会造成 Adaptor 1 过剩(图 3c), 形成接头自连条带. 因此, 最佳连接比例的确定至关重要. 另外, 研究发现, 50  $\mu\text{L}$  体积为不影响回收效率的最佳连接体系.

### 2.4 Adaptor 2 的连接及目的条带的扩增和纯化

研究中发现, Adaptor 2 连接的关键影响因素是连接温度, 我们用同一批样品分别在不同的温度下做 Adaptor 2 的连接, 出现三种情况: 一是连接温度在 20~22 °C, 在 100 bp 和 90 bp 处各出现一条带, 经检测发现位于 100 bp 处的是目的条带(图 4a); 二是连接温度在 25 °C, 出现一条 100 bp 的条带, 经检测全部为目的条带(图 4b); 三是连接温度在 28~30 °C, 出现一条 90 bp 的条带, 经进一步检测这条带并不是目的条带, 是两个接头的自连带(图 4c). 因此确定连接温度 25 °C 最为合适. Adaptor 1 和 Adaptor 2 连接之后, 文库的目的片段在 100 bp 左右, 在 PCR 扩增时过多的循环数会导致引物二聚体的产生, 影响 DHSs 文库的质量, 一般采用 12 个循环, 如果获得的文库质量较低可以将循环数增至 13 个, 以确保获得足够的 DHSs 文库浓度. 为了达到测序所需要的浓度(3 nM), 可以减少回收液 Tris-HCl 的体积以增大产物浓度, 根据测序要求产

物最低体积不能低于  $9 \mu\text{L}$ . 用 Nanodrop 2000 初步检测文库质量.



(a) 表示 Adaptor 1 和目的片段连接比例为 1 : 8; (b) 为 1 : 15; (c) 为 1 : 3.

图 3 Adaptor 1 和目的片段连接比例的确定

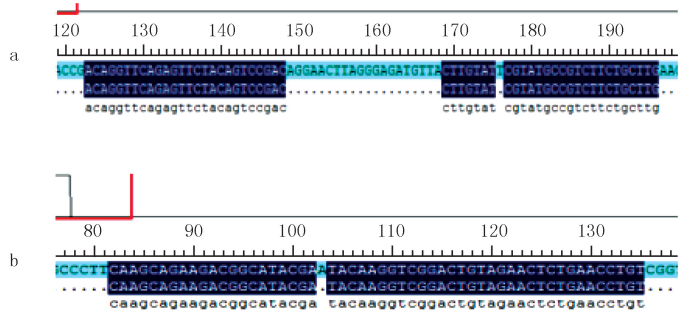


(a) 表示连接温度在  $20 \sim 22 \text{ }^{\circ}\text{C}$ ; (b) 表示连接温度在  $25 \text{ }^{\circ}\text{C}$ ; (c) 表示连接温度在  $28 \sim 30 \text{ }^{\circ}\text{C}$ .

图 4 最佳连接温度的确定

## 2.5 DHSs 文库的检测与保存

将符合图 4b 所回收纯化的高质量样品做单克隆, 测序结果经 DNAMAN 软件比对, 结果表明, 目的片段的两端分别连接上了 Adaptor1 和 Adaptor2, 中间部分为预期长度的目的片段(图 5a). 同时, 将图 4a 和 c 中所获得的 90 bp 的片段做单克隆后测序, 没有出现连接产物中间预期的目的片段(图 5b). 连接后的检测是非常必要的, 一是因为文库片段较小, 在分辨率



(a) 连接上目的片段的克隆; (b) 无目的片段克隆.

图 5 文库质量的检测

不高的情况下, 不能准确辨认出中间是否连接有目的片段; 二是由于初次在雷蒙德氏棉中构建 DHSs 文库, 需要准确确定目的片段连接成功后, 才能进行后续实验, 获得预期结果. 文库建成后, 必须将文库保存在无 RNA 的管中, 并应立即放于  $-80 \text{ }^{\circ}\text{C}$  的超低温冰箱中, 以保持文库的稳定性, 避免污染.

## 3 结 论

DHSs 文库构建后进行测序和生物信息学分析, 成为一种高效快捷鉴定基因组中顺式作用元件的方法, 除了在一些动物和模式植物上有较为成熟的研究应用外<sup>[8-12]</sup>, 在基因组较大的多倍体植物中研究应用还具有一定的难度. 棉花是全球最重要的经济作物之一, 也是研究植物单细胞发育最理想的模式植物. 雷蒙德氏棉是公认的异源四倍体 D 亚组供体种之一<sup>[17]</sup>, 具有提高纤维品质、抗旱、抗病及抗虫的潜质<sup>[18]</sup>. 随着雷蒙德氏棉基因组测序的完成<sup>[19-20]</sup>, 构建雷蒙德氏棉 DHSs 图谱, 开展重要农艺性状相关基因的调控机理研究显

得非常重要。本研究就雷蒙德氏棉 DHSs 文库的构建,形成了一套较为成熟的技术方法,为深入进行雷蒙德氏棉功能基因挖掘,乃至其他植物的相关研究提供借鉴。

### 参 考 文 献

- [1] Wu C, Wong Y C, Elgin S C. The chromatin structure of specific genes; II. Disruption of chromatin structure during gene activity[J]. Cell, 1979, 16(4): 807-814.
- [2] Wu C. The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I[J]. Nature, 1980, 286(5776): 854-860.
- [3] Gross D S, Garrard W T. Nuclease hypersensitive sites in chromatin[J]. Annual Review of Biochemistry, 1988, 57(57): 159-197.
- [4] Thurman R E, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome[J]. Nature, 2012, 489(7414): 75-82.
- [5] Scott W A, Wigmore D J. Sites in simian virus 40 chromatin which are preferentially cleaved by endonucleases[J]. Cell, 1978, 15: 1511-1518.
- [6] Varshavsky A J, Sundin O, Bohn M. A stretch of 'late' SV40 viral DNA about 400 bp long which includes the origin of replication is specifically exposed in SV40 minichromosomes[J]. Cell, 1979, 16: 453-466.
- [7] Weisbrod S, Weintraub H. Isolation of a subclass of nuclear proteins responsible for conferring a DNase I-sensitive structure on globin chromatin[J]. Proc Natl Acad Sci USA, 1979, 76: 630-634.
- [8] Wu C, Bingham P M, Livak K J, et al. The chromatin structure of specific genes; I. Evidence for higher order domains of defined DNA sequence[J]. Cell, 1979, 16: 797-806.
- [9] Hesselberth J R, Chen X, Zhang Z, et al. Global mapping of protein-DNA interactions *in vivo* by digital genomic footprinting[J]. Nat Methods, 2009, 6: 283-289.
- [10] Kharchenko P V, Alekseyenko A A, Schwartz Y B, et al. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster* [J]. Nature, 2011, 471: 480-485.
- [11] Boyle A P, Davis S, Shulha H P, et al. High-resolution mapping and characterization of open chromatin across the genome[J]. Cell, 2008, 132: 311-322.
- [12] Zhang W L, Wu Y F, Schnable J C, et al. High-resolution mapping of open chromatin in the rice genome[J]. Genome Research, 2012, 22(1): 151-162.
- [13] Zhang W L, Zhang T, Wu, Y F, et al. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis* [J]. Plant Cell, 2012, 24(7): 2719-2731.
- [14] Qiu Z K, Li R, Zhang S B, et al. Identification of regulatory DNA elements using genome-wide mapping of DNase I hypersensitive sites during tomato fruit development[J]. MPBJ, 2016, 9(8): 1168-1182.
- [15] Wang M J, Tu L L, Li M, et al. Asymmetric subgenome selection and cis- regulatory divergence during cotton domestication[J]. Nature Genetics, 2017, 49: 579-587.
- [16] Zhang H B, Zhao X P, Ding X L, et al. Preparation of megabase-size DNA from plant nuclei[J]. The Plant Journal, 1995, 17(1): 175-184.
- [17] 王坤波, 刘旭. 棉属多倍化研究进展[J]. 中国农业科技导报, 2013, 15(2): 20-27.
- [18] 黄滋康, 季道藩, 潘家驹. 中国棉花遗传育种学[M]. 济南: 山东科学技术出版社, 2003: 8-32.
- [19] Paterson A H, Wendel J F, Gundlach H, et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres[J]. Nature, 2012, 492(7429): 423-427.
- [20] Wang K B, Wang Z W, Li F G. The draft genome of a diploid cotton *Gossypium raimondii* [J]. Nature Genetics, 2012, 44(10): 1098-1103.

## Construction of *Gossypium raimondii* DNase I hypersensitive site library

Wang Huimin<sup>1,2</sup>, Liu Yuling<sup>2</sup>, Wei Yangyang<sup>2</sup>, Zhang Shulin<sup>2</sup>, Liu Zhen<sup>2</sup>, Lu Quanwei<sup>2</sup>, Yang Taiyou<sup>1</sup>, Peng Renhai<sup>1,2</sup>

(1. College of Life Sciences, Henan Normal University, Xinxiang 453007, China;

2. College of Biology and Food Engineering, Anyang Institute of Technology, Anyang 455000, China)

**Abstract:** DNase I hypersensitive sites (DHSs) usually are locations of the cis-regulatory elements on chromatin, such as promoters, enhancers, regulators and attenuators. Study on the number and dynamic changes of DHSs among the whole genome is an important auxiliary means to understand the function of cis-regulatory elements, reveal gene expression regulation mechanism, and even create new genetic material using genetic editing. Based on the construction of DHSs library in rice and *Arabidopsis thaliana*, DHSs library construction about diploid cotton *Gossypium raimondii* ( $D_5$ ) was established, which will lay the foundation for further research on functional genomics in cotton.

**Keywords:** *Gossypium raimondii*; DNase I hypersensitive sites (DHSs); cis-regulatory elements; library construction

[责任编辑 王凤产]