

亚硫酸氢钠测序法检测水稻 *FIE* 基因 CpG 岛甲基化状态

汪艳杰, 龙 鸿, 姚家玲*

(华中农业大学生命科学技术学院, 武汉 430070)

摘 要: 建立了适用于水稻基因组特定基因甲基化检测的亚硫酸氢钠测序法, 并利用此方法对 *FIE2A* 基因 CpG 岛部分片段的甲基化差异进行了研究。采用 CTAB 法提取水稻叶片和胚乳细胞的基因组 DNA, 经亚硫酸氢钠化学修饰后, 针对已修饰的 *FIE* 基因序列设计特异引物并结合巢式 PCR 扩增, TA 载体克隆、测序, 最后对测序结果进行分析。结果表明巢式 PCR 能够增加特异性产物的产生, *FIE* 基因 CpG 岛在对称的 CG 和 CNG 位点甲基化水平较高, 而在非对称 CNN 位点甲基化水平最低, 此外在叶片中的平均甲基化水平较高。由此表明本研究建立的亚硫酸氢钠测序法适用于水稻基因组特定基因甲基化状态的检测。

关键词: 亚硫酸氢钠测序法; DNA 甲基化; 水稻; *FIE* 基因; CpG 岛

中图分类号: Q944.46

文献标识码: A

文章编号: 2095-0837(2011)01-0134-06

Analysis of Methylation at CpG Island of *FIE* in Rice by Bisulfite Sequencing

WANG Yan-Jie, LONG Hong, YAO Jia-Ling*

(College of Life Science & Technology, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: To test the feasibility of bisulfite sequencing in methylation analysis at a specific gene of rice, we analyzed DNA methylation at fragment of CpG island of *FIE2A*. In the present paper, genomic DNA was extracted from leaf and endosperm of rice with the CTAB method. After treated with bisulfite sodium, we amplified the CpG island of *FIE* gene with designed primers using nest-PCR, sequencing results after cloning in TA-vector. Results showed that nest-PCR increased the specificity of the PCR products, and symmetric sequences of CG, CNG were hypermethylated, whereas asymmetric sequences of CNN were hypomethylated. Wild leaf displayed a high average level of methylation. Our results indicated that bisulfite sequencing could be used in analysis of methylation at CpG island of specific genes in rice.

Key words: Bisulfite sequencing; DNA methylation; Rice; *FIE*; CpG island

在真核生物中, DNA 甲基化是基因表达调控研究中的热点。DNA 甲基化是通过 DNA 甲基转移酶 (DNA methyltransferase) 的催化作用从甲基供体 S-腺苷甲硫氨酸 (SAM) 中转移甲基到 C5 位胞嘧啶 (^{5m}C)、N6 位腺嘌呤 (^{6m}A) 和 N7 位鸟嘌呤 (^{7m}G) 上。^{5m}C 是最先发现的发生修饰的碱基^[1]。在 CG 二核苷酸中胞嘧啶发生甲基化是许多真核生物基因组的共同特征, 而且 5-甲基胞嘧啶是唯一存在的化学性修饰碱基。哺乳类动物的甲基化多发生在对称的 CG 序列处, 而植物的 DNA 甲基化多发生在 CG、CNG (N 表示任何碱基) 和 CNN (不对称的, N

表示 A、C、T) 等位点^[2,3]。

重亚硫酸氢钠反应原理于 1970 年提出^[4], 但当时并没有用于甲基化检测, 直到 90 年代 Susan 等和 Frommer^[5,6]才将此方法用于甲基化检测。该方法相对于以前的甲基化检测方法有很大改进, 检测效率和精度均有提高。亚硫酸氢盐法测序 (bisulfite sequencing) 的原理是, 亚硫酸氢盐处理后, DNA 序列中的胞嘧啶 (C) 发生磺化、脱氨基和脱磺基反应, 转化为尿嘧啶 (U), 而 5-甲基胞嘧啶 (^{5m}C) 保持不变, 转换后的序列经 PCR 扩增后, DNA 甲基化差异就转变为序列差异, 再对测序结果

收稿日期: 2010-03-31, 修回日期: 2010-09-09。

基金项目: 国家重点基础研究发展计划 (973 计划; 2007CB108704); 国家自然科学基金项目 (30971551)。

作者简介: 汪艳杰 (1983-), 女, 硕士研究生, 研究方向为植物分子生物学。

* 通讯作者 (Author for correspondence. E-mail: yaojlm@mail. hzau. edu. cn)。

进行分析。该方法是检测 DNA 序列中 5-甲基胞嘧啶含量应用最广泛的方法,可以在单碱基分辨率水平检测基因或基因组的甲基化状态^[7],是 DNA 序列甲基化状态分析的“金标准”。由于这种方法实验流程长,反应条件剧烈,后续扩增较困难,引物设计和 PCR 条件等在实验中均需进一步摸索。

拟南芥的 *FIE* (fertilization independent endosperm) 基因表达产物与其它几个 *FIS* (fertilization independent seed) 类基因的表达产物形成 PcG (polycomb group) 复合物,在拟南芥种子早期发育过程中调控细胞增殖,主要是通过抑制受精前中央细胞的分裂来控制胚乳的起始发育。随后遗传、分子和生化分析证实,拟南芥的 *FIE* 蛋白是多个不同复合物的重要成员,分别参与调控春化诱导的成花途径和种子发育过程^[8-10]。玉米中有 2 个 *FIE* 基因的同源物 *ZmFie1* 和 *ZmFie2*^[11],其中 *ZmFie1* 在胚乳中特异表达,而且在胚乳发育过程中一直呈现印记 (imprinting) 表达特点,其印记受 DNA 甲基化调控^[12]。水稻基因组中也有 2 个 *FIE* 基因同源物,并且与拟南芥、玉米的 *FIE* 基因编码相同的 WD40 结构域^[13],但其功能和作用机制尚不明确。我们利用亚硫酸氢盐方法检测水稻叶片和胚乳基因组 *FIE* 基因 CpG 岛部分片段的甲基化状态,对目的片段的扩增条件进行了一些优化,为进一步研究水稻 PcG 类基因的功能和作用机制奠定基础。

1 材料与方法

1.1 材料

以 Dongjin 品种水稻为材料,水稻小花颖壳张开后,用记号笔将颖壳张开的小穗点颖标记,取开花后 5 d 的小穗,将胚乳与果皮、种皮分离开来,用液氮冷冻,并保存于 -80°C 。取分蘖前期叶片,同样保存于 -80°C 。

1.2 基因组 DNA 的亚硫酸氢盐转化与转化效率的验证

1.2.1 基因组 DNA 的亚硫酸氢盐处理

(1) CTAB 法提取基因组 DNA 并纯化。取 DNA 溶液 (内含 $1\sim 2\text{ }\mu\text{g}$ DNA), PCR 仪中 94°C 变性 4 min,立即取出置冰上。将 $2\text{ }\mu\text{L}$ 3 mol/L NaOH 加入上述体系, 42°C 温浴 20~30 min。(2) 温浴期间配制亚硫酸氢钠修饰溶液。亚硫酸氢钠溶液:称取 5.4 g 亚硫酸氢钠,依次加入 8 mL 灭菌水、

400 μL 10 mol/L NaOH,避光放置。对二苯酚溶液:称取 0.088 g 对二苯酚,加入到 40 mL 水中,震荡混匀。(3) 将 340 μL 对二苯酚溶液加入到亚硫酸氢钠溶液中混匀。(4) 将上述 210 μL 混合修饰液加入到 DNA 溶液中,用矿物油覆盖, 55°C 温浴约 16 h,每 3 h 变性 4 min。

1.2.2 亚硫酸氢盐处理后 DNA 的回收、脱磺基、中和、纯化

用柱式回收试剂盒 (上海生工 UNIQ-10) 回收修饰后的 DNA。将 4.5 μL 新鲜配制的 3 mol/L NaOH 加入到上述回收的 DNA 溶液中, 37°C 脱磺基 20 min,加入 28 μL 5 mol/L 醋酸铵溶液,加入 $1\sim 2\text{ }\mu\text{L}$ $10\text{ mg}/\mu\text{L}$ 的糖原做沉淀指示剂,加入无水乙醇, -20°C 沉淀过夜,离心回收 DNA,70% 乙醇洗涤沉淀,溶于 20 μL 水中。可立即使用或保存于 -20°C 备用。

1.3 亚硫酸氢盐处理后 DNA 的扩增、转化和克隆测序

1.3.1 *FIE2A* 基因生物信息学分析

在 TIGR (<http://rice.plantbiology.msu.edu/osa1.shtml>) 数据库中,用拟南芥中已知的 *FIE* 基因为靶基因在水稻基因组数据库中调出 *FIE* 基因序列以及该序列的相关信息;*FIE* 基因有 2 个拷贝,分别为 *FIE2A* 和 *FIE2B*。应用 EMBOSS 在线软件包 (<http://www.ebi.ac.uk/Tools/emboss/cpgplot/>) 分析水稻 *FIE2A* 基因序列的 CpG 岛的数量和位置,CpG 岛的评价标准是,a: 200 bp,b:C + G 含量 $>50\%$,c:CpG 频率观察值/CpG 频率期望值 $>60\%$ ^[14]。Primer premier 5.0 软件设计引物扩增 CpG 岛的部分片段,设计引物的模板已经过 methyl primer express software v 1.0 软件的转换,转换的标准是 CG 的胞嘧啶不变,仍为胞嘧啶,非 CG 的胞嘧啶转换为 T。

1.3.2 转化效率的验证

引物序列 (武汉思特进生物工程公司合成) 如下,F-L2: 5-AATCGGATTCGGTGTAAGT-3; F-R2: 5-CTGGAACGTGTTACACTGT-3。

降落 PCR 扩增程序如下,外轮扩增: 94°C 4 min, 94°C 45 s,退火温度从 67°C $\rightarrow$ 57°C 降落 (每 2 个循环降退火温度 2°C),60 s, 72°C 60 s,每一退火温度循环 2 次,再以 57°C 循环 25 次, 72°C 10 min。

1.3.3 目标片段的扩增

引物序列如下, CG3-L1: 5'-TAAGTAG-GAAGAGGGAGTATAAGGC-3'; CG3-L2: 5'-GG-GAGTATAAGGCGTGTATAAGTTT-3'; CG3-R: 5'-ATCCACAACCTACTAAATCAAAACCC-3'。

巢式 PCR 扩增程序如下,外轮扩增:94 °C 4 min, 94 °C 45 s,60 °C 60 s,72 °C 90 s,35 个循环,72 °C 10 min。内轮扩增:94 °C 4 min,94 °C 45 s,59 °C 40 s,72 °C 30 s,40 个循环,72 °C 10 min。内轮扩增反应的模板是外轮扩增的产物稀释 100 倍。

本研究通过巢式 PCR 扩增,内轮反应采用热启动和非热启动的方法扩增。热启动方法是:在 94 °C 预变性 2 min 后,加入 *Taq* 酶,继续进行后续反应即可。

1.3.4 转化、克隆与测序

氯化钙法制备大肠杆菌 DH5 α 感受态细胞(参考《分子克隆实验指南》第三版)^[15],连接体系如下:T₄ DNA ligase buffer 1 μ L, PMD18-T 载体 (Takara 公司) 0.5 μ L, DNA 5 μ L, T₄ DNA ligase 1 μ L,灭菌水补足至 10 μ L。16 °C 连接过夜,并将连接产物转化进 DH5 α 感受态细胞。复苏后摇菌, M13F 和 M13R 引物测序鉴定(南京金斯瑞生物科技公司),每个样品挑选 6 个克隆。

1.3.5 测序结果分析

应用 NCBI 在线分析软件(<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>) 将每个克隆的测序结果去除载体和引物后,以文本格式保存,每个片段的序列信息组成独立的数据库。本文应用 Hetzl 等^[16]提出的 CyMATE 在线软件分析方法分析所测克隆信息生成的甲基化谱。由于判断胞嘧啶是否发生甲基化是分析测序结果,不能只通过一个克隆测序的结果就判定是否甲基化,每次均挑选 6 个以上的克隆测序,通过计算克隆中转化或者未转化的胞嘧啶所占的比例来表示甲基化的比例,用以表示每个样品的各个核苷酸位点的甲基化状况。采用的计算方法是:胞嘧啶甲基化比例(%) = 甲基化的胞嘧啶数目/(甲基化的胞嘧啶数目 + 非甲基化的胞嘧啶数目)。

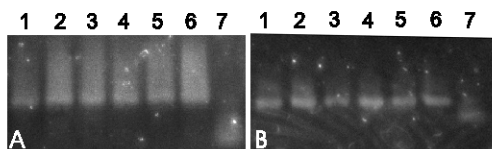
2 结果与分析

2.1 目标片段的 PCR 扩增

2.1.1 巢式 PCR 扩增结果

外轮扩增产物呈现弥散状态,没有特异性条带

的产生,而以外轮产物稀释 100 倍后作为模板,可以得到特异性条带。因此,巢式 PCR 相对于常规 PCR,可以增加产物特异性(图 1)。



A:外轮扩增产物;B:内轮扩增产物。1~3 泳道:叶片扩增产物;4~6 泳道:胚乳扩增产物;7 泳道:对照(模板用水代替)。

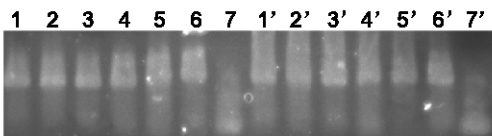
A: Amplified product of first round; B: Amplified product of second round. Line 1~3: Amplified product of leaf; Line 4~6: Amplified product of endosperm; Line 7: Control (water instead of template).

图 1 巢式 PCR 两轮扩增产物凝胶电泳分析

Fig. 1 Agarose gel electrophoresis of nest two rounds PCR amplification results

2.1.2 热启动 PCR 扩增结果

热启动可以防止在 PCR 反应第一步因引物的错配或引物二聚体的形成而导致非特异性扩增,从而提高目的片段的扩增效率。由 PCR 扩增图中可以看出(图 2),主带特异性增强,但是非特异性条带并未减少,可见是否采用热启动的方法对扩增结果并没有明显的促进作用。



1~7 泳道为没有采用热启动的扩增结果。1~3 泳道:叶片扩增产物;4~6 泳道:胚乳扩增产物;7 泳道:对照(模板用水代替)。1'~7' 泳道为采用热启动的扩增结果。1'~3' 泳道:叶片扩增产物;4'~6' 泳道:胚乳扩增产物;7' 泳道:对照(模板用水代替)。

Line 1~7: Amplified product without hot-start PCR; Line 1~3: Amplified product of leaf; Line 4~6: Amplified product of endosperm; Line 7: Control (water instead of template). Line 1'~7': Amplified product with hot-start PCR; Line 1'~3': Amplified product of leaf; Line 4'~6': Amplified product of endosperm; Line 7': Control (water instead of template).

图 2 PCR 扩增产物凝胶电泳分析

Fig. 2 Agarose gel electrophoresis of PCR

2.2 亚硫酸氢钠测序法检测不同水稻样品 *FIE2A* 基因的 CpG 岛部分片段的甲基化状态

Cymate 在线软件分析亚硫酸氢钠测序结果如图 3 所示,据此分析样品的甲基化水平、甲基化变化(表 1)。本研究所分析片段长度为 243 bp,胞嘧啶

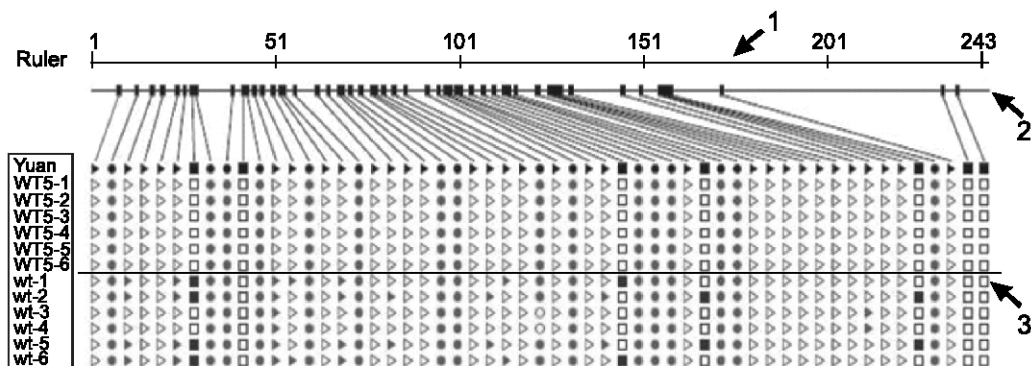


图3 Cymate 软件输出结果

Fig.3 Graphical output of CyMATE

总数 55 个,其中 CG 类型的胞嘧啶 16 个,CNG 类型的胞嘧啶 9 个,CNN 类型的胞嘧啶 30 个。

表 1 不同水稻样品 *FIE2A* 基因 CpG 岛部分片段的甲基化水平

Table 1 Percent of methylation at CG,CNG,CNN and average level of cytosine methylation of different samples in rice

材料 Samples	叶片 Leaf	授粉后 5 d 的胚乳 5 day after pollination endosperm
CGN(%)	97.92	100
CNG(%)	23.81	0
CNN(%)	15.63	0
C 甲基化平均水平 Average level of cytosine methylation	40.00	29.09

实验结果表明,CG 二核苷酸发生甲基化的比例很大,在叶片中可以达到 97.92%,授粉后 5 d 的胚乳中达到 100%,此基因片段的 CG 都处于甲基化状态;除 CG 位点之外的 C 发生甲基化的比例很小,CNG 在两种材料中分别为 23.81%、0。CNN 发生的甲基化比例更小,分别为 15.63%、0。且 CNG、CNN 中的 C 没有完全转化为 T,说明这个片段仅发生了部分甲基化。由以上结果可见,叶片和授粉后 5 d 的胚乳中 *FIE2A* 基因的部分片段 CG 胞嘧啶甲基化多发生在 CGN、CNG 以及 CNN 处,此基因片段在 CGN 处胞嘧啶发生甲基化的比例最

大,比 CNG 以及 CNN 发生的比例都要高,而在 CNN 处胞嘧啶发生甲基化的比例最低。叶片中胞嘧啶甲基化水平比授粉后 5 d 胚乳中胞嘧啶甲基化水平高,表明生殖器官中胞嘧啶表达水平比营养器官中高。

3 讨论

近年来亚硫酸氢盐测序法在技术上不断改进,包括亚硫酸氢盐的处理时间、温度等几个参数以及后期修饰产物扩增问题等;使用巢式 PCR 引物,降低模板降解,减少 PCR 误差等方法使得这项技术变得更简单^[17-19],它能够揭示每一个 DNA 序列中每一个胞嘧啶位点的甲基化状态,为表观遗传多样性图谱提供了一份高分辨率的数据。

有研究表明,基因组 DNA 在高盐高碱性高温度的处理条件下,80%~96% 的 DNA 都发生了降解^[17],模板变得很复杂,这使得引物设计变得很难,常规 PCR 扩增也困难。一般要设计多对引物进行扩增,并结合适合的 PCR 条件,最后选择出较好的引物。本研究针对扩增目的片段进行了几个方面的探索:采用巢式 PCR 扩增,增加特异性产物的产生;采用热启动的方法以观察能否增加产物的特异性。结果表明巢式 PCR 能够较好地扩增出目的片段,而采用热启动并没有增强 PCR 反应的特异性,效果欠

佳。姚群峰等^[20]也应用巢式 PCR 对扩增区域进行两次扩增,提高了检测的特异性和灵敏度。

本研究以水稻叶片和授粉后 5 d 的胚乳为材料,对 *Fie2A* 基因 CpG 岛的 DNA 序列进行甲基化差异分析,同时已有研究表明,玉米叶片中 *Fie1* DNA 序列中 CG 位点的甲基化水平为 58% ~ 97%,在 CNG 位点可以达到 50% ~ 57%^[21],本研究中水稻两种材料中 CG 位点的胞嘧啶甲基化水平达到 97% 以上,而另外两种类型的胞嘧啶甲基化水平比玉米中的同种胞嘧啶类型甲基化水平低。水稻营养器官的 *Fie2A* 基因甲基化水平比生殖器官的甲基化水平高,表明 *Fie2A* 基因在胚乳中的表达量高,这可能与 *FIE* 基因的功能相关。华中农业大学芯片数据库中显示 *Fie2A* 基因在胚乳中表达量比营养组织高,上述结果与之吻合。

由表 1 可知,营养器官中胞嘧啶甲基化水平为 40%,比生殖器官胚乳中胞嘧啶甲基化水平 29.09% 高,此外本实验室对水稻 *FIE* 突变体的研究结果发现结实率降低(数据未发表),由于 DNA 甲基化对基因表达起调控作用,可以推断,*FIE* 基因的甲基化水平与结实率降低有一定的关系。此外,在拟南芥中已经发现了 3 个 *E(z)* 同源基因: *CLF*、*MEA* 和 *SWN/EzA1*,其中 *CLF* 基因突变后导致拟南芥茎生叶和花的形态发生变化,表明 *CLF* 基因与其它多梳家族基因 *FIE* 共同调控此途径。我们对水稻中 *Fie2A* 基因的研究表明,多梳家族 *Ez1* 基因突变可以对 *Fie2A* 基因的表达产生影响,影响程度随着基因突变程度不同而不同(数据未发表),从而可以预测 *Ez1* 与 *Fie2A* 等多梳家族基因之间共同作用可能控制胚乳的发育。

甲基化研究是表观遗传学的一个重要方面,是一个复杂的调控网络。本文的甲基化研究结果对进行水稻胚乳发育过程中的 *Fie2A* 基因的表达以及与其它基因的相互作用的研究具有一定的指导作用,具体机制有待于进一步研究。

参考文献:

- [1] Hotchkiss R D. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography[J]. *J Biol Chem*, 1948, 175(1): 315–332.
- [2] Chan S W L, Henderson I R, Jacobsen S E. Gardening the genome: DNA methylation in *Arabidopsis thaliana* [J]. *Nature Reviews Genetics*, 2005, 6: 351–360.
- [3] Oakeley E J. DNA methylation analysis: A review of current methodologies [J]. *Pharmacol Ther*, 1999, 84(3): 389–400.
- [4] Shapiro R, Servis R E, Welcher M. Reactions of uracil and cytosine derivatives with sodium bisulfite: A specific deamination method [J]. *J Am Chem Soc*, 1970, 92(2): 422–424.
- [5] Susan J C, Harrison J, Paul C L, Frommer M. High sensitivity mapping of methylated cytosines [J]. *Nucl Acid R*, 1994, 22(15): 2990–2997.
- [6] Frommer M, McDonald L E, Millar D S, Collis C M, Watt F, Grigg G W, Molloy P L, Paul C L. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands [J]. *Proc Natl Acad Sci*, 1992, 89: 1827–1831.
- [7] Lister R, O'Malley R C, Tonti-Filippini J, Gregory B D, Berry C C, Millar A H, Ecker J R. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis* [J]. *Cell*, 2008, 133(3): 523–536.
- [8] Sung S, Amasino R M. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3 [J]. *Nature*, 2004, 427: 159–164.
- [9] Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon Y, Sung Z R, Goodrich J. Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis* [J]. *Development*, 2004, 131: 5263–5276.
- [10] Wang X, Ma L. Polycomb-group (Pc-G) proteins control seed development in *Arabidopsis thaliana* [J]. *J Integrat Plant Bio*, 2007, 49(1): 52–59.
- [11] Danilevskaya O N, Hermon P, Hantke S, Muszynski M G, Kollipara K, Ananiev E V. Duplicated *fie* genes in maize: Expression pattern and imprinting suggest distinct functions [J]. *Plant Cell*, 2003, 15(2): 425–438.
- [12] Gutierrez-Marcos J F, Costa L M, Pra M D, Scholten S, Kranz E, Perez P, Dickinson H G. Epigenetic asymmetry of imprinted genes in plant gametes [J]. *Nat Genet*, 2006, 38(8): 876–878.
- [13] Luo M, Platten D, Chaudhury A, Peacock W J, Dennis E S. Expression, imprinting, and evolution of rice homologs of the polycomb group genes [J]. *Molecular Plant*, 2009, 2(4): 1–13.
- [14] Gardiner-Garden M, Frommer M. CpG islands in

- vertebrate genomes[J]. *J Mol Biol*, 1987, 196(2): 261–282.
- [15] Sambrook J, Russell D W, 著. 分子克隆实验指南[M]. 第 3 版. 黄培堂, 等译. 北京: 科学出版社, 2005.
- [16] Hetzl J, Foerster A M, Raidl G, Mittelsten Scheid O. CyMATE: A new tool for methylation analysis of plant genomic DNA after bisulphite sequencing[J]. *Plant J*, 2007, 51(3): 526–536.
- [17] Grunau C, Clark S J, Rosenthal A. Bisulfite genomic sequencing: Systematic investigation of critical experimental parameters[J]. *Nucl Acid R*, 2001, 29(13): e65.
- [18] Tetzner R, Dietrich D, Distler J. Control of carry-over contamination for PCR-based DNA methylation quantification using bisulfite treated DNA[J]. *Nucl Acid R*, 2007, 35(1): e4.
- [19] Hayatsu H, Negishi K, Shiraishi M, Tsuji K, Moriyama K. Chemistry of bisulfite genomic sequencing, advances and issues[J]. *Nucl Acid Symp Ser*, 2007, 51: 47–48.
- [20] 姚群峰, 康新江, 宁勇, 唐朝晖, 周宜开. 巢式甲基化特异性聚合酶链反应在 *p16* 基因启动子过甲基化检测中的应用[J]. 中华检验医学杂志, 2004, 27(2): 89–91.
- [21] Hermon P, Srilunchang K O, Zou J, Dresselhaus T, Danilevskaya O N. Activation of the imprinted polycomb group *Fie1* gene in maize endosperm requires demethylation of the maternal allele[J]. *Plant Mol Biol*, 2007, 64: 387–395.

(责任编辑: 张 平)