

Hi-Meth: 特定位点 DNA 甲基化高通量检测平台

李慧颖^{1,2}, 刘庆², 郭旻¹, 王克剑², 严长杰¹, 王春²

- 1 扬州大学农学院 江苏省粮食作物现代产业技术协同创新中心 教育部植物功能基因组学重点实验室
江苏省作物基因组学与分子育种实验室, 江苏 扬州 225009
- 2 中国农业科学院 中国水稻研究所 水稻生物学重点实验室, 浙江 杭州 310006

李慧颖, 刘庆, 郭旻, 王克剑, 严长杰, 王春. Hi-Meth: 特定位点 DNA 甲基化高通量检测平台. 生物工程学报, 2022, 38(8): 3049-3061.

LI HY, LIU Q, GUO M, WANG KJ, YAN CJ, WANG C. Hi-Meth: a platform for high-throughput detection of site-specific DNA methylation. Chin J Biotech, 2022, 38(8): 3049-3061.

摘要: 胞嘧啶甲基化是 DNA 表观遗传修饰的主要类型之一, 在维持正常细胞功能和调控基因表达中具有重要作用。重亚硫酸盐测序法 (bisulfite sequencing PCR, BSP) 是特异性位点 DNA 甲基化检测的通用方法, 能明确目的片段中每一个 CpG 位点的甲基化状态, 但此方法需要大量的单克隆测序, 操作过程较繁琐、成本昂贵。因此, 开发准确、高效、便捷的 DNA 甲基化检测技术对提升表观遗传研究效率具有重要意义。基于本课题组开发的高通量突变类型检测平台 Hi-TOM (high-throughput tracking of mutations), 我们进一步建立了特定位点 DNA 甲基化高通量检测平台 Hi-Meth (high-throughput detection of DNA methylation)。DNA 样品通过重亚硫酸盐处理之后, 仅需一轮 PCR 扩增即可通过 Hi-Meth 平台获得特定位点 DNA 甲基化分析结果。利用 Hi-Meth 平台, 对水稻不同基因启动子区域进行了 DNA 甲基化检测分析, 并与基于 BSP 方法获得的结果进行了比较。结果表明, Hi-Meth 策略与 BSP 策略检测结果基本一致。而且通过 Hi-Meth 平台可以更准确、便捷地获得特异性位点 DNA 甲基化分析结果。综上所述, Hi-Meth 为特定 DNA 区域提供了重要的甲基化检测平台, 对表观遗传研究具有重要意义。

关键词: 5-甲基胞嘧啶; 甲基化特异位点检测; 高通量测序平台; Hi-Meth

Received: March 14, 2022; **Accepted:** May 16, 2022; **Published online:** May 19, 2022

Supported by: National Natural Science Foundation of China (31871703, 31970519); Central Public-interest Scientific Institution Basal Research Fund, China (CPSIBRF-CNRRI-202203)

Corresponding authors: WANG Chun. Tel/Fax: +86-571-63370204; E-mail: wangchun@caas.cn

YAN Changjie. Tel/Fax: +86-514-87979218; E-mail: cjyan@yzu.edu.cn

基金项目: 国家自然科学基金 (31871703, 31970519); 中央级公益性科研院所基本科研业务费专项 (CPSIBRF-CNRRI-202203)

Hi-Meth: a platform for high-throughput detection of site-specific DNA methylation

LI Huiying^{1,2}, LIU Qing², GUO Min¹, WANG Kejian², YAN Changjie¹, WANG Chun²

1 Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Jiangsu Key Laboratory of Crop Genetics and Physiology/Key Laboratory of Plant Functional Genomics of the Ministry of Education, Agricultural college of Yangzhou University, Yangzhou University, Yangzhou 225009, Jiangsu, China

2 State Key Laboratory of Rice Biology, China National Rice Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou 310006, Zhejiang, China

Abstract: Cytosine methylation is one of the major types of DNA epigenetic modifications and plays an important role in maintaining normal cell function and regulating gene expression. Bisulfite sequencing PCR (BSP) based cloning and sequencing is a general method for detecting DNA methylation at specific sites, which can clarify the methylation status of each CpG site in the target fragment. However, this method requires large amounts of single-clonal sequencing, which is complicated to operate, time consuming and expensive. Therefore, the development of an accurate, efficient and convenient DNA methylation detection technology is of great significance to improve the efficiency of epigenetic research. Based on the high-throughput mutation detection platform Hi-TOM (high-throughput tracking of mutations) developed by our group, we further established a site-specific DNA methylation high-throughput detection platform Hi-Meth (High-throughput Detection of DNA Methylation). After bisulfite treatment of DNA samples, the specific site-specific DNA methylation analysis results could be obtained through the Hi-Meth platform by performing only one round of PCR amplification. Using the Hi-Meth platform, the DNA methylation status of two promoter regions of rice were detected. The DNA methylation results from Hi-Meth were consistent with the results from BSP-based method. Thus, site-specific DNA methylation analysis results could be obtained accurately and conveniently through the Hi-Meth platform. In conclusion, Hi-Meth provides an important methylation detection platform for specific DNA regions, which has important significance for epigenetic research.

Keywords: 5-methylcytosine; detection of methylation specific site; high-throughput sequencing platform; Hi-Meth

DNA 甲基化是指将甲基选择性地添加到胞嘧啶上形成 5-甲基胞嘧啶 (5-methylcytosine, 5mC) 的过程, 其在调控基因表达、维持染色体结构和基因印记等生物学过程中都发挥着重要作用, 是一种重要的表观遗传修饰方式^[1-2]。其中, 有效的甲基化水平检测手段是表观遗传研究的关键, 可靠的甲基化信息对进一步研究表观遗传具有重要意义。

近年来, 随着人们对甲基化研究的不断深入, 根据研究内容和实验目的不同, 可以分为全基因组甲基化检测和特异位点甲基化检测两类方法^[3-4]。其中全基因组甲基化检测主要是宏观、系统地检测 DNA 的甲基化情况^[5]。特异位点甲基化检测技术则是微观、精确地检测目标片段每一个胞嘧啶的甲基化情况。目前, 特异位点甲基化检测技术主要是基于重亚硫酸盐对

DNA 的化学转化^[6], 重亚硫酸盐能将基因组中未发生甲基化的胞嘧啶 (cytosine, C) 转化为尿嘧啶 (uracil, U), 而被甲基化的胞嘧啶不能被转化, 再通过 PCR、测序或荧光等技术检测 C/T 的比例来判断 DNA 甲基化水平^[7]。其中应用最为广泛的是重亚硫酸盐测序法 (bisulfite sequencing PCR, BSP), 首先利用重亚硫酸盐处理 DNA, 再利用甲基化特异性引物对处理后的 DNA 进行 PCR 扩增, 扩增产物通过分子克隆获得多个亚克隆, 进一步通过 Sanger 测序获得特定区域中各胞嘧啶的甲基化情况^[8]。重亚硫酸盐测序法既可以定性分析也可以定量分析, 应用比较广泛^[9]。但使用这个方法需要挑选较多的单克隆, 操作繁琐, 效率低, 成本也较为昂贵^[10]。随着二代测序 (next generation sequencing, NGS) 技术的不断发展, 又开发出了重亚硫酸盐处理结合 NGS 进行甲基化水平检测的方法, 可以更精准地进行定量分析。但是, NGS 文库构建步骤繁杂、等待测序时间长以及后期数据分析复杂而阻碍了该方法的进一步应用。

Hi-TOM (high-throughput tracking of mutations) 平台^[11]是由本实验室开发的基于 NGS 的基因编辑突变检测平台, 结合了 NGS 混合文库构建的通用策略及后期数据分析, 具有操作简单、准确度高、通量高、成本低等优点。基于 Hi-TOM 平台的高通量测序方法, 结合 DNA 甲基化测序数据特点, 本研究建立了一个兼顾实验操作性和结果稳定性的甲基化检测平台 Hi-Meth (high-throughput detection of DNA methylation)。该方法使用重亚硫酸盐处理后的 DNA, 通过两轮 PCR 构建测序文库并进行 NGS 测序, 进一步通过 Hi-Meth 网站进行详细的数据分析, 最终获得甲基化水平分析结果。本文通过 Hi-Meth 法和 BSP 法对水稻基因启动子区域的 DNA 甲基化进行检测, 证明了 Hi-Meth 方法的准确性。

1 材料与方法

1.1 材料

本研究所用材料为中粳 3037 背景的突变体凤凰稻 (*phoenix, pho*)^[12]和粳稻日本晴 (*Oryza sativa* L. cv. Nipponbare)。

1.2 方法

1.2.1 特定位点甲基化基因片段的选取和引物设计

选择已发表的具有 DNA 甲基化的基因片段进行甲基化水平检测: 凤凰稻 *OsMADS1* 基因^[12]启动子区域 (LOC_Os03g11614, -1525/-1189 th, 336 bp) 和 *OsICE1* 基因^[13]启动子区域 (LOC_Os11g32100, -876/-677 th, 199 bp)。

为了测定目标区域的甲基化情况, 我们对 *OsMADS1* 基因和 *OsICE1* 基因启动子区域设计了甲基化特异性引物对 (methylation-specific primer, MP) 和甲基化不敏感引物对 (methylation-insensitive primer, MIP)。MP 的设计原则如下^[14]: (1) 选择引物时尽量避开 CG 位点。(2) 正反向引物 Tm 值尽量趋于一致, 65 °C 以上有利于特异性扩增。(3) 重亚硫酸盐转化的过程会使 DNA 片段化且形成单链, 设计扩增片段在 100–250 bp 之间。(4) 将正向引物中的 C 替换为 T, 反向引物中的 G 替换为 A。

MIP 在 MP 引物设计的基础上, 将正向引物中的 C 替换为 C/T 的简并碱基 Y, 将反向引物中的 G 替换为 G/A 的简并碱基 R^[15]。

同时, 为了适用 Hi-Meth 平台检测, 在上述引物对 5'端分别加上搭桥序列 (5'-ggagtgtacgggtgtgc-3'和 5'-gagttggatgctggatgg-3'), 用于后续 NGS 建库。

1.2.2 DNA 提取

取凤凰稻和日本晴的叶片, 采用植物基因组 DNA 提取试剂盒 (DP305, 天根生化科技有限公司) 提取高质量 DNA。取约 1 g 叶片, 加入液氮

充分研磨。将加过 0.1% 的 β -巯基乙醇的 GP1 取 700 μ L 加入叶片粉末中颠倒混匀, 65 $^{\circ}$ C 加热 20 min。等体积加入氯仿充分混匀, 12 000 r/min 室温离心 5 min, 移上清至含有 700 μ L GP2 的离心管中, 混匀。将混匀的液体移入吸附柱中, 12 000 r/min 室温离心 30 s。分别依次加入一次 500 μ L GD 和两次 600 μ L PW, 12 000 r/min 室温下空离 2 min, 室温放置数分钟晾干, 用 80 μ L 洗脱液 TE 进行洗脱。基因组 DNA 提取后用 Nano drop 测浓度, OD_{260}/OD_{280} 在 1.7–1.9 之间, 并通过 1.5% 琼脂糖凝胶电泳检测, 大小在 15 000 bp 左右的单一条带。

1.2.3 重亚硫酸盐转化及纯化回收

采用 DNA 重亚硫酸盐转化试剂盒 (DP215, 天根生化科技有限公司) 进行重亚硫酸盐转化并纯化回收。取 1 μ g 制备好的基因组 DNA, 加入试剂盒中缓冲液 DP 10 μ L, Bisulfite Mix 90 μ L, 用灭菌水补至 120 μ L 颠倒混匀放入 PCR 仪中 (95 $^{\circ}$ C 10 min, 64 $^{\circ}$ C 60 min, 4 $^{\circ}$ C 30 s) 进行重亚硫酸盐处理。处理后 DNA 按重亚硫酸盐转化试剂盒说明书进行纯化。首先, 向吸附柱 CB1 中加入 500 μ L 的平衡液 BL, 12 000 r/min 室温离心 1 min, 倒掉废液。重亚硫酸盐处理程序结束后, 短暂离心将管中反应体系转移至干净的离心

管中。加入 600 μ L 的结合液 PB, 充分混匀后加入吸附柱 CB1 中, 室温放置 2 min, 12 000 r/min 室温离心 30 s, 倒废液。加入 600 μ L 漂洗液 PW, 12 000 r/min 室温离心 30 s, 倒废液。加入 600 μ L 溶液 DB, 室温放置 15 min, 12 000 r/min 室温离心 30 s, 倒废液。加入 600 μ L 漂洗液 PW, 12 000 r/min 室温离心 30 s, 倒废液。重复上述步骤。12 000 r/min 室温下空离 2 min。室温放置数分钟至彻底晾干。取出吸附柱 CB1, 放入一个干净的离心管中, 向吸附膜中间悬空滴加 20 μ L 洗脱缓冲液 EB, 室温放置 2 min。12 000 r/min 室温离心 2 min, 收集溶液。

1.2.4 对转化并纯化回收后的目标片段进行 PCR 扩增

以重亚硫酸盐处理回收后的 DNA 为模板, 进行 PCR 扩增。反应体系为 50 μ L: 包括 2 \times Rapid Taq Master Mix (P222, 南京诺唯赞生物科技股份有限公司) 25 μ L, 10 μ mol/L 正向反向引物 (引物序列见表 1, 由浙江尚亚生物技术有限公司合成) 各 1.5 μ L, 重亚硫酸盐转化后的 DNA 500 ng 和灭菌水 20 μ L。扩增反应程序: 95 $^{\circ}$ C 3 min; 95 $^{\circ}$ C 15 s, 50 $^{\circ}$ C 15 s, 72 $^{\circ}$ C 10 s, 34 个循环; 最后 72 $^{\circ}$ C 3 min。取 5 μ L 用 1.5% 琼脂糖凝胶电泳检测 PCR 扩增条带。

表 1 引物序列

Table 1 Primes used in this study

Primer names	Primer sequences (5'→3')	Size (bp)
MADSI-MP-F	ggagtgtacggtgtgcAATATTAGTGGATGTATGAAGGAGT	43
MADSI-MP-R	gagttggatgctggatggAACTCTACCATAACCAACCAA	39
MADSI-MIP-F	ggagtgtacggtgtgcAAYAYTAGYGGATGTATGAAGGAGT	43
MADSI-MIP-R	gagttggatgctggatggAACTCTRCCATRCCAARCCAA	39
MADSI-MIP-DF-F	ggagtgtacggtgtgcTTYAATGYATGAAGGYGATYAAG	41
MADSI-MIP-DF-R	gagttggatgctggatggTTCCACRRCTCTATCTCRATTAT	43
MADSI-MIP-DR-F	ggagtgtacggtgtgcATAATYGAGATAGAGGYGGTGGAA	43
MADSI-MIP-DR-R	gagttggatgctggatggCACCARCACCRCCTTCTTACCTA	44
ICE1-MIP-F	ggagtgtacggtgtgcAAYGAATYGAYGGAYATYAA	39
ICE1-MIP-R	gagttggatgctggatggCTCCACTTAATCATRATCTCATT	41

Note: gene-MP-F/R: methylation-specific primer; gene-MIP-F/R: methylation-insensitive primer; primer sequence lowercase letters: bridging sequence for Hi-TOM platform; capital letters for primer sequences: site-specific sequence for PCR.

1.2.5 亚克隆构建及 BSP 分析

取 35 μL 上述方法 1.2.4 中的 PCR 产物用凝胶回收试剂盒进行回收, 回收片段与 1 μL 5 min TA/Blunt-Zero Cloning Kit (C601-01, 南京诺唯赞生物科技股份有限公司) 中的克隆载体进行连接, 将重组载体转化至大肠杆菌 DH5 α 菌株, 培养 14–16 h, 挑取 11–20 个阳性克隆培养后, 使用质粒小提试剂盒 (DP103, 天根生化科技公司) 提取质粒, 送往擎科生物科技有限公司进行测序。将所有测序正确的质粒整理为 FASTA 格式, 将参考序列和测序结果上传至甲基化水平在线分析工具 Kismeth (<http://katahdin.mssm.edu/kismeth>) 进行分析^[16]。

1.2.6 Hi-Meth 平台登记及测序

将 1.2.4 中获得的 PCR 产物取 10 μL 送至 Hi-TOM 平台, 第二轮 PCR 及后续测序由测序组完成。登记过程如下: (1) 进入 Hi-TOM 主页面 (<http://www.hi-tom.net/hi-tom/>), 如图 1A 所示, 快递信息及寄样时间选填, 样品及个人信息依次填好, 测序结果将发送至所填写的电子邮箱。样品数量为需要检测样品数量, 自动从 1 开始编号; 片段长度为第一轮 PCR 扩增产物的长度, 过滤阈值选择 1%, 在甲基化检测中需要选择 5 000 reads 及以上的数据量, 以保证足够的数用于分析。(2) 待检测样品的参考序列如图 1B 所示, 模板可在图 1A 中黑框标注处下载。将 FASTA 格式的参考序列上传至主页面, 确认提交。确认提交后即可获得本次任务的任务号 (USR-xxxx)。7–10 d 后可获得 NGS 结果。

在邮件收到测序结果后, 不能直接使用 Hi-TOM 分析的结果, 需要用 Hi-Meth 程序对原始测序数据进行解析。首先进入 Hi-TOM 主页面 (<http://www.hi-tom.net/hi-tom/>) 选中如图 1A 中黄框部分: 点击查看已登录样品信息, 在样品信息中选中如图 1C 黑框部分, 按任务号查

看已递交的样品信息, 找到对应的 USR-xxxx 并点击如图 1C 黄框部分中的原始数据进行下载, 通过邮件的方式将原始数据发送至登记的邮箱。如图 1D 黑框部分所示, 将后缀为 R1 和 R2 对应的原始数据在 Hi-Meth 页面 (http://www.hi-tom.net/Hi-TOM_meth/) 分别对应地上传至 Upload read 1 和 Upload read 2 中。并将参考序列上传至 Hi-Meth 对应位置中, 提交后仅需几分钟即可获得以 Excel 文件格式的甲基化分析结果。此处提交的参考序列必须严格地以特异扩增部分正向引物 5'端第一个碱基为起始、以特异扩增部分反向引物 5'端第一个碱基为结束的整个目标片段整理为 FASTA 格式, 参考序列和分析过程紧密相关, 一旦出现错误会导致分析测序结果完全错误。

2 结果与分析

2.1 Hi-Meth 平台的建立

在之前的研究中, 为了方便基因编辑突变的检测, 我们建立了一套基于 NGS 高通量突变检测技术及配套分析网站 Hi-TOM, 极大地推进了 DNA 突变情况的检测^[11]。其策略是利用两次常规 PCR 完成高通量测序文库的快速构建, 并配套在线网站自动解析出详细核苷酸序列变异信息。其中, 第一轮 PCR 主要目的是扩增目的条带, 并在目标特异性引物上加了一段搭桥序列 (正向引物 5'端加搭桥序列 5'-ggagtgtacggtgtgc-3', 反向引物 5'端加搭桥序列 5'-gagttgatgctggatgg-3'), 用于第二轮测序建库; 第二轮 PCR 则主要完成二代高通量测序的建库, 通用引物包括搭桥序列、固定条形码序列、测序平台特异性适配序列, 使得每个样本的数据都可以通过固定条形码序列进行准确地拆分分析。为了进一步方便科研人员使用, 我们对 Hi-TOM 服务平台升级优化: 用户只需

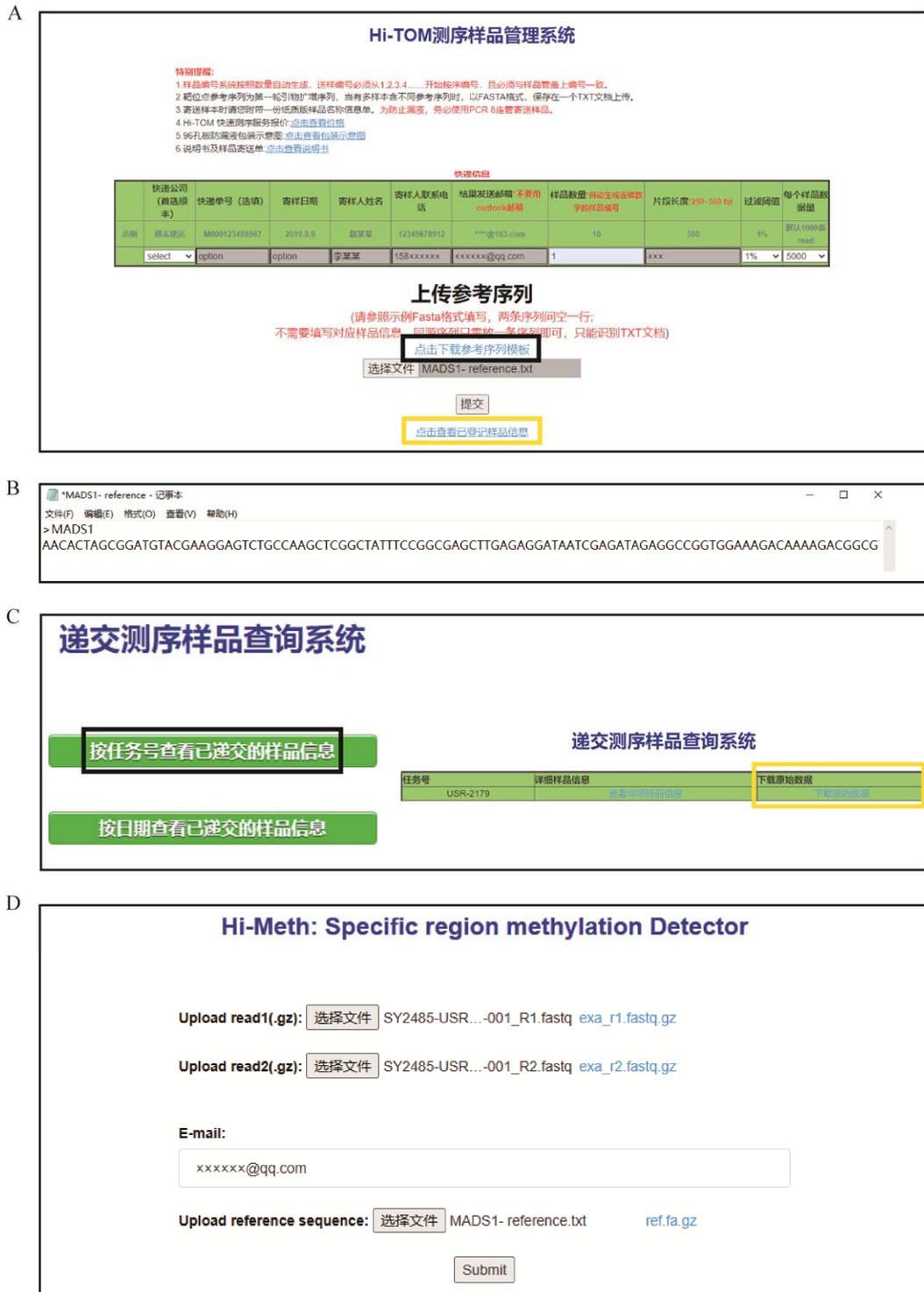


图 1 Hi-TOM 和 Hi-Meth 登记页面

Figure 1 The home page of Hi-TOM and Hi-Meth. (A) The home page of Hi-TOM. (B) The reference sequence which should be uploaded in FASTA format. (C) Download raw data according to the task number. (D) The home page of Hi-Meth, upload reads represent uploading the forward reads (named_r1.fastq.gz) and reverse reads (named_r2.fastq.gz).

要完成第一轮 PCR, 将第一轮 PCR 产物寄给 Hi-TOM 测序组, 就可以通过邮件获取分析结果, 极大地提高了 DNA 突变的鉴定效率。

在 Hi-TOM 的高通量测序平台基础上, 我们希望为特定位点甲基化检测配套对应的分析

程序, 建立一个准确、高效的特定位点 DNA 甲基化检测流程, 以提高 DNA 甲基化检测效率, 降低检测成本。

如图 2 所示, 第一步同样是用重亚硫酸盐处理 DNA, 在甲基化胞嘧啶状态不改变的情况

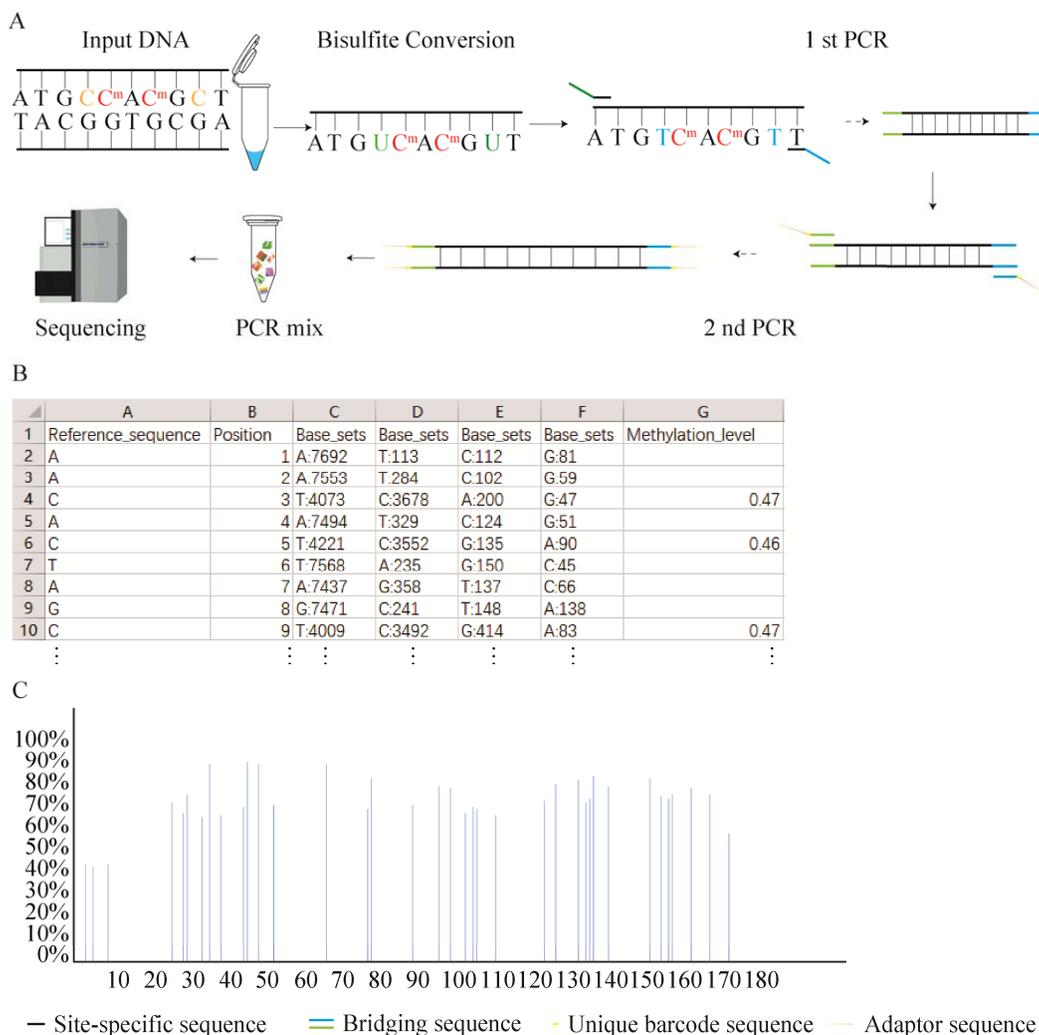


图 2 Hi-Meth 流程图

Figure 2 Schematic illustration of the workflow of Hi-Meth. (A) After DNA extraction and bisulfite treatment, the target sequences were amplified using site-specific primers (first-round PCR). The products of first-round PCR were used as templates for second-round PCR (barcoding PCR) in the 96-well plate kit. By barcoding PCR, the products of each sample were barcoded. All products of second-round PCR were pooled in equimolar amounts in a single tube and sent for NGS. (B) Example of Hi-Meth partial analysis results. Column A represents the submitted reference sequence; Column B has the first base in alignment in Position 1; Columns C–F are the number of bases measured at each position A, T, C, and G, respectively. Column G is the methylation level of the corresponding site C. Ellipses in the figure indicate sites not shown. (C) Map of DNA methylation levels. The X axis represents cytosine positions in the analyzed region, and the Y axis represents methylation ratios.

下,使未甲基化的胞嘧啶转变为尿嘧啶。针对需要检测的区间,设计特异性引物检测靶标区域的甲基化状态;并在引物上添加特定的搭桥序列,用于后续的高通量测序建库。所获得的PCR产物,经琼脂糖凝胶电泳检测确保扩增获得正确的目的条带后,将剩余的PCR产物寄往Hi-TOM测序组。Hi-TOM测序组按流程进行第二轮PCR的高通量测序建库及测序。

基于重亚硫酸盐转化后,大量C/T间的转变会导致与递交的基因组参考序列差异较大,无法直接通过Hi-TOM分析系统进行数据分析。因此,我们通过Perl脚本重新编程,对甲基化检测数据进行分析,最终以Excel表格及图输出分析结果。如图2B中为Hi-Meth分析后邮件发送的测序结果:第A列呈现的是提交的参考序列;第B列以比对上正向搭桥序列5'-ggagtgtacgggtgtgc-3'的每条reads后的第一个碱基为位置1(Position 1),以比对上反向搭桥序列5'-gagttggatgctggatgg-3'后的第一个碱基为结束;第C到第F列是每个位置A、T、C、G各碱基所测到的数目,所呈现的方式以碱基数由高到低依次排列,其中单个碱基在5%以内的占比为测序误差造成,可忽略不计。第G列为C对应位点的甲基化水平。限于篇幅,图2B中只显示了示例的前10个碱基情况。图2C中将每个位点甲基化水平可视化。其中,横坐标为参考序列对应的碱基位置,纵坐标为相应位点的甲基化水平。

2.2 Hi-Meth 甲基化检测效果验证

本研究选择凤凰稻*OsMADS1*基因^[12]启动子区为特异位点进行甲基化检测用于检验Hi-Meth方法的准确性。以*OsMADS1*启动子区中的180 bp为目标区域,分别设计了MP和MIP进行扩增(图3A, 3B),以比较不同引物设计对检测特定区域甲基化是否存在差异。其中,

MP和MIP遵循方法1.2.1中引物设计原则设计,图3A中绿色标识的为搭桥序列,黄色标识的为目标片段中的C,蓝色标识的为MP引物中C→T及G→A的变化,紫色标识的为MIP引物中C→Y及G→R的变化。

在按方法1.2.2及1.2.3对凤凰稻DNA进行重亚硫酸盐处理后,用MP(表1中*MADS1-MP-F/R*)和MIP(表1中*MADS1-MIP-F/R*)引物分别扩增*OsMADS1*基因的启动子区域得到PCR产物,经1.5%琼脂糖凝胶电泳检测,条带扩增大小正确(图3C)。取10 μL PCR产物送至Hi-Meth平台,按方法1.2.6获得甲基化分析结果。另外,取35 μL PCR产物进行亚克隆构建,挑选13个正确的单克隆,结果上传至Kismeth(<http://katahdin.mssm.edu/kismeth>)进行BSP分析。

通过Hi-Meth和BSP对目标片段中35个C位点的甲基化水平进行比较。如图3D所示,在MP扩增的目标片段中,位点3、5、9、164、169和174的C甲基化水平为0,这是由于在MP引物中已经将C都转变为T造成的,不能作为甲基化水平检测的结果。除去引物部分,其余29个C位点都检测到存在甲基化修饰,其中,Hi-Meth和BSP检测方法检测到的甲基化水平相差在0%–10%之间的共有16个位点,甲基化水平相差在10%–20%之间的有10个位点,甲基化水平相差在20%–30%之间的仅有3个位点,证明甲基化检测结果基本一致。在用MIP引物扩增的目标片段甲基化检测中,如图3E所示,由于MIP引物使用了简并碱基,这35个C位点都能检测到存在甲基化修饰,Hi-Meth和BSP检测方法检测到共25个位点的甲基化水平相差在0%–10%之间,8个位点的甲基化水平相差在10%–20%之间,2个位点的甲基化水平相差在20%–30%之间,甲基化检测结果基本一致。

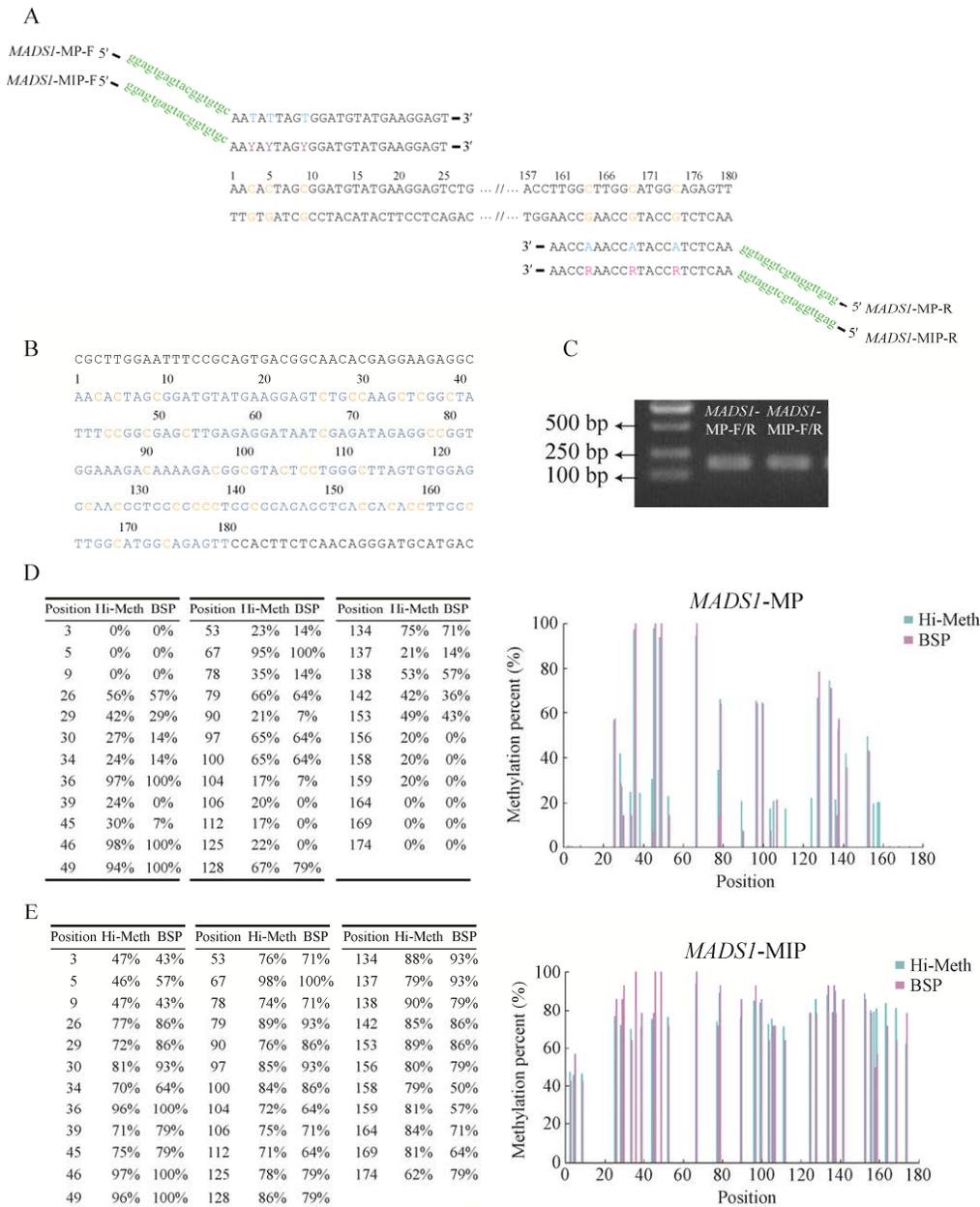


图 3 *OsMADS1* 基因启动子区域的甲基化检测

Figure 3 Methylation detection in the promoter region of *OsMADS1* gene. (A) Primer design of methylation site in *OsMADS1* promoter. The green letters represent adaptor. The yellow letter C is cytosine, which is the target of methylation. MP primer: replace all the C (yellow letter) with T (blue letter) in the primer. MIP primer: replace all the C (yellow letter) with Y (degenerate bases, purple letter) in the primer. (B) The sequences marked from 1 to 180 in *OsMADS1* promoter is the reference sequence (blue letters). The yellow letter C shows all cytosines in this region. (C) Agarose gel showing amplicons using MP or MIP primers. (D) Profiles of DNA methylation in 180 bp region using MP primer. X axis represents cytosine positions in the analyzed region, and Y axis represents methylation ratios. (E) Profiles of DNA methylation in 180 bp region using MIP primer. X axis represents cytosine positions in the analyzed region, and Y axis represents methylation ratios.

此外,将 MP 引物与 MIP 引物扩增产物的甲基化检测结果比对发现,使用 MIP 引物较 MP 引物检测到的特异位点甲基化水平高。这是由于 MIP 会匹配上更多的重亚硫酸盐处理后的 DNA 模板造成的,该结论与 Fuso 等提出的与 MIP 相比,MP 的检测能力较低相符^[17]。因此,使用 MIP 引物获得的甲基化检测结果更接近真实的甲基化水平。

2.3 MIP 引物区段甲基化检测结果不适合作为可信结果

为了验证 2.2 部分利用 MIP 引物所检测的引物部分甲基化水平是否准确,我们将 MIP 引物部分作为目标片段分别设计了检测引物对(表 1 中 *MADSI-MIP-DF-F/R* 和 *MADSI-MIP-DR-F/R*)进行扩增(图 4A、4B 和 4C),获得 PCR 产物(图 4D),同时利用 BSP 和 Hi-Meth 方法来检测位点 3、5、9、164、169 和 174 C 的甲基化水平。

如图 4E 所示,无论作为引物还是作为目标片段,Hi-Meth 与 BSP 甲基化检测结果基本一致,甲基化水平最多相差在 10%–20%之间。但是,当 *MADSI-MIP-F* 作为引物进行甲基化检测时,C3、C5 和 C9 的甲基化程度在 40%–60%之间,而当 *MADSI-MIP-F* 作为目标区段进行甲基化检测时,C3、C5 和 C9 的甲基化程度在 70%–99%之间,远远高于作为引物时的甲基化水平。同样的,当 *MADSI-MIP-R* 作为引物进行甲基化检测时,C164、C169 和 C174 的甲基化程度在 60%–80%之间,而当 *MADSI-MIP-F* 作为目标区段进行甲基化检测时,C164、C169 和 C174 的甲基化程度在 40%–75%之间,也与作为引物时的甲基化水平不符(图 4E)。根据以上结果,作为 MIP 引物部分的甲基化检测结果不能作为可信结果,仅供参考。

2.4 Hi-Meth 适用于目标片段甲基化检测

为了进一步检测 Hi-Meth 方法的可靠性,我们又以水稻日本晴 *OsICE1* 启动子区 199 bp 为目标区域设计了 MIP 引物(表 1 *ICE1-MIP-F/R*),以重亚硫酸盐处理后的 DNA 为模板进行 PCR 扩增(图 5A、5B 和 5C),分别用 Hi-Meth 和 BSP 两种方法对去除引物部分的目标片段甲基化水平进行检测,同 2.3。

OsICE1 基因启动子该目标区域一共检测到 22 个 C 位点存在甲基化修饰(图 5D)。对 Hi-Meth 和 BSP 两种检测方法进行比较发现,其中,13 个位点甲基化水平相差在 0%–10%之间,7 个位点甲基化水平相差在 10%–20%之间,2 个位点甲基化水平相差 20%–30%之间。总体可以看出 Hi-Meth 和 BSP 法两种检测结果基本一致。

3 讨论

DNA 甲基化水平的测定是表观遗传学相关研究的基础,可靠的甲基化信息对研究维持正常细胞的功能、遗传印记和胚胎发育等,具有至关重要的作用^[18]。为了能够准确、便捷、高通量地对特定区域胞嘧啶甲基化水平进行精准检测,我们基于 Hi-TOM 高通量测序平台建立了一个可以精准检测目标片段 DNA 甲基化水平的平台 Hi-Meth。

相较于动物甲基化位点仅分布于 CpG 位点,植物甲基化情况更为复杂,甲基化位点除了分布于 CpG 位点,还分布于 CHG、CHH 位点(H 代表 A、C 或 T)^[19]。为了验证 Hi-Meth 平台的可靠性,我们以植物水稻为例对水稻基因启动子区进行甲基化检测,并以具有代表性的检测方法 BSP 进行验证。我们发现二者都可以很好地检测特定位点 DNA 甲基化水平,并且两种检测结果基本一致。相较于 BSP 法得到

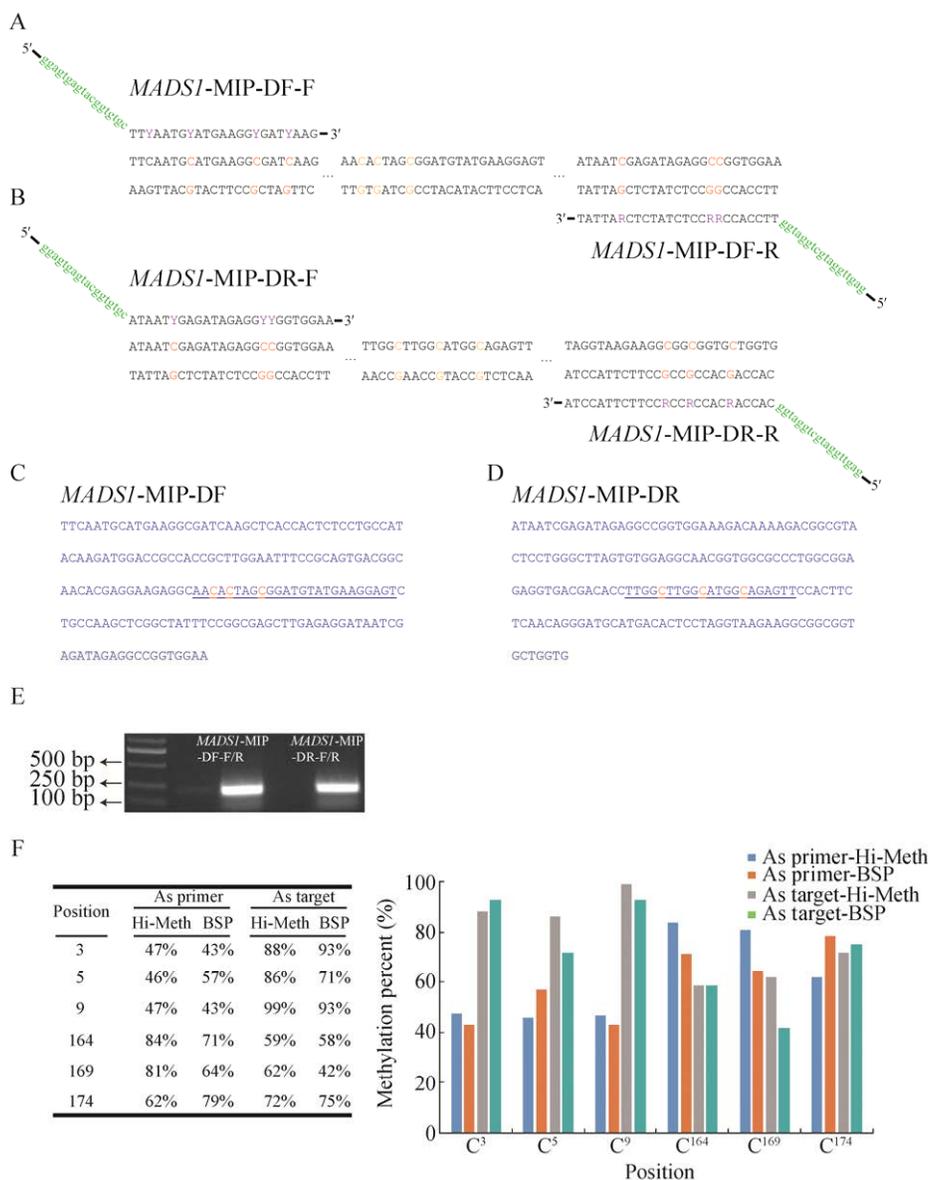


图 4 MIP 引物部分甲基化水平检测

Figure 4 Comparison of methylation level of MIP primers. (A) Schematic design of primers for methylation detection of the genome sequence of left primers for methylation detection of *OsMADS1* promoter region. The adaptor is highlighted in green, and the cytosine is marked in yellow. MIP primer: replace all the C (yellow letter) with Y (degenerate bases, purple letter) in the primer. (B) Schematic design of primers for methylation detection of the genome sequence of right primers for methylation detection of *OsMADS1* promoter region. The adaptor is highlighted in green, and the cytosine is marked in yellow. MIP primer: replace all the C (yellow letter) with Y (degenerate bases, purple letter) in the primer. (C) The amplified sequences for methylation detection of the genome sequence of left primers for methylation detection of *OsMADS1* promoter region. The targeted cytosines in left primer for methylation detection of *OsMADS1* promoter region are highlighted in yellow. (D) The amplified sequences for methylation detection of the genome sequence of right primers for methylation detection of *OsMADS1* promoter region. The targeted cytosines in right primer for methylation detection of *OsMADS1* promoter region are highlighted in yellow. (E) Agarose gel showing amplicons from MIP-DF/DR primer. (F) Methylation of MIP-DF/DR primer, which was used for methylation detection of *OsMADS1* promoter region, was detected by Hi-Meth and BSP respectively.

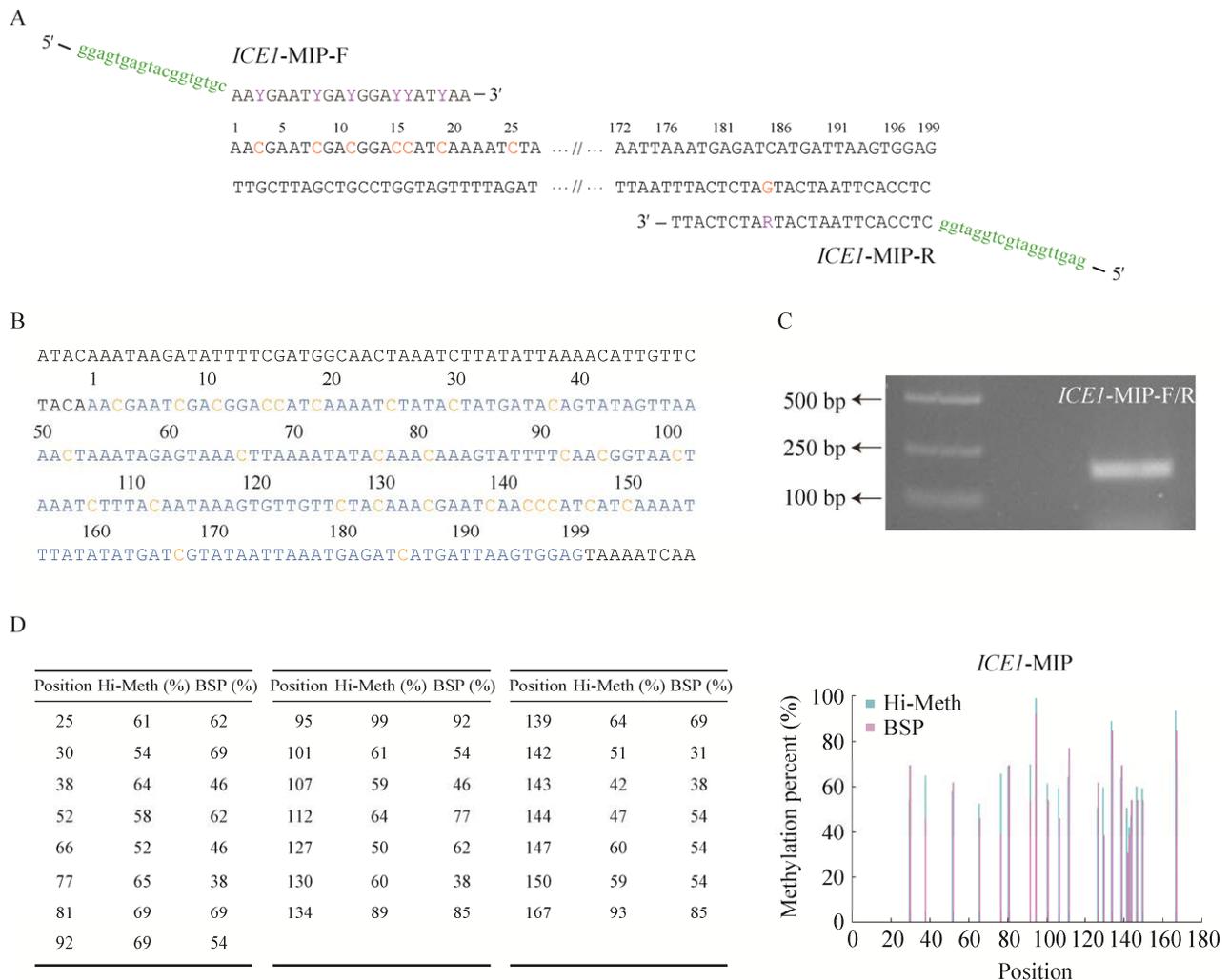


图 5 *OsICE1* 基因启动子区域的甲基化检测

Figure 5 Methylation detection in the promoter region of *OsICE1* gene. (A) Primer design of methylation site in *OsICE1* promoter. The green letters represented adaptor. The yellow letter C is cytosine, which is the target of methylated. MIP primer: replace all the C (yellow letter) with Y (purple letter) in the primer. (B) The sequences marked from 1 to 199 in *OsICE1* promoter is the reference sequence (blue letters). The yellow letter C shows all cytosines in this region. (C) Agarose gel showing amplicons from MP and MIP primer. (D) Profiles of DNA methylation in 199 bp region using MIP primer. X axis represents cytosine positions in the analyzed region, and Y axis represents methylation ratios.

PCR 产物后还需要进行亚克隆、手动整理数据等步骤，Hi-Meth 法在得到 PCR 产物后只需送至测序平台即可获得分析后的数据结果，大大减少了科研人员的工作量。此外，BSP 法的甲基化水平依赖于挑选的单克隆数目，如果挑选的单克隆数目少，会随机影响位点的甲基化水

平检测，对于低频率甲基化位点可能会有一定遗漏，需要挑选较多的克隆进行比对分析，这不仅增加了科研人员的工作量，也增加了检测成本；而 Hi-Meth 法通过高通量测序产生大量的数据，不受随机性影响，完全可以检测出低频甲基化位点的甲基化水平，仅需一次测序即

可获得所有位点可靠的甲基化水平, 因此对于大批量或低频甲基化水平的检测, Hi-Meth 的性价比更高。此外, Hi-Meth 分析程序解决了高通量测序数据分析的问题, 为不熟悉生信分析的实验人员提供了简洁清晰的甲基化检测结果。

本研究还测试了 MP 和 MIP 对甲基化水平检测的影响, 相较于 MP 引物, 使用 MIP 引物不仅可以更好地对目标片段进行扩增, 而且检测到的甲基化水平也更高、更接近于真实的甲基化水平, 建议设计甲基化检测引物时采用 MIP 引物设计策略。

Hi-TOM 测序平台每周进行两次高通量测序上机, 运行稳定, 用户包括先正达、西湖大学等在内的多家科研单位及公司。依托 Hi-TOM 测序平台, Hi-Meth 的高通量测序频率同样为每周两次, 测序速度得以保障, 无论单个还是大批量样品, 经过 7 d 即可获得真实可靠的测序结果, 为位点特异性甲基化研究提供了极大的便利。为了更好地呈现分析结果, 未来我们也会根据用户的需求继续完善和优化 Hi-Meth, 力求为科研人员提供最简洁清晰的分析结果。

REFERENCES

- [1] Meng H, Cao Y, Qin JZ, et al. DNA methylation, its mediators and genome integrity. *Int J Biol Sci*, 2015, 11(5): 604-617.
- [2] Pajares MJ, Palanca-Ballester C, Urtasun R, et al. Methods for analysis of specific DNA methylation status. *Methods*, 2021, 187: 3-12.
- [3] Li SZ, Tollefsbol TO. DNA methylation methods: global DNA methylation and methylomic analyses. *Methods*, 2021, 187: 28-43.
- [4] Khodadadi E, Fahmideh L, Khodadadi E, et al. Current advances in DNA methylation analysis methods. *Biomed Res Int*, 2021, 2021: 8827516.
- [5] Yong WS, Hsu FM, Chen PY. Profiling genome-wide DNA methylation. *Epigenetics Chromatin*, 2016, 9: 26.
- [6] Ruan QY, Zou FX, Wang Y, et al. Sensitive, rapid, and automated detection of DNA methylation based on digital microfluidics. *ACS Appl Mater Interfaces*, 2021, 13(7): 8042-8048.
- [7] Gouil Q, Keniry A. Latest techniques to study DNA methylation. *Essays Biochem*, 2019, 63(6): 639-648.
- [8] Hernández HG, Tse MY, Pang SC, et al. Optimizing methodologies for PCR-based DNA methylation analysis. *BioTechniques*, 2013, 55(4): 181-197.
- [9] Akika R, Awada Z, Mogharbil N, et al. Region of interest methylation analysis: a comparison of MSP with MS-HRM and direct BSP. *Mol Biol Rep*, 2017, 44(3): 295-305.
- [10] 姚燕丽, 黄宇哲, 韩芝斌, 等. 特异位点 DNA 甲基化常用检测技术及应用进展. *分子诊断与治疗杂志*, 2021, 13(7): 1192-1196.
Yao YL, Huang YZ, Han ZB, et al. Common detection techniques and application progress of DNA methylation at specific sites. *J Mol Diagn Ther*, 2021, 13(7): 1192-1196 (in Chinese).
- [11] Liu Q, Wang C, Jiao XZ, et al. Hi-TOM: a platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. *Sci China Life Sci*, 2019, 62(1): 1-7.
- [12] Wang KJ, Tang D, Hong LL, et al. DEP and AFO regulate reproductive habit in rice. *PLoS Genet*, 2010, 6(1): e1000818.
- [13] 韩一豪, 李新月, 解洪杰, 等. 水稻品种耐冷适应性与 ICE1 基因甲基化变异关系的研究. *农业生物技术学报*, 2017, 25(9): 1381-1390.
Han YH, Li XY, Xie HJ, et al. Adaptability of rice (*Oryza sativa*) varieties to cold tolerance associated with methylation variation of ICE1 gene. *J Agric Biotechnol*, 2017, 25(9): 1381-1390 (in Chinese).
- [14] Henderson IR, Chan SR, Cao XF, et al. Accurate sodium bisulfite sequencing in plants. *Epigenetics*, 2010, 5(1): 47-49.
- [15] Candiloro ILM, Mikeska T, Dobrovic A. Assessing alternative base substitutions at primer CpG sites to optimise unbiased PCR amplification of methylated sequences. *Clin Epigenetics*, 2017, 9: 31.
- [16] Gruntman E, Qi YJ, Slotkin RK, et al. Kismeth: analyzer of plant methylation states through bisulfite sequencing. *BMC Bioinformatics*, 2008, 9: 371.
- [17] Fuso A, Ferraguti G, Scarpa S, et al. Disclosing bias in bisulfite assay: MethPrimers underestimate high DNA methylation. *PLoS One*, 2015, 10(2): e0118318.
- [18] Moore LD, Le T, Fan GP. DNA methylation and its basic function. *Neuropsychopharmacology*, 2013, 38(1): 23-38.
- [19] Wang J, Wang CN, Long Y, et al. Universal endogenous gene controls for bisulphite conversion in analysis of plant DNA methylation. *Plant Methods*, 2011, 7: 39.

(本文责编 郝丽芳)