

DOI: 10.13376/j.cblls/2018046

文章编号: 1004-0374(2018)04-0374-09



汪海林, 理学博士, 中国科学院生态环境研究中心研究员、博士生导师, 中国科学院“百人计划”入选者, 国家“杰出青年科学基金”获得者。主要研究方向: 高灵敏 DNA 损伤和核酸表观修饰分析; 单分子水平蛋白质-DNA 相互作用分析及 DNA 同源重组机制研究; DNA 甲基化与去甲基化调控以及相关的环境健康效应。发现高等真核生物基因组 N<sup>6</sup>-甲基腺嘌呤新修饰。现已发表 140 篇学术论文, 包括 *Cell*、*Cell Stem Cell*、*Mol Cell*、*Proc Natl Acad Sci USA*、*J Am Chem Soc*、*Cell Res*、*Cell Discov*、*Nucleic Acids Res*、*Anal Chem*、*Environ Sci Technol* 等国际著名学术期刊。曾获得中国科学院院长特别奖 (1997)、中国分析测试协会科学技术特等奖 (2015), 以及中国科学院“优秀研究生导师奖” (2012、2015)、“优秀研究生指导教师奖” (2013)、“杰出成就奖”(主要完成者) (2013)。2011 年, 中国科学院“百人计划”终期考核被评为“优秀”; 2016 年, “国家杰出青年基金”结题考核被评为“优秀”。培养 2 名研究生获中国科学院研究生院院长奖学金特别奖, 1 名获中国科学院优秀博士学位论文奖。

## 真核生物基因组 DNA 甲基化和去甲基化分析

刘保东<sup>1</sup>, 汪海林<sup>1,2\*</sup>

(1 中国科学院生态环境研究中心环境化学与生态毒理学国家重点实验室, 北京 100085; 2 中国科学院大学, 北京 100049)

**摘要:** DNA 甲基化是真核生物的重要表观遗传修饰, 如胞嘧啶 C<sup>5</sup> 位甲基化 5-甲基胞嘧啶 (5mC) 和腺嘌呤 N<sup>6</sup> 位甲基化 6-甲基腺嘌呤 (6mA)。DNA 5mC 可经 Tet 双加氧酶催化氧化形成 5-羟甲基胞嘧啶 (5hmC)、5-醛甲基胞嘧啶 (5fC) 和 5-羧基胞嘧啶 (5caC)。这些氧化产物不仅是去甲基化过程的中间体, 而且也可能存在各自特有的表观调控功能。其中, 5hmC 异常可能和癌症相关, 有可能成为疾病诊断的生物标志物。发展可靠、高灵敏和抗干扰能力强的 DNA 甲基化和去甲基化检测技术和方法至关重要, 有助于理解甲基化和去甲基化的分子机制以及提高肿瘤的诊断水平。现针对 DNA 甲基化和去甲基化检测技术进行简要介绍。

**关键词:** DNA 甲基化; DNA 去甲基化; 检测方法; 5mC; 6mA; 5hmC; 5fC; 5caC; 5hmU

**中图分类号:** Q523.03; Q71 **文献标志码:** A

## Sensitive analysis of DNA methylation and demethylation intermediates for eukaryotes

LIU Bao-Dong<sup>1</sup>, WANG Hai-Lin<sup>1,2\*</sup>

(1 State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China;  
2 University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Eukaryotic DNA methylations, including methylation of C<sup>5</sup> position of cytosine 5-methylcytosine (5mC)

收稿日期: 2017-12-30

基金项目: 国家自然科学基金项目(21435008, 9174321)

\*通信作者: E-mail: hlwang@rcees.ac.cn

and N<sup>6</sup> position of adenine (6-methyladenine, 6mA), are critical and important epigenetic markers. DNA 5mC can be oxidized to form three intermediates, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), 5-carboxylcytosine (5caC), as catalyzed by Tet family dioxygenases. In addition to being as demethylation intermediates, 5hmC content is negatively correlated with tumor occurrence, and thus is a potential biomarker indicating occurrence and development of tumors. To understand these events and underlying mechanisms, it is essential to develop reliable, sensitive and interference-free methods for detection of DNA methylation and demethylation. Here we briefly review and discuss analytical and sequencing technologies for detection of epigenetic modifications involved in DNA methylation and demethylation.

**Key words:** DNA methylation; DNA demethylation; detection method; 5mC; 6mA; 5hmC; 5fC; 5caC; 5hmU

遗传信息载体 DNA 由 4 种碱基 (A、G、C、T) 组成。其中, 胞嘧啶和腺嘌呤可发生由酶催化产生的甲基化, 是表观遗传学研究的重要对象。胞嘧啶甲基化可发生在 C<sup>5</sup> 位和 N<sup>4</sup> 位, 分别形成 5-甲基胞嘧啶 (5mC) 和 4-甲基胞嘧啶 (4mC); 腺嘌呤甲基化发生在 N<sup>6</sup> 位, 形成 N<sup>6</sup>-甲基腺嘌呤 (6mA)<sup>[1-3]</sup>。5mC 广泛存在于真核生物基因组, 而 6mA 和 4mC 则主要存在于原核生物 DNA。早期研究者只在低等真核生物小球藻、衣藻和四膜虫等基因组中检测到 6mA<sup>[4-7]</sup>, 而 2015 年以来在果蝇、线虫、猪、斑马鱼等基因组中相继检测到 6mA<sup>[8-12]</sup>, 表明 6mA 可能广泛存在于真核生物基因组。目前尚无真核生物 DNA 4mC 存在的报道。

5mC 拥有相对稳定的化学结构, 如果没有经过维持性 DNMT1 的甲基化复制, 可实现被动去甲基化<sup>[13]</sup>。但是, 长期以来存在一个重要的科学问题: 是否存在主动去甲基化途径。布氏锥虫 (*Trypanosoma brucei*) 中 JBP1 (J binding protein 1) 和 JBP2 能够将 T 氧化为 5hmU, 其辅助因子为 Fe<sup>2+</sup> 和  $\alpha$  酮戊二酸<sup>[14-15]</sup>。Tahiliani 等<sup>[16]</sup>发现, 小鼠中 JBP1、JBP2 的同源蛋白 TET1 (ten-eleven translocation 1) 能将 5mC 氧化为 5hmC。TET 还能将 5hmC 继续氧化生成 5fC 和 5caC<sup>[17-18]</sup>, 从而被 TDG (thymine DNA glycosylase) 从 DNA 中去除<sup>[17]</sup>。Yin 等<sup>[19]</sup>发现, 维生素 C 在体外和体内均显著增强 TET 酶活性, 是潜在的 TET 酶辅助因子。Pfaffeneder 等<sup>[20]</sup>通过 [<sup>13</sup>C,<sup>D</sup><sub>3</sub>]-Methionine 示踪发现仅有 0.06% 的 5hmU 可能来自 5hmC 脱氨, 随后采用 [<sup>13</sup>C,<sup>15</sup>N<sub>2</sub>]-T 示踪证实 75% 的 5hmU 来自 TET 酶氧化, 另外 25% 可能由 T 和 ROS (reactive oxygen species) 反应生成。这和先前 Nabel 等<sup>[21]</sup>未检测到 AID (activation-induced cytidine deaminase) 和 APOBEC (apolipoprotein B mRNA editing enzyme catalytic subunit) 对 5hmC 的脱氨活性相一致。5hmU 则可以通过 SMUG1 (single-strand-selective mono-

functional uracil-DNA glycosylase 1)、MBD4 (methyl-CpG binding domain 4) 和 TDG 去除<sup>[22-23]</sup>。现已在线虫中发现腺嘌呤的甲基化酶 DAMT-1 (DNA N<sup>6</sup> adenine methyltransferase 1), 而在果蝇中发现的 TET 同源蛋白 DMAD (DNA 6mA demethylase) 则发挥 6mA 去甲基化作用。此外, 在小鼠胚胎干细胞 (mESCs) 中 ALKBH1 (AlkB homolog 1, histone H2A dioxygenase) 调控 6mA 的去甲基化<sup>[8,10,12]</sup>, 但体外并未重复出 ALKBH1 的 6mA 去甲基化活性。值得注意的是, 体外反应中, 如果使用过高浓度 Fe<sup>2+</sup> (1 mmol/L), 易造成假阳性结果。目前所发现的主要 DNA 甲基化和去甲基化途径如图 1 所示。

DNA 甲基化和去甲基化修饰在生物体内可能发挥各自独立的功能。哺乳动物中, 5mC 主要富集在基因组胞嘧啶-鸟嘌呤二核苷酸 (CpG) 位点<sup>[24-25]</sup>, 参与基因沉默、X 染色体失活、细胞分化和胚胎发育等过程<sup>[25-27]</sup>。6mA 在细菌限制修饰系统 (restriction-modification system)、细胞周期、DNA 复制和修复等过程发挥重要作用<sup>[28-30]</sup>。异常的 DNA 甲基化和羟甲基化通常和疾病的发生密切相关, 如急性髓细胞样白血病<sup>[31]</sup>、畸胎瘤<sup>[32]</sup>、癌症<sup>[33-35]</sup>等。而基于以下因素, 5hmC、5fC 和 5caC 可能不仅是 5mC 的氧化产物, 而且可能拥有各自特定的表观遗传功能。第一, 对 DNA 可塑性的影响 5fC>5hmC>5caC>5mC<sup>[36]</sup>, 可能通过改变 DNA 双螺旋构象, 发挥对染色体和基因表达的调控作用; 第二, 结合蛋白种类 5hmC<5mC<5fC<sup>[37]</sup>, 可能通过不同的结合蛋白来调控不同的信号通路; 第三, 5hmC、5fC、5caC 和 5mC 在基因组 DNA 上的分布存在显著差异<sup>[38-41]</sup>; 另外, 5hmC 对 RNA 聚合酶 2 的转录速率没有明显影响<sup>[42-43]</sup>, 而 5fC 和 5caC 能够抑制 RNA 聚合酶 2 的转录<sup>[42]</sup>。值得注意的是, 5hmC 脱氨或 T 经 Tet 酶氧化均可形成 5hmU, 但其表观遗传功能尚不清楚<sup>[44-45]</sup>。

对以上 DNA 修饰在基因组的丰度以及空间分

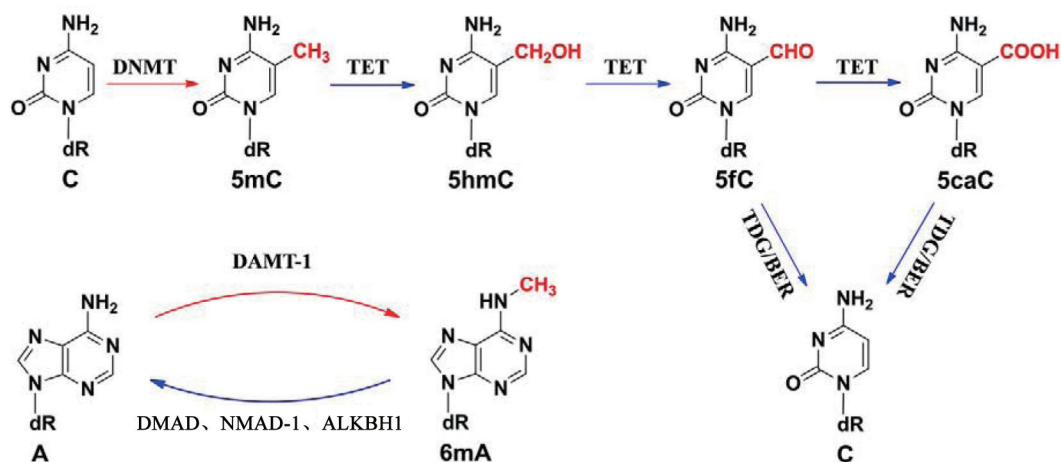


图1 真核生物基因组DNA胞嘧啶、腺嘌呤的甲基化和去甲基化

布的分析,则是深入理解其功能的重要基础。这些分析方法包括超高效液相色谱-串联质谱法(UHPLC-MS/MS)、气相色谱-串联质谱法(GC-MS/MS)、斑点杂交(dot blot)、免疫荧光成像和DNA测序技术等。限于篇幅,本文仅对胞嘧啶甲基化和去甲基化的主要检测方法进行简要介绍,重点介绍5hmU和6mA的检测方法。

## 1 胞嘧啶甲基化和去甲基化的检测方法

胞嘧啶甲基化和去甲基化的总量测定方法包括 $^{32}\text{P}$ 标记-纸层析法<sup>[16,46]</sup>、毛细管电泳法<sup>[47]</sup>、GC-MS/MS<sup>[48]</sup>、酶联免疫法<sup>[46,49]</sup>、UHPLC-MS/MS等<sup>[19,50-52]</sup>,而绘制甲基化和去甲基化修饰图谱技术主要包括重亚硫酸盐测序(bisulfite sequencing)<sup>[53-54]</sup>、氧化重亚硫酸盐测序(oxidative bisulfite sequencing, oxBS-Seq)<sup>[39]</sup>、单分子实时测序(single-molecule real-time sequencing, SMRT)等<sup>[39,52,55-58]</sup>。其中UHPLC-MS/MS具有灵敏度高、特异性好和能够提供物质的结构信息的特点,能够对DNA修饰进行定性和定量分析,而DNA测序技术能够绘制特定DNA修饰的精细图谱。下面简要介绍这两种技术。

利用UHPLC-MS/MS法测定基因组修饰的整体水平,首先需要将DNA酶解为单核苷,并根据不同核苷修饰在色谱柱上的保留能力差异进行分离,然后在离子源内离子化为带电荷粒子,在电场的作用下根据荷质比大小发生迁移,最终被高灵敏的质谱检测器接收。由于该技术利用核苷修饰的特有离子对进行检测,因而赋予了其强大的定性和定量能力。本实验室开发了一种快速、高灵敏度的UHPLC-MS/MS分析方法,在人类精子DNA中检

测到5mC水平为3.79%~4.65%,且日内和日间精确度良好(1.50%~5.70%)<sup>[59]</sup>。研究结果还发现, $\text{NH}_4\text{HCO}_3$ 能够通过增强5mC、5hmC和5fC的电喷雾离子化效率,从而显著提升检测方法的灵敏度<sup>[19,51,60]</sup>。采用 $\text{NH}_4\text{HCO}_3$ 增强的UHPLC-MS/MS法,本实验室还对环境因子(酮类、重金属离子等)对基因组5mC、5hmC水平的调控进行了研究<sup>[61-64]</sup>。Liu等<sup>[65]</sup>开发的UHPLC-ESI<sup>+</sup>-串联离子阱质谱对5hmC、5fC、5caC的定量限分别为0.056、0.098、0.14 fmol,并应用到人和鼠组织DNA去甲基化的检测。

高通量测序技术(high-throughput sequencing)具有单碱基分辨、成本相对便宜的特点,因而基于高通量测序的重亚硫酸盐测序技术(BS-Seq)被广泛用于绘制5mC的基因组分布图谱<sup>[53-54]</sup>。其原理是重亚硫酸盐能够将C脱氨转化为U,而5mC保持不变,测序时C读作T,而5mC读作C,通过与不做处理的DNA序列测序比对,从而绘制5mC在基因组上的分布图谱。随着5hmC、5fC和5caC的发现,研究证实BS-Seq不能区分5mC和5hmC,并且中间产物CMS(cytosine 5-methylenesulfonate)阻碍DNA聚合酶链式反应<sup>[66]</sup>。基于BS-Seq原理,Booth等<sup>[39]</sup>采用 $\text{KRuO}_4$ 将5hmC氧化为5fC,而5mC不被氧化,然后采用BS-Seq绘制5mC和5hmC的单碱基分辨基因图谱。Yu等<sup>[40]</sup>采用葡萄糖基化酶对5hmC进行糖基化保护,然后采用TET将5mC和5fC氧化为5caC,最后采用BS-Seq绘制5hmC的单碱基分辨基因图谱。此外,还有CAB-Seq法和redBS-Seq法,如表1所示。

此外,SMRT是一种非常有应用前景的单碱基测序技术,其原理是四种荧光标记的dNTP分子在

DNA 聚合酶催化下于纳米孔中高速延伸, 标记在磷酸上的荧光基团随着 DNA 链的延伸而脱落, 根据 dNTP 与不同修饰形式碱基在聚合反应中的动力学差异, 检测荧光持续时间, 从而直接对碱基修饰进行测序。因此, SMRT 具有超长的读长、准确度高和不需要进行 PCR 扩增等优势。2010 年, Flusberg 等<sup>[58]</sup>首次使用 SMRT 技术对实际样本 DNA 中的 6mA、5mC 和 5hmC 修饰进行了测序, 但 5mC 和 5hmC 测定不够准确。Song 等<sup>[69]</sup>和 Chavez 等<sup>[70]</sup>采用葡萄糖基化酶将 5hmC 转化为 HS-N<sub>3</sub>-5ghmC, 极大地提高了 5hmC 在 SMRT 测序中的灵敏度。Clark 等<sup>[71]</sup>发现, TET1 将 5mC 氧化为 5caC 能够加快对 5mC 的读取。SMRT 还能够用于 DNA 甲基化酶识别位点的分析<sup>[72-73]</sup>以及对高 GC 含量 DNA 的测序<sup>[74]</sup>。未来, SMRT 仍有很大发展空间。

## 2 5hmU 的检测

小鼠胚胎干细胞中, 5hmU 主要来源于 T 的 TET 氧化以及细胞内 ROS 的攻击<sup>[20]</sup>。5hmU 的检测方法主要包括 UHPLC-MS/MS 和 DNA 测序技术。Li 等<sup>[75]</sup>采用 UHPLC-ESI-MS/MS 检测了小鼠不同组织(心、肝、脾、肺等)基因组, 均未检测到 5hmU 的存在, 方法检出限为 3.5 pmol。Zarakowska 等<sup>[76]</sup>采用更高灵敏度的 UHPLC-ESI<sup>+</sup>-MS/MS 法, 定量限为 5 fmol, 在猪的肝和脑基因组中检测到 5hmU, 含量分别为 (0.6 ± 0.2) 5hmU/10<sup>6</sup> dN 和 (1.3 ± 0.3) 5hmU/10<sup>6</sup> dN。Pfaffeneder 等<sup>[20]</sup>采用 UHPLC-ESI<sup>-</sup>-MS/MS 法, 能够对低至 6.4 fmol 的 5hmU 进行定量, 约 5 个 5hmU/10<sup>7</sup> dN。Bullard 等<sup>[77]</sup>采用 Base J [5-(β-glucopyranosyl) hydroxymethyluracil] 糖基转移酶将 *Trypanosome* 基因组中 5hmU 转化为 Base J, 然后采用 JBP1 抗体富集含有 Base J 的 DNA 片段进行测序, 发现 5hmU 和转录阻遏相关。Kawasaki 等<sup>[44]</sup>采用特异性化学标记法富集 *Leishmania* 基因组含有

5hmU 的 DNA 片段进行高通量测序, 发现 5hmU 集中在链开关区域、端粒区域和基因间隔区, 而 Base J 主要富集在链开关区域, 并且二者重合度大于 90%。由于 5hmU 的 SMRT 检测信号较弱<sup>[78]</sup>, 可以将 5hmU 转化为信号更强的 Base J 进行 SMRT 测序<sup>[79]</sup>。

## 3 腺嘌呤的甲基化和去甲基化检测

1973 年, Gorovsky 等<sup>[7]</sup>采用纸层析色谱 - 放射性 <sup>3</sup>H 标记法在 *Tetrahymena pyriformis* 核 DNA 检测到 6mA。1981 年, Pratt 等<sup>[6]</sup>采用 HPLC-UV 法检测 *Tetrahymena pyriformis* 中的 6mA, 发现 6mA 主要存在于其大核 DNA, 不是随机分布, 而是主要分布在接头 DNA 上。Harrison 等<sup>[80]</sup>利用限制性内切酶 *DpnI* 能够酶切全甲基化的 G(6mA)TC 位点的特性, 发现 *Tetrahymena pyriformis* 核 DNA 上存在 GATC 特异性甲基化位点。此外, 还有毛细管电泳 - 激光诱导荧光法等 6mA 检测方法<sup>[81-82]</sup>。这些早期 6mA 研究的检测方法存在放射性危害、样品需求量大、检测灵敏度低(检测限 0.01%~0.1% 6mA/dA) 和无法提供 6mA 结构信息的缺点。2017 年, Hong 等<sup>[83]</sup>发现 Ag<sup>+</sup> 能够使 DNA 聚合酶复制时有效地发生 A-C 错配, 而 6mA-C 错配不稳定, 导致 DNA 扩增终止。他们证实在单链和双链 DNA 探针上能够单碱基水平区分 A 和 6mA; 在 Ag<sup>+</sup> 存在时, 可以通过滚环扩增实现 6mA 的定性和定量分析。当前 6mA 的检测方法主要包括抗体捕获法、UHPLC-MS/MS 法和 DNA 测序技术。

### 3.1 6mA 特异性抗体依赖的定性定量方法

6mA 抗体的点杂交<sup>[8,10,12,84-85]</sup>能够对 DNA 中 6mA 进行初步定性和半定量分析。Dabe 等<sup>[86]</sup>采用 ELISA 法对 3 种栉水母门动物 *Mnemiopsis leidyi*、*Beroe abyssicola* 和 *Pleurobrachia* 的基因组 6mA 进行了检测, 含量分别为基因组的 0.02%、0.01% 和

表1 基于高通量测序技术的5mC、5hmC、5fC和5caC的单碱基测序技术

Base	Seq	BS-Seq	oxBS-Seq	TAB-Seq	CAB-Seq	redBS-Seq
C	C	T	T	T	T	T
5mC	C	C	C	T	C	C
5hmC	C	C	T	C	C	C
5fC	C	T	T	T	T	C
5caC	C	T	T	T	C	T
	C	5mC+5hmC	5mC	5hmC	5mC+5hmC+5caC	5mC+5hmC+5fC
化合物	-	Bisulfite	KRuO <sub>4</sub>	β-GT+TET	EtONH <sub>2</sub>	NaBH <sub>4</sub>
参考文献	-	[67]	[39]	[40]	[68]	[38]

0.025%。6mA 免疫荧光成像<sup>[8,10-11,85]</sup>能够对细胞中 6mA 进行原位检测,尤其适用于少量样本分析,如斑马鱼、猪的配子以及受精卵的早期发育。上述 6mA 抗体依赖的实验对 6mA 抗体的特异性要求高,并且不能对 6mA 进行精确的定量分析。此外,6mA 抗体不能区分 RNA 的 m6A,所以必须用 RNase 充分消化,以去除 DNA 或者细胞中的 RNA。鉴于 6mA 抗体不能很好地区分 6mA 和 1mA,其仍需质谱检测来辅助定性。

### 3.2 UHPLC-MS/MS法

2015年,本课题组采用自主开发的超高灵敏度的 UHPLC-MS/MS 分析方法<sup>[8]</sup>,在模式生物果蝇基因组 DNA 中检测到 6mA (0.001%~0.07%)<sup>[8]</sup>。随后,采用高分辨飞行时间质谱对富集的 6mA 馏分进行鉴定,通过比对 6mA 馏分和 6mA 标准品的母离子、碎片离子的精确分子质量(偏差 1.02 ppm),最终确认在果蝇基因组中检测到的甲基化腺嘌呤的确是 6mA。采用 UHPLC-ESI-MS/MS 技术,在线虫、小鼠胚胎干细胞和非洲爪蟾中检测到的 6mA 含量分别为 0.01%~0.4%<sup>[10]</sup>、0.0001%~0.0005%<sup>[12]</sup>和 0.00009%<sup>[84]</sup>。此外,UHPLC-MS/MS 技术还被应用到斑马鱼、猪、人细胞、大鼠和小鼠脑等的 6mA 测定<sup>[11,87-89]</sup>。UHPLC-ESI-MS/MS 拥有极高的灵敏度,能够对痕量水平的 6mA 进行快速、准确的定量分析,能够通过色谱保留和分子质量差异区分 m6A、1mA 和 6mA,但是要防止原核生物 DNA 的卷入。

### 3.3 6mA免疫沉淀测序法(6mA-immunoprecipitation sequencing, 6mA-DIP-Seq)

此方法的原理是基于 6mA 抗体的特异性,富集 DNA 文库中含有 6mA 修饰的 DNA 片段,采用高通量测序对富集的 DNA 片段进行测序,将测序结果比对 DNA 参比序列,从而绘制 6mA 在整个基因组上的富集区域和频数图谱。而对富集序列的深度分析,还能够得出 DNA 腺嘌呤甲基转移酶的识别位点。采用 6mA-DIP-Seq 技术,本课题组发现在果蝇卵巢基因组上 6mA 倾向于富集在转座子元件区域<sup>[8]</sup>;Greer 等<sup>[10]</sup>发现线虫基因组上 6mA 均匀分布;Liu 等<sup>[11]</sup>发现在斑马鱼早期胚胎基因组 6mA 主要分布在重复序列上。由于 6mA 抗体的非特异性,如能够识别 RNA 的 m6A 以及 DNA 的 1mA 等,这些因素均会导致 6mA 基因组分布的失真。

### 3.4 限制性内切酶介导6mA测序(restriction enzyme-based 6mA sequencing, 6mA-RE-Seq)

其原理是限制性内切酶对识别序列中腺嘌呤甲

基化敏感,当该位点腺嘌呤甲基化时 DNA 不被切开,而当该位点腺嘌呤未甲基化时能够被切断,通过高通量测序可以实现甲基化位点的单碱基分辨。Luo 等<sup>[90]</sup>发现 *DpnI* 除了能够识别 G(6mA)TC 外,还能够较弱地识别 C(6mA)TC 和 G(6mA)TG,高通量测序结果显示 *Chlamydomonas*、*Plasmodium* 和 *Penicillium* 基因组中 6mA 主要以全甲基化的形式存在于回文序列中。6mA-RE-Seq 具有高灵敏、快速和低成本的特点,但是只能识别特定的限制性内切酶位点,而对其他位点的 6mA 则无法检测。

### 3.5 SMRT技术

Greer 等<sup>[10]</sup>采用 SMRT 测序,发现在线虫基因组中存在 AGAA 和 GAGG 两种甲基化特异性位点;Wu 等<sup>[12]</sup>采用 SMRT-ChIP 测序发现,小鼠胚胎干细胞基因组上 6mA 主要富集在长 LINE 元件上;Mondo 等<sup>[91]</sup>采用 SMRT 测序,发现真菌基因组中 6mA 对称分布在 ApT 二核苷酸上,并且集中在基因表达的转录起始位点周围,这和 5mC 的分布恰恰相反;而在四膜虫基因组中,6mA 大多分布在接头 DNA 的 AT 位点上<sup>[85]</sup>。然而,昂贵的测序成本限制了 SMRT 的广泛应用。另外,SMRT 无法有效区分 1mA 和 6mA,需要 UHPLC-MS/MS 来确认检测信号归属于 6mA 或 1mA。

### 3.6 稳定同位素核苷<sup>15</sup>N<sub>5</sub>-dA示踪联合UHPLC-MS/MS法

真核生物痕量 6mA 的 LC-MS 分析和抗体依赖的实验面临无处不在的细菌污染难题<sup>[84,92]</sup>。DNA 测序虽然能够识别 6mA,但是其价格昂贵,操作繁琐,不适宜快速、大量的样品分析。本课题组设想 <sup>15</sup>N<sub>5</sub>-dA 通过细胞嘌呤补救合成途径掺入 DNA,进一步发生甲基化生成 <sup>15</sup>N<sub>5</sub>-6mA,而培养基中潜在的灭活原核 DNA 以及细胞 DNA 提取、DNA 酶解过程中的原核 DNA 则不会存在 <sup>15</sup>N<sub>5</sub>-6mA 标记。利用这一原理即可通过检测 <sup>15</sup>N<sub>5</sub>-6mA 来对细胞自身产生的 6mA 进行定性和定量分析。在实验过程中,本课题组发现在 293T 细胞中 <sup>15</sup>N<sub>5</sub>-dA 完全以 <sup>15</sup>N<sub>4</sub>-dA 而不是 <sup>15</sup>N<sub>5</sub>-dA 的形式掺入基因组 DNA,意味着甲基化会生成 <sup>15</sup>N<sub>4</sub>-6mA。因此,可以通过检测 <sup>15</sup>N<sub>4</sub>-6mA 来对细胞自身产生的 6mA 进行定性和定量检测。本课题组通过 siRNA 敲低 ADA (adenosine deaminase) 基因表达水平以及 EHNA [erythro-9-(2-hydroxy-3-nonyl) adenine] 抑制 ADA 酶活性实验,证明了 ADA 介导的嘌呤补救合成途径是 <sup>15</sup>N<sub>5</sub>-dA 转化为 <sup>15</sup>N<sub>4</sub>-dA 的主要通路(图 2),暗示 <sup>15</sup>N<sub>5</sub>-dA 暴露

ADA 活性不足的原核细胞会部分或全部以  $^{15}\text{N}_5\text{-dA}$  的形式掺入 DNA。的确如此, 采用  $^{15}\text{N}_5\text{-dA}$  暴露感染猪鼻支原体的 293T 细胞, 在提取的二者 DNA 中检测到大量的 6mA 和  $^{15}\text{N}_5\text{-6mA}$ 。因此, 该方法还可以通过监测细胞中  $^{15}\text{N}_5\text{-6mA}$  的生成来指示细胞发生支原体等 ADA 活性缺乏的原核生物污染。

鉴于 RNA 中 m6A 的去甲基化形式包括 hm6A ( $\text{N}^6\text{-hydroxymethyladenosine}$ ) 和 f6A ( $\text{N}^6\text{-formyladenosine}$ )<sup>[94]</sup>, 那么真核生物 DNA 是否也存在 6mA 的去甲基化形式 6hmA 和 6fA, 这还有待解答。

#### 4 总结与展望

DNA 甲基化和去甲基化是表观遗传调控的重要组成部分, 而这些修饰在基因组 DNA 上总量和空间分布的动态变化, 是理解这些修饰发挥生理功能的基础。本文介绍了胞嘧啶甲基化 5mC 和去甲

基化 (5hmC、5fC 和 5caC)、5hmU 和 6mA 的重要检测技术, 包括总量检测技术和 DNA 测序技术两类。针对一些实验样本不足的特点, 亟需开发更高灵敏度的超高效液相色谱-质谱检测方法以及 DNA 修饰的原位检测方法, 如免疫荧光法, 以及更高灵敏度的单碱基分辨 DNA 修饰测序技术。针对不同 DNA 修饰可能具有迥异的功能及其潜在的相关性, 需要开发同时检测 DNA 修饰在基因组的丰度和分布特征的分析方法。因为一些 DNA 修饰稳定性差, 如 5hmU 和 5caC, 可以尝试将其转化为稳定的化合物形式后再分析。而一些 DNA 修饰可能存在背景干扰的情况, 需要开发特异性更好的分析方法。鉴于一些修饰既来源于酶的催化, 又来源于细胞内 ROS 产生, 开发同时检测两种来源修饰的分析方法很重要。相信随着这些样本需求少、检测指标丰富、灵敏度高和抗干扰能力强的分析方法的开发, 必将促进 DNA 表观遗传学的发展。

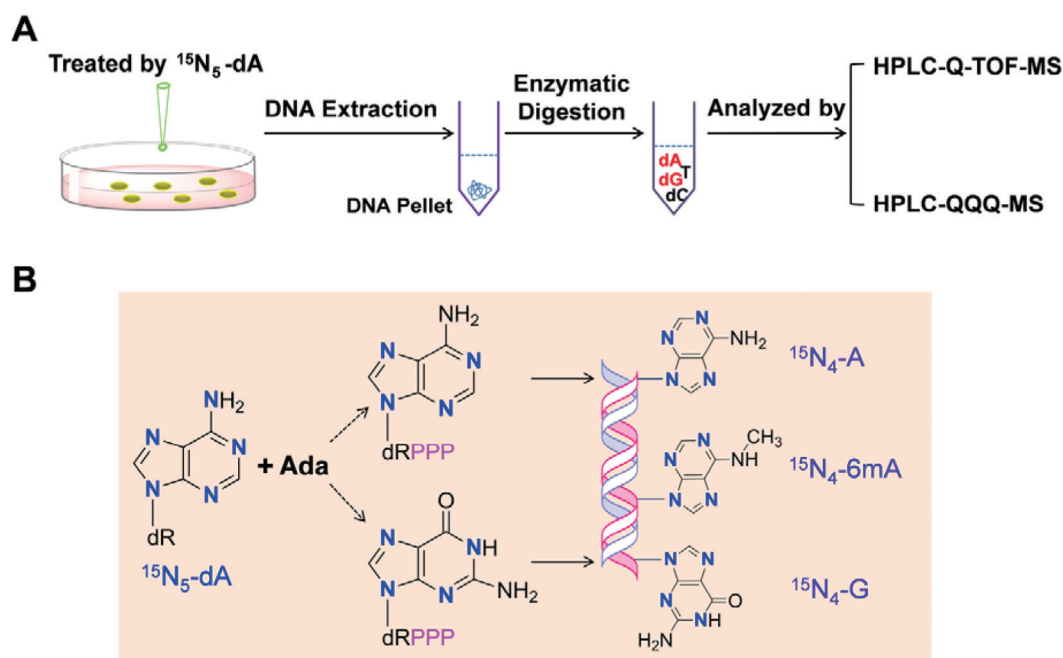


图2  $^{15}\text{N}_5\text{-dA}$  示踪检测真核细胞基因组 6mA 的流程和标记原理<sup>[93]</sup>

#### [参 考 文 献]

- [1] Cheng X. Structure and function of DNA methyltransferases. *Annu Rev Biophys Biomol Struct*, 1995, 24: 293-318
- [2] Wion D, Casades J.  $\text{N}^6\text{-methyl-adenine}$ : an epigenetic signal for DNA-protein interactions. *Nat Rev Microbiol*, 2006, 4: 183-92
- [3] Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*, 2005, 74: 481-514
- [4] Cummings DJ, Tait A, Goddard JM. Methylated bases in DNA from *Paramecium aurelia*. *Biochim Biophys Acta*, 1974, 374: 1-11
- [5] Hattman S, Kenny C, Berger L, et al. Comparative study of DNA methylation in three unicellular eucaryotes. *J Bacteriol*, 1978, 135: 1156-7
- [6] Pratt K, Hattman S. Deoxyribonucleic acid methylation and chromatin organization in *Tetrahymena thermophila*. *Mol Cell Biol*, 1981, 1: 600-8
- [7] Gorovsky MA, Hattman S, Pleger GL. [ $^{15}\text{N}$ ]methyl adenine

- in the nuclear DNA of a eucaryote, *Tetrahymena pyriformis*. *J Cell Biol*, 1973, 56: 697
- [8] Zhang G, Huang H, Liu D, et al. *N*<sup>6</sup>-methyladenine DNA modification in *Drosophila*. *Cell*, 2015, 161: 893-906
- [9] Fu Y, Luo GZ, Chen K, et al. *N*<sup>6</sup>-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. *Cell*, 2015, 161: 879-92
- [10] Greer EL, Blanco MA, Gu L, et al. DNA methylation on *N*<sup>6</sup>-Adenine in *C. elegans*. *Cell*, 2015, 161: 868-78
- [11] Liu J, Zhu Y, Luo GZ, et al. Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. *Nat Commun*, 2016, 7: 13052
- [12] Wu TP, Wang T, Seetin MG, et al. DNA methylation on *N*<sup>6</sup>-adenine in mammalian embryonic stem cells. *Nature*, 2016, 532: 329-33
- [13] Rougier N, Bourc'his D, Gomes DM, et al. Chromosome methylation patterns during mammalian preimplantation development. *Gene Dev*, 1998, 12: 2108-13
- [14] Cliffe LJ, Kieft R, Southern T, et al. JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res*, 2009, 37: 1452-62
- [15] Yu Z, Genest PA, Riet BT, et al. The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase. *Nucleic Acids Res*, 2007, 35: 2107-15
- [16] Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 2009, 324: 930-5
- [17] He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 2011, 333: 1303-7
- [18] Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 2011, 333: 1300-3
- [19] Yin R, Mao SQ, Zhao B, et al. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc*, 2013, 135: 10396-403
- [20] Pfaffeneder T, Spada F, Wagner M, et al. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat Chem Biol*, 2014, 10: 574-81
- [21] Nabel CS, Jia H, Ye Y, et al. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol*, 2012, 8: 751-8
- [22] Robertson AB, Dahl JA, Klungland A. DNA metabolism: bases of DNA repair and regulation. *Nat Chem Biol*, 2014, 10: 487-8
- [23] Sjolund AB, Senejani AG, Sweasy JB. MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. *Mutat Res*, 2013, 743-744: 12-25
- [24] Bird AP. CpG-rich islands and the function of DNA methylation. *Nature*, 1986, 321: 209-13
- [25] Wu H, Zhang Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell*, 2014, 156: 45-68
- [26] Fuks F. DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev*, 2005, 15: 490-5
- [27] Mohandas T, Sparkes R, Shapiro L. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science*, 1981, 211: 393-6
- [28] Loenen WA, Dryden DT, Raleigh EA, et al. Highlights of the DNA cutters: a short history of the restriction enzymes. *Nucleic Acids Res*, 2014, 42: 3-19
- [29] Collier J, McAdams HH, Shapiro L. A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proc Natl Acad Sci USA*, 2007, 104: 17111-6
- [30] Messer W, Noyer-Weidner M. Timing and targeting: the biological functions of Dam methylation in *E. coli*. *Cell*, 1988, 54: 735-7
- [31] Delhommeau F, Dupont S, Valle VD, et al. Mutation in *TET2* in myeloid cancers. *New Engl J Med*, 2009, 360: 2289-301
- [32] Koh KP, Yabuuchi A, Rao S, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*, 2011, 8: 200-13
- [33] Chen RZ, Pettersson U, Beard C, et al. DNA hypomethylation leads to elevated mutation rates. *Nature*, 1998, 395: 89-93
- [34] Ehrlich M. DNA methylation in cancer: too much, but also too little. *Oncogene*, 2002, 21: 5400-13
- [35] Kim YI, Giuliano A, Hatch KD, et al. Global DNA hypomethylation increases progressively in cervical dysplasia and carcinoma. *Cancer*, 1994, 74: 893-9
- [36] Ngo TT, Yoo J, Dai Q, et al. Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability. *Nat Commun*, 2016, 7: 10813
- [37] Iurlaro M, Ficiz G, Oxley D, et al. A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol*, 2013, 14: R119
- [38] Booth MJ, Marsico G, Bachman M, et al. Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat Chem*, 2014, 6: 435-40
- [39] Booth MJ, Branco MR, Ficiz G, et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, 2012, 336: 934-7
- [40] Yu M, Hon GC, Szulwach KE, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, 2012, 149: 1368-80
- [41] Lu X, Han D, Zhao BS, et al. Base-resolution maps of 5-formylcytosine and 5-carboxylcytosine reveal genome-wide DNA demethylation dynamics. *Cell Res*, 2015, 25: 386
- [42] Kellinger MW, Song CX, Chong J, et al. 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol*, 2012, 19: 831-3
- [43] Williams K, Christensen J, Pedersen MT, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, 2011, 473: 343-8
- [44] Kawasaki F, Beraldi D, Hardisty RE, et al. Genome-wide

- mapping of 5-hydroxymethyluracil in the eukaryote parasite *Leishmania*. *Genome Biol*, 2017, 18: 23
- [45] Olinski R, Starczak M, Gackowski D. Enigmatic 5-hydroxymethyluracil: Oxidatively modified base, epigenetic mark or both? *Mutat Res Rev Mutat Res*, 2016, 767: 59-66
- [46] Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, 2009, 324: 929-30
- [47] Wang X, Song Y, Song M, et al. Fluorescence polarization combined capillary electrophoresis immunoassay for the sensitive detection of genomic DNA methylation. *Anal Chem*, 2009, 81: 7885-91
- [48] Tang Y, Gao XD, Wang Y, et al. Widespread existence of cytosine methylation in yeast DNA measured by gas chromatography/mass spectrometry. *Anal Chem*, 2012, 84: 7249-55
- [49] Zhong S, Li Z, Jiang T, et al. Immunofluorescence imaging strategy for evaluation of the accessibility of DNA 5-hydroxymethylcytosine in chromatin. *Anal Chem*, 2017: 5702-6
- [50] Zhang N, Liu B, Wang H. Technology and methodology of DNA adducts analysis. *Sci Sin Chim*, 2017, 47: 1432-44
- [51] Yin R, Mo J, Lu M, et al. Detection of human urinary 5-hydroxymethylcytosine by stable isotope dilution HPLC-MS/MS analysis. *Anal Chem*, 2015, 87: 1846-52
- [52] Zhao C, Wang H, Zhao B, et al. Boronic acid-mediated polymerase chain reaction for gene- and fragment-specific detection of 5-hydroxymethylcytosine. *Nucleic Acids Res*, 2014, 42: e81
- [53] Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, 2008, 454: 766-70
- [54] Cokus SJ, Feng S, Zhang X, et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, 2008, 452: 215-9
- [55] Zhao C, Wang H. The progress on sequencing and detection of hydroxymethylated DNA. *Acta Chim Sin*, 2013, 71: 26-35
- [56] Song CX, Szulwach KE, Dai Q, et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, 2013, 153: 678-91
- [57] Shen L, Wu H, Diep D, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell*, 2013, 153: 692-706
- [58] Flusberg BA, Webster DR, Lee JH, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods*, 2010, 7: 461-5
- [59] Wang X, Suo Y, Yin R, et al. Ultra-performance liquid chromatography/tandem mass spectrometry for accurate quantification of global DNA methylation in human sperms. *J Chromatogr B*, 2011, 879: 1647-52
- [60] Yin R, Liu S, Zhao C, et al. An ammonium bicarbonate-enhanced stable isotope dilution UHPLC-MS/MS method for sensitive and accurate quantification of acrolein-DNA adducts in human leukocytes. *Anal Chem*, 2013, 85: 3190-7
- [61] Yin R, Mo J, Dai J, et al. Nickel(II) inhibits Tet-mediated 5-methylcytosine oxidation by high affinity displacement of the cofactor iron(II). *ACS Chem Biol*, 2017, 12: 1494-8
- [62] Li C, Liu B, Zhong S, et al. MEK inhibitor PD0325901 and vitamin C synergistically induce hypomethylation of mouse embryonic stem cells. *Oncotarget*, 2016, 7: 39730-9
- [63] Li C, Wang F, Wang H. Tetrachloro-1,4-benzoquinone induces apoptosis of mouse embryonic stem cells. *J Environ Sci (China)*, 2017, 51: 5-12
- [64] Yin R, Zhang D, Song Y, et al. Potent DNA damage by polyhalogenated quinones and H<sub>2</sub>O<sub>2</sub> via a metal-independent and Intercalation-enhanced oxidation mechanism. *Sci Rep*, 2013, 3: 1269
- [65] Liu S, Wang J, Su Y, et al. Quantitative assessment of Tet-induced oxidation products of 5-methylcytosine in cellular and tissue DNA. *Nucleic Acids Res*, 2013, 41: 6421-9
- [66] Huang Y, Pastor WA, Shen Y, et al. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One*, 2010, 5: e8888
- [67] Meissner A, Gnirke A, Bell GW, et al. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*, 2005, 33: 5868-77
- [68] Lu X, Song CX, Szulwach K, et al. Chemical modification-assisted bisulfite sequencing (CAB-Seq) for 5-carboxylcytosine detection in DNA. *J Am Chem Soc*, 2013, 135: 9315-7
- [69] Song CX, Clark TA, Lu XY, et al. Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. *Nat Methods*, 2011, 9: 75-7
- [70] Chavez L, Huang Y, Luong K, et al. Simultaneous sequencing of oxidized methylcytosines produced by TET/JBP dioxygenases in *Coprinopsis cinerea*. *Proc Natl Acad Sci USA*, 2014, 111: E5149-58
- [71] Clark TA, Lu X, Luong K, et al. Enhanced 5-methylcytosine detection in single-molecule, real-time sequencing via Tet1 oxidation. *BMC Biol*, 2013, 11: 4
- [72] Fang G, Munera D, Friedman DI, et al. Genome-wide mapping of methylated adenine residues in pathogenic *Escherichia coli* using single-molecule real-time sequencing. *Nat Biotechnol*, 2012, 30: 1232-9
- [73] Zhu L, Zhong J, Jia X, et al. Precision methylome characterization of *Mycobacterium tuberculosis* complex (MTBC) using PacBio single-molecule real-time (SMRT) technology. *Nucleic Acids Res*, 2016, 44: 730-43
- [74] Shin SC, Ahn DH, Kim SJ, et al. Advantages of single-molecule real-time sequencing in high-GC content genomes. *PLoS One*, 2013, 8: e68824
- [75] Li W, Liu M. Distribution of 5-hydroxymethylcytosine in different human tissues. *J Nucleic Acids*, 2011, 3: 870726
- [76] Zarakowska E, Gackowski D, Foksinski M, et al. Are 8-oxoguanine (8-oxoGua) and 5-hydroxymethyluracil (5-hmUra) oxidatively damaged DNA bases or transcription (epigenetic) marks? *Mutat Res Genet Toxicol Environ Mutagen*, 2014, 764-765: 58-63
- [77] Bullard W, Kieft R, Sabatini R. A method for the efficient and selective identification of 5-hydroxymethyluracil in



- genomic DNA. *Biol Methods Protoc*, 2017, 2: bpw006-bpw
- [78] Clark TA, Spittle KE, Turner SW, et al. Direct detection and sequencing of damaged DNA bases. *Genome Integr*, 2011, 2: 1-9
- [79] Genest PA, Baugh L, Taipale A, et al. Defining the sequence requirements for the positioning of base in DNA using SMRT sequencing. *Nucleic Acids Res*, 2015, 43: 2102-15
- [80] Harrison GS, Findly RC, Karrer K. Site-specific methylation of adenine in the nuclear genome of a eucaryote, *Tetrahymena thermophila*. *Mol Cell*, 1986, 6: 2364-70
- [81] Kraiss AM, Cornelius MG, Schmeiser HH. Genomic  $N^6$ -methyladenine determination by MEKC with LIF. *Electrophoresis*, 2010, 31: 3548-51
- [82] Dohno C, Shibata T, Nakatani K. Discrimination of  $N^6$ -methyl adenine in a specific DNA sequence. *Chem Commun (Camb)*, 2010, 46: 5530-2
- [83] Hong T, Yuan Y, Wang T, et al. Selective detection of  $N^6$ -methyladenine in DNA *via* metal ion-mediated replication and rolling circle amplification. *Chem Sci*, 2017, 8: 200-5
- [84] Koziol MJ, Bradshaw CR, Allen GE, et al. Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications. *Nat Struct Mol Biol*, 2016, 23: 24-30
- [85] Wang Y, Chen X, Sheng Y, et al.  $N^6$ -adenine DNA methylation is associated with the linker DNA of H2A. Z-containing well-positioned nucleosomes in Pol II-transcribed genes in *Tetrahymena*. *Nucleic Acids Res*, 2017, 45: 11594-606
- [86] Dabe EC, Sanford RS, Kohn AB, et al. DNA methylation in basal metazoans: insights from ctenophores. *Integr Comp Biol*, 2015, 55: 1096-110
- [87] Yao B, Cheng Y, Wang Z, et al. DNA  $N^6$ -methyladenine is dynamically regulated in the mouse brain following environmental stress. *Nat Commun*, 2017, 8: 1122
- [88] Wang Y, Wang Y, Sheng Y, et al. A comparative study of genome organization and epigenetic mechanisms in model ciliates, with an emphasis on *Tetrahymena*, *Paramecium* and *Oxytricha*. *Eur J Protistol*, 2017, 61: 376-87
- [89] Liang D, Wang H, Song W, et al. The decreased  $N^6$ -methyladenine DNA modification in cancer cells. *Biochem Biophys Res Commun*, 2016, 480: 120-5
- [90] Luo GZ, Wang F, Weng X, et al. Characterization of eukaryotic DNA  $N^6$ -methyladenine by a highly sensitive restriction enzyme-assisted sequencing. *Nat Commun*, 2016, 7: 11301
- [91] Mondo SJ, Dannebaum RO, Kuo RC, et al. Widespread adenine  $N^6$ -methylation of active genes in fungi. *Nat Genet*, 2017, 49: 964-8
- [92] Luo GZ, Blanco MA, Greer EL, et al. DNA  $N^6$ -methyladenine: a new epigenetic mark in eukaryotes? *Nat Rev Mol Cell Biol*, 2015, 16: 705-10
- [93] Liu B, Liu X, Lai W, et al. Metabolically generated stable isotope-labeled deoxynucleoside code for tracing DNA  $N^6$ -methyladenine in human cells. *Anal Chem*, 2017, 89: 6202-9
- [94] Fu Y, Jia G, Pang X, et al. FTO-mediated formation of  $N^6$ -hydroxymethyladenosine and  $N^6$ -formyladenosine in mammalian RNA. *Nat Commun*, 2013, 4: 1798