

DOI: 10.13376/j.cblls/2017040

文章编号: 1004-0374(2017)03-0302-08



钱伟强, 北京大学生命科学学院研究员, 博士生导师。2007年在中国科学院遗传与发育生物学研究所获得博士学位, 并先后在康奈尔大学、加州大学河滨分校和普渡大学从事博士后研究。2013年入选中组部“青年千人计划”, 2015年获得国家自然科学基金委“优秀青年科学基金”。主要从事植物基因组DNA主动去甲基化的分子机制及植物基因反沉默的表观遗传学机制研究。相关研究发表在 *Science*、*Molecular Cell*、*PLoS Genetics* 等杂志。

植物中DNA甲基化及去甲基化研究进展

李 艳, 钱伟强*

(北京大学生命科学学院, 北京 100871)

摘 要: DNA 甲基化及去甲基化是生物体调控基因表达的重要机制。在植物中, DNA 甲基化的建立主要通过依赖于小 RNA 的甲基化途径来实现, 最近的研究发现了 DNA 甲基化最初起始的新机制。DNA 主动去甲基化由 ROS1/DME 家族介导并通过碱基切除修复的通路来实现。最新的研究发现了 ROS1 下游参与碱基切除修复的多个关键因子。更重要的是, 越来越多的参与调控 DNA 主动去甲基化的特异性的因子及其作用机制被发现。现介绍植物体中 DNA 甲基化建立和维持的机制, 并着重讨论植物体中 DNA 主动去甲基化方向的最新进展。

关键词: DNA 甲基化; DNA 主动去甲基化; RdDM; ROS1

中图分类号: Q943.2 **文献标志码:** A

Mechanisms of DNA methylation and demethylation in plants

LI Yan, QIAN Wei-Qiang*

(School of Life Sciences, Peking University, Beijing 100871)

Abstract: DNA methylation and demethylation are important mechanisms regulating gene expression. In *Arabidopsis thaliana*, the establishment of DNA methylation is achieved by the RNA-directed DNA methylation (RdDM) pathway. Recent studies revealed mechanisms that trigger the initiation of the canonical RdDM pathway. Active DNA demethylation in *A. thaliana* is mediated by the ROS1 subfamily of 5-methylcytosine-specific DNA glycosylases via the base excision repair mechanism. Recent studies identified key factors that catalyze the downstream steps of active DNA demethylation. More importantly, multiple factors that regulate site-specific targeting of ROS1 were identified. This review summarizes mechanisms of DNA methylation and also discusses recent advances in active DNA demethylation.

Key words: DNA methylation; active DNA demethylation; RdDM; ROS1

收稿日期: 2016-04-11

基金项目: 国家自然科学基金项目(31571326, 31522005)

*通信作者: E-mail: wqqian@pku.edu.cn; Tel: 010-62768230

DNA 甲基化发生于胞嘧啶第五位碳原子上, 是一种很重要的表观遗传学修饰。在哺乳动物细胞中, DNA 甲基化主要发生于 CG 位点; 而在植物中, DNA 甲基化发生于 CG、CHG、CHH 位点 (H 代表除鸟嘌呤之外的其他碱基)。在哺乳动物细胞中, 60%~90% 的 CG 位点被甲基化^[1]; 而在高等植物细胞中, 被甲基化的胞嘧啶最多可以占到总胞嘧啶数的 50%^[2]。发生于基因启动子区域的 DNA 甲基化通常起到抑制基因转录的作用^[3]。DNA 甲基化在维持基因组稳定性方面起到重要作用, 同时, DNA 甲基化调控很多关键的生物学过程, 如抑制印记基因表达、使 X 染色体失活、控制基因的组织特异性表达等^[4]。

1 植物体中DNA甲基化的建立及维持

在植物中, DNA 在各个位点的甲基化的建立是通过一条依赖于小 RNA 的甲基化途径来实现的^[4]。在这条途径中, 植物特有的 RNA 聚合酶 Pol IV 首先转录出 30~40 nt 长度的转录本^[5], 然后, RNA 聚合酶 RDR2 以之为模板合成双链的 RNA^[6]。之前人们设想 RDR2 能够指数型扩增 Pol IV 转录本, 从而以很少数量的 Pol IV 转录本产生很多的双链 RNA。但最近的研究证实, Pol IV 的活性与 RDR2 的活性高度一致, 并且一条 Pol IV 转录本对应于一条双链 RNA^[5,7]。切割酶 DCL3 切割 RDR2 产生的双链 RNA 产生长度为 24 nt 的小 RNA。双链小 RNA 的一条链结合 AGO4 并与 Pol V 产生的长链非编码转录本发生碱基配对^[8]。通过碱基配对, AGO4 和小 RNA 形成的复合体招募 DNA 甲基转移酶 DRM2 (哺乳动物细胞中 DNMT3 的同源蛋白)

到靶位点, 催化 DNA 甲基化的建立^[9-10]。

由于这条依赖于小 RNA 的甲基化途径最初由 Pol IV 和 Pol V 介导, DNA 甲基化建立的特异性取决于 Pol IV 和 Pol V 的特异性招募。之前的研究显示两个 Pol IV 的相互作用蛋白 CLSY1 和 DTF1 能够识别第 9 位赖氨酸被甲基化的组蛋白 H3, 并帮助 Pol IV 定位到这个位点^[11-13]。SUVH2 和 SUVH9 通过与 Pol V 的相互作用蛋白复合体 DMS3/DRD1/RDM1 (DDR) 结合, 介导 Pol V 被招募到已经有较低程度 DNA 甲基化的位点^[14-17]。这样的话, 这条经典的依赖于小 RNA 的甲基化途径其实是依赖于 DNA 序列上一些已经建立好的表观遗传学标记。但究其根本, 那些已经建立好的 DNA 甲基化修饰最初是如何建立起来的, 这个问题是研究 DNA 甲基化的难点。由于一些转座子元件的 DNA 甲基化依赖于 RNA 聚合酶 Pol II 以及 RDR6^[18-19], 人们认为 Pol II 和 RDR6 产生的 RNA 前体可以被 DCL2 和 DCL4 切割生成 21~22 nt 小 RNA。这些小 RNA 负责起始较低水平的 DNA 甲基化, 从而激发经典的依赖于 24 nt 小 RNA 的甲基化途径^[19-21] (图 1)。Mari-Ordonez 等^[22] 的报道表明, DCL3 可以切割 Pol II 和 RDR6 产生的 RNA 前体产生 24 nt 小 RNA, 这些小 RNA 起始 DNA 甲基化。2016 年, Ye 等^[23] 研究发现, 部分不依赖于 DCL 蛋白切割而产生的小 RNA 有可能参与 DNA 甲基化的起始。虽然大多数不依赖于 DCL 蛋白切割而产生的小 RNA 由 Pol IV 转录产生, 参与 DNA 甲基化起始的这类小 RNA 的前体是由 Pol II 转录, 然后 Pol II 转录本可能被 RDR6 转变成双链 RNA。双链 RNA 中的一条链与 AGO4 结合并被一类有削减活性的 3'-5' 核酸外切酶

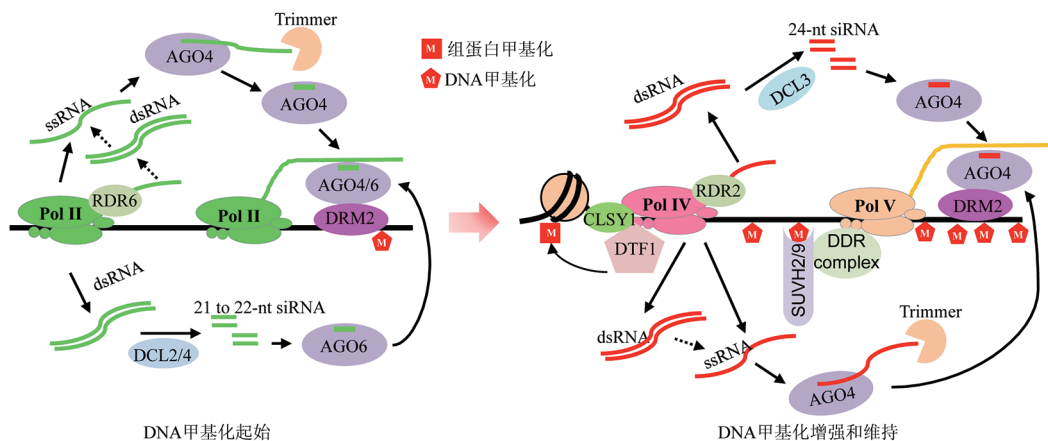


图1 植物中DNA甲基化的起始和维持

削减，从而产生长度呈梯度分布的 20~60 nt 小 RNA^[23]。这些不依赖于 DCL 蛋白切割产生的小 RNA 在没有任何甲基化修饰的位点从头建立 DNA 甲基化，为之后参与经典的甲基化建立途径的蛋白质的招募提供标记 (图 1)。

一旦 DNA 甲基化建立，在 CG 位点发生的甲基化由 DNA 甲基转移酶 MET1 (哺乳动物细胞中 DNMT1 的同源蛋白) 来维持，在 CHG 位点的甲基化由 DNA 甲基转移酶 CMT3 来维持，而在 CHH 位点发生的甲基化不能被维持，必须在每个细胞周期中被重新建立^[18]。当 CHH 甲基化的靶位点被建立较低程度的 DNA 甲基化之后，Pol IV、RDR2 以及 Pol V 被认为可以被招募到这些已经建立表观遗传学标记的位点。Pol IV 和 RDR2 产生双链或单链的 RNA 转录本。双链的转录本被 DCL3 切割产生 24 nt 的小 RNA，通过经典的依赖于小 RNA 的途径加强 CHH 甲基化 (图 1)。而单链的转录本结合 AGO4 被 RNA 外切酶削减产生不依赖于 DCL 蛋白的长度呈梯度分布的 20~60 nt 的小 RNA^[23-24]。虽然与经典的 24 nt 小 RNA 的产生途径不同，这些不依赖于 DCL 蛋白的小 RNA 的产生同样促进 CHH 甲基化水平的提高 (图 1)。除了 DNA 甲基转移酶 DRM2 之外，最近的报道揭示 CMT2 也参与催化 CHH 位点发生 DNA 甲基化^[25]。

2 植物体中 DNA 去甲基化途径

基因组中 DNA 甲基化的水平及模式不仅取决于 DNA 甲基化的建立与维持，也取决于 DNA 去甲基化的发生。DNA 去甲基化可以分为被动去甲基化和主动去甲基化。被动去甲基化指在 DNA 复制后，由于维持型 DNA 甲基转移酶失活，新合成的 DNA 链不能被甲基化的情形。主动去甲基化不依赖于 DNA 复制，是由一系列的酶催化发生的^[26]。在植物体中，DNA 主动去甲基化是由 ROS1 家族转葡糖基酶来起始的。ROS1 家族成员包括 ROS1、DME、DML2 和 DML3，它们都具有切除 5- 甲基胞嘧啶的活性，不论甲基化处于 CG、CHG 还是 CHH 序列中^[27]。ROS1、DML2 和 DML3 负责植物不同组织体细胞中的 DNA 去甲基化，而 DME 特异地在之后发育成胚乳的雌配子体中心细胞中表达并负责中心细胞及胚乳中母系印记基因的 DNA 去甲基化^[28]。当去甲基化发生时，ROS1、DME、DML2 和 DML3 不仅能够水解 DNA 链脱氧核苷酸和碱基之间的糖苷键，去除甲基化的胞嘧啶，同时，也能

够利用它们的裂解酶活性在 DNA 链上去掉碱基的位点产生切口^[27,29-30]。因此，ROS1、DME、DML2 和 DML3 被称为双功能转葡糖基酶，不同于只能切除碱基的单功能转葡糖基酶。当它们的裂解酶活性在 DNA 链上产生一个切口时，这样的反应被称为 β 消除反应。 β 消除反应的产物为一段有缺口的 DNA，缺口的 3' 端为 3' 磷酸- α,β - 不饱和醛 (3'-PUA)^[31-32]。当它们的裂解酶活性在 DNA 链上产生两个切口时，这样的反应被称为 β,δ 消除反应。 β,δ 消除反应的产物同样为一段有缺口的 DNA，但缺口的 3' 端为 3' 磷酸基^[31-32]。无论是进行 β 消除反应，还是进行 β,δ 消除反应，DNA 缺口的 3' 端都需要转变为 3' 羟基。这样，下游的 DNA 聚合酶才能识别这个 3' 末端并在此加上一个非甲基化的胞嘧啶，然后 DNA 连接酶进一步进行连接反应完成 DNA 去甲基化的过程。

ROS1 家族蛋白介导的 DNA 主动去甲基化过程类似于 DNA 损伤修复中的碱基切除修复，但是在很长一段时期内，ROS1 家族蛋白下游的其他负责参与 DNA 去甲基化过程的酶都没有被发现及研究。最新的研究显示，拟南芥中 ZDP 基因编码一个 N 端为 DNA 结合区域、C 端为 3' 磷酸酶结构域的蛋白质。由于哺乳动物细胞中 PNKP 蛋白在碱基切除修复中负责把 β,δ 消除反应产物 3' 磷酸基转变为 3' 羟基^[33]，而 ZDP 与之有同源序列，研究者推测 ZDP 很可能参与 DNA 主动去甲基化过程中 ROS1 β,δ 消除反应产物的转变。进一步的研究证实，ZDP 能够特异性地把 ROS1 β,δ 消除反应产物转变为 3' 羟基，而不能把 ROS1 β 消除反应产物 3'-PUA 转变为 3' 羟基^[34] (图 2)。在细胞中 ZDP 与 ROS1 共定位，ZDP 的突变导致几百个位点甲基化水平升高，这些位点主要位于转座子元件，并且这些位点在 ROS1 的突变体中也表现为甲基化水平升高^[34]。

然而，在 *RD29A* 这个转基因的启动子区域，ZDP 突变导致的甲基化水平升高程度比 ROS1 突变导致的甲基化水平升高程度小；另外，有一些位点的甲基化水平在 *ros1* 突变体中升高，但在 *zdp* 突变体中不变。这些事实说明，ROS1 下游有另一条不依赖于 ZDP 的去甲基化分支通路。这条分支通路很可能是由于 ROS1 进行了 β 消除反应，产生的 3'-PUA 被不依赖于 ZDP 的酶转变为 3' 羟基。由于哺乳动物细胞中 APE1 蛋白在碱基切除修复中行使这种催化功能，推测在植物体中 APE 蛋白很可能通过转变 β 消除反应产物参与到 DNA 去甲基化途径。

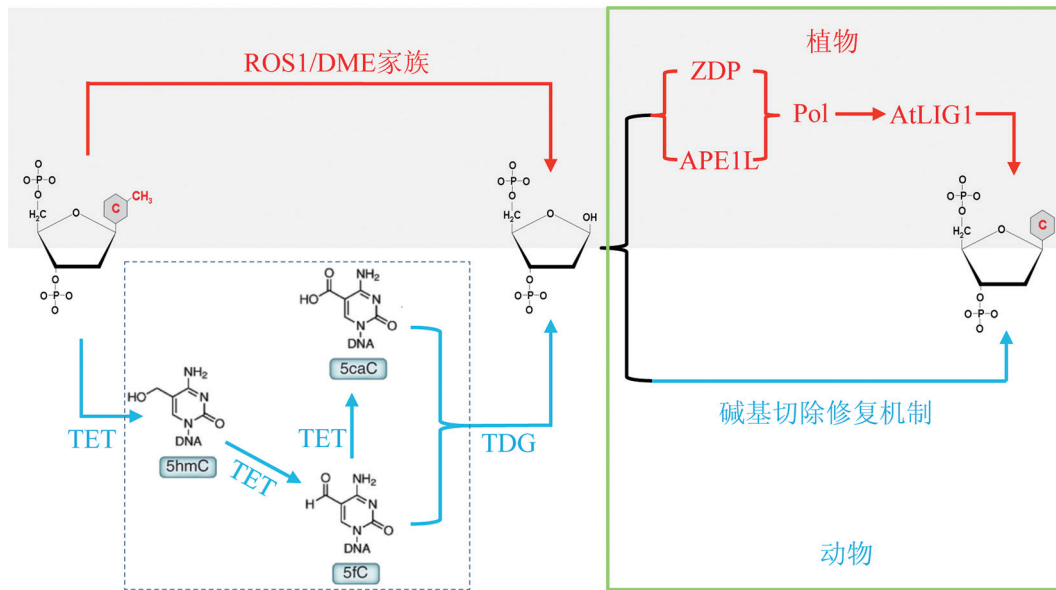


图2 植物和动物中DNA主动去甲基化途径

植物体中编码3个APE蛋白,分别是APE1L、APE2、ARP^[35]。Li等^[36]的生化实验证实,APE1L具有很强的把3'-PUA转变为3'羟基的活性,同时具有微弱的把3'磷酸基转变为3'羟基的活性;ARP具有微弱的把3'-PUA转变为3'羟基的活性;而APE2则不具有这种活性;APE1L在体外与ROS1相互结合,并在细胞内与ROS1部分共定位;APE1L的突变导致上千个位点DNA甲基化水平升高,与ZDP突变导致的甲基化水平升高不同的是,这些甲基化水平升高的位点更多地集中于编码基因的区域。这样,APE1L和ZDP参与的反应代表了DNA去甲基化途径的两个不同分支(图2)。

由于植物中存在多个DNA聚合酶并且一些DNA聚合酶的突变是胚胎致死的,目前无法确定哪一个DNA聚合酶参与DNA去甲基化途径中非甲基化胞嘧啶的掺入;但是,2015年,Li等^[37]研究证实,在AtLIG1、AtLIG1a、AtLIG4、AtLIG6这4个DNA连接酶中,AtLIG1与ROS1、APE1L、ZDP在细胞内共定位并负责去甲基化完成后DNA链切口的连接。至此,植物体中整个DNA去甲基化通路及除DNA聚合酶之外的负责完成各个步骤的关键因子基本清晰。

在哺乳动物细胞中,ROS1和DME的同源蛋白未被发现,5-甲基胞嘧啶也不能直接被切除。5-甲基胞嘧啶可以被AID/APOBEC蛋白通过脱氨反应转变成胸腺嘧啶,然后,被胸腺嘧啶转葡萄糖基酶通过G/T错配修复的方式去除^[38]。或者,5-甲基

胞嘧啶可以被TET家族蛋白氧化,生成5-羟甲基胞嘧啶、5-乙酰胞嘧啶及5-羧基胞嘧啶^[39]。TDG转葡萄糖基酶具有去除这些修饰后的胞嘧啶的活性,之后碱基切除修复通路的因子负责完成整个DNA去甲基化过程^[39](图2)。由此可见,尽管动物和植物中DNA主动去甲基化在起始的时候采用的机制不同,但最终都是通过碱基切除修复机制来完成的。

3 植物体中DNA主动去甲基化的调控

植物体中DNA主动去甲基化的调控首先在于ROS1的表达水平受到其启动子区域一段序列甲基化水平的调控^[40-41]。这段序列位于ROS1 5'UTR及启动子上的一个转座子元件之间^[40],当细胞中依赖于小RNA的甲基化途径活性升高时,这段序列被甲基化。这样的甲基化促进ROS1表达水平升高以及DNA主动去甲基化活动的增强,最终甲基化途径的活性被拮抗,细胞中一些位点的甲基化水平达到平衡^[40-41]。反之,当细胞中依赖于小RNA的甲基化途径活性降低,其靶位点不需要去甲基化时,这段序列被去甲基化,ROS1表达水平下降。这段序列被认为是细胞中甲基化水平的一个感知器,这个感知器起到调节去甲基化途径活性及一些位点最终甲基化水平的重要作用。

另一个调控植物体中DNA主动去甲基化的方式在于植物体对ROS1这个去甲基化酶的活性的调节。2015年,报道显示ROS1蛋白具有一段保守的铁硫结构单元。MET18,又被称为MMS19,能够

结合 ROS1, 并通过转运铁硫簇到 ROS1 蛋白上来调节 ROS1 酶的活性^[42-43]。当 *MET18* 突变后, 细胞中多个位点甲基化水平升高, 这些位点与 *ros1* 突变体中甲基化水平升高的位点在很大程度上重合。

植物体中 DNA 去甲基化的发生并不是全局性的, 而存在位点特异性, 因此, 植物体中 DNA 主动去甲基化的调控还在于 ROS1 等去甲基化酶如何被特异性地招募到其靶位点进行去甲基化。之前的研究显示 ROS3 与 ROS1 共定位, ROS3 通过与小 RNA 结合来指导 ROS1 被招募到特定位点。2012年, Qian 等^[44] 和 Li 等^[45] 利用一种特殊的 Chop PCR 反应进行的筛选及遗传学筛选发现, 一个组蛋白乙酰转移酶 IDM1 能够调控 ROS1 的招募; *IDM1* 又被称为 *ROS4*, 编码一个含有 N 端乙酰转移酶结构域、MBD 结构域及 PHD 指状结构域的蛋白; IDM1 能够识别组蛋白 H3 第 4 位赖氨酸不被双甲基化或三甲基化但是 DNA 的 CG 序列被甲基化的位点, 并在这些位点催化组蛋白 H3 第 18 位及第 23 位赖氨酸的乙酰化。这些表观遗传学标记的建立不仅有利于 ROS1 等去甲基化酶的识别, 也有利于创造更宽松的染色质环境使 ROS1 等去甲基化酶能够接近这些位点。当 *IDM1* 突变后, 大约有 1 000 个位点 DNA 甲基化水平升高, 这些位点的甲基化水平在 *ros1* 突变体中也表现为升高, 但是在 *ros1 idm1* 双突变体中, 这些位点的甲基化水平不会继续升高, 说明 IDM1 和 ROS1 处于同一条 DNA 去甲基化信号通路^[44]。利用同样的 Chop PCR 进行筛选, 研究者发现 IDM2 能够通过调节 IDM1 对组蛋白 H3 第 18 位赖氨酸乙酰化的活性来参与 DNA 去甲基化^[46-47]。*IDM2*, 又被称为 *ROS5*, 编码一个具有 α 晶体蛋白结构域的非典型热激蛋白。*IDM2* 与 IDM1 直接相互作用, 并在细胞内部分共定位^[46-47]。当 *IDM2* 突变后, IDM1 对组蛋白 H3 第 18 位赖氨酸乙酰化的活性降低。与 *IDM1* 突变一样, *IDM2* 的突变导致 1 000 多个位点 DNA 甲基化水平增高, 进一步的分析显示 *IDM2* 与 IDM1、ROS1 处于同一条去甲基化信号通路^[46], 但是 IDM1 与 *IDM2* 调控的区域更集中于基因编码区域^[44,46]。利用遗传学筛选及亲和纯化蛋白质复合体的方法, 研究者进一步发现 MBD7 和 IDM3 (又被称为 *IDL1*) 与 IDM1 和 *IDM2* 相互作用并参与 DNA 去甲基化途径^[48-49]。*IDM3* 是一个和 *IDM2* 序列高度类似的非典型热激蛋白。MBD7 具有结合甲基化的 CG 序列的结构域, 因此, 能够把 IDM 蛋白复合体招募到 CG 甲基化

密集的区域进行组蛋白 H3 的乙酰化^[48-50]。

总之, IDM 蛋白复合体的发现大大拓展了人们对于植物体中 DNA 去甲基化调控机制的了解, 但是 IDM 蛋白复合体只能解释 ROS1 是怎样被招募到它的一部分靶位点的。希望未来的研究能够发现更多的负责招募 ROS1 到特定染色质区域的蛋白质。

4 植物体中DNA主动去甲基化的生物学功能

在植物体中, 依赖于小 RNA 的甲基化途径介导其同源序列的转录水平沉默。DNA 主动去甲基化对于防止转基因及内源基因被甲基化沉默有重要意义。*ROS1* 基因的发现源于其突变导致 *RD29A* 启动子驱动的荧光素酶 (LUC) 转基因及其驱动的内源基因沉默^[51]。*ROS1* 对于 *35S* 启动子驱动的编码卡那霉素抗性的 *NPTII* 基因的表达也是必需的。ZDP 作用于 ROS1 等去甲基酶下游, ZDP 对于 *RD29A* 驱动的转基因及内源基因的表达也是必需的^[34]。相反, *APEIL* 的突变不影响 *RD29A-LUC* 的表达^[36]。由于 ZDP 和 *APEIL* 控制两条不同的去甲基化分支通路, ZDP 和 *APEIL* 被认为负责防止不同内源基因的沉默^[34,36], 但是, 目前还不清楚这两个蛋白具体控制哪些位点的表达。由于 IDM1、IDM2、IDM3 和 MBD7 起到重要的调节 ROS1 定位的作用, 它们的突变也都分别导致基因组多个位点的沉默。和 ROS1 不同的是, 它们对于 *35S-NPTII* 基因的表达是必需的, 但是并不控制 *RD29A-LUC* 的表达^[44,46,48-50]。

DNA 主动去甲基化对印记基因在特定时空表达有重要意义。*DME* 被发现对于母系印记基因 *MEA* 和 *FWA* 在胚乳中的特异性表达起决定性作用, *DME* 的突变导致 *MEA* 和 *FWA* 等母系印记基因被沉默, 胚胎母系致死^[28-29]。*ZDP* 和 *APEIL* 的单突变不影响母系印记基因 *MEA* 和 *FWA* 在胚乳中的特异性表达, 但有趣的是, *ZDP* 和 *APEIL* 的双突变导致 *MEA* 和 *FWA* 等母系印记基因被沉默, 最终导致和 *DME* 突变后一样的胚胎母系致死表型^[36]。负责 DNA 去甲基化最后一步连接反应的 *AIL1G1* 突变后, 同样导致 *MEA* 和 *FWA* 等母系印记基因被沉默^[37]。

转座子和反转座子在通常情况下被甲基化及其他抑制型组蛋白修饰所沉默。它们的沉默防止了转座事件影响正常基因的表达, 避免了基因组不稳定。长期以来的研究发现, 有些转座子其实有很低水平的表达, DNA 去甲基化是维持这些转座子表达的重要因素^[52]。另外, DNA 去甲基化有时足以引发

转座子或反转座子的跳跃^[53]。例如, 2011年, La等^[54]研究发现, 在水稻中, *Tos17* 反转座子的跳跃被DNA去甲基化酶DNG701所调控。在长期的进化过程中, 转座子的跳跃对于形成基因组多样性, 或进化出有利于宿主细胞的基因有重要意义。

DNA去甲基化还可以响应非生物或生物胁迫, 起到重要的激活特定群体基因表达的作用。最早在玉米中进行的研究发现, 冷刺激会诱导玉米根部细胞基因组去甲基化, 这样的去甲基化对于多个基因同时被诱导表达以响应冷刺激以及引起一些可遗传至后代的表观遗传学改变具有重要意义^[55]。2014年, Le等^[56]研究发现, 在拟南芥中DNA去甲基化能够在植物体被真菌病原体处理后诱发多个抗病基因的激活。而在番茄中进行的研究发现, ROS1的过表达能够调控响应盐刺激的类黄酮合成及抗氧化通路基因的表达^[57]。

总之, 随着在各个物种中参与DNA去甲基化的因子被鉴定出来, 以及各种条件下基因组DNA去甲基化的情况被研究, DNA去甲基化的更多功能将被揭示。

5 总结和展望

模式植物拟南芥以其基因组小、生活周期短、便于遗传筛选等特点, 成为研究DNA主动去甲基化机制及功能的理想实验材料。尽管人们发现DNA去甲基化在玉米、水稻及其他植物中发挥着重要生物学功能, 这些物种中DNA去甲基化的途径及负责关键反应的酶还都未知。今后的研究当以拟南芥中DNA去甲基化的知识为基础, 进一步研究一些重要作物中DNA去甲基化的机制。这样的研究或将有利于提高农作物抗病抗逆性。另外, 尽管人们对于ROS1介导的DNA去甲基化途径相对清楚, 植物体中很可能存在其他独立于ROS1的DNA主动去甲基化机制, 但是这些机制是什么及究竟如何发挥作用目前尚不清楚。另外, 更多的组蛋白修饰与DNA去甲基化途径之间的关系还很不清楚。环境和表观遗传又是怎样的关系, 这些问题都需要在接下来的研究中解决。

[参 考 文 献]

- [1] Ehrlich M, Gama-Sosa MA, Huang LH, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res*, 1982, 10: 2709-21
- [2] Doerfler W. DNA methylation and gene activity. *Annu Rev Biochem*, 1983, 52: 93-124
- [3] Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*, 2002, 16: 6-21
- [4] Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*, 2010, 11: 204-20
- [5] Zhai J, Bischof S, Wang H, et al. A one precursor one siRNA model for Pol IV-dependent siRNA biogenesis. *Cell*, 2015, 163: 445-55
- [6] Haag JR, Ream TS, Marasco M, et al. *In vitro* transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol Cell*, 2012, 48: 811-8
- [7] Li S, Vandivier LE, Tu B, et al. Detection of Pol IV/RDR2-dependent transcripts at the genomic scale in *Arabidopsis* reveals features and regulation of siRNA biogenesis. *Genome Res*, 2015, 25: 235-45
- [8] Wierzbicki AT, Haag JR, Pikaard CS. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, 2008, 135: 635-48
- [9] Wierzbicki AT, Ream TS, Haag JR, et al. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet*, 2009, 41: 630-4
- [10] Zhong X, Du J, Hale CJ, et al. Molecular mechanism of action of plant DRM *de novo* DNA methyltransferases. *Cell*, 2014, 157: 1050-60
- [11] Law JA, Vashisht AA, Wohlschlegel JA, et al. SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet*, 2011, 7: e1002195
- [12] Law JA, Du J, Hale CJ, et al. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature*, 2013, 498: 385-9
- [13] Zhang H, Ma ZY, Zeng L, et al. DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. *Proc Natl Acad Sci USA*, 2013, 110: 8290-5
- [14] Johnson LM, Du J, Hale CJ, et al. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature*, 2014, 507: 124-8
- [15] Liu ZW, Shao CR, Zhang CJ, et al. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet*, 2014, 10: e1003948
- [16] Law JA, Ausin I, Johnson LM, et al. A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in *Arabidopsis*. *Curr Biol*, 2010, 20: 951-6
- [17] Zhong X, Hale CJ, Law JA, et al. DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat Struct Mol Biol*, 2012, 19: 870-5
- [18] Stroud H, Greenberg MV, Feng S, et al. Comprehensive analysis of silencing mutants reveals complex regulation

- of the *Arabidopsis* methylome. *Cell*, 2013, 152: 352-64
- [19] Nuthikattu S, McCue AD, Panda K, et al. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiol*, 2013, 162: 116-31
- [20] Panda K, Slotkin RK. Proposed mechanism for the initiation of transposable element silencing by the RDR6-directed DNA methylation pathway. *Plant Signal Behav*, 2013, 8: e25206
- [21] McCue AD, Panda K, Nuthikattu S, et al. ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *EMBO J*, 2015, 34: 20-35
- [22] Mari-Ordonez A, Marchais A, Etcheverry M, et al. Reconstructing *de novo* silencing of an active plant retrotransposon. *Nat Genet*, 2013, 45: 1029-39
- [23] Ye R, Chen Z, Lian B, et al. A Dicer-independent route for biogenesis of siRNAs that direct DNA methylation in *Arabidopsis*. *Mol Cell*, 2016, 61: 222-35
- [24] Yang DL, Zhang G, Tang K, et al. Dicer-independent RNA-directed DNA methylation in *Arabidopsis*. *Cell Res*, 2016, 26: 66-82
- [25] Zemach A, Kim MY, Hsieh PH, et al. The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*, 2013, 153: 193-205
- [26] Zhu JK. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet*, 2009, 43: 143-66
- [27] Penterman J, Zilberman D, Huh JH, et al. DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci USA*, 2007, 104: 6752-7
- [28] Choi Y, Gehring M, Johnson L, et al. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell*, 2002, 110: 33-42
- [29] Gehring M, Huh JH, Hsieh TF, et al. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell*, 2006, 124: 495-506
- [30] Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, et al. Demeter repressor of silencing 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci USA*, 2006, 103: 6853-8
- [31] Agius F, Kapoor A, Zhu JK. Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci USA*, 2006, 103: 11796-801
- [32] McCullough AK, Dodson ML, Lloyd RS. Initiation of base excision repair: glycosylase mechanisms and structures. *Annu Rev Biochem*, 1999, 68: 255-85
- [33] Jilani A, Ramotar D, Slack C, et al. Molecular cloning of the human gene, *PNKP*, encoding a polynucleotide kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. *J Biol Chem*, 1999, 274: 24176-86
- [34] Martinez-Macias MI, Qian W, Miki D, et al. A DNA 3' phosphatase functions in active DNA demethylation in *Arabidopsis*. *Mol Cell*, 2012, 45: 357-70
- [35] Murphy TM, Belmonte M, Shu S, et al. Requirement for abasic endonuclease gene homologues in *Arabidopsis* seed development. *PLoS One*, 2009, 4: e4297
- [36] Li Y, Cordoba-Canero D, Qian W, et al. An AP endonuclease functions in active DNA demethylation and gene imprinting in *Arabidopsis* [corrected]. *PLoS Genet*, 2015, 11: e1004905
- [37] Li Y, Duan CG, Zhu X, et al. A DNA ligase required for active DNA demethylation and genomic imprinting in *Arabidopsis*. *Cell Res*, 2015, 25: 757-60
- [38] Rai K, Huggins IJ, James SR, et al. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell*, 2008, 135: 1201-12
- [39] Hill PW, Amouroux R, Hajkova P. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. *Genomics*, 2014, 104: 324-33
- [40] Lei M, Zhang H, Julian R, et al. Regulatory link between DNA methylation and active demethylation in *Arabidopsis*. *Proc Natl Acad Sci USA*, 2015, 112: 3553-7
- [41] Williams BP, Pignatta D, Henikoff S, et al. Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. *PLoS Genet*, 2015, 11: e1005142
- [42] Duan CG, Wang X, Tang K, et al. MET18 connects the cytosolic iron-Sulfur cluster assembly pathway to active DNA demethylation in *Arabidopsis*. *PLoS Genet*, 2015, 11: e1005559
- [43] Han YF, Huang HW, Li L, et al. The cytosolic iron-sulfur cluster assembly protein MMS19 regulates transcriptional gene silencing, DNA repair, and flowering time in *Arabidopsis*. *PLoS One*, 2015, 10: e0129137
- [44] Qian W, Miki D, Zhang H, et al. A histone acetyltransferase regulates active DNA demethylation in *Arabidopsis*. *Science*, 2012, 336: 1445-8
- [45] Li X, Qian W, Zhao Y, et al. Antisilencing role of the RNA-directed DNA methylation pathway and a histone acetyltransferase in *Arabidopsis*. *Proc Natl Acad Sci USA*, 2012, 109: 11425-30
- [46] Qian W, Miki D, Lei M, et al. Regulation of active DNA demethylation by an alpha-crystallin domain protein in *Arabidopsis*. *Mol Cell*, 2014, 55: 361-71
- [47] Zhao Y, Xie S, Li X, et al. REPRESSOR OF SILENCING5 encodes a member of the small heat shock protein family and is required for DNA demethylation in *Arabidopsis*. *Plant Cell*, 2014, 26: 2660-75
- [48] Lang Z, Lei M, Wang X, et al. The methyl-CpG-binding protein MBD7 facilitates active DNA demethylation to limit DNA hyper-methylation and transcriptional gene silencing. *Mol Cell*, 2015, 57: 971-83
- [49] Li Q, Wang X, Sun H, et al. Regulation of active DNA demethylation by a methyl-CpG-binding domain protein in *Arabidopsis thaliana*. *PLoS Genet*, 2015, 11: e1005210
- [50] Wang C, Dong X, Jin D, et al. Methyl-CpG-binding domain protein MBD7 is required for active DNA demethylation in *Arabidopsis*. *Plant Physiol*, 2015, 167: 905-14
- [51] Gong Z, Morales-Ruiz T, Ariza RR, et al. ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*,

- encodes a DNA glycosylase/lyase. *Cell*, 2002, 111: 803-14
- [52] Zhu J, Kapoor A, Sridhar VV, et al. The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. *Curr Biol*, 2007, 17: 54-9
- [53] Saze H, Tsugane K, Kanno T, et al. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. *Plant Cell Physiol*, 2012, 53: 766-84
- [54] La H, Ding B, Mishra GP, et al. A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon Tos17 and promotes its transposition in rice. *Proc Natl Acad Sci USA*, 2011, 108: 15498-503
- [55] Steward N, Ito M, Yamaguchi Y, et al. Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J Biol Chem*, 2002, 277: 37741-6
- [56] Le TN, Schumann U, Smith NA, et al. DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. *Genome Biol*, 2014, 15: 458
- [57] Bharti P, Mahajan M, Vishwakarma AK, et al. AtROS1 overexpression provides evidence for epigenetic regulation of genes encoding enzymes of flavonoid biosynthesis and antioxidant pathways during salt stress in transgenic tobacco. *J Exp Bot*, 2015, 66: 5959-69