

DOI: 10.13376/j.cblls/2017124

文章编号: 1004-0374(2017)10-0926-08



朱冰, 中国科学院生物物理研究所研究员, 博士生导师, 入选中组部“万人计划”, 国家“杰出青年科学基金”获得者。获霍华德·休斯医学院国际青年科学家奖、药明康德生命化学奖、谈家桢生命科学创新奖。现任 *J Biol Chem* 等杂志编委, 中国细胞生物学会染色质生物学分会会长。课题组的主要研究方向是表观遗传信息建立与维持的分子机制和新型表观调控因子的鉴定与理解。课题组在复制偶联的核小体组蛋白分配方式、组蛋白甲基化修饰的继承方式、染色质修饰酶的活性调节机制等问题上取得了重要成果。一系列研究成果发表于 *Science*、*Mol Cell*、*Genes Dev*、*Cell Res*、*PLoS Genet*、*J Biol Chem* 等学术期刊。

TET家族蛋白介导的DNA氧化的调控与其生物学功能

熊俊, 朱冰*

(中国科学院生物物理研究所生物大分子国家重点实验室, 北京 100101)

摘要: 发生在 DNA 胞嘧啶上的甲基化 (5mC) 是哺乳动物细胞基因组上最主要的 DNA 修饰形式, 其形成的碳碳键具有较高键能, 不易被破坏。TET 家族蛋白可以催化 5mC 逐渐氧化成羟甲基胞嘧啶 (5hmC)、醛基胞嘧啶 (5fC) 和羧基胞嘧啶 (5caC), 再通过细胞分裂过程中 DNA 复制, 或者利用碱基切除修复途径, 最终实现 DNA 去甲基化。过去几年表观基因组学和结构生物学的研究都表明, 在不同细胞、不同的基因组位点, 5mC 的氧化反应受到严格的调控, 主要表现在两个方面: 5mC 氧化反应发生的基因组范围和 5mC 逐步氧化反应的进行程度。以国家自然科学基金委重大研究计划“细胞编程和重编程的表观遗传机制”为依托, 朱冰实验室发现了胚胎干细胞的多能性转录因子 SALL4A 与 TET 家族蛋白共同调节远端调控区域 5mC 的氧化过程。首先, 将介绍 5mC 的不同氧化产物在小鼠基因组上的分布和动态变化, 进而讨论 TET 家族蛋白催化 5mC 氧化反应的调控机制, 最后, 探讨 5mC 氧化参与调节基因组转录的可能的生物学功能。

关键词: DNA 氧化; DNA 羟甲基化; TET 蛋白

中图分类号: Q523; Q554 **文献标志码:** A

TET-mediated DNA oxidation and its biological functions

XIONG Jun, ZHU Bing*

(National Key Laboratory of Biomacromolecules, Institute of Biophysics,
Chinese Academy of Sciences, Beijing 100101, China)

Abstract: 5-methylcytosine is the most abundant modified nucleotide in mammalian genome, which is hard to remove from the genome due to the carbon-carbon bond between the methyl group and cytosine. TET family proteins successively catalyze 5mC to oxidized forms including 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), which may result in eventual demethylation through passive dilution during mitosis or with the help of thymine DNA glycosylase (TDG) and the base excision repair machinery.

收稿日期: 2017-09-05

基金项目: 国家自然科学基金项目(91419305)

*通信作者: E-mail: zhubing@ibp.ac.cn

Recent genomic mapping and structural studies have revealed that 5mC oxidation in the genome is strictly regulated at multiple layers. In this review, we first describe the distribution and dynamics of different 5mC oxidative products in mouse genomes. Then we summarize current knowledge on how the oxidation reaction is controlled. Finally, we discuss the possible role of TET-mediated DNA oxidation in transcription regulation.

Key words: DNA oxidation; DNA hydroxymethylation; TET proteins

哺乳动物基因组 DNA 中胞嘧啶的甲基化修饰被认为是最稳定的表观遗传修饰, 其通常发生在 CpG 二核苷酸回文序列, 形成对称的甲基化修饰 (fully methylation)。DNA 复制后, 带有甲基化修饰的亲代基因组 DNA 母链与没有修饰的子链配对, 形成半甲基化状态的 CpG 序列 (hemi-methylation)。DNA 甲基转移酶 1 (DNMT1) 以半甲基化状态的 CpG 为底物, 催化子链 CpG 甲基化, 从而以半保留复制的方式, 在子代细胞中维持与亲代细胞基因组 DNA 一致的甲基化修饰^[1-3]。近年来的研究发现, TET 家族蛋白能够催化 5mC 的第 5 位甲基逐步氧化成羟甲基、醛基和羧基, 催化生成 5hmC 的反应速率远高于催化生成 5fC 和 5caC^[4-8]。5mC 的氧化会在 DNA 复制后生成不利于 DNMT1 催化的半修饰状态 CpG, 因而会导致数次细胞分裂后, 发生 5mC 氧化反应的位点在子代细胞中失去甲基化修饰^[9]。另一方面, TDG (thymine DNA glycosylase) 可以高效地切除基因组 DNA 上的 5fC 和 5caC, 产生脱碱基位点, 继而在碱基切除修复 (base excision repair, BER) 途径的作用下修复缺口, 回复到无修饰的状态, 这一过程即 DNA 氧化去甲基化^[7,10-11]。TDG 的这种活性也在一定程度上使得基因组上的 5fC 和 5caC 含量远低于 5hmC 的含量。DNA 上修饰状态的动态变化拓展了 DNA 甲基化所承载的表观遗传信息的可塑性, 越来越多的研究开始发现, TET 家族蛋白催化的 5mC 氧化反应在维持细胞转录组和细胞状态稳定方面的重要作用。

1 5mC氧化产物在基因组上的分布与动态变化

基因组上的 DNA 氧化反应受到严格的调控, 一方面 DNA 序列特异的转录因子或辅因子可以招募 TET 蛋白或 TDG 到特定的基因组位点; 另一方面, 结构生物学的研究发现, TET 家族蛋白催化 5mC 氧化的能力逐级减弱, 氧化反应持续进行的能力可以被精细地调控^[8]。已经有研究表明, 基因组不同区域的 5hmC 被进一步氧化的可能性不尽相同, 在一些区域, 5hmC 可以稳定存在, 而在别的区域

会被迅速地进一步氧化, 导致 DNA 去甲基化。在这些调控机制的共同作用下, 5mC 各种不同氧化产物在基因组上的分布也各有差异。

在小鼠胚胎干细胞中, 活跃转录基因的启动子和转录起始位点附近通常缺少 5hmC, 以及进一步的氧化产物 5fC 和 5caC^[12-18]。这些启动子上 CpG 序列的分布密度一般较高, 且通常是未甲基化的, 同时分布有带组蛋白 H3K4 三甲基化修饰的核小体, 能够抑制 DNA 甲基转移酶的活性^[19], 因而 DNA 氧化反应的底物相对较少。TET1 通过 CXXC 结构域与其他蛋白相互作用, 也主要分布在这类启动子上, 可以将错误 DNA 甲基化迅速氧化来维持这些启动子的低甲基化状态^[20], 也有可能从空间上阻止 DNA 甲基化酶与其底物 DNA 结合, 防止启动子被错误地甲基化^[21]。胚胎干细胞中有一类所谓的 bivalent 启动子, 这类启动子既带有抑制性组蛋白修饰 H3K27me3, 又富集激活性组蛋白修饰 H3K4me3, 其所调控的基因多与胚胎发育、组织分化等有关, 所以在胚胎干细胞中多处于转录被抑制的状态, 而在随后的分化过程中将被激活表达^[22]。5mC 的各种氧化产物都在 bivalent 启动子区域分布较多, 而 bivalent 启动子要维持其已经完成初始活化的状态, 则有可能需要活跃的 DNA 氧化反应^[12,14-15,23]。

利用亲和富集的方法, 可以在基因内部 (gene body), 尤其是活跃转录的基因外显子上检测到相对较多的 5hmC^[12,24]。这一现象在人或小鼠的其他细胞类型和原代组织中也被相继发现, 且在多数细胞中, 基因内部的 5hmC 富集程度与基因转录水平呈正相关。基因内部的 5mC 主要由 DNA 甲基转移酶 DNMT3B 催化产生^[25], 较高度度的 DNA 甲基化可以抑制从基因内部起始的错误转录的发生^[26]。目前尚不清楚基因内部的 DNA 氧化是如何发生的, 然而, 较丰富的 5mC 底物是导致氧化反应活跃的先决条件之一。5hmC 在外显子和内含子的边界区域有更高的富集, 暗示 TET 蛋白催化的 5mC 氧化可能参与调节外显子的选择和 mRNA 剪切^[27-29]。

基因的远端调控区域 (或增强子) 通常为

DNase I 敏感区域, 其范围内 CpG 二核苷酸序列含量较少, DNA 甲基化水平较低 (low-methylated regions, LMRs), 但一般富集较高水平的组蛋白 H3K4me1 修饰^[30-31]。虽然 TET1 在远端调控区域分布的富集程度不高, 明显低于其在启动子和 CpG 密集区域的定位, 但 5hmC 的丰度在远端调控区域却相对较高^[14-15,23,32-33]。Yu 等^[14]在单碱基精度下分析 5hmC 的分布发现, 5hmC 主要富集在转录因子结合位点两侧, 而在结合位点中心区域则没有 5hmC 分布。在缺失 TDG 的小鼠胚胎干细胞中, 5fC 和 5caC 会在基因组上逐渐积累。通过分析这些积累下来的 5fC 和 5caC 在基因组上的分布, 发现 5fC 和 5caC 都倾向于积累在基因远端调控位点上, 表明这些区域有较高的 5mC 氧化反应活性^[16-18]。在缺失 TET2 或者所有 TET 蛋白的小鼠胚胎干细胞基因组上, 也可以检测到远端调控位点上 DNA 甲基化水平升高^[18,34], 说明这些位点发生了 TET 蛋白介导的 DNA 去甲基化事件。转录因子结合在远端调控区域, 通过 DNA 成环 (DNA looping), 促使启动子与远端调控区域相互接触, 并使得结合在启动子的 TET1 能够催化远端调控区域的 DNA 氧化反应。在基金委重大研究计划“细胞编程和重编程的表观遗传机制”项目的支持下, 本实验室发现, TET1 在远端调控区域的催化活性有限, 其催化的 5mC 氧化反应多停留在 5hmC 阶段, 不能完全介导 DNA 去甲基化, 而 TET2 特异性地被招募到远端调控区域, 催化 5hmC 的进一步氧化产生 5fC 或 5caC, 继而促使 DNA 氧化去甲基化过程得以完成^[35]。

2 TET蛋白介导的DNA氧化去甲基化的调节机制

TET 蛋白是依赖于二价铁离子和 α -酮戊二酸的双加氧酶, 将铁离子作为氧化还原的辅因子, 以 α -酮戊二酸、 O_2 和 5mC 为底物, 催化生成琥珀酸、 CO_2 和 5hmC。细胞代谢状态和所处微环境的改变会影响一系列代谢产物和电解质浓度, 从而影响细胞内相关的酶的生化活性, 甚至改变酶催化反应的性质 (图 1)^[36]。 α -酮戊二酸是三羧酸循环的中间产物, 主要由异柠檬酸脱氢酶 (isocitrate dehydrogenase, IDH) 催化产生。在胶质瘤和恶性血液系统肿瘤里常见 *IDH1* 和 *IDH2* 基因的功能获得型突变, 这种突变使得 IDH 将 α -酮戊二酸催化生成 *R*-2-羟戊二酸 ((*R*)-2-hydroxyglutarate, *R*-2HG)^[37-38], 一方面降低细胞内 α -酮戊二酸的浓度; 另一方面, *R*-2HG

能够竞争性结合并抑制依赖于 α -酮戊二酸的双加氧酶活性, 抑制 TET 蛋白介导的基因组 DNA 氧化去甲基化^[39]。因此, *IDH* 突变的组织细胞常表现出基因组 DNA 甲基化水平上升、5mC 氧化产物减少及某些组蛋白甲基化修饰水平上升等特征^[36,40-42]。此外, 琥珀酸脱氢酶和延胡索酸水化酶的突变或缺失导致细胞内琥珀酸和延胡索酸浓度升高, 这两种三羧酸循环的中间产物与 α -酮戊二酸的化学结构类似, 也可以竞争性结合并抑制依赖于 α -酮戊二酸的双加氧酶活性 (图 1)^[43]。细胞内的溶氧量也会影响 TET 蛋白的活性。在缺氧条件下, 某些体外培养的肿瘤细胞基因组上 5hmC 的含量显著下降, 缺氧降低 TET 蛋白活性, 导致肿瘤患者体内局部缺氧的肿瘤组织细胞 DNA 甲基化水平升高^[44]。此外, 维生素 C 可以直接与 TET 蛋白的催化结构域结合, 通过帮助 TET 蛋白正确折叠, 或者循环利用辅因子二价铁离子来提高 TET 蛋白的催化活性^[45]。

通过改变细胞代谢产物来影响 TET 蛋白活性的效果是广泛而缺少特异性的。调节 TET 蛋白对 DNA 底物的选择性和结合能力, 以及对 TET 蛋白进行翻译后修饰则可实现对 TET 蛋白介导的 DNA 氧化反应的精细调控。如前所述, TET1 利用其 CXXC 结构域结合在 CpG 序列含量较高的基因启动子附近, 催化启动子上错误出现的 5mC 迅速地被氧化去甲基化, 也有可能通过 DNA 成环介导的远端调控区域与启动子的相互接触, 催化远端调控区域的 DNA 氧化。TET2 的 CXXC 结构域在进化过程中与 TET2 蛋白分离, 由于发生染色体倒位, 编码 TET2 CXXC 结构域的外显子形成一个独立基因, 表达名为 IDAX 的蛋白^[46]。IDAX 通过两种可能的途径调节 TET2 的活性: 其一是 IDAX 与 TET2 结合形成蛋白质复合物, 将 TET2 招募到其结合区域; 其二是 IDAX 可激活 CASPASE, 导致 TET2 被降解, 从而降低基因组 DNA 氧化反应活性^[46]。与之类似地, 多种与 TET 相互作用的蛋白相继被报道能够调节 TET 蛋白介导的 DNA 氧化反应。母体效应蛋白 DPPA3 可与 TET2 和 TET3 结合, 在体外反应实验中抑制这两个蛋白的催化活性^[47]。TET3 催化受精之后雄原核与雌原核基因组上 5mC 的氧化, 帮助原核基因组实现全基因组范围的 DNA 去甲基化^[48-50]。多种研究手段都发现, 在这一过程中, 雌原核上 DNA 氧化反应的水平远远低于雄原核^[49-52], 其原因一方面是因为受精前成熟卵细胞基因组的甲基化水平就低于精子, 另一方面则可能是由于 DPPA3

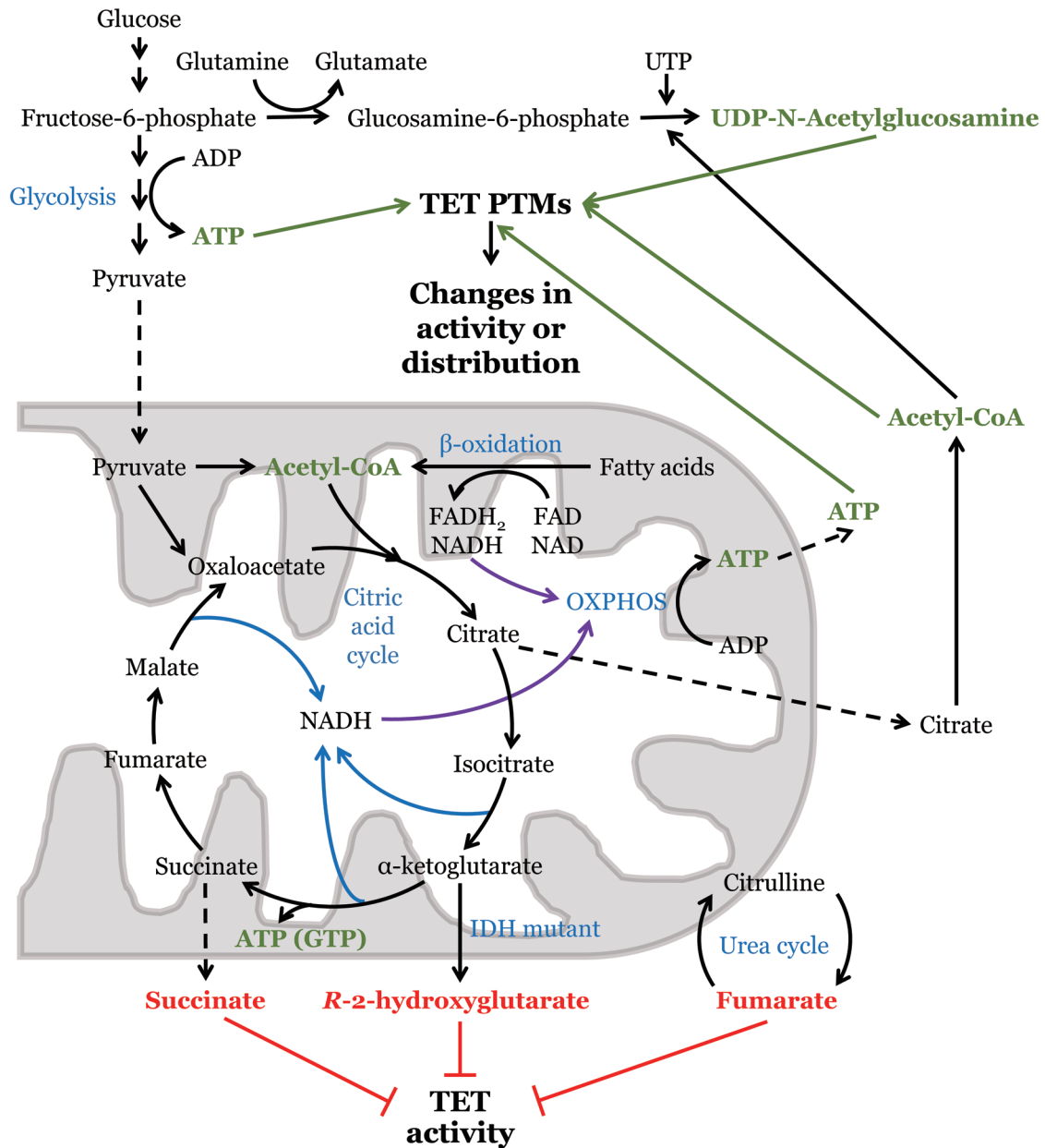


图1 TET蛋白的活性受到细胞内多种代谢产物的调节

被特异地招募到雌原核上, 对 TET3 的活性有所抑制。此外, CRL4 泛素连接酶复合物可以促进 TET 家族蛋白与基因组 DNA 的结合, 提高催化反应效率^[53-54]。序列或位点专一性的 DNA 结合蛋白可以将不同的 TET 蛋白招募到基因组特定位置, 如小鼠胚胎干细胞中 NANOG^[55]、PRDM14^[56] 和 LIN28A^[57], 以及急性髓系白血病细胞中转录因子 WT1^[58] 和小鼠视网膜细胞中的转录因子 REST^[59] 等。

对 TET 蛋白进行翻译后修饰可以调节 TET 蛋白的细胞定位、蛋白稳定性以及酶催化活性。O-GlcNAc 糖基化修饰是最早被报道的 TET 蛋白翻

译后修饰, 多个研究组报道了乙酰氨基葡萄糖转移酶 (O-linked β -N-acetylglucosamine transferase, OGT) 与 TET 蛋白相互作用, 催化 TET 蛋白 O-GlcNAc 修饰^[60-61]。然而, 目前还缺少确凿的证据来证明 O-GlcNAc 修饰影响 TET 蛋白功能的分子机制, 文献中已报道的方式包括阻止 TET1 蛋白被泛素化-蛋白酶体途径降解^[60]、调节 TET3 蛋白在细胞内的定位分布^[61]等。此外, 有研究报道, CRL4 泛素连接酶复合物催化的 TET 蛋白单泛素化和乙酰转移酶 P300 催化的乙酰化都能够提高 TET 蛋白与染色质的结合能力, 在体外实验中利用 P300 乙酰化

TET2 蛋白可以促进其催化活性^[54,62]。这些翻译后修饰参与调控 TET 蛋白介导的 DNA 氧化反应,甚至有可能在基因组上这些修饰酶结合的区域,就地 (on-site) 修饰邻近的 TET 蛋白,调节其催化活性。糖基化、乙酰化以及磷酸化^[63]都以细胞内小分子代谢物作为反应底物,因此,细胞内能量平衡和代谢状态的变化不仅会影响 TET 蛋白催化反应所需的小分子底物和辅因子浓度,还可能通过影响 TET 蛋白的翻译后修饰来调节其催化活性。研究 TET 蛋白参与响应胞内外环境变化过程中的作用及其分子机制,将有利于更好地理解 TET 蛋白与 DNA 氧化的生物学功能。

3 TET蛋白与DNA氧化的生物学功能

关于 TET 蛋白与 5mC 氧化的生物学功能的讨论,主要包括两个方面:其一是通过氧化去甲基化实现 DNA 去甲基化;其二是作为一类特殊的核苷酸,招募能与之特异性结合的蛋白质来行使功能。综合近几年的研究发现,在胚胎发育早期的两个重要的全基因组范围 DNA 去甲基化事件中,TET 蛋白介导的 DNA 氧化去甲基化都发挥了重要的作用^[21]。2016 年,Dai 等^[64]报道了 TET 蛋白介导的 DNA 氧化去甲基化在小鼠胚胎原肠运动中,与 DNA 甲基转移酶一起调节 Lefty-Nodal 信号通路,确保胚胎正常发育。如前所述,基因组上 5fC 和 5caC 的含量远低于 5hmC,其原因一是 TET 蛋白催化 5hmC 进一步氧化成为 5fC 和 5caC 的活性相对较弱,二是 5fC 和 5caC 会被 TDG 迅速地切除掉,所以,这两种修饰的核苷酸能否作为一种稳定的表观遗传修饰发挥功能还有待进一步研究。Vermeulen 实验室利用蛋白质质谱等高通量技术,系统性地分析了多种细胞和组织中选择性结合 5hmC 的蛋白,发现不同细胞和组织中具有其组织细胞特有的结合 5hmC 的蛋白^[65]。他们还发现在神经细胞中 UHRF2 和 THY28 表达量较高,能够特异性地结合 5hmC,并且过表达 UHRF2 可以显著提高 TET1 在细胞中的催化活性。Zhou 等^[66]解析了 UHRF2 与羟甲基化修饰的 DNA 的共结晶晶体结构,并结合体外生化实验,发现 UHRF2 以二聚体的形式结合带有对称羟甲基化修饰的 CpG 二核苷酸序列。神经元细胞因为不再进行细胞分裂,基因组上的 5hmC 不会被 DNA 复制所稀释,所以,神经元基因组上的 5hmC 含量非常丰富,例如浦肯野细胞中 5hmC 约占所有 5mC 的 40%^[67],完全羟甲基化的 CpG 比例也相应

较高。如前所述,Mellen 等^[68]报道神经元的基因组上,基因内部的羟甲基化水平与基因转录呈正相关,UHRF2 在神经元中结合完全羟甲基化的 CpG 后,如何招募或促进 TET 蛋白的催化活性,以及影响基因转录的分子机制还有待更深入的研究。值得注意的是,胚胎干细胞基因组上仅有 21% 的羟甲基化修饰是在 CpG 二核苷酸序列上对称存在的^[14];Wen 等^[29]也报道在人的大脑皮层组织细胞中,基因内部 DNA 的有义链上含有更多的 5hmC,无义链上则有更多的 5mC;此外,Wu 等^[69]利用 MAB-seq (M.SssI methylase-assisted bisulfite sequencing) 技术发现在 TDG 表达被抑制的小鼠胚胎干细胞基因组上,绝大部分的 5fC/5caC 也是非对称存在的。综合以上研究成果,可以提出一个有趣的假设:活跃转录基因内部的有义链上有较多的 DNA 氧化反应发生,而借助碱基切除修复途径完成去甲基化的过程,将有可能抑制以有义链为模板进行的反义转录。进一步分析 5fC/5caC 和 BER 活性在有义链和无义链上的差异,以及其对 DNA 转录的影响,将为理解在基因内部 TET 蛋白介导 DNA 氧化反应的生物学意义提供更多的线索。

基因远端调控区域富集 5hmC,也有较多的 DNA 氧化去甲基化事件发生。因为没有证据证明 5hmC 可以像 5mC 那样在细胞分裂过程中也复制到新合成的 DNA 子链上,所以一般认为基因组上的 5hmC 一直处于动态平衡的状态。在基金委重大研究计划“细胞编程和重编程的表观遗传机制”项目的支持下,本实验室利用稳定同位素标记的细胞培养 (SILAC) 联合亲和纯化与蛋白质定量质谱技术,发现胚胎发育早期重要的多能性转录因子 SALL4A 倾向于结合含有 5hmC 修饰的 DNA^[35]。在小鼠胚胎干细胞中,SALL4A 蛋白主要定位于远端调控区域,而且其与染色质的结合在很大程度上依赖于 TET1 蛋白。SALL4A 可以稳定 TET2 在其结合的远端调控区域上的定位,通过影响 TET2 与其催化底物 DNA 的相互作用,促使 5hmC 被进一步氧化^[35]。据此,本实验室提出了 5mC 的协同递进氧化概念,结合基因组特定功能区域的蛋白,能够招募 TET 蛋白催化其结合位点附近的 5hmC 进一步氧化,从而介导去甲基化 (图 2)。在没有转录因子结合的情况下,染色质处于开放状态的基因组调控区域 DNA 将会以所谓被动填充 (passive fill-in) 的方式被甲基化^[70-71]。本实验室的研究揭示,在不同发育阶段不同组织细胞中,有可能存在类似的利用染色质

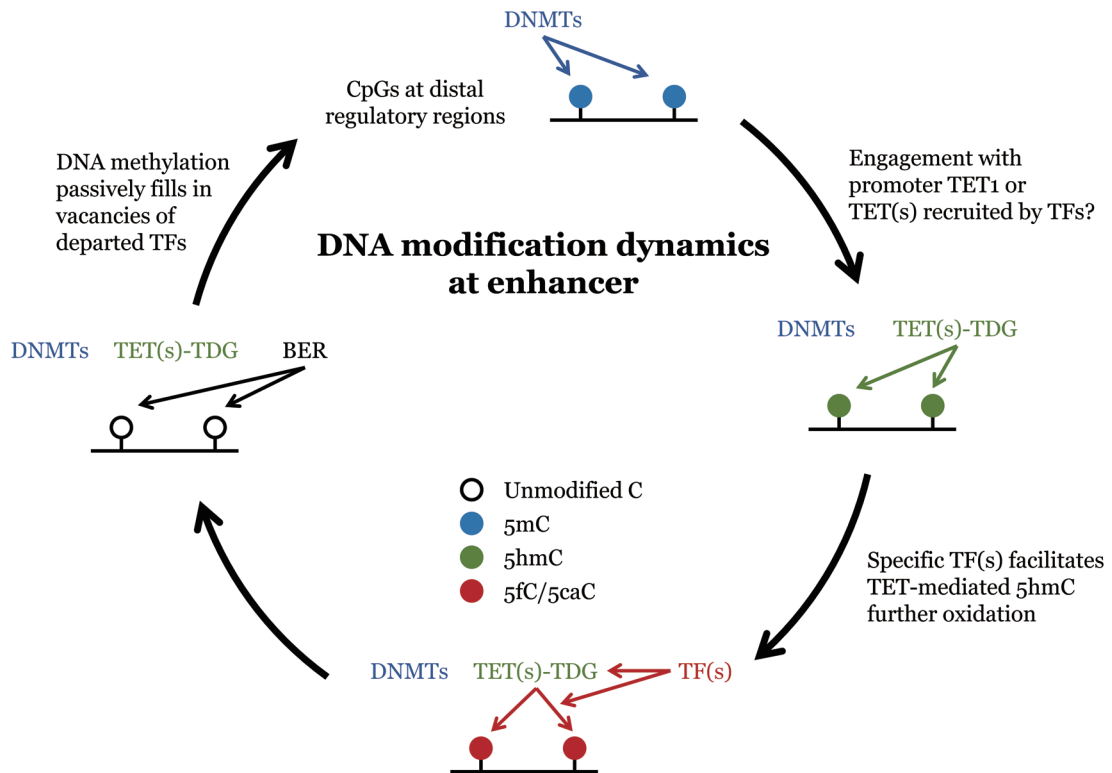


图2 基因远端调控区域DNA修饰的动态变化

或 DNA 结合蛋白, 通过调节 TET 蛋白介导的 DNA 氧化反应来维持调控区域特定功能的机制。

4 结语

DNA 氧化去甲基化的发现, 首次证实了在哺乳动物细胞中 DNA 去甲基化可以经由酶催化的生化反应来实现。近几年在研究 DNA 氧化去甲基化分子机制、生化原理以及生物学意义方面取得了丰硕的成果, 然而, 还有一些问题有待将来的研究解答: 第一, 5caC 能否通过脱羧反应生成没有修饰的胞嘧啶, 实现 DNA 去甲基化; 第二, 越来越多的表观遗传相关的修饰酶被报道具有不依赖酶催化活性的功能, 而 TET 蛋白不依赖催化活性的功能还没有被明确地解析; 第三, 5hmC 能否像 5mC 招募 MECP2 等那样作为稳定的表观遗传学标记招募特异结合蛋白。未来在这些方面的研究, 将使人们对 TET 家族蛋白介导的 DNA 氧化的机制和生物学意义有更深入的认识。

[参 考 文 献]

[1] Hermann A, Goyal R, Jeltsch A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA

processively with high preference for hemimethylated target sites. *J Biol Chem*, 2004, 279: 48350-9

[2] Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science*, 1975, 187: 226-32

[3] Riggs AD. X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet*, 1975, 14: 9-25

[4] 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

[5] Ito S, D'Alessio AC, Taranova OV, et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, 2010, 466: 1129-33

[6] 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

[7] 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

[8] Hu L, Lu J, Cheng J, et al. Structural insight into substrate preference for TET-mediated oxidation. *Nature*, 2015, 527: 118-22

[9] Hashimoto H, Liu Y, Upadhyay AK, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res*, 2012, 40: 4841-9

[10] Maiti A, Drohat AC. Thymine DNA glycosylase can

- rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem*, 2011, 286: 35334-9
- [11] Weber AR, Krawczyk C, Robertson AB, et al. Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat Commun*, 2016, 7: 10806
- [12] Xu Y, Wu F, Tan L, et al. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell*, 2011, 42: 451-64
- [13] Williams K, Christensen J, Pedersen MT, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, 2011, 473: 343-8
- [14] Yu M, Hon GC, Szulwach KE, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, 2012, 149: 1368-80
- [15] Wu H, D'Alessio AC, Ito S, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev*, 2011, 25: 679-84
- [16] 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
- [17] 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
- [18] Lu F, Liu Y, Jiang L, et al. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev*, 2014, 28: 2103-19
- [19] Ooi SK, Qiu C, Bernstein E, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature*, 2007, 448: 714-7
- [20] Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep*, 2011, 13: 28-35
- [21] Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet*, 2017, 18: 517-34
- [22] Bernstein BE, Mikkelsen TS, Xie X, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 2006, 125: 315-26
- [23] Pastor WA, Pape UJ, Huang Y, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*, 2011, 473: 394-7
- [24] Ficz G, Branco MR, Seisenberger S, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*, 2011, 473: 398-402
- [25] Baubec T, Colombo DF, Wirbelauer C, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*, 2015, 520: 243-7
- [26] Neri F, Rapelli S, Krepelova A, et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature*, 2017, 543: 72-7
- [27] Khare T, Pai S, Konciewicz K, et al. 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nat Struct Mol Biol*, 2012, 19: 1037-43
- [28] Huang Y, Chavez L, Chang X, et al. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc Natl Acad Sci USA*, 2014, 111: 1361-6
- [29] Wen L, Li X, Yan L, et al. Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol*, 2014, 15: R49
- [30] Stadler MB, Murr R, Burger L, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*, 2011, 480: 490-5
- [31] Shen Y, Yue F, McCleary DF, et al. A map of the *cis*-regulatory sequences in the mouse genome. *Nature*, 2012, 488: 116-20
- [32] Stroud H, Feng S, Morey Kinney S, et al. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol*, 2011, 12: R54
- [33] Szulwach KE, Li X, Li Y, et al. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet*, 2011, 7: e1002154
- [34] Hon GC, Song CX, Du T, et al. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol Cell*, 2014, 56: 286-97
- [35] Xiong J, Zhang Z, Chen J, et al. Cooperative action between SALL4A and TET proteins in stepwise oxidation of 5-methylcytosine. *Mol Cell*, 2016, 64: 913-25
- [36] Kaelin WG Jr, McKnight SL. Influence of metabolism on epigenetics and disease. *Cell*, 2013, 153: 56-69
- [37] Rakheja D, Konoplev S, Medeiros LJ, et al. IDH mutations in acute myeloid leukemia. *Hum Pathol*, 2012, 43: 1541-51
- [38] Dang L, White DW, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature*, 2009, 462: 739-44
- [39] Xu W, Yang H, Liu Y, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α ketoglutarate-dependent dioxygenases. *Cancer Cell*, 2011, 19: 17-30
- [40] Lu C, Ward PS, Kapoor GS, et al. IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature*, 2012, 483: 474-8
- [41] Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*, 2010, 18: 553-67
- [42] Lian CG, Xu Y, Ceol C, et al. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell*, 2012, 150: 1135-46
- [43] Xiao M, Yang H, Xu W, et al. Inhibition of α -KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev*, 2012, 26: 1326-38
- [44] Thienpont B, Steinbacher J, Zhao H, et al. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature*, 2016, 537: 63-8
- [45] 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
- [46] Ko M, An J, Bandukwala HS, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature*, 2013, 497: 122-6
- [47] Bian C, Yu X. PGC7 suppresses TET3 for protecting DNA methylation. *Nucleic Acids Res*, 2014, 42: 2893-905
- [48] Gu TP, Guo F, Yang H, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, 2011, 477: 606-10
- [49] Guo F, Li X, Liang D, et al. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell*, 2014, 15: 447-58
- [50] Shen L, Inoue A, He J, et al. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell*, 2014, 15: 459-70
- [51] Wang L, Zhang J, Duan J, et al. Programming and inheritance of parental DNA methylomes in mammals. *Cell*, 2014, 157: 979-91
- [52] Peat JR, Dean W, Clark SJ, et al. Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation. *Cell Rep*, 2014, 9: 1990-2000
- [53] Yu C, Zhang YL, Pan W, et al. CRL4 complex regulates mammalian oocyte survival and reprogramming by activation of TET proteins. *Science*, 2013, 342: 1518-21
- [54] Nakagawa T, Lv L, Nakagawa M, et al. CRL4(VprBP) E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. *Mol Cell*, 2015, 57: 247-60
- [55] Costa Y, Ding J, Theunissen TW, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, 2013, 495: 370-4
- [56] Okashita N, Kumaki Y, Ebi K, et al. PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development*, 2014, 141: 269-80
- [57] Zeng Y, Yao B, Shin J, et al. Lin28A binds active promoters and recruits Tet1 to regulate gene expression. *Mol Cell*, 2016, 61: 153-60
- [58] Wang Y, Xiao M, Chen X, et al. WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol Cell*, 2015, 57: 662-73
- [59] Perera A, Eisen D, Wagner M, et al. TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. *Cell Rep*, 2015, 11: 283-94
- [60] Shi FT, Kim H, Lu W, et al. Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *J Biol Chem*, 2013, 288: 20776-84
- [61] Zhang Q, Liu X, Gao W, et al. Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked β -N-acetylglucosamine transferase (OGT). *J Biol Chem*, 2014, 289: 5986-96
- [62] Zhang YW, Wang Z, Xie W, et al. Acetylation enhances TET2 function in protecting against abnormal DNA methylation during oxidative stress. *Mol Cell*, 2017, 65: 323-35
- [63] Bauer C, Gobel K, Nagaraj N, et al. Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J Biol Chem*, 2015, 290: 4801-12
- [64] Dai HQ, Wang BA, Yang L, et al. TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. *Nature*, 2016, 538: 528-32
- [65] Spruijt CG, Gnerlich F, Smits AH, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*, 2013, 152: 1146-59
- [66] Zhou T, Xiong J, Wang M, et al. Structural basis for hydroxymethylcytosine recognition by the SRA domain of UHRF2. *Mol Cell*, 2014, 54: 879-86
- [67] Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, 2009, 324: 929-30
- [68] Mellen M, Ayata P, Dewell S, et al. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*, 2012, 151: 1417-30
- [69] Wu H, Wu X, Shen L, et al. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nat Biotechnol*, 2014, 32: 1231-40
- [70] Thurman RE, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome. *Nature*, 2012, 489: 75-82
- [71] Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell*, 2013, 49: 825-37