

## 表观遗传新标志物 5hmC 的研究进展\*

葛苗苗 蔡春\*\*

(广东医学院湛江科研中心, 湛江 524023)

**摘要** 5-羟甲基胞嘧啶(5-hydroxymethylcytosine, 5hmC)作为表观遗传的新标志物, 已引起人们的极大兴趣. 5hmC 由 TET 家族酶催化氧化 5-甲基胞嘧啶(5-methylcytosine, 5mC)产生, 被称为高等生物基因组 DNA 的“第六碱基”. 5hmC 不仅可以影响基因组结构及功能, 还在早期胚胎发育中发挥重要的作用. 本文综述了 5hmC 的代谢通路、生物学功能、在基因组的分布及分析方法的研究进展.

**关键词** 5-羟甲基胞嘧啶, TET 蛋白, 早期胚胎发育

**学科分类号** Q523, R34

**DOI:** 10.3724/SP.J.1206.2013.00400

哺乳动物 DNA 主要存在两种胞嘧啶修饰方式, 即 5-甲基胞嘧啶(5-methylcytosine, 5mC)与 5-羟甲基胞嘧啶(5-hydroxymethylcytosine, 5hmC). 最近研究发现, 5hmC 可以被进一步氧化为 5-甲酰胞嘧啶(5-formylcytosine, 5fC)及 5-羧基胞嘧啶(5-carboxylcytosine, 5caC). 5mC 作为表观遗传的重要标志物, 已有较深入的研究, 而 5hmC 最初并不被重视, 直到近年发现可由 TET (ten-eleven translocation) 家族酶催化 5mC 产生, 才引起人们的关注. 1972 年, Penn 等<sup>[1]</sup>在成年大鼠、小鼠和青蛙的脑组织 DNA 中发现 5hmC. 不过, 当时人们认为羟甲基胞嘧啶可能来源于处理样品中 DNA 的自发性氧化损伤, 因而未引起重视. 直到 2009 年, 两个实验室同时证实 5hmC 在小鼠脑细胞与胚胎干细胞中的存在<sup>[2-3]</sup>, 才引起人们的关注. 研究表明, 与 5mC 一样, 5hmC 可能是另一个重要表观遗传学标记, 它除了参与 DNA 的主动脱甲基化过程, 还对基因的表达具有双重调控作用. 本文主要从 5hmC 在 DNA 甲基化动态变化中的作用, 5hmC 及 TET 蛋白在胚胎发育中的生物学功能, 5hmC 的基因组分布, 肿瘤组织中羟化酶基因 TET 及异柠檬酸脱氢酶基因 IDH 的改变, 以及 5hmC 的分析方法等方面综述了 5hmC 的最新研究进展.

### 1 DNA 甲基化的体内代谢及代谢中间体的相关蛋白

在细胞发育及分化中, 由 DNA 甲基转移酶 (DNA methyltransferase, Dnmt) 催化的甲基化, TET 介导的 5mC 氧化, 以及随后碱基切除修复 (base-excision repair, BER) 介导的脱甲基化建立的动态平衡, 可以转移及重建新的表观基因组状态. 在这个转变中, 5fC、5caC 及 5hmC 可能与不同的转录因子、聚合酶、潜在的识别蛋白以及其他细胞表观遗传组件相互作用, 进而影响 DNA 的甲基化和脱甲基化的改变.

#### 1.1 5hmC 与 DNA 甲基化的动态变化

在哺乳动物体内, DNA 的甲基化水平发生动态变化, 并且在分裂细胞与非分裂细胞中的脱甲基化途径不同. 研究发现, DNA 脱甲基化存在两个机制: 主动脱甲基化及被动脱甲基化. 在非分裂细

\* 国家自然科学基金资助项目(21375029)和广东医学院研究团队基金资助项目(TD1125).

\*\* 通讯联系人.

Tel: 0759-2388591, E-mail: caichun2006@tom.com

收稿日期: 2013-12-19, 接受日期: 2014-02-26

胞中，主要发生主动的脱甲基化。在此过程中，5hmC 及 TET 蛋白存在调节 DNA 甲基化动态变化的多个途径(图 1)<sup>[4]</sup>。被动脱甲基化主要与 Dnmt1、Dnmt3a 及 Dnmt3b 的表达降低或失活有关。在分裂细胞中，即便在 DNA 甲基转移酶存在的情况下，依然可以发生 DNA 的被动脱甲基化，这与 DNA 甲基转移酶 Dnmt1 对 5hmC 的识别性很差有关，5hmC 可以阻止由 Dnmt1 介导的胞嘧啶的甲基化，干扰细胞分裂中甲基化的维持，从而导致被动的脱甲基化<sup>[5]</sup>。在低 S-腺苷甲硫氨酸(S-adenosylmethionine, SAM)的条件下，Dnmts 参与 DNA 的脱甲基化，Dnmts 可以使 5mC 脱氨基生成胸腺嘧啶 T，形成 C/T 错配修复，然后通过错配修复机制(mismatch repair, MMR)纠正<sup>[6]</sup>。5mC 及

5hmC 均可经 AID (activation-induced deaminase)/APOBEC (apolipoprotein B mRNA-editing enzyme catalytic polypeptide)家族成员脱氨基，分别生成胸腺嘧啶 T 和 5-羟甲基尿嘧啶(5-hydroxymethyluracil, 5hmU)。Song 等<sup>[7]</sup>认为 5hmC 比 5mC 更易被 AID/APOBEC 家族成员脱氨基，激活 BER 介导的 DNA 脱甲基化。TET 蛋白可以将 5mC 氧化为 5hmC，并且还可以将 5hmC 进一步氧化为 5fC 及 5caC，最终经历如 TDG(thymine DNA glycosylase)及 SMUG1 (single strand selective monofunctional Uracil DNA glycosylase 1)的 BER 途径切除<sup>[8-9]</sup>。因而，在 DNA 主动脱甲基化中，5hmC 是一个中间体，而 TET 蛋白则可能是 DNA 脱甲基化的调控者。

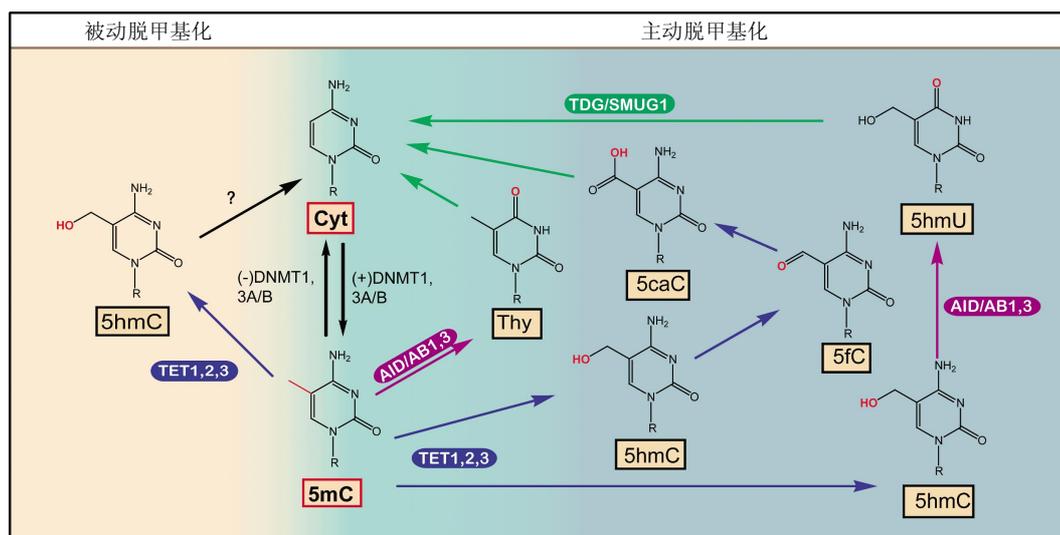


Fig. 1 DNA demethylation pathways

图 1 DNA 脱甲基化途径

此外，在图 1 中还存在 5hmC 向胞嘧啶直接转化的 DNA 被动脱甲基化。Chen 等<sup>[10]</sup>研究发现，在体外，哺乳动物的 Dnmt3a 与 Dnmt3b 具有双向作用，既可以发挥 DNA 甲基转移酶的作用，又可以作为 DNA 脱羟甲基化酶，发挥脱羟甲基化的作用。与 TET 酶介导的 5mC 向 5hmC 的转化不同，Dnmt3a 及 Dnmt3b 催化 5hmC 直接转化为胞嘧啶的脱羟甲基化反应，既不需要 Fe(II)的参与，也不需要 α-酮戊二酸(α-ketoglutaric acid, α-KG)的参与。并且，Dnmt3a 与 Dnmt3b 的完整催化位点是其发挥脱羟甲基化活性所必需的，Dnmt3a 与

Dnmt3b 的甲基化催化位点的变异，可以导致它们脱羟甲基化活性减少 50%~60%。

有研究表明 DNA 主动脱甲基化的氧化产物 5fC 及 5caC 可以降低 RNA 聚合酶 II 的转录速率及底物特异性<sup>[11]</sup>，5fC 倾向于出现在基因调控元件中非活性状态的增强子上，并且 5fC 有可能通过招募转录激活因子 p300 调控基因激活<sup>[12]</sup>。未来的研究将会集中于阐明 5fC 及 5caC 在 DNA 主动脱甲基化中的作用和机制。通过研究 5fC 及 5caC 对 RNA 聚合酶 II 转录的作用机制，或许可以为胞嘧啶的甲基化状态及转录调节之间的关系探讨提供新视点。

另外, Schiesser 等<sup>[13]</sup>发现, 在小鼠胚胎干细胞中存在 5caC 的脱羧反应. 因而, 5caC 是否可以通过一个假定的脱羧酶直接转化为胞嘧啶, 有待进一步研究.

## 1.2 5mC 氧化产物的相关蛋白

DNA 表观遗传动态调控系统中的 5mC 氧化产物 5hmC、5fC、5caC 以及有关蛋白 TET 和 TDG 的关系, 已有不少报道. 这些 5mC 的氧化产物是 DNA 主动脱甲基化的中间体, 对于发育期间的编程与重编程很重要. 有研究显示, 这些新的碱基修饰物可能拥有表观调控功能. 5hmC、5fC 及 5caC 或许也可以通过募集结合蛋白进而影响染色质的结构以及基因的表达.

在 DNA 的主动脱甲基化中, TET 与 TDG 调控 5mC 的连续氧化, BER 与不同的蛋白质共同作用, 通过 p300 或是其他的 5fC/5caC 结合蛋白来调控主动的 DNA 脱甲基化. Song 等<sup>[12]</sup>研究发现, 5fC 及 5caC 在不存在 TDG 的转录调控元件的远端富集, 并且与转录辅助激活物 p300 结合, p300 是转录因子与转录元件的桥梁, 可以促进染色质重塑及激活基因转录. Spruijt 等<sup>[14]</sup>研究发现, 在小鼠胚胎干细胞中, 5mC 的氧化产物 5hmC、5fC 及 5caC 具有富集不同转录因子、DNA 糖基化酶及 DNA 修复蛋白的作用, 且不同胞嘧啶修饰的识别蛋白很少重叠. Thy28、C3orf37 及 Neil1 蛋白可以识别结合 5mC 的所有氧化物. 5hmC 的识别蛋白 DNA 糖基化酶 Mpg、Neil3 及解旋酶 Recql 可能是通过碱基切除修复参与 DNA 的主动脱甲基化, 最终将 5hmC 转化为胞嘧啶. Iurlaro 等<sup>[15]</sup>对小鼠胚胎干细胞的 DNA 进行启动子测序, 发现 5hmC 只有少数识别结合蛋白, 如 RPL26、PRP8 和 DNA 错配修复蛋白 MHS6, 但是 5fC 有很多较强识别结合能力的蛋白质, 并且大多是修复蛋白与转录调控因子, 且 5fC 结合蛋白与转录调控及染色质调控有关, 尤其是与叉头框(forkheadbox, Fox)结构域转录因子及 NuRD(nucleosome remodeling deacetylase)复合物有关, 可见 5fC 既是 DNA 主动脱甲基化的中间体, 又可以作为一个表观遗传标记发挥作用. 5caC 不仅是 DNA 主动脱甲基化的中间产物, 并且可能在基因的转录调控中也发挥作用. Alioui 等<sup>[16]</sup>研究发现, 5hmC 及 5caC 以相似的空间分布模式存在于蝶螈滤泡细胞核的常染色质基因富集区. Kellinger 等<sup>[17]</sup>研究发现, 在小鼠胚胎干细胞中, 5caC 及 5fC 均可以降低 RNA 聚合酶 II 的转录率及

底物特异性.

总之, 胞嘧啶的氧化产物可能通过募集转录因子以及其他不太可能与 DNA 主动脱甲基化相关的蛋白, 诱导 DNA 的损伤应答, 引发碱基切除修复, 最终引起 DNA 的主动脱甲基化. Spruijt 等<sup>[14]</sup>的研究显示, Uhrf 2 及 TET1 催化区的共表达会引起 5hmC、5fC 及 5caC 水平的瞬间上调, 提示 Uhrf 2 通过 TET1 促进 5mC 的连续氧化. 因而, 弄清调控 TET 催化氧化 5mC 为 5hmC 以及进一步氧化 5hmC 为 5fC 及 5caC 的机制将是很有意义的. 另外, 许多 5hmC、5fC 及 5caC 的识别蛋白与癌症发病相关, 包括 UHRF2、CARF、p53 及 HELLS. 调节 5mC 及其氧化物的水平对于正常细胞的内稳态至关重要, 这些表观遗传标记物的失调会影响发育期间细胞增殖及分化的平衡. 未来的研究将会集中于揭示每个碱基修饰的具体作用及作用机制.

## 2 5hmC 及 TET 蛋白在胚胎干细胞分化与发育中的作用

虽然 DNA 甲基化在体细胞中具有相对的稳定性及遗传特性, 但在发育的原始生殖细胞及受精卵中会发生全基因组的 DNA 脱甲基化. DNA 甲基化的动态变化在早期胚胎发育中, 尤其是在胚胎基因组的活化、X 染色体的失活以及分化中起重要作用, 并且对于多能性因子通过脱甲基化建立多能性也是必需的. 5mC 的氧化是受精后父源基因组重编程的第一步, 是哺乳动物早期生命周期中重要的一部分.

### 2.1 5hmC 在受精卵中的含量变化

5hmC 不仅是 DNA 主动脱甲基化的一个中间产物, 还是一个影响基因组结构及功能的表观遗传学标志. 在胚胎的早期发育中, 5hmC 参与主动脱甲基化过程, 5hmC 可能在建立及维持胚胎多能性状态中起重要作用, 因为它在转录活化基因起始位点和 Polycomb 抑制基因启动子延伸区富集.

在哺乳动物的早期胚胎发育中, 全基因组的甲基化会发生明显变化. 在胚胎受精之前, 雌雄生殖细胞具有较高的甲基化水平, 受精后父源基因组迅速发生主动脱甲基化, 而母源基因组仍然保持甲基化状态. 随着卵裂的进行, 母源基因组发生被动的脱甲基化, 直到桑葚胚阶段达到最低水平. Defossez 等<sup>[17]</sup>的研究显示, 在小鼠早期胚胎发育中, 父源遗传部分与母源遗传部分的 DNA 甲基化

呈现非对称的动态变化, 并且持续很长一段时间. 5hmC 在小鼠受精卵的雄性原核中含量很高, 在雌性原核中几乎不存在, 而 5mC 存在于雌性原核中, 但不存在于雄性原核. 另外, 通过 TET3 介导的 5mC 向 5hmC 的氧化可能是启动受精卵通过重新甲基化进行重编程的一个关键步骤, 而随后 5hmC 在卵裂期通过 DNA 复制进行被动稀释可能是最终父源基因组 5mC 消除的一个机制. Inoue 等<sup>[18]</sup>发现, 在受精卵分裂中, 5hmC 同它的衍生物 5fC 和 5caC 可以通过 DNA 复制的被动稀释简单消除. 此外, 在胚胎发育中, DNA 甲基化模式具有组织特异性, 这可以解释其与细胞分化间的关系. 细胞分化程度与 5hmC 的含量之间存在很强的相关性, 终末分化细胞中 5hmC 的含量最高, 可见, 5hmC 对胚胎干细胞性状维持的重要性. Szwagierczak 等<sup>[19]</sup>发现, 5hmC 在小鼠胚胎干细胞中含量很高, 但胚胎干细胞分化后, 5hmC 的含量明显减少, 后又在终末分化细胞中再次增加. 尽管人们对 5hmC 在胚胎干细胞中的分布做了大量研究, 但 5hmC 在胚胎干细胞分化中的作用及机理仍有待进一步的研究.

## 2.2 TET1、TET2 及 TET3 在胚胎发育中的作用

TET 蛋白家族不仅可以催化 5mC 转化为 5hmC, 还在基因转录调控中具有一定作用. 小鼠胚胎干细胞中 TET 蛋白和 5hmC 分布的研究发现, TET 蛋白和 5hmC 可能在调控与多能性及细胞分化相关基因的表达中起一定作用. 目前认为, TET1 对胚胎干细胞多能性的维持具有重要作用, TET2 与机体正常造血系统的调控密切相关, 而 TET3 主要与分化有关. Zhang 等<sup>[20]</sup>的研究显示, TET1 与 TET2 在小鼠胚胎干细胞中高度表达, 并且在分化中迅速下调, 而 TET3 在胚胎干细胞中表达水平很低, 但在胚胎细胞分化中表达上调. Nanog 是近年来在胚胎干细胞中发现的一个重要的转录因子, 它在胚胎干细胞自我更新及多能性的维持中起关键作用. Costa 等<sup>[21]</sup>发现, Nanog 蛋白在维持胚胎干细胞多能性的细胞重编程中, 需要 TET1 及 TET2 蛋白的协同作用, 并且这种协同作用依赖于 TET1 及 TET2 的催化活性.

关于 TET1 在胚胎发育中的作用存在一些争论. Zhang 等<sup>[20]</sup>及 Freudenberg 等<sup>[22]</sup>认为 TET1 对调控与胚胎干细胞自我更新及谱系决定有关基因的表达是至关重要的. 然而, 也有研究表明 TET1 不是维持小鼠胚胎干细胞自我更新所必需的, Jaenisch 等<sup>[23]</sup>的研究显示, TET1 的缺失只是引起 5hmC 水

平的部分降低, 不影响小鼠胚胎干细胞的多能性及自新能力, 并且小鼠的胚胎发育及产后发育正常. 因此, 只有弄清 5hmC 及 TET 调控这些至关重要基因的机制, 才能确定它们在小鼠胚胎干细胞的自我更新及多能性维持中的作用. TET2 及 5hmC 与产后造血系统的发育有关, 并且 TET2 表达的改变或 5hmC 水平的改变, 可能会破坏细胞的动态平衡, 最终导致造血细胞的增殖及成熟异常. Shide 等<sup>[24]</sup>研究发现, TET2 在造血干细胞自我更新的调控中起重要作用, 与野生型小鼠相比, TET2 功能丢失小鼠的造血干细胞增殖快, 而造血干细胞的增殖异常与髓系恶性血液病的发病有关. 在受精卵的发育中, 母源的 TET3 是催化父源基因组中 5mC 氧化为 5hmC 的关键酶. Iqbal 等<sup>[25]</sup>发现, 在卵母细胞及受精卵中, TET1 及 TET2 的表达水平很低, 而 TET3 在受精卵及卵母细胞中高度表达, 但在卵裂期迅速消失. 研究显示, TET3 与受精卵的发育有关. TET3 的纯合突变可以导致新生儿死亡, 突出 TET3 在胚胎发育期多个器官发育中的重要作用<sup>[26]</sup>. 另外, 有研究显示, TET 蛋白似乎不仅仅是 5mC 羟化酶及转录调节因子, 它除参与多能性的维持及胚胎发育, 还可能参与 RNA 的加工<sup>[11]</sup>.

至今为止, 为何 TET1 的表达及 5hmC 的含量在胚胎干细胞较高, 以及 TET1、TET2、TET3 与 5hmC 在着床后胚胎发育的生物学作用还不清楚. 因而, 5hmC 及 TET 在胚胎干细胞、胚胎发育中的作用及作用机制有待深入研究.

## 3 5hmC 在基因组的分布

TET1 及 5hmC 在干细胞中对基因表达具有双重调节作用, 一方面促进多能性基因的转录, 另一方面抑制发育调节基因的表达, 这两种作用可以有效维持干细胞的多能性状态.

TET1 及 5hmC 均在含有转录活化标记 H3K4me3(histone H3 lysine 4 trimethylation)及富含 PRC2(Polycomb repression complex 2)的转录抑制标记 H3K27me3(histone H3 lysine 27 trimethylation)二价结构的基因启动子区高度富集. Ito 等<sup>[27]</sup>的研究显示, TET1 倾向于结合在高度转录基因及 PRC2 抑制基因的启动子, 它在胚胎干细胞基因转录调控中具有双重作用, 不仅能调控 CpG 富集启动子处的 DNA 甲基化水平, 还能促进干细胞中与多能性相关因子的转录, 并参与 Polycomb 靶向的发育调控因子的抑制. Neri 等<sup>[28]</sup>研究发现, 在胚胎干细胞

中, PRC2 将 TET1 募集到转录抑制标记 H3K27me3 区域, 且 TET1 存在两个富集区, 一个是与 Sin3A 共存于非羟甲基化的转录起始位点, 另一个是与 PRC2 及 5hmC 共存于转录起始位点下游. Wu 等<sup>[29]</sup>对小鼠胚胎干细胞全基因组范围内 5hmC 分布的研究发现, 5hmC 既在转录活化的基因体内存在, 也在 Polycomb 抑制的发育调节基因启动子区富集.

5hmC 片段主要集中在外显子和转录起始位点 (transcriptional start site, TSS) 附近. Wu 等<sup>[29]</sup>对 5hmC 的分布进行分析, 发现 5hmC 优先富集于基因富集区, 近 60% 的 5hmC 富集在基因体, 包括 5' 非翻译区 (5' untranslated regions, 5' UTR)、编码的外显子、内含子及 3' 非翻译区 (3' untranslated regions, 3' UTR). 5hmC 与基因表达的调控元件可能存在一定的联系, 它通常富集于 CpG 富集的基因启动子、增强子、多能性转录因子结合位点及绝缘子的 CTCF (CCCTC-binding factor, CCCTC-结合因子) 结合位点<sup>[30]</sup>. Sun 等<sup>[31]</sup>通过限制性内切酶 *AbaSI* 结合测序的方法绘制 E14 小鼠胚胎干细胞 5hmC 的分布, 经分析发现 5hmC 一般在一些重复子区域富集, 在 CTCF 结合位点及带有 H3K4me1 (monomethylated histone H3 lysine 4) 标记的非活性状态的增强子观察到 5hmC 的富集. 5hmC 倾向于结合在高度转录基因的基因体 (尤其是外显子) 以及 PRC2 抑制基因的启动子, 它在基因转录调控中的具体作用取决于存在位置. Zhang 等<sup>[32]</sup>分析含有 CpG 岛的基因启动子, 发现 5hmC 在 Polycomb 抑制基因启动子区, 尤其是 TSS 的近端上游及基因体的 5' 端高度富集, 而在转录活化且仅含 H3K4me3 的基因内含子, 尤其是基因体的 3' 端富集. 此外, 5hmC 倾向于在中等 CpG 密度的 TET1 结合位点富集, 这与 TET1 的酶活性有关<sup>[29, 33]</sup>. 5hmC 在高 CpG 密度的 TET1 结合位点含量很低, 这是因为 TET1 可以快速催化 5mC 发生羟甲基化, 形成 5hmC, 进而转化为未甲基化的胞嘧啶, 而且 CpG 富集区高水平的 H3K4me3, 可以有效地阻止胚胎干细胞中的 Dnmt3a2/Dnmt3b/Dnmt3L 与这些区域的 DNA 序列结合, 从而抑制 5mC 的累积. 因而, 5hmC 在高转录活化标志 H3K4me3 基因启动子含量很低, 甚至不存在, 但在单价组蛋白抑制标记 H3K27me3 基因启动子富集.

然而, Pastor 等<sup>[30]</sup>与 Williams 等<sup>[33]</sup>的研究没有发现 5hmC 与基因的转录水平或是 RNA 聚合酶 II

的活性有明显的关系. Song 等<sup>[34]</sup>的研究也没有发现 5hmC 对 RNA 聚合酶 II 的转录速率及底物特异性有明显影响. 因而, 5hmC 与基因转录活性之间是否存在直接的关系有待于进一步研究.

#### 4 5hmC 的环境调节

新陈代谢在调节基因组 DNA 甲基化及染色质结构方面的作用已经引起人们的重视, 研究发现, 环境因素可以通过干扰代谢来影响全基因组的表观遗传模式, 进而影响基因的表达. Fang 等<sup>[34]</sup>研究发现, 饮食、遗传基因多态性和环境中的化学物质的作用, 均可导致 DNA 甲基化状态的改变. 因此, 研究环境因素对代谢的影响, 也许可以更好地理解环境因素对染色体、DNA 甲基化以及羟甲基化模式的影响, 加速认识表观遗传学改变在人类生长、发育及疾病发展过程中的作用机制.

在 CpG 二核苷酸中, 胞嘧啶是 DNA 甲基化的主要碱基, 而鸟嘌呤是 DNA 氧化损伤的位点, 8-羟基脱氧鸟苷 (8-hydroxy-2'-deoxyguanosine, 8-OHdG) 是 DNA 氧化损伤的敏感生物标志物. CpG 二核苷酸中鸟嘌呤的氧化会减少此位点的甲基结合结构域 (methyl binding domain, MBD), 从而抑制邻近胞嘧啶的甲基化. Zawia 等<sup>[35]</sup>研究发现, 环境毒素如重金属铅, 在大脑发育中能够抑制 DNA 甲基转移酶的活性, 并且可以通过改变 CpG 二核苷酸的甲基化或氧化, 促进与阿尔茨海默病 (Alzheimer's disease, AD) 发病有关的基因表达, 促进大脑中 DNA 的氧化损伤. Pilsner 等<sup>[36]</sup>在子宫内铅暴露对脐带血基因组 DNA 甲基化影响的研究显示, 产前胚胎发育暴露于铅环境会影响新生儿的表观基因组, 并且可能会影响整个生命进程中的长期表观遗传编程, 增加疾病易感性.

此外, 发育暴露于环境毒素中, DNA 甲基化模式的改变在一定程度上可能是由氧化应激介导的代谢改变来调控的. Sirtuin 是 NAD<sup>+</sup> 依赖性哺乳动物去乙酰酶, Sirt3 是去乙酰化酶 Sirtuin 家族的一员, 它可以使线粒体中的 IDH2 通过去乙酰化而活化, IDH2 在线粒体内可以将 NADP<sup>+</sup> 转化为 NADPH, 引起 NADPH 水平的增加. NADPH 及 NADH 是启动合成代谢的生化反应及提供还原能力对抗氧化应激所需的辅因子. Wang 等<sup>[37]</sup>研究发现, 环境毒素可以引起氧化应激, 而氧化应激降低 NADPH 及 NADH 的水平, 并且伴随着线粒体中 NAD<sup>+</sup> 含量的增加, 这可以激活 NAD<sup>+</sup> 依赖性去乙

酰酶 Sirtuin. 因而, 他们推测 Sirt3 可以通过激活线粒体中的 IDH2, 增加  $\alpha$ -KG 的生成量, 激活将 5mC 氧化为 5hmC 的 TET 酶, 进而引起基因组中 5mC 与 5hmC 模式的改变(图 2). Thomson 等<sup>[38]</sup>研究发现, 发育暴露于非基因毒性致癌物苯巴比妥中, 小鼠肝脏中一些转录上调的基因近端启动子区 5mC 及 5hmC 的模式发生相反的动态改变, 但 5hmC 及 5mC 在转录起始位点附近均减少, 提示这些位点的 DNA 可能通过中间体 5hmC 发生主动的脱甲基化, 因而 5hmC 及 5mC 的水平可以作为药物反应的检测指标.

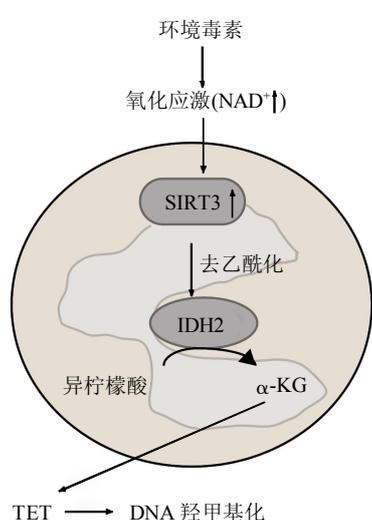


Fig. 2 Regulation of TET by environmental toxins

图 2 环境毒素对 TET 酶的调控

环境对人类及哺乳动物模型 DNA 甲基化模式的改变, 在一定程度上是通过氧化应激介导的代谢改变来调节的. 特定基因启动子的 5hmC 与 5mC 水平可以作为暴露于环境毒素的生物标志物. 每种环境毒素以其独特的方式影响代谢与染色质结构, 这可以作为反映环境毒素剂量及类型的“指纹”. 因而, DNA 甲基化与羟甲基化的分析, 将成为环境毒理学家研究环境毒素对人体健康影响的极为重要手段. 然而, 环境毒素通过影响代谢改变表观遗传修饰的具体作用机制还有待于深入的研究.

## 5 肿瘤组织中代谢酶基因 TET 及 IDH 的异常表达与 5hmC 水平的变化

研究发现, 与癌旁正常组织相比, 肿瘤组织中 5hmC 的含量显著降低. Jin 等<sup>[39]</sup>通过液相色谱串联质谱法以及基于 5hmC 抗体的免疫斑点印迹和免疫

组化分析方法, 发现肺癌、脑癌、乳腺癌、肝癌以及黑色素瘤等癌症中存在 5hmC 含量的大幅下降.

肿瘤中 5hmC 含量的显著下降可以通过三个方面来解释. a. 细胞的过度增殖可以引起 5hmC 含量的下降, 这是因为 5mC 向 5hmC 的氧化形成半羟甲基化的 CpG 岛, 在 DNA 复制期间, Dnmt1 不能识别半羟甲基化的 CpG 岛, 因而在细胞分裂中, 不能维持 DNA 的羟甲基化, 随着细胞的分裂, 其中 5hmC 的含量逐渐下降. b. 肿瘤组织中 TET 酶的表达水平较正常组织低, TET 酶表达水平的下降会引起 5mC 向 5hmC 的转化减少. c. 肿瘤组织中 5hmC 含量下降的原因还可解释为产生 TET 酶发挥催化氧化活性所需辅因子  $\alpha$ -KG 的代谢途径受损, 因而 TET 酶不能很好地发挥其羟化酶活性, 5mC 向 5hmC 的转化受影响.  $\alpha$ -KG 是三羧酸循环 (tricarboxylic acid cycle, TCA) 中异柠檬酸的代谢产物, 异柠檬酸在产生  $\alpha$ -KG 的代谢中需要 IDH1 及 IDH2 的催化.

哺乳动物细胞中 NADPH 产生的一个机制是在异柠檬酸向  $\alpha$ -KG 转化的反应中产生的, 由 NADP 依赖性酶 IDH1 及 IDH2 催化. IDH1 与 IDH2 是同源酶, 分别在细胞质及线粒体中发挥作用. IDH1 与 IDH2, 在结构及功能上与 NAD 依赖性酶 IDH3 不同, IDH3 在 TCA 中发挥作用, 产生氧化及磷酸化所必需的 NADH. 肿瘤中 IDH1 或 IDH2 的衍生突变, 会导致从产生 NADPH 向消耗 NADPH 的转变, 这可能对细胞氧化还原平衡的维持有重要影响, 并且会使其失去正常的催化产生  $\alpha$ -KG 的功能, 取而代之的是 2-羟戊二酸 (2-hydroxyglutarate, 2-HG) (图 3)<sup>[40-41]</sup>. Xu 等<sup>[42]</sup>研究发现, 2-HG 是包括组蛋白脱甲基化酶与 5mC 羟化酶 TET 家族在内的多个  $\alpha$ -KG 依赖性双加氧酶的竞争性拮抗剂.  $\alpha$ -KG 及 Fe(II) 依赖性双加氧酶 TET1、TET2 及 TET3 负责将 5mC 氧化为 5hmC, 2-HG 抑制 TET1、TET2 以及 TET3 的羟化酶活性, 从而导致 5hmC 的减少. 5mC 向 5hmC 转化的减少必然会引起 CpG 岛 DNA 甲基化的累积. 因此, IDH1/2 的突变与 DNA 甲基化的增加及 5hmC 水平的降低有关.

IDH 突变与许多肿瘤的发病具有相关性. 肿瘤中 IDH 的突变有很好的预后判断及诊断作用, 并且可以为临床上某些肿瘤提供新的治疗靶点. 许多肿瘤中存在 IDH1 及 IDH2 的突变, 包括神经胶质瘤、急性髓系白血病 (acute bone marrow of leukemia,

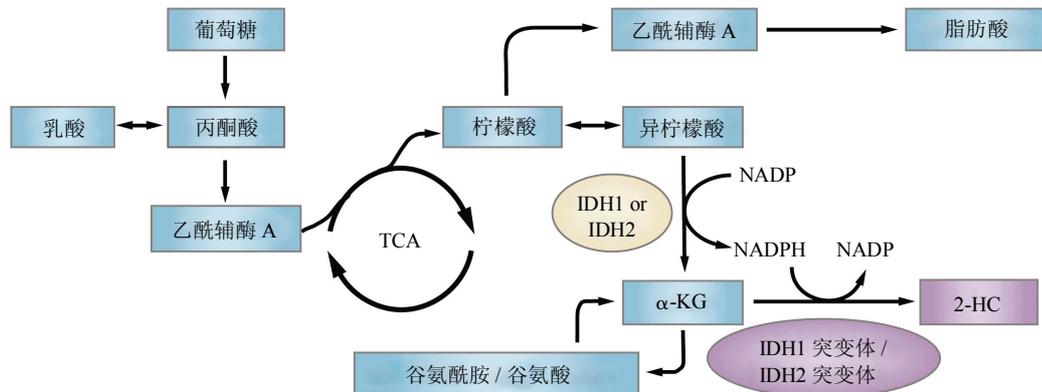


Fig. 3 IDH1 and IDH2 mutations cause an oncometabolic gain of function

图3 IDH1及IDH2的突变导致肿瘤代谢物的产生

AML)、急性淋巴细胞白血病(acute lymphocytic leukemia, ALL)、骨髓纤维化症、肝内胆管癌、黑色素瘤、软骨样肿瘤、少数的结肠癌及前列腺癌甚至一些副神经节瘤<sup>[43]</sup>,其中以多形性成胶质细胞瘤及软骨肉瘤中最为突出<sup>[44]</sup>。

在神经胶质瘤中, IDH 基因的突变是一个很好的预后生物标志物, 并且有助于临床上的诊断分类。Pan 等<sup>[45]</sup>在对神经胶质瘤中 IDH1 及 IDH2 的突变状态与临床病理特征的相关性研究中发现, 在早期阶段, IDH 的体细胞杂合突变频率很高, 并且 p53、表皮生长因子受体(epidermal growth factor receptor, EGFR)表达的检测与 IDH 突变的分析结合有助于神经胶质瘤的鉴别诊断。Zhang 等<sup>[46]</sup>的研究显示, 在 IDH 突变的退行性神经胶质瘤中, 有较高比例的 O6-甲基鸟嘌呤-DNA 甲基转移酶(O6-methylguanine-DNA methyltransferase, MGMT)启动子甲基化, 并且肿瘤更多地发生在大脑额叶, 更少地发生在颞叶。不过, 这种相关性还需要进一步的研究确证。Frenel 等<sup>[47]</sup>研究发现, 结合 IDH 突变及染色体 1p/19q 杂合性缺失, 可以将退行性神经胶质瘤分为预后特征不同的三类。此外, Yao 等<sup>[48]</sup>研究发现, IDH1 的突变与神经胶质瘤的恶化无关, 但可以作为有效的诊断标志物, 带有 IDH1 突变的病人拥有更长的无恶化存活期及总生存期。在 AML 中, IDH 突变不是独立的生存预测指标, 它与预后的相关性很复杂, 可能取决于等位基因、病人年龄及治疗方案等。带有 R132H IDH1 突变的 AML 病人拥有较低的总生存期, 带有 R172K IDH2 突变的病人拥有较低的完全缓解率, 且较带有

R140Q IDH2 突变的病人预后差。在儿童 AML 中, IDH 的突变与总生存期的增加有关。在大剂量化疗的年轻人中, R140Q IDH2 的突变与好的预后有关, 但在老年 AML 患者中, IDH 的突变, 尤其是 R172K IDH2 突变与不良预后相关<sup>[49]</sup>。在 ALL 中, IDH 的突变可能有不良的预后。Zhang 等<sup>[50]</sup>发现在 ALL 中, IDH1 R132 的突变可能是一个复发性标志, 带有 IDH1 R132 突变的病人呈现异常的髓系抗原表达, 但 IDH 在 ALL 发病中的具体作用还需要深入研究。此外, 在甲状腺癌中也发现 IDH1 的突变, Hemeryly 等<sup>[51]</sup>的研究显示, 24% 的甲状腺滤泡癌, 18% 的滤泡型乳头状癌以及 8% 的甲状腺乳头状癌存在 IDH1 的突变。目前还没有在甲状腺癌中发现 IDH2 的突变。

由于 IDH1/2 突变提供的信息有助于肿瘤的临床诊断及治疗效果判断, 因而, 发展 IDH1/2 可靠、灵敏及高通量的分析方法, 具有重要的应用价值。

5mC 向 5hmC 的转化需要 TET 蛋白的催化, 因而 TET 酶基因表达或功能发挥的异常, 会引起 5hmC 含量的下降, 并且还与其他某些肿瘤的发病有关。TET1 最初发现是染色体易位基因, 在某些白血病患者体内与 MLL(mixed lineage leukemia)基因融合, 但其在造血系统的发育及白血病形成中的作用还是有争议的, 需要进一步的研究。有研究显示, TET1 在肝癌及结直肠癌等实体瘤中的表达水平明显下降, 并且 TET1 可能是通过去甲基化来促进金属蛋白酶组织抑制因子(tissue inhibitors of metalloproteinase, TIMP)家族蛋白 TIMP2 及 TIMP3 的表达, 以抑制癌细胞的侵袭<sup>[52-54]</sup>。神经胶质瘤的

发病大多与 IDH1 及 IDH2 的突变有关, IDH1 在 II 级及 III 级神经胶质瘤中的突变很频繁. 而 Müller 等<sup>[55]</sup>研究发现, 细胞核 TET1 排斥的高频发生可能是与 IDH1 突变与否无关的神经胶质瘤细胞中 5hmC 减少的一个可替代途径. 另外, Kim 等<sup>[56]</sup>发现, 在野生型 IDH1/2 低级弥散性星形细胞瘤中, TET2 基因启动子的甲基化或许是其一个致病机制. TET2 是骨髓增生异常综合症 (myelodysplastic syndrome, MDS) 及其他类型白血病中研究最多的 TET 蛋白酶家族基因, 并且是髓系恶性血液病中突变频率最高的基因之一. TET2 参与造血干细胞的调节机制, 对正常骨髓细胞发育至关重要, 其酶活性的失活可能会导致髓系肿瘤的发生<sup>[24]</sup>. 在 MDS 病人中, TET2 蛋白表达的下调与原癌基因 miR-22(microRNA-22) 的过表达有关. Song 等<sup>[57]</sup>研究发现, miR-22 通过作用于 TET2 发挥致癌作用, miR-22 转基因小鼠表现出全基因组 5hmC 的低水平以及造血干细胞自我更新的增加与异常分化, 进而发展为 MDS 及血液恶性肿瘤. 在髓系肿瘤中, TET2 功能的丢失与大范围的低甲基化有关. Ko 等<sup>[58]</sup>发现 TET2 突变的多种髓系恶性血液病患者的骨髓样本显示全基因组 5hmC 的低水平, 并且许多与催化结构域有关的 TET2 位点的突变表现出与其酶活性无关. 因此, 白血病形成中 TET2 功能丢失的分子机制目前仍不清楚, 并且需要深入的研究, 检测髓系恶性血液病中 5hmC 的含量可能是诊断及预后的一个很有价值的手段, 并且对评估抗癌药物的疗效及确定治疗方法也有一定的价值. 另外, Lian 等<sup>[59]</sup>在黑色素瘤的研究中发现, 羟化酶 TET2 在黑色素瘤中的表达水平明显下调, 并且 TET2 的重新引入可以恢复 5hmC 的水平, 降低黑色素瘤细胞的代谢能力, 抑制黑色素瘤的生长. 至今, TET3 与癌症的直接相关性还没有报道, 可能是由于在三个 TET 蛋白中, TET3 是维持人体正常生理活动最重要的调控者, 至关重要的 TET3 的突变或调节异常可能会导致死亡.

人们对 5hmC 及 TET 蛋白与癌症的发生已进行了大量的研究, 但它们在癌症中的具体作用及作用机制还不清楚. 因而分析与癌症相关的全基因组羟甲基化区域及 TET 蛋白的结合位点将有助于确定它们在癌症发展中的作用. 此外, 需要对 TET 蛋白的催化活性依赖功能及非催化活性依赖功能进行区分, 以确定其催化结构的破坏是否影响其发挥某些功能.

## 6 5hmC 的主要分析方法

作为第六种碱基, 5hmC 近年来受到了广泛的关注, 但 5hmC 在基因组范围内的分布, 目前仍了解不多. DNA 特定区域和位点羟甲基化的分析对于 5hmC 生物学功能的研究具有重要意义. 一些 DNA 甲基化常用的定性及定量分析方法, 比如以亚硫酸盐测序及甲基化敏感酶为基础的方法, 不能区分 5hmC 与 5mC, 导致 5hmC 的分析仍存在很多困难. 目前 5hmC 的分析方法主要有以下几种:

### 6.1 限制性内切酶法

对于既含有胞嘧啶, 又含有 5- 甲基胞嘧啶及 5- 羟甲基胞嘧啶的基因位点, 可以运用 *Msp* I / *Hpa* II 酶切的 DNA 葡糖化方法来分析其中的 5hmC. 虽然 *Msp* I 和 *Hpa* II 识别相同的核苷酸序列(CCGG), 但是对不同的甲基化状态敏感. *Msp* I 识别并切割未修饰的胞嘧啶及 5mC, 对 5hmC 则无影响, *Hpa* II 则识别并切割未修饰的胞嘧啶. 经内切酶处理之后, 可开展多种下游分析, 如 qPCR、新一代测序、DNA 印迹、芯片等. 此法操作简便, 主要缺点是它只能用于分析 CCGG 序列中的 5hmC<sup>[60]</sup>. 此外, Sun 等<sup>[61]</sup>在测定小鼠胚胎干细胞 5hmC 全基因组分布的研究中, 建立了限制性内切酶 *Aba*S I 结合测序的方法, 该法灵敏度高、特异性强、分辨率高.

### 6.2 酶放射性糖基化标记法

以酶结合放射性标记的葡萄糖进入基因组 5hmC 为基础, Szwagierczak 等<sup>[62]</sup>设计了一种通过放射性计数进行定量的方法, 将含有放射性标记的葡萄糖残基 UDP-[<sup>3</sup>H]-Glu 转移至 hmC 形成 [<sup>3</sup>H]-ghmC. 该方法灵敏度较高, 并且不需要特别专业及昂贵的设备, 但是酶促反应一般不彻底. Song 等<sup>[61-62]</sup>先用化学标记取代了放射性标记, 通过 T4 噬菌体的  $\beta$ - 葡糖基转移酶(T4  $\beta$ -GT)将含化学修饰叠氮的葡萄糖残基 UDP-6-N<sub>3</sub>-Glu 转移至 5hmC 形成 N<sub>3</sub>-5ghmC, 后又通过  $\beta$ -GT 将一个酮取代的葡萄糖分子转移至 5hmC 上, 比先前方法中用到的叠氮取代的葡萄糖转移途径更有效, 并且提供了一个独立于 DNA 序列的单碱基分辨率的 5hmC 检测方法.

### 6.3 沉淀法

利用一些酶和化学反应的组合来分离出包含单个 5hmC 的 DNA 片段, 然后采用沉淀或免疫沉淀技术可达到对 5hmC 的特异性分析. 主要包括糖基

化、高碘酸氧化和生物素化(glucosylation, periodate oxidation, biotinylation, GLIB)处理, 抗 5- 亚甲基磺酸胞嘧啶(antiserum to cytosine 5-methylenesulphonate, anti-CMS)、JBPI(J-binding protein)沉淀及 5hmC 免疫沉淀(5hmC immunoprecipitation, 5hmC-IP)等。

5hmC-IP 是目前研究 5hmC 较为常用的方法之一, 也是较其他沉淀技术更为直接的方法。5hmC-IP 具有良好的灵敏性, 但不适用于单碱基分辨率的 5hmC 分析, 以及 5mC 与 5hmC 靠得很近的 DNA 序列中 5hmC 的分析<sup>[29]</sup>。

Anti-CMS 是基因组 DNA 经亚硫酸氢钠处理, 将 5hmC 转换成 5- 亚甲基磺酸胞嘧啶(cytosine 5-methylenesulphonate, CMS), 然后利用 CMS 特异抗体免疫沉淀含 CMS 的 DNA 序列。该法具有很高的灵敏性, 可以检测 DNA 中极少量的 5hmC。对于 CMS 富集的基因组 DNA, 使用了 Illumina 测序仪, 其较长的读长适合亚硫酸氢盐处理的 DNA 与基因组的高效比对<sup>[30, 58]</sup>。

JBPI 沉淀是通过 T4  $\beta$ -GT 选择性糖基化 5hmC, 形成  $\beta$ -glu-5hmC, 提供一个可以被 JBPI 蛋白有效选择性沉淀到磁珠上的靶点。沉淀的 DNA 适用于 qPCR 分析、基因芯片分析及测序分析。该法简便、经济且可以有效地选择性识别基因组 DNA 中的 5hmC<sup>[63]</sup>。

GLIB 具有很高的灵敏性, 可以分离低至仅含一个 5hmC 的 DNA 片段。对于 GLIB 处理的 DNA, 采用 Helicos 单分子 DNA 测序, 这种方法不需要扩增步骤, 从而避免了 PCR 偏向。与 5hmC-IP 相比, 选择性标记 5hmC 后再对标记物进行免疫沉淀则具有不依赖 5hmC 的浓度和更高灵敏度的特点<sup>[30]</sup>。

#### 6.4 液相色谱串联质谱法

近年来, 建立了 5mC 及 5hmC 的液相色谱串联质谱分析方法, 该法快速、灵敏、准确性高、检测限低, 可以同时测定基因组 DNA 中 5mC 与 5hmC 的水平, 比现有的 5hmC 定量方法, 如末端标记的薄层色谱法及放射性标记的糖基化法, 更灵敏简便。Munzel 等<sup>[64]</sup>首先通过 LC-MS (liquid chromatography mass spectrometry) 联用精确定量了小鼠脑组织内 5hmC 的分布情况, Le 等<sup>[65]</sup>运用反相色谱柱建立了一种更快速、稳定且准确的 5hmC 的 LC-MS/MS (liquid chromatography tandem mass spectrometry) 分析方法, 测定了小鼠胚胎干细胞中 5hmC 的含量、人类诱导性多能干细胞及成纤维细

胞中 5mC 与 5hmC 的水平。Munzel 等及 Le 等都是利用酶将 DNA 水解为核苷酸, 存在酶解不完全而影响准确度的可能性, 且耗时长, 成本高。本实验室 Zhang 和 Chen 等<sup>[66-68]</sup>采用亲水色谱柱建立了测定组织中全基因组 5hmC 水平的 LC-MS/MS 方法, 该方法的 DNA 样品用甲酸裂解, 缩短了样品处理时间, 提高了分析灵敏度及效率, 有效解决了 DNA 可能因酶解不完全而影响准确度的问题。LC-MS/MS 在定量方面有许多优势, 已经得到广泛应用, 具有可以扩展到其他罕见核苷酸测定的灵活性, 能满足生物样品中 5hmC 定量分析的要求。

尽管目前已经发展了多种 5hmC 的分析方法, 但是仍有许多有待解决的问题。主要集中在特定区域 5hmC 的分析方面。如在免疫沉淀测序中, 由于缺乏可特异性沉淀 5hmC 的抗体, 使 DNA 的沉淀率较低。此外, 5hmC 高密度区域的测序也比较困难。未来将会集中于特定位点 5hmC 分析方法的研究, 发展灵敏、高通量和低成本的 5hmC 分析方法。

## 7 展 望

5hmC 作为胞嘧啶修饰过程的中间产物, 与 5mC 一样是重要的表观遗传学标记物, 这些发现丰富了我们对 DNA 甲基化的表观遗传调控过程的认识。TET 蛋白和 5hmC 的研究为 DNA 甲基化 / 脱甲基化及其生物学功能研究提供了新观点, 为表观遗传学的研究开辟了全新的研究领域。5hmC 是一种很有前景的生物标志物, 与人类多种疾病的发生及发展密切相关。未来的研究可能集中于更好地阐明 5hmC 和 TET 蛋白的生物学功能, 探讨 TET 蛋白在癌症发展中的作用和原理。另外, 饮食是影响健康的重要因素, 饮食中的营养素提供了表观修饰作用所需的反应底物, 其供应量也对表观修饰作用有影响, 饮食及各种环境因素是否影响 TET 酶的活性, 进而诱导不同的 DNA 羟甲基化, 有望成为以后的重要研究方向。

## 参 考 文 献

- [1] Penn N W, Suwalski R, O'Riley C, *et al.* The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J*, 1972, **126**(4): 781-790
- [2] Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science*, 2009, **324**(5929): 929-930
- [3] Tahiliani M, Koh K P, Shen Y, *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 2009, **324**(5929): 930-935

- [4] Bhutani N, Burns David M, Blau Helen M. DNA demethylation dynamics. *Cell*, 2011, **146**(6): 866–872
- [5] Valinluck V, Sowers L C. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Research*, 2007, **67**(3): 946–950
- [6] Szyf M. DNA methylation and demethylation probed by small molecules. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 2010, **1799**(10–12): 750–759
- [7] Guo Junjie U, Su Y, Zhong C, *et al.* Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, 2011, **145**(3): 423–434
- [8] He Y F, Li B Z, Li Z, *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 2011, **333**(6047): 1303–1307
- [9] Ito S, Shen L, Dai Q, *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 2011, **333**(6047): 1300–1303
- [10] Chen C C, Wang K Y, Shen C K J. The mammalian *de novo* DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *J Biol Chem*, 2012, **287**(40): 33116–33121
- [11] Kellinger M W, Song C X, 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**(8): 831–833
- [12] Song C X, Szulwach K E, Dai Q, *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, 2013, **153**(3): 678–691
- [13] Schiesser S, Hackner B, Pfaffeneder T, *et al.* Mechanism and stem-cell activity of 5-carboxylcytosine decarboxylation determined by isotope tracing. *Angew Chem Int Ed Engl*, 2012, **51**(26): 6516–6520
- [14] Spruijt C G, Gnerlich F, Smits A H, *et al.* Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*, 2013, **152**(5): 1146–1159
- [15] 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**(10): R119
- [16] Alioui A, Wheldon L M, Abakir A, *et al.* 5-Carboxylcytosine is localized to euchromatic regions in the nuclei of follicular cells in axolotl ovary. *Nucleus*, 2012, **3**(6): 565–569
- [17] Defossez P A, Salvaing J, Aguirre-Lavin T, *et al.* 5-methylcytosine and 5-hydroxymethylcytosine spatiotemporal profiles in the mouse zygote. *PLoS One*, 2012, **7**(5): e38156
- [18] Inoue A, Shen L, Dai Q, *et al.* Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res*, 2011, **21**(12): 1670–1676
- [19] Szwagierczak A, Bultmann S, Schmidt C S, *et al.* Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*, 2010, **38**(19): e181
- [20] Ito S, D'Alessio A C, Taranova O V, *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, 2010, **466**(7310): 1129–1133
- [21] Costa Y, Ding J, Theunissen T W, *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, 2013, **495**(7441): 370–374
- [22] Freudenberg J M, Ghosh S, Lackford B L, *et al.* Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. *Nucleic Acids Research*, 2011, **40**(8): 3364–3377
- [23] Dawlaty M M, Ganz K, Powell B E, *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell*, 2011, **9**(2): 166–175
- [24] Shide K, Kameda T, Shimoda H, *et al.* TET2 is essential for survival and hematopoietic stem cell homeostasis. *Leukemia*, 2012, **26**(10): 2216–2223
- [25] Iqbal K, Jin S G, Pfeifer G P, *et al.* Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci USA*, 2011, **108**(9): 3642–3647
- [26] Gu T P, Guo F, Yang H, *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, 2011, **477**(7366): 606–610
- [27] Wu H, D'Alessio A C, Ito S, *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*, 2011, **473**(7347): 389–393
- [28] Neri F, Incarnato D, Krepelova A, *et al.* Genome-wide analysis identifies a functional association of Tet1 and Polycomb PRC2 in mouse embryonic stem cells. *Genome Biol*, 2013, **14**(8): R91
- [29] Wu H, D'Alessio A C, Ito S, *et al.* Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Gene Develop*, 2011, **25**(7): 679–684
- [30] Pastor W A, Pape U J, Huang Y, *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*, 2011, **473**(7347): 394–397
- [31] Sun Z, Terragni J, Borgaro J G, *et al.* High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell Rep*, 2013, **3**(2): 567–576
- [32] Wu H, Zhang Y. Tet1 and 5-hydroxymethylation: A genome-wide view in mouse embryonic stem cells. *Cell Cycle*, 2011, **10**(15): 2428–2436
- [33] Williams K, Christensen J, Pedersen M T, *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, 2011, **473**(7347): 343–348
- [34] Fang M, Chen D, Yang C S. Dietary polyphenols may affect DNA methylation. *J Nutr*, 2007, **137**(1 Suppl): 223S–228S
- [35] Zawia N H, Lahiri D K, Cardozo-Pelaez F. Epigenetics, oxidative stress, and Alzheimer disease. *Free Radic Biol Med*, 2009, **46**(9): 1241–1249
- [36] Pilsner J R, Hu H, Ettinger A, *et al.* Influence of prenatal lead exposure on genomic methylation of cord blood DNA. *Environ Health Perspect*, 2009, **117**(9): 1466–1471

- [37] Wang L, Chia N C, Lu X, *et al.* Hypothesis: Environmental regulation of 5-hydroxymethylcytosine by oxidative stress. *Epigenetics*, 2011, **6**(7): 853–856
- [38] Thomson J P, Lempiainen H, Hackett J A, *et al.* Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol*, 2012, **13**(10): R93
- [39] Jin S G, Jiang Y, Qiu R, *et al.* 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Research*, 2011, **71**(24): 7360–7365
- [40] Cairns R A, Harris I S, Mak T W. Regulation of cancer cell metabolism. *Nature Reviews Cancer*, 2011, **11**(2): 85–95
- [41] Gross S, Cairns R A, Minden M D, *et al.* Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *Journal of Experimental Medicine*, 2010, **207**(2): 339–344
- [42] Xu W, Yang H, Liu Y, *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell*, 2011, **19**(1): 17–30
- [43] Horbinski C. What do we know about IDH1/2 mutations so far, and how do we use it?. *Acta Neuropathol*, 2013, **125**(5): 621–636
- [44] Losman J A, Kaelin W G, Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes Dev*, 2013, **27**(8): 836–852
- [45] 潘怡, 齐雪岭, 王雷明, 等. 胶质瘤患者异柠檬酸脱氢酶基因突变分析. *中华病理学杂志*, 2013, **42**(5): 292–298  
Pan Y, Qi X L, Wang L M, *et al.* *Chin J Pathology*, 2013, **42**(5): 292–298
- [46] Zhang C B, Bao Z S, Wang H J, *et al.* Correlation of IDH1/2 mutation with clinicopathologic factors and prognosis in anaplastic gliomas: a report of 203 patients from China. *J Cancer Res Clin Oncol*, 2013, **140**(1): 45–51
- [47] Frenel J S, Leux C, Loussouarn D, *et al.* Combining two biomarkers, IDH1/2 mutations and 1p/19q codeletion, to stratify anaplastic oligodendroglioma in three groups: a single-center experience. *J Neurooncol*, 2013, **114**(1): 85–91
- [48] Yao Y, Chan A K, Qin Z Y, *et al.* Mutation analysis of IDH1 in paired gliomas revealed IDH1 mutation was not associated with malignant progression but predicted longer survival. *PLoS One*, 2013, **8**(6): e67421
- [49] McKenney A S, Levine R L. Isocitrate dehydrogenase mutations in leukemia. *J Clin Invest*, 2013, **123**(9): 3672–3677
- [50] Zhang Y, Wei H, Tang K, *et al.* Mutation analysis of isocitrate dehydrogenase in acute lymphoblastic leukemia. *Genet Test Mol Biomarkers*, 2012, **16**(8): 991–995
- [51] Hemerly J P, Bastos A U, Cerutti J M. Identification of several novel non-p.R132 IDH1 variants in thyroid carcinomas. *Eur J Endocrinol*, 2010, **163**(5): 747–755
- [52] Hsu C H, Peng K L, Kang M L, *et al.* TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. *Cell Rep*, 2012, **2**(3): 568–579
- [53] Liu C, Liu L, Chen X, *et al.* Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1. *PLoS One*, 2013, **8**(5): e62828
- [54] Yang H, Liu Y, Bai F, *et al.* Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene*, 2013, **32**(5): 663–669
- [55] Müller T, Gessi M, Waha A, *et al.* Nuclear exclusion of TET1 is associated with loss of 5-hydroxymethylcytosine in IDH1 wild-type gliomas. *Amer J Pathol*, 2012, **181**(2): 675–683
- [56] Kim Y H, Pierscianek D, Mittelbronn M, *et al.* TET2 promoter methylation in low-grade diffuse gliomas lacking IDH1/2 mutations. *J Clin Pathol*, 2011, **64**(10): 850–852
- [57] Song S J, Ito K, Ala U, *et al.* The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. *Cell Stem Cell*, 2013, **13**(1): 87–101
- [58] Ko M, Huang Y, Jankowska A M, *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, 2010, **468**(7325): 839–843
- [59] Lian Christine G, Xu Y, Ceol C, *et al.* Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell*, 2012, **150**(6): 1135–1146
- [60] Kinney S M, Chin H G, Vaisvila R, *et al.* Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. *J Biol Chem*, 2011, **286**(28): 24685–24693
- [61] Song C X, Szulwach K E, Fu Y, *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*, 2010, **29**(1): 68–72
- [62] Song C X, Sun Y, Dai Q, *et al.* Detection of 5-hydroxymethylcytosine in DNA by transferring a keto-glucose by using T4 phage  $\beta$ -glucosyltransferase. *Chem Bio Chem*, 2011, **12**(11): 1682–1685
- [63] Robertson A B, Dahl J A, Vagbo C B, *et al.* A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. *Nucl Acid Res*, 2011, **39**(8): e55–e55
- [64] Munzel M, Globisch D, Bruckl T, *et al.* Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angewandte Chemie*, 2010, **49**(31): 5375–5377
- [65] Le T, Kim K P, Fan G, *et al.* A sensitive mass spectrometry method for simultaneous quantification of DNA methylation and hydroxymethylation levels in biological samples. *Anal Biochem*, 2011, **412**(2): 203–209
- [66] 陈丽宇, 张立坚, 张良滔, 等. 液相色谱-串联质谱分析组织中基因组脱氧核糖核酸羟甲基胞嘧啶水平. *色谱*, 2012, **30**(5): 533–537  
Chen L Y, Zhang L J, Zhang L T, *et al.* *Chin J Chromatography*, 2012, **30**(5): 533–537
- [67] Zhang L, Zhang L, Zhou K, *et al.* Simultaneous determination of global DNA methylation and hydroxymethylation levels by hydrophilic interaction liquid chromatography-tandem mass spectrometry. *J Biomol Screen*, 2012, **17**(7): 877–884
- [68] Zhang L T, Zhang L J, Zhang J J, *et al.* Quantification of the sixth DNA base 5-hydroxymethylcytosine in colorectal cancer tissue and C-26 cell line. *Bioanalysis*, 2013, **5**(7): 839–845

## Research Advances on The New Epigenetic Marker: 5-Hydroxymethylcytosine\*

GE Miao-Miao, CAI Chun\*\*

(Guangdong Medical College, Zhanjiang 524023, China)

**Abstract** Recently, the new epigenetic marker, 5-hydroxymethylcytosine (5hmC) has been widely studied. 5hmC derived from the enzymatic oxidation of 5-methylcytosine (5mC) by TET(ten-eleven translocation) protein family and called as the "six base" of higher organism genomic DNA. It is found that 5hmC not only influence the structure and function of genome, but also play an important role during early embryonic development. This review mainly discussed the research advances of 5hmC in metabolic pathways, biological function, genome distribution and analysis methods.

**Key words** 5-hydroxymethylcytosine, TET proteins, early embryonic development

**DOI:** 10.3724/SP.J.1206.2013.00400

---

\* This work was supported by grants from The National Natural Science Foundation of China (21375029) and The Research Team Fund of Guangdong Medical College (TD1125).

\*\*Corresponding author.

Tel: 86-759-2388591, E-mail: caichun2006@tom.com

Received: December 19, 2013 Accepted: February 26, 2014