

cal potential supply the energy for the secretion. The molecule which has secretion signal on its C terminal is helped by the Hly family members to pass the inner and outer mem-

brane at the same time while omitting the periplasmic space.

Key words protein secretion, nascent peptide, protein factors, secretion signal, *E. coli*

CpG 甲基化与基因调控

康毅滨 吴晓晖 魏 勇 柴建华

(复旦大学遗传学研究所, 上海 200433)

摘要 CpG 双核苷酸中的胞嘧啶甲基化和去甲基化在哺乳动物的基因表达中有重要的调控作用. 哺乳动物基因组中有两类启动子: CpG 岛启动子和 CpG 缺乏启动子. 两种蛋白质因子通过与甲基化 CpG 的相互作用影响基因表达. CpG 岛在基因组分析中也有广泛的用途.

关键词 甲基化, 基因表达, CpG 岛, 基因组分析

CpG 双核苷酸中胞嘧啶环第 5 位碳原子的甲基化是哺乳动物最为重要的一种 DNA 修饰, 约 75% 的 CpG 位点是甲基化的. 哺乳动物的基因组根据 DNA 甲基化程度的不同可以分为两个部分^[1]: 约 98% 的 DNA 中, CpG 双核苷酸大约每 50~100 bp 出现一个, 而且甲基化程度很高. 其余约 2% 的 DNA 中, 低甲基化的 CpG 双核苷酸约 10 bp 出现一个. 后者虽然只占基因组的一小部分, 却引起人们广泛的兴趣. 这一部分 DNA 以 45 000 多个约 1 kb 大小的片段散布于基因组中, 称为 CpG 岛. 它们与基因的 5' 端密切相关^[2], 包括了基因的启动子和 5' 端的一个或数个内含子. 在人类基因组中大约 60% 的基因 (包括全部持家基因和 40% 的组织特异性表达基因) 含有 CpG 岛^[3]. DNA 的甲基化对哺乳动物的正常发育有重要的调控作用, 并且这种调控很大程度上是通过 CpG 的甲基化与去甲基化起作用的. 因此, DNA 甲基化和 CpG 的研究将增进我们对基因调控和人类基因组的认识.

1 DNA 甲基化在哺乳动物发育和进化中的意义

甲基化的 DNA 是诱发突变的主要因素之

一, 尽管 CpG 在哺乳动物基因组中的出现频率只有预期频率的 1/5, 但约有 1/3 人类遗传病可归因于由 CpG 到 TpG 的颠换所引起的点突变, CpG 的甲基化可大大增加突变的概率. 为了逃避这种危险, 大部分动物的基因组甲基化程度很低, 而高等脊椎动物则不然, 甲基化 DNA 广泛分布于它们的基因组中, 包括许多编码基因^[4]. 因此 DNA 甲基化必然对哺乳动物的正常发育起着不可缺少的作用以致于能抵消由突变概率增大带来的遗传负荷.

早在 1975 年, Riggs 就推测甲基化可能是基因表达调节系统的基础. 此后, 在 1979 年 Taylor 等人的实验首先证实 DNA 甲基化可能在哺乳动物的发育中扮演一定的角色, 他们发现 10T1/2 细胞用甲基化抑制剂 5-氮胞苷处理后能分化成肌肉细胞和一些其它的类型细胞. 之后的许多研究成果进一步证实甲基化对基因表达有重要的调控作用, 这些成果主要有: a. 许多在分化细胞中特异性表达的基因, 它们的转录调控序列是非甲基化的, 而在其它类型细胞中相应的这些序列是甲基化的; b. 在特定

位点甲基化模式的改变相应地会引起基因表达活力的改变; c. 甲基化位点可在许多次细胞分裂后得以维持并可能形成后生突变 (epigenetic mutations); d. CpG 岛的甲基化引起的基因关闭可能是细胞在体外培养中失去类型特征的原因; e. 甲基化会改变 DNA 分子的空间构型从而影响其与一些蛋白质因子的结合。

但以上这些实验一般是在体外培养的细胞中进行的, 不能完全反映基因在体内的情况, 此外 5-氮胞苷除抑制甲基化外还有其它一些副作用。并且许多低等动物的正常发育看来并不依赖甲基化作用。甲基化到底在哺乳动物基因调控和发育中起多大作用? E. Li 等^[5]的实验回答了这一问题, 他们用基因靶技术使小鼠生殖细胞的甲基转移酶发生突变, 由此育成的纯系突变小鼠胚胎发育迟缓, 并在胚胎发育的一定时期死亡。这一结果清楚地说明: 甲基化作用的缺乏会导致哺乳动物发育异常并有致死作用。

甲基化对哺乳动物的重要性可以从两个方面进行解释: 首先, 已知 DNA 甲基化可以抑制转录, 因此通过对转座因子、前病毒等有害序列表达的抑制可以保护哺乳动物细胞免受侵害, 这和原核生物的限制-修饰系统有异曲同工之处。其次, 甲基化与哺乳动物发育的许多现象密切相关, 如 X 染色体失活、基因组印迹以及抗体编码基因的 V (D) J 重排等。失活 X 染色体上基因的启动子部位往往是甲基化的, 而非失活 X 染色体上的相应位置则是非甲基化的。Sapienza 等^[5]还发现子代从亲代得到一些等位基因一个是高甲基化的, 另一个为低甲基化, 这可能是基因组印迹的形成原因。最近的一些研究成果还表明: 基因组印迹可能被哺乳动物用来防止单性生殖的发生。此外, 由过高或过低甲基化引起的不正常基因组印迹与癌症的发生有关^[7]。但甲基化究竟是基因失活的原因还是结果? Lock 等研究了 X 染色体失活与甲基化的时序, 发现甲基化发生在染色体失活之后, Enver 等也对 γ 珠蛋白基因进行了这一研究并得到相同的结果。Bird^[2]由此认为: 甲基

化可能被细胞用来在抑制基因表达的初始信号消失之后维持基因的失活状态。

DNA 的 CpG 甲基化可以显著降低幼淋巴细胞内的 V (D) J 重组^[8], 这可能是由于甲基化使 DNA 具有抵抗 V (D) J 重组酶的能力。由此我们可以猜想, 淋巴细胞基因组中的 CpG 甲基化可以防止由于重组酶错误作用于其它位点而造成的突变。从进化的意义上讲, 由于哺乳动物寿命较长且基因组最为复杂, 在 DNA 复制、重排等过程中因 DNA 内切酶的错误作用而发生突变的可能性很大, 甲基化 CpG 可以防止这种情况的发生, 因而成为脊椎动物 (特别是哺乳动物) 基因组的特有现象。当然, 甲基化 CpG 同时又可以增加突变的概率, 因此哺乳动物中的 mCpG 的量应处于尽可能低的水平上, 这便可以解释为何哺乳动物基因组中 CpG 双核苷酸的出现频率远低于预期值。

2 DNA 甲基化和 CpG 调控基因表达的机理

近年来由于发现了两种与 mCpG 结合的蛋白: MeCP1 和 MeCP2, 人们对 DNA 甲基化对基因表达所起的调控作用有了一些初步的认识。MeCP1 能和至少包含 12 个 mCpG 的 DNA 片段结合。前面提到 DNA 甲基化还可能改变 DNA 分子的构型, 从而通过影响 DNA 与转录因子的结合来抑制其转录活性, 但 Boyes 等人发现如果 MeCP1 的浓度很低, 基因往往不会因甲基化而失活^[9]。因此 MeCP1 与甲基化失活有密切的联系。MeCP2 可以与一个 mCpG 结合, Lewis^[10]通过原位免疫荧光的方法证实 MeCP2 一般集中于富含甲基化 CpG 的染色体异染色质部位, 但 MeCP2 的生物学作用仍不很明了。

DNA 甲基化对转录的影响机理如图 1 所示, 调节作用与启动子附近 CpG 岛浓度和启动子强弱有关。稀疏分布的 CpG 的甲基化可以关闭弱启动子所控制的基因但不能关闭强启动子所控制的基因。如果高密度的 CpG 被甲基化则即使强启动子也不能使基因获得表达。这一调

节系统尤如一个开关, 它的关闭或开启取决于 mCpGs 对 MeCP1 的亲合力与启动子强度之间的平衡.

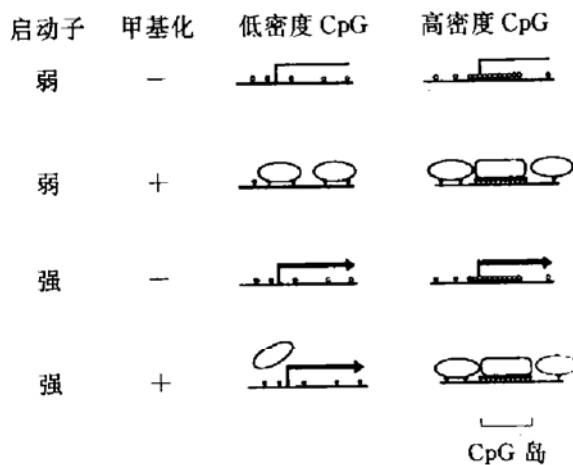


图 1 DNA 甲基化控制基因转录的作用模型^[11]

转录进行与否与 3 个因素有关: 1. mCpGs 与启动子的距离; 2. mCpG 的密度; 3. 启动子的强弱. 椭圆框指与 mCpGs 松散结合的 MeCP1, 方框为与 mCpGs 紧密结合的 MeCP1. 粗细箭头分别指出强弱启动子控制的转录. 白圆圈为非甲基化的 CpG 双核苷酸, 黑圆圈为甲基化的 CpG.

哺乳动物基因组中有两类启动子: 第一类为富含 CpG, 恒定非甲基化的启动子, 称为 CpG 岛启动子 (CpG-island promoters). 另一类启动子所含 CpG 密度较低且在大部分细胞中是高度甲基化的, 称为 CpG 缺乏启动子 (CpG-deficient promoters), 这类启动子仅发现于组织特异性表达基因中.

第一类启动子所控制的大多数基因在一般情况下处于持续开放状态, 这些基因即所谓持家基因 (housekeeping genes). 但在一些情况下 (如失活 X 染色体、基因组印迹以及脆性 X 染色体等), CpG 岛会被甲基化并能结合 MeCP1 而后使基因持久关闭. 第二类启动子中的 CpG 在大多数细胞中是甲基化的, 基因被关闭, 而在进行特异性表达的细胞中, 这些启动子附近通过重排等过程引入一个增强子, 这样即使 CpG 是甲基化的也不能阻止转录进行. 前面已经提到过组织特异性表达基因在被激活之后它所含的 CpG 往往失去甲基化, Sullivan 通过基

因活化和去甲基化的时序研究认为甲基化是基因活化的结果. 去甲基化一方面可能作为基因活化的印记, 另一方面可能进一步提高基因组特异性表达的水平.

3 CpG 岛在基因组分析中的应用

根据 Larsen 等人在 1992 年对欧洲分子生物学实验室数据库 (EMBL database) 所作的调查, 约 57% 的人类基因包含 CpG 岛, 这其中有一半是持家基因, 另有一半是组织特异性表达基因. 根据已经发表的一些基因顺序, CpG 岛几乎总是包括它们所联系的基因 5' 端的几个外显子. 而 CpG 岛又能被某些稀切点酶检出, 所以这一方法可以用来快速寻找大片段 DNA 中的基因. CpG 两侧的 DNA 可以作为探针通过 Northern 杂交的方法从高表达的组织 (如脑) 中得到与此 CpG 岛相关的基因所转录出的 mRNA, 也可用来筛选 cDNA 文库. 利用这种方法能够检出即使在细胞中只有微量表达的基因所相应的 mRNA 和 cDNA, 而其它手段如动物杂交 (zoo blotting) 便不能作到这一点.

由于 CpG 岛一般位于基因的 5' 端因此用这种方法得到的大部分 cDNA 将是完整的; 此外, cDNA 序列在基因组中常分布在数十到数百 kb 的范围中, 而 CpG 岛则只有约 1 kb 大小, 因此利用 CpG 岛可以准确地将基因定位在基因组上; 最后由于 CpG 岛大多为包含一些稀切点酶识别位点的单拷贝顺序, 因而是构建基因组物理图谱的一种理想的位标. 当然, 这种方法也存在一些缺点, 其中主要一点是: 由于这一 CpG 岛只能涉及 60% 的基因, 其余 40% 的基因将不能用这种方法进行研究.

近年来人们一般用稀切点酶寻找 CpG 岛, 并将它应用于基因组分析中. 它唯一的缺点是必须用脉冲电泳来分析酶切结果并且 CpG 岛两侧的 DNA 序列需要进一步克隆. Patel 等^[12]设计了一种新的克服这些缺点的方法, 它是将大片段 DNA 用 Not1 酶切, 然后接上一个接头, 以 Alu 顺序和接头顺序为引物进行 PCR 扩增, PCR 产物再克隆进载体.

最近 Cross 等人^[13]发展了一种新的方法, 他们将 MeCP2 改造成只保留与 mCpG 结合的位点 (MBD), 用 MBD 制成的层析柱可以吸附一定甲基化程度的 CpG 双核苷酸. 利用这种层析柱并结合使用甲基化酶, 他们从人总 DNA 中 (已用 MseI 切成约占 760 bp 的片段) 得到了 CpG 岛库, 由于 MseI 的作用位点一般不位于 CpG 岛内, 因此所得到的 CpG 岛大部分将是完整的. 当然这种方法仍需进行改进 (如改用别的酶, 从而既能不破坏 CpG 岛又能将总 DNA 切成约 1 kb 大小). 可以相信 CpG 岛库的建立将有助于人类基因组的研究工作.

参 考 文 献

- 1 Bird A P, Taggart M, Frommer M *et al.* Cell, 1985; **40**: 91
- 2 Bird A. Trends Genet, 1987; **3**: 342
- 3 Larsen F, Gundersen G, Lopez R *et al.* Genomic, 1992; **13**: 1095
- 4 Bird A P. Nature, 1986; **321**: 209
- 5 Li E, Bestor T H, Jaenisch R. Cell, 1992; **69**: 915
- 6 Sapienza C, Peterson A C, Rossaut J *et al.* Nature, 1987; **328**: 251
- 7 Feinberg A P. Nature Genet, 1994; **4**: 110
- 8 Lin C, Lieber M R. EMBO J, 1992; **11**: 315
- 9 Boyes J, Bird A. Cell, 1991; **64**: 1123
- 10 Lewis J D, Meehan R R, Henzel W J *et al.* Cell, 1992; **68**: 905
- 11 Bird A. Cell, 1992; **70**: 5
- 12 Patel K, Cox R, Shipley J *et al.* Nucl Acids Res, 1991; **19**: 4371
- 13 Cross S H, Charlton J A, Nan X *et al.* Nature Genet, 1994; **6**: 236

CpG Methylation and Gene Regulation. Kang Yibin, Wu Xiaohui, Wei Yong, Chai Jianhua (*Institute of Genetics, Fudan University, Shanghai 200433, China*).

Abstract The methylation and demethylation of cytosine in CpG dinucleotide plays an important role in regulation of mammalian gene expression. There are two kinds of promoters in mammalian genome: CpG-island promoters and CpG-deficient promoters. Two protein factors influence gene expression by interacting with methylated CpGs. CpG islands also have useful applications in genome analysis.

Key words methylation, gene expression, CpG island, genome analysis

载脂蛋白 A- I 基因表达调控的研究进展

尹银亮 王克勤 陈保生

(中国医学科学院基础医学研究所, 北京 100005)

摘要 载脂蛋白 (apolipoprotein, apo) A- I 是一种重要的血浆蛋白, 其基因表达具有明显的组织特异性、种属特异性和发育阶段特异性. Apo A- I 基因 5' 调控区的肝细胞特异性增强子等顺式作用元件同 ARP-1、HNF-4、RXR α 等反式作用因子相作用, 调控该基因的特异性表达.

关键词 apo A- I 基因, 基因表达, 调控

Apo A- I 是血浆高密度脂蛋白的主要载脂蛋白, 对机体脂质运输和代谢起着重要作用. 目前, 人、猴、狒狒、猪、大鼠、兔和鸡等动物的 apo A- I 基因已先后被克隆, 各种动物的

apo A- I 基因主要在肝脏和小肠中表达, 呈现出明显的组织特异性. 另一方面, 不同种属的