

表观遗传学 技术前沿

主编 姜怡邓 杨晓玲 张慧萍



科学出版社

感谢国家自然科学基金(81360052, 81460121、81560086, 81670416, 81660258)、宁夏自然科学基金(NZ15209、NZ15066、NZ15211)和宁夏科技领军创新人才(2015, 2016)项目的资助

表观遗传学技术前沿

主 编 姜怡邓 杨晓玲 张慧萍

副主编 马胜超 杨安宁

编 委 (以姓氏笔画为序)

丁 宁	马晓莉	马燕琼	王艳华
王 楠	毛彩艳	邓 梅	朱光荣
孙 岳	李 凡	李淑强	杨松昊
张 辉	和杨杨	赵清国	侯苗苗
徐灵博	高婷婷	郭 伟	谢 琳

科 学 出 版 社

北 京

内 容 简 介

生命是一个复杂的过程,“DNA-RNA-蛋白质”这一中心法则不能解释所有生命现象。表观遗传学是在这些例外的基础之上发展而成的,是基于非基因序列改变所致基因表达水平变化,如DNA甲基化和染色质重塑等。表观遗传学已经部分颠覆了我们对生物结构和行为方式的理解。本书共分为七篇,主要内容涵盖DNA甲基化、组蛋白修饰、miRNA调控研究前沿和实验技术。在主要阐述了表观遗传学的基本知识的基础上,结合本领域国内外最新科研成果,系统总结介绍了当前表观遗传学领域的前沿进展和最新研究方法。

图书在版编目(CIP)数据

表观遗传学技术前沿 / 姜怡邓, 杨晓玲, 张慧萍主编. —北京: 科学出版社, 2017.1

ISBN 978-7-03-051495-0

I. ①表… II. ①姜… ②杨… ③张… III. ①表观遗传学-研究 IV. ①Q3

中国版本图书馆 CIP 数据核字 (2017) 第 002456 号

责任编辑: 王 颖 / 责任校对: 郭瑞芝

责任印制: 徐晓晨 / 封面设计: 陈 敬

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2017年1月第一版 开本: 787×1092 1/16

2017年1月第一次印刷 印张: 16 3/4

字数: 394 000

定价: 98.00 元

(如有印装质量问题, 我社负责调换)



目 录

第一篇 表观遗传学概述及分析方法

第一章 表观遗传学概述····· 1	第三章 组蛋白修饰分析方法····· 15
第二章 DNA 甲基化分析方法····· 8	第四章 非编码 RNA 相关分析方法····· 19

第二篇 基因表达检测

第五章 基因表达检测概述····· 26	第三节 原位杂交技术····· 35
第六章 基因表达检测方法····· 27	第四节 定点突变技术····· 41
第一节 RT-PCR 技术····· 27	第五节 基因芯片····· 43
第二节 荧光定量 PCR 技术····· 29	第六节 启动子区活性分析····· 50

第三篇 miRNA 检测技术

第七章 miRNA 概述····· 52	第三节 荧光定量 PCR 技术检测 miRNA····· 62
第八章 miRNA 检测技术概述····· 56	第四节 荧光素酶活性检测技术····· 68
第一节 生物信息学····· 56	第九章 miRNA 分析的应用····· 70
第二节 miRNA 芯片技术····· 59	

第四篇 蛋白质检测技术

第十章 蛋白检测概述····· 82	第三节 免疫荧光技术····· 95
第十一章 蛋白检测技术概要····· 87	第四节 免疫组化技术····· 100
第一节 蛋白芯片技术····· 87	第五节 ELISA 技术····· 103
第二节 蛋白印迹技术····· 91	第六节 激光共聚焦技术····· 115

第五篇 DNA 甲基化与组蛋白修饰检测

第十二章 DNA 甲基化与组蛋白修饰 概述····· 119	甲基化芯片技术····· 145
第十三章 DNA 甲基化与组蛋白修饰方法 概述····· 136	第五节 DNA 甲基化检测——质谱 技术····· 147
第一节 生物信息学分析····· 136	第六节 组蛋白甲基化检测····· 150
第二节 DNA 甲基化检测——MSP 法····· 138	第七节 组蛋白乙酰化检测····· 155
第三节 DNA 甲基化检测——焦磷酸测 序法····· 141	第八节 组蛋白糖基化检测····· 158
第四节 DNA 甲基化检测——DNA	第九节 组蛋白泛素化检测····· 162
	第十节 组蛋白甲基化和组蛋白乙酰化 位点特异性分析····· 163

第六篇 基因与蛋白相互作用检测

第十四章 基因与蛋白相互作用检测		第五节 电泳迁移率变动实验	182
概述	166	第六节 Southern 印迹杂交	185
第十五章 基因与蛋白相互作用检测方法		第七节 表面等离子共振	189
概述	167	第八节 酵母单杂交系统	194
第一节 生物信息学时代	167	第九节 噬菌体展示技术	198
第二节 ChIP 技术	172	第十节 核酸适体技术	207
第三节 DNase I 足迹实验	180	第十一节 拉下实验	219
第四节 甲基化干扰试验	181		

第七篇 蛋白质与蛋白质相互作用

第十六章 蛋白与蛋白相互作用概述	221	第四节 GST 融合蛋白纯化技术	245
第十七章 蛋白与蛋白相互作用检测方法		第五节 羟基蛋白质足迹法	249
概述	229	第六节 细菌双杂交系统	252
第一节 生物信息学软件预测	237	第七节 酵母双杂交技术	255
第二节 染色质免疫共沉淀	242	第十八章 蛋白质与蛋白质相互作用	
第三节 印迹叠加或 Far-Western 印迹	245	应用	259

参考文献	260
------	-----

第一篇 表观遗传学概述及分析方法

第一章 表观遗传学概述

表观遗传调控是当前分子遗传学研究的热点之一，同时也是基因表达调控的重要组成部分。目前其研究重点主要集中在组蛋白共价修饰、DNA 甲基化、基因组印记、染色质重塑和非编码 RNA 等方面。目前，针对表观遗传修饰的研究方法取得了突飞猛进的进展，其表现在两方面：一方面，方法的特异性和灵敏度都在不断得到提高；另一方面，检测表观修饰的分析方法正在逐步从个别位点向高通量检测的方向发展，同时由定性检测向定量分析的方向发展。除此之外，随着研究的发现，新一代测序技术的应用将大大推动表观遗传研究的发展，这些技术包括单分子纳米孔测序法、单分子实时测序法等。近年来，随着不断深入研究，研究人员开发出各种研究方法以满足不同类型研究的需要。本篇目的在于主要介绍目前已有的表观基因组学研究方法并对其进行简要总结和分析。

一、表观遗传学发展历史

早在 2000 多年前古希腊哲学家亚里士多德在 *On the Generation of Animals* 一书中首先提出后生理理论 (the theory of epigenesis)，它主要相对于先成论，其内容是描述新器官的发育由未分化的团块逐渐形成的。生物学家 Waddington CH 于 1939 年首先在《现代遗传学导论》中提出了“epigenetics”这一术语，并于 1942 年将表观遗传学划为生物学的一大分支，其是研究基因与决定表型的基因产物之间的因果关系。

科学家 Holliday R 于 1975 年较为准确地描述了表观遗传学的基本信息。他认为表观遗传学不仅存在于发育过程中，而且可以在成体阶段研究可遗传的基因表达改变，这些信息能经过减数分裂和有丝分裂在个体和细胞间世代传递，而不借助于 DNA 序列的改变，简而言之，表观遗传是非 DNA 序列差异的核遗传。

二、表观遗传学概念

所谓表观遗传学是指研究的生物通过调控基因表达，同时却不改变基因序列所导致的遗传现象的理论。表观遗传学研究的重点是基因表达的调控机制方式，包括组蛋白共价修饰、DNA 甲基化、基因组印记、染色质重塑和非编码 RNA 等调控方式。该理论最早是在 1939 年由英国学者 Waddington CH 在其著作《现代遗传学导论》中提出并形成的初步理论。表观遗传学是对经典遗传学的补充与进一步的发展，涉及何时、何地以何种方式去应用遗传学信息。我们认识的基因组包括两类遗传信息：即表观遗传学信息及 DNA 序列遗传信息。细胞及人体正常功能的维持均为这两种信息互相影响、保持动态平衡的结果，如果这两种因素的任何一种出现表达失衡，都有可能导导致人体疾病的发生。因此，表观遗传学研究是生命科学中一个普遍而又至关

重要、不可忽视的新的研究领域,其不仅在基因的表达、遗传、调控方面发挥重要作用,而且在生命发育、免疫、炎症、血管新生、衰老及再生医学、肿瘤发生、变异性疾病的发生与防治中起着极其重要的作用(表 1-1-1)。综上所述,研究表观遗传学的重要性不亚于 20 世纪 50 年代沃森和克里克发现 DNA 双螺旋结构所引发的关于染色体上基因的研究。

表 1-1-1 遗传学与表观遗传学的比较

项目	表观遗传学	遗传学
两者关系	表观遗传学是遗传学的分支学科	
研究基础	基因序列不改变	基因序列发生改变
研究内容	DNA 甲基化、组蛋白修饰、RNA 调控、染色体重塑等	基因突变、基因杂合丢失等

三、表观遗传学的主要调节机制

表观遗传学的主要调节机制方式包括 DNA 甲基化、组蛋白甲基化与乙酰化及非编码 RNA 等三种方式。然而研究发现,机体生活的环境与这些调节机制的改变密切相关,每个生物个体都有特定的属于自己的基因组与表观基因组,表观基因组是指在不改变 DNA 序列的情况下抑制或激活基因的表达,而这种由表观基因组所调控的基因表达往往又受多种环境因素的影响。换句话说,机体日常呼吸的空气、饮用的水、食用的食物、所处的环境当中所潜在因素的影响均可对基因表达的关闭或开启产生巨大的影响和作用。因此,我们可以知道表观遗传学更加强调生活的环境对人体表观遗传因素的影响。

四、表观遗传学对基因的表达调控

(一) 基因选择性表达的调控

基因选择性表达的调控包括 DNA 的甲基化、组蛋白的共价修饰、染色体重塑、X 染色体失活和基因印记等。

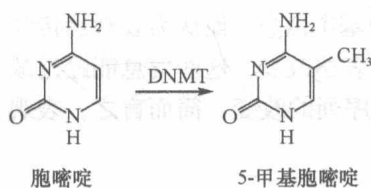


图 1-1-1 甲基胞嘧啶结构图

1. DNA 甲基化 (DNA methylation) 是指在 DNA 甲基转移酶 (DNA methyltransferases, DNMTs) 的作用下将甲基添加到 DNA 分子中的碱基上,这是一种极为常见且重要的表观遗传现象。我们常见的 DNA 甲基化通常发生在甲基位上,同时也可以与 DNA 链上的胞嘧啶 (C) 第 5 位碳原子间的共价结合形成 5-甲基胞嘧啶 (5mC) (图 1-1-1)。除此以外,在腺嘌呤 (A) N6 位置可以引入一个甲基形成 N6-甲基腺嘌呤 (N6-mA),甲基化也可发生在此位点。

研究证实, DNA 甲基化在真核生物中主要发生在 CpG 二联核苷酸上,也常在植物基因组 CNN 和 CNG 序列上发现,在生物体基因序列的某些区域, CpG 序列达到很高的密度,可达均值的 5 倍以上,成为含有 C 和 G 的富集区,我们称之为 DNA 序列海洋中的 CpG 岛 (CpG islands)。DNA 甲基化作为表观遗传学最重要的调控现象之一,对基因的表达发挥着重要的调控功能,涉及转座子扩散防御、异染色质形成、源基因表达调控、基因印迹、细胞分化和转基因沉默等,对生物的生命活动、新陈代谢有着重要的作用。

2. 组蛋白共价修饰 (histone modification) 方式包括:乙酰化 (acetylation)、甲基化 (methylation)、核糖化 (ribosylation)、磷酸化 (phosphorylation)、类泛素化 (sumoylation)

和泛素化 (ubiquitylation) 等 (图 1-1-2)。这些修饰共同构成了可以通过特殊解码蛋白解读的组蛋白密码,并在解码过程中发挥作用,从而影响基因表达。在这么多的组蛋白的修饰方式中乙酰化和甲基化主要发生在组蛋白赖氨酸 (-lys) 残基上,研究发现乙酰化与转录激活有相关性;随之发现的组蛋白的甲基化既可能与转录抑制相关,也可能与转录激活相关,抑制或激活取决于所添加的甲基基团数目以及存在的甲基化 lys 位点,其中每个 lys 残基可以形成单甲基化、二甲基化、三甲基化三种形式 (加上 1~3 个甲基基团) 中的任意一种,每种形式都发挥其固有的、独特的功能。

组蛋白的磷酸化是另外一种极其重要的调控方式,在 DNA 修复、基因转录、染色质凝聚及细胞凋亡等过程中发挥着重要作用。核糖化是指组蛋白 H1、H2A、H2B 及 H3 和多聚 ADP-核糖的共价结合,研究者发现 ADP-核糖基化是启动真核细胞内复制过程的扳机;而泛素化是指将被降解的蛋白质连接上泛素标记,事实证明一些可以诱导基因启动子区域的 H2B 组蛋白会被修饰以启动基因表达;类泛素蛋白 (small ubiquitin-like modifiers, SUMO) 对蛋白质的翻译后修饰与蛋白质的细胞内定位、转录活性和稳定性均有关,相关研究表明,其也可能参与异染色质结构的调节。近些年来,人们逐渐认识到表观遗传在基因表达调控方面的重要作用,并研究出一系列检测表观遗传修饰的方法,尤其是针对 DNA 甲基化和组蛋白修饰检测方法这方面取得了较大进展,到目前为止,检测方法的特异性和灵敏度都在不断得到提高;同时,表观修饰的检测正在逐步从个别位点向高通量检测的方向发展,逐步从定性检测向定量分析转向发展。

3. 染色质重塑 是指染色质结构和位置发生了变化,其主要涉及核小体在能量驱动下的重新排列或置换,结果便是它改变了基因启动子区内核小体的排列方式,增加了启动子和基因转录装置的可接近结合性。

真核生物遗传学过程的物质基础均是染色质,染色质构型整体和局部的动态改变是基因功能调控的相关因素。研究表明,组蛋白 N 端尾巴修饰和染色质重塑的发生密切相关,尤其是对组蛋白 H3 和 H4 的修饰调节。组蛋白的修饰可以直接影响核小体的结构,同时也为其他蛋白提供了 DNA 作用的结合位点。目前,染色质重塑主要包括两种类型:一类是依赖 ATP 的物理修饰,即利用 ATP 水解后释放的能量将 DNA 和组蛋白的结合分开,使转录能够顺利进行下去;另一类是含有组蛋白脱乙酰酶和乙酰转移酶的化学修饰类型,即依靠水解 ATP 来提供所需的能量,使得染色质重塑复合物可完成染色质结构的改变,根据水解 ATP 的亚基不同,可将复合物分为 ISW 复合物、SWI/SNF 复合物及其他类型的复合物。研究证实,这些复合物及相关的蛋白均与转录的抑制和激活、细胞周期、DNA 修复及 DNA 的甲基化呈一定的相关性。

4. 基因印记 又称遗传印记,是指基因的表达与否取决于它们是在母源染色体上,还是在父源染色体上以及母源染色体或父源染色体上的基因是否发生沉默,继而是如何表达出来。基因印记是通过生化途径,在某个基因或基因组域上标记其双亲来源信息的生物学过程。研究表明,有些印记基因只从父源染色体上表达,相应地有些则只从母源染色体上表达。基因组印记可以是非共价标记 (如核基因组定位, DNA-蛋白质相互作用及 DNA-RNA 相互作用),也可以是共价标记 (DNA 甲基化) 上的,印记的方法包括维持整个细胞周期中双亲表观记号的、

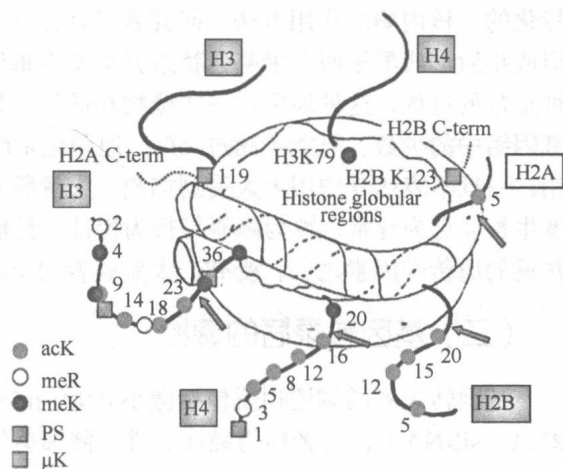


图 1-1-2 组蛋白修饰方式

特化的、核内酶的作用方法。研究者发现在胚胎中，来自不同亲本的 IGF2 甲基化状态不同，但成年后形成配子时，甲基化状态就会发生重设，形成新的状态。研究发现，基因组印记是一种正常的过程，这种现象在一些植物和低等动物中已被发现多年。通常，印记的基因只占人类基因组中的少数，可能不超过 5%，但印记基因在胎儿的行为和生长发育中起着至关重要的作用，一旦不表达其作用大大受到抑制。大量研究发现，基因组印记病主要表现为生长迟缓、过度生长、行为异常、智力障碍等行为特征。目前，在肿瘤的研究中认为印记缺失是引起肿瘤最常见的遗传学因素之一，受到广大实验者的重视。

(二) 基因转录后的调控

基因转录后的调控物质包括微小 RNA(microRNA, miRNA)和小干扰 RNA(small interfering RNA, siRNA)，两者均为翻译后进行修饰剪切形成的产物。

miRNA 和 siRNA 均为非编码的 RNA。它们是两种序列特异性转录后基因表达的调节因子。miRNA 的长短与 siRNA 相似，均为 22bp 左右，多数 miRNA 具有高度的保守性、基因表达组织特异性和时序性三大特点。miRNA 的研究已成为当前生物医学研究的前沿热点之一，尤其是 miRNA 与肿瘤的关联日渐成为人们关注的重点方向，针对 miRNA 与肿瘤的治疗产生了特异性的靶点。如果某一 miRNA 特异性表达于某一肿瘤，那么该 miRNA 就可用于肿瘤的早期预测，也可特异性导入某一 miRNA 或特异性敲除某一 miRNA 来用于肿瘤的治疗，这些目前均已开始应用到临床，如图 1-1-3 所示为 miRNA 的研究思路。除此之外，DNA 甲基化、组蛋白乙酰化等方面也与 miRNA 之间存在相互影响，因而这些构成了细胞基因调控系统最为复杂的调控网络。自从科学家沃森和克里克对 DNA 结构进行了完美的阐释后，从某种疾病发病机制的研究到疾病治疗方案或方法的确定，使生命科学从此踏入了基因时代，引来了时代的新潮。随之人类基因组计划(human genome project, HGP)成果的公布，把全世界几十年的目光都聚焦在以中心法则为核心的染色体基因组学上。直到表观遗传的提出，这无疑是为基因组学的研究开拓了一个新的领域。目前，表观遗传学已成为最新、最炙手可热的研究热点。研究人员发现，表观遗传学可以从更微观的视角、从更深的层次来解释许多目前基因组学无法解释的难题，攻克更多的问题，如同卵双生的兄弟日后为何会产生如此大的差异，环境因素究竟如何影响性状等方面。尽管迄今为止对表观遗传的研究仍然十分短暂，但人类已经取得了一定的成就。



图 1-1-3 miRNA 的研究思路

五、表观遗传学研究的应用进展

先前的研究发现多种疾病与表观遗传调控及表观特征的变化息息相关,这些疾病的遗传特点不能够用精确的遗传方式来解释,同时随着年龄的增长,相关疾病会出现由于表观遗传变化的持续传递,随之出现长期积累而患病率增高的现象,这些会严重影响人们的生活质量及水平。由于表观遗传具有可逆性这一重要特点,因此从表观遗传学角度研究疾病的预防、诊断和治疗等受到人们的广泛关注,成为新的研究热点。常见的与表观遗传相关的疾病包括心血管疾病、癌症、纤维化、自身免疫性疾病等。

(一) 表观遗传与肿瘤

研究发现,肿瘤的发生与 DNA 损伤修复缺陷、原癌基因的激活、抑癌基因的沉默等机制密切相关。近年来,大量的 DNA 甲基化研究证据发现不同类型的肿瘤癌细胞存在不同程度的异常 DNA 甲基化行为,包括某些抑癌基因及修复基因等的高甲基化和整个基因组的低甲基化等共同调控着癌症的发生发展过程。肿瘤与基因组 DNA 低甲基化的因果关系不明显,目前对于基因组 DNA 低甲基化是肿瘤发展中的促进和维持因素,还是肿瘤发生的起始事件,这些仍存在争议。研究者 Schulz WA 等发现基因组 DNA 低甲基化多发生在转移性前列腺癌,且与染色体的不稳定性相关。目前,我们已知癌症发生发展的重要因素之一是癌症相关基因启动子区及其附近 CpG 岛的异常高甲基化,目前已经在肿瘤中发现了众多由于高甲基化导致表观沉默的基因,也就是高甲基化低表达的现象,包括细胞周期调控相关基因 (*cyclin D1*、*cyclin D2*、*Rb*、*p16*、*p53*、*p73* 等)、信号转导相关基因 (*RASSF1*、*LKB1 / STK11* 等)、凋亡相关基因 (*DAPK*、*CASP8* 等)、DNA 修复基因 (*MGMT*、*hMLH1*、*hMLH2*、*BRCA1* 等)等。科学家针对人类 12 种肿瘤的 70 个肿瘤细胞系中,15 个基因的启动子甲基化状态进行系统分析后,Paz 等发现每种肿瘤至少有 1 个基因的启动子区发生高甲基化,同时这些基因启动子甲基化具有肿瘤类型的特异性。例如,结直肠癌细胞系中经常出现金属蛋白酶组织抑制剂 *TIMP-3* 以及错配修复基因 *hMLH1* 的高甲基化,而 *CDH1* 基因的启动子区高甲基化,因此肿瘤的发生发展与基因和组蛋白的异常修饰呈现密切相关的关系。目前,组蛋白修饰的研究主要集中于组蛋白 H3、H4 的甲基化和乙酰化修饰的研究。研究表明,肿瘤细胞的组蛋白大多呈现高甲基化和低乙酰化状态,能够活化组蛋白的修饰行为从而启动基因的表达,引起其表达的改变;反之,抑制组蛋白活性的修饰行为可引起相关基因的沉默。随之在胃癌的研究中,研究人员 Lee 等发现,组蛋白 H3 的低乙酰化以及组蛋白 H3 第 9 位赖氨酸残基高甲基化均可导致抑癌基因 *Runx* 相关的转录因子 3 位基因的表达水平的抑制。随之在对肺癌的研究中,林刚等发现抑癌基因的表达可在组蛋白 H4 第 3 位精氨酸残基甲基化水平增高后受到抑制,进而导致肺癌发生;研究还发现组蛋白甲基化水平的变化同时受到蛋白质精氨酸甲基转移酶 (*PRMT5*) 的调控。因此,我们可以推测肿瘤的发生发展也可能与组蛋白修饰相关酶的功能的紊乱息息相关,而特定的启动子区域与组蛋白去乙酰化酶发生异常结合而抑制功能基因的转录可能也是肿瘤发生的机制之一,早期在急性早幼粒细胞白血病的研究中,Peterson CL 等发现染色体易位形成维甲酸受体 α 融合蛋白,与含有组蛋白去乙酰化酶的辅抑制复合物相互作用,造成维甲酸受体 α 的靶基因的转录抑制,导致粒细胞成熟障碍引起白血病。

研究肿瘤形成中出现的表观遗传的改变,对肿瘤的发生机制、早期诊断及预防治疗大有裨益,发现 DNA 的高甲基化是肿瘤发生的早期事件,可以预测肿瘤的发生,因而高灵敏度的甲

基化状态检测技术对肿瘤的早期诊断有意义。Widschwendter 等在研究结肠腺瘤性息肉病基因启动子甲基化现象时发现, DNA 甲基化对恶性肿瘤的预后和转移监测具有一定的指导意义和实践意义。以上所阐述的表观遗传变化的可逆性特征, 为肿瘤的治疗提供了一个新的思路: 疾病状态下异常的表观遗传变化可通过新的分子靶向药物被逆转, 从而达到所需治疗的目的。目前, 肿瘤表观遗传治疗的主要研究方向为针对组蛋白去乙酰化抑制剂和 DNA 甲基化抑制剂, 现已有多个去甲基化药物进入了临床试验并取得了初步疗效, 特别是对血液系统恶性肿瘤的治疗, 这些为我们诊断治疗提供了新的思路。即便如此, 将甲基化应用于肿瘤的早期诊断在充分认识甲基化发生的规律之前仍然尚早, 将组蛋白乙酰化抑制剂及去甲基药物安全用于临床仍有一定的距离, 对药物引起的不良反应还需要长期细致的观察, 这些均有待于进一步的研究与发现。

(二) 表观遗传与心血管疾病

冠心病、心绞痛等心血管疾病的始动病因是动脉粥样硬化 (atherosclerosis, AS) 的发生, 受环境因素和遗传因素的共同影响。近年来研究发现, AS 的发生与 DNA 的异常甲基化密切相关。20 世纪 90 年代末, Newman 提出了 DNA 甲基化促进 AS 的假说, 即维生素 B₁₂ 和叶酸缺乏可以促使同型半胱氨酸 (Hcy) 聚集, 同时引起 S-腺苷甲硫氨酸 (SAM) 水平下降, 导致 DNA 甲基化水平降低, 进而 DNA 低甲基化可促使纤维沉积、血管平滑肌细胞增殖、使外周血参加炎性和免疫反应的细胞过度增殖, 从而促进了 AS 的发生发展。在 AS 发生发展过程中出现整体基因组低甲基化的同时, 某些特定基因启动子区 CpG 岛也存在异常高甲基化现象。目前, 已报道的特定基因异常甲基化的研究包括抑癌基因 *p53*、细胞外超氧化物歧化酶 (EC-SOD)、雌激素受体 (ER) 等少数几个研究对象。雌激素受体 A (ERA) 是参与细胞分化调控的抗生长基因, Leader JE 等研究发现, 与健康对照组相比重度 AS 患者 ERA 基因甲基化的水平明显升高, ERA 高度甲基化后导致基因沉默, 降低了对血管平滑肌细胞异常生长的抑制能力, 从而发生 ERA 蛋白表达降低, 保护性功能相应降低, 从而导致动脉硬化的形成。Mercer J 等研究发现 *p53* 基因高甲基化与 AS 密切相关, 故 *p53* 基因也是 AS 的一个保护因素, 值得深入研究。动物研究也证实, 在患有 AS 的家兔中, EC-SOD 出现低甲基化, 影响其转录表达, 提示其与 AS 的发生有关。

(三) 表观遗传与自身免疫性疾病

自身免疫疾病是涉及全身免疫缺陷的疾病, 该疾病的形成涉及自身反应性淋巴细胞的激活, 进而开启各种基因表达的表观遗传的改变。目前, 研究较多的自身免疫疾病是系统性红斑狼疮 (systemic lupus erythematosus, SLE), 组蛋白修饰与 DNA 甲基化在 SLE 的发病中发挥着重要作用, 也是我们目前研究的热点。研究表明, DNA 低甲基化的成熟 T 细胞在体外具有自身反应性, 可以复制出红斑狼疮样模型, 供于实验的研究。最初, 红斑狼疮 DNA 甲基化研究的证据来源于药物性红斑狼疮的研究, 研究发现服用胍屈嗪和普鲁卡因胺的患者均出现 DNA 低甲基化现象。在对系统性红斑狼疮动物模型和患者的研究中, Pan 等发现 T 淋巴细胞 DNA 的甲基化状态在系统性红斑狼疮发病中发挥明显的作用。Lu 等用 DNA 甲基化转移酶抑制剂 5-氮胞苷处理正常 T 细胞, 基因芯片结果发现 100 多种 T 细胞基因的转录增加到 2 倍以上。其中包括 Perforin (细胞毒分子), IgE γ C γ 1 (TCR ζ 链替代物), CD11a (黏附分子) 及 CD70 (B 细胞共刺激分子) 等。通过建立 CD4⁺ T 细胞低甲基化模型, 李尊忠等发现 T 淋巴细胞一些共刺激因子和黏附因子的表达, 使 T 细胞具有自身反应性。一方面, 通过大量杀伤巨噬细胞, 增加血液循环中的自身抗原; 另一方面, 促使 B 细胞活化并分泌免疫球蛋白。另外, 单核/巨

噬细胞的杀伤及人类基因组中 HERV 序列的活化使自身抗原增加,共同参与系统性红斑狼疮病程中的免疫紊乱,导致多器官、多组织的损伤。另外,微小 RNA 在系统性红斑狼疮中也有研究。DaiY 等通过 miRNA 芯片筛选出系统性红斑狼疮患者外周血单个核细胞中差异表达的 16 个 miRNA,发现系统性红斑狼疮疾病活动性与部分 miRNA 显著相关,为今后的研究提供了又一方面。目前,表观遗传学在系统性红斑狼疮中的研究多局限在机制研究方面,随着 DNA 甲基化等表观遗传研究的深入,将有望更深入地了解病因而开创新的系统性红斑狼疮治疗方法。

(四) 表观遗传与纤维化疾病

目前,纤维化疾病与表观遗传的研究也越来越受到关注,但目前多局限于动物和细胞的实验水平,研究发现纤维化与表观遗传变化的发生密切相关。多种基因的表达在肝脏纤维化的过程中都会发生改变,一方面,抑制纤维化的基因表达减弱;另一方面,促进肝纤维化的基因表达增强。肝星状细胞活化增殖是纤维化的中心环节,其凋亡抑制或活化增殖均能促进纤维化的发生,研究发现基因的甲基化与肝星状细胞的活性数量密切相关。肺纤维化表观遗传学方面的研究也有开展,肺纤维化形成过程中涉及细胞外基质大量沉积,该过程中主要有肌成纤维细胞和成纤维细胞参与。研究发现,肺纤维化组织的 *p16* 基因启动子出现甲基化,高甲基化使 *p16* 基因表达水平下降,使肺成纤维化细胞的细胞周期调控紊乱,导致成纤维细胞的过度增殖,引起肺纤维化。通过高通量筛选差异表达的基因研究人员吕长俊等寻找与纤维化相关的基因,为进一步研究这些基因的表观遗传状态提供了依据,同时也为纤维化的诊疗指明新的方向。这些研究与纤维化相关的疾病(如特发性纤维化、肝癌、肺尘埃沉着病等)在预防、诊治、监测等方面将会有新的突破,更值得我们去深入研究。

此外,表观遗传调控在其他领域也进行了较多研究,如代谢综合征、糖尿病等。总而言之,表观遗传学已成为生命科学研究的新焦点,它为人类疾病指明了新的研究方向,弥补了经典遗传学的不足,正因为如此,表观遗传学相关技术也颇受关注,近年来也成为人类研究的重点和热点。

第二章 DNA 甲基化分析方法

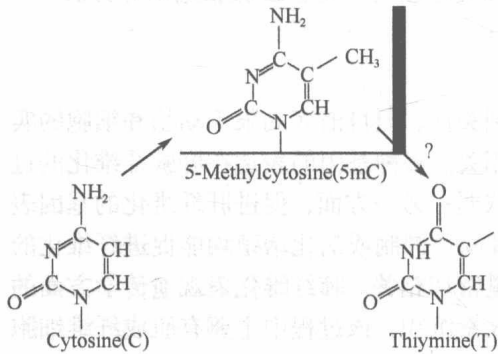


图 1-2-1 甲基化过程

目前, 接近全基因组或全基因组甲基化检测技术较多, 主要分为三类: 第一类技术以限制性内切酶为基础, 用一个或多个酶限制性切割甲基化 DNA (McrBC) 或未甲基化 DNA (如 *Not I* 和 *Hpa II*)。这个方法结合毛细管测序、芯片等技术已经检测了多种生物的全基因组甲基化水平, 但这些仅限于内切酶能够识别的 CpG 位点。第二类技术依赖于全基因组 DNA 重亚硫酸盐转换, 经重亚硫酸盐处理后基因组 DNA 未甲基化胞嘧啶 (C) 转换为尿嘧啶 (U) (经扩增后最终为 T), 甲基化 C 保持不变(图 1-2-1)。

此方法可以进行单 CpG 位点解析并且可以结合芯片

或高通量测序。然而, 此方法也有一定缺陷: 序列在经过重亚硫酸盐转换后, 其特异性便会降低, 在芯片上很难设计足够的特异性探针进行全基因组分析。此外, 如果此方法用于较大基因组日常分析, 其价格将十分昂贵。第三类技术是以免疫学为基础, 用含有甲基结合结构域的蛋白或者用 5-甲基胞嘧啶特异性抗体通过免疫沉淀富集基因组未甲基化或甲基化片段, 称为甲基化 DNA 免疫共沉淀 (methylated DNA immunoprecipitation, MeDIP 或 mDIP), 目前发现 MeDIP 结合芯片 (MeDIP-chip) 技术可以对任何物种进行高通量全基因组 DNA 甲基化作图。以上三类技术可以与芯片或测序技术组合形成多种方法: 如酶切结合芯片技术、酶切结合测序、亚硫酸氢钠结合芯片技术、亚硫酸氢钠结合测序、免疫沉淀结合芯片技术和免疫沉淀结合测序等。

一、以限制性内切酶为基础的 DNA 甲基化检测技术

(一) 甲基化敏感扩增多态性

甲基化敏感扩增多态性 (methylation sensitive amplification polymorphism, MSAP) 是研究人员在扩增片段长度多态性 (amplified fragment length polymorphism, AFLP) 技术的基础上建立起了本方法。其基本流程是: 提取基因组 DNA, 根据 *Msp I*、*Hpa II* 两种酶对甲基化敏感程度不同, 分别用 *EcoR I* / *Msp I*、*EcoR I* / *Hpa II* 两组酶组合对基因组 DNA 进行双酶切, 连上接头后, 用按照接头序列设计的引物加上一个选择性碱基进行 PCR 扩增, 称为预扩增。将预扩增产物稀释后, 再加入另外两个选择性碱基引物进行第二次 PCR 扩增, 此过程称为选择性扩增。扩增产物变性后在 6% 的 PAGE 胶上电泳, 最后采用同位素放射自显影或银染方法处理 PAGE 胶, 统计并分析 DNA 条带即可得出基因组 CpG 位点甲基化状态。还可以对差异扩增片段回收克隆测序, 通过比对可鉴定甲基化差异位点。相对其他 DNA 甲基化测定技术而言, MSAP 技术有如下优点: 可在全基因组范围检测 CCGG 位点的胞嘧啶甲基化变化; 操作相对简便, 在 AFLP 基础上可直接操作; 不需知道被测 DNA 的序列信息。但 MSAP 技术仍存在其自身的局限性: 只能检测 CCGG 位点的胞嘧啶甲基化。目前, 此方法常用于研究

动植物全基因组甲基化状态。

(二) McrBC 酶切法

McrBC 是一个 DNA 甲基化特异性限制性内切酶，它能够识别 (A/G)C 基序中甲基化的 C，切割一条或两条链上含有甲基化胞嘧啶的 DNA。研究发现 McrBC 不作用于非甲基化的 DNA。McrBC 的 DNA 识别位点包含两个 (G/A)mC 形式的半位点。最适宜切割的半位点的间距是 55~103bp，而该间距最长可达 3kb。McrBC 切割时需要鸟苷三磷酸 (GTP) 参与，该酶在有 GTP 的非水解类似物存在时可以特异性地结合到甲基化的 DNA 上，而不产生切割反应。McrBC 在一对 P_{um}CG 序列元件上进行切割，因此可以检测相当大一部分的甲基化 CpG。在现在的研究中，该方法通常结合芯片技术来检测甲基化 DNA，其基本程序是：首先用 McrBC 酶对全基因组 DNA 进行酶切，然后将芯片与酶切片段杂交，对数据统计分析后检测甲基化 DNA。该方法的局限性在于：McrBC 不能识别内部胞嘧啶甲基化了的 *Hpa* II/*Msp* I 位点 (CCGG)，而且该酶在每一对半位点之间便切割一次，而切割位点靠近这一对半位点的一端，但通常距最近的甲基化碱基大约 30bp。当 DNA 片段含有多个 McrBC 识别半位点时 (如含有甲基化胞嘧啶的基因组 DNA)，因该酶的切割位点具有不确定性，极易导致切割位点重叠。

二、重亚硫酸盐法

(一) 重亚硫酸盐修饰法

重亚硫酸盐修饰法的原理是在 HSO_3^- 存在的条件下，单链 DNA 能有效地将未甲基化 C 转变成 U，两轮 PCR 扩增后，m5C 则保持不变，仍然为 C，而该位点则变为 T。通过扩增克隆测序比对，可以鉴别待测位点是否发生了甲基化改变。该技术很容易确定个体基因组 DNA 分子中各个单链的甲基化位置及其甲基化状态，如图 1-2-2 所示。

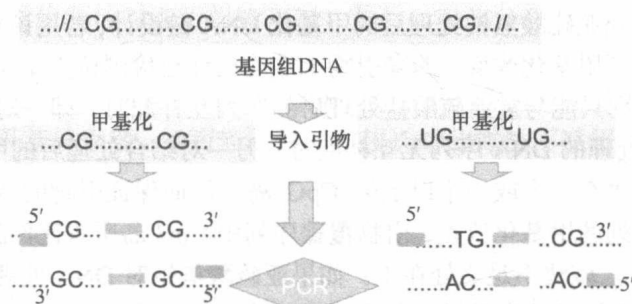


图 1-2-2 重亚硫酸盐修饰法

在此基础上，该法可以结合限制性内切酶分析、DNA 直接测序、甲基化敏感的单核苷酸引物延伸法、甲基化特异性 PCR 扩增、锁式探针技术及高效液相色谱法等方法测出相应序列的甲基化情况。

(二) 重亚硫酸盐结合 DNA 直接测序法

重亚硫酸盐结合 DNA 直接测序法由研究者 Formmer 等提出，其基本过程是：提取高质量基因组 DNA，平均分为两组，一组作为对照；另一组经重亚硫酸盐处理后发生转换。

用引物进行扩增后,对分别获得的两组扩增片段进行测序对比,如果处理后一组的 C 变为 T,则这些转变的位点为没有发生甲基化的位点;若处理后仍然为 C,则这些位点发生了甲基化改变。该方法适用于特定基因的甲基化检测,如测定某个基因在某种状态下的甲基化状态。

(三) 联合重亚硫酸盐限制性内切酶分析法

联合重亚硫酸盐限制性内切酶分析法 (combined bisulfite restriction analysis, COBRA) 由 Peter 和 Xiong 报道,其基本步骤是:首先用重亚硫酸盐对所需样本的 DNA 进行处理,用 PCR 扩增目的片段,随后用限制性内切酶消化,此酶识别序列中需包含有 CG 的序列,如 *Bst*UI (CGCG)。如果在待测序列中, C 未发生甲基化,则 PCR 后转变为 TGTG,那么 *Bst*UI 识别位点丢失,丧失了切割功能;若其识别序列中的 C 发生完全甲基化 (5mCG5mCG),则 PCR 扩增后保留为 CGCG, *Bst*UI 能够识别并进行切割位点。然后,酶切产物再经电泳分离、探针杂交、扫描定量后即可得出原样本中甲基化的比例。研究发现,用 Agilent 2100 Bio-analyzer 和 COBRA 对酶切产物进行直接分析,使 COBRA 的定量无放射性污染,而且更准确、快速。这种方法相对比较简便,不需要研究者预先得知样本序列及 CpG 位点;同时,所需要的样本量少,可适用于石蜡包埋样本的分析检测,可对 DNA 甲基化水平进行定量研究。但是本法存在局限性:只能获得特殊酶切位点的甲基化情况,因此检测结果若为阴性并不能完全排除样品 DNA 中甲基化存在的可能性。所以,此方法常用于样本量比较少, DNA 甲基化定量检测的试验中。

(四) 重亚硫酸盐结合甲基化特异性 PCR 法

研究者 Herman 等首先提出了甲基化特异性 PCR,这是一种检测基因组 DNA 甲基化水平的常用方法。该法是在用重亚硫酸盐处理样品 DNA 后,甲基化的胞嘧啶不变,而非甲基化的胞嘧啶则被尿嘧啶替代。在 PCR 反应时,设计两套不同的引物对:一对引物针对经亚硫酸氢钠处理后的非甲基化 DNA 链设计,若用该对引物能扩增出片段,说明该检测位点没有甲基化;另一对引物序列针对经亚硫酸氢钠处理后的甲基化 DNA 链设计,若用该对引物能扩增出片段,说明该检测位点发生了甲基化改变。两套引物末端均设计至检测位点结束:引物都具有很高的特异性;两套引物分别只能与重亚硫酸盐处理后的序列互补配对,即一对结合处理后的非甲基化 DNA 链,与未经处理的 DNA 序列无互补配对,另一对结合处理后的甲基化 DNA 链; MSP 引物覆盖序列中必须含有一个或一个以上的 CpG 岛,从而保证引物的特异性,经克隆测序就检测出引物所覆盖序列的甲基化位点,引物覆盖序列中 CpG 岛所占比例越高,甲基化 DNA 检出率越高 (图 1-2-3)。该法不足之处在于:如果亚硫酸氢钠对 DNA 处理不完全,则容易导致假阳性;引物的选择和设计非常关键,若选择或设计不当容易出现假阳性。可通过限制性内切酶酶解法检验 PCR 产物做进一步判断。

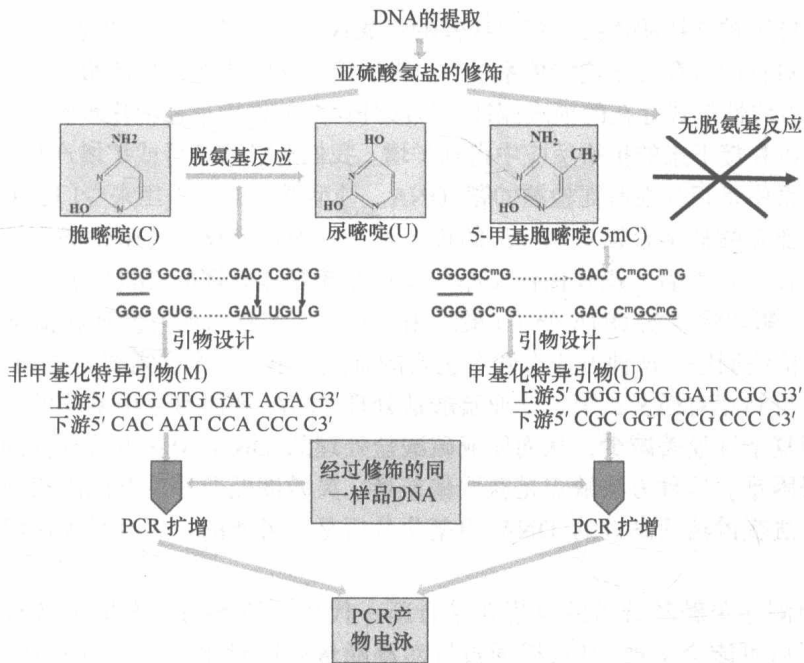


图 1-2-3 亚硫酸盐结合甲基化特异性 PCR 法

(五) 甲基化敏感的单核苷酸引物延伸法

甲基化敏感的单核苷酸引物延伸法 (methylation-sensitive single nucleotide primer extension, Ms-SnuPE) 由 Gonzalgo 和 Jones 提出, 可用于快速判断 DNA 片段中某些具体位点的甲基化状况。该法的基本原理是: 用重亚硫酸盐处理样本 DNA 后, 再用 PCR 扩增目的片段, 产物电泳分离后平均分为两份作为 Ms-SnuPE 模板。然后, 分别针对所检测的 CpG 位点设计上游引物, 使 3' 端引物紧邻 C, 扩增时加 ^{32}P 标记的 dTTP 或 dCTP 进行单核苷酸延伸反应。若是未被甲基化, 则标记的 dTTP 参与反应; 反之, 如果待测位点被甲基化, 则同位素标记的 dCTP 会在反应延伸时连于引物末端。最后, 电泳分离产物并成像。根据某一 CpG 位点的 C/T 信号强度比, 可定量其甲基化程度。此外, 也可以不电泳, 将 PCR 产物直接转移到尼龙膜上, 成像后即可得到所测的多个 CpG 位点的平均甲基化程度。Ms-SnuPE 可快速定量多个 CpG 位点的甲基化程度, 然而当待测序列较长时此方法不适用, 且此方法具有一定放射性。

(六) 高效液相色谱法 (high performance liquid chromatography, HPLC)

Oefner 等首先提出变性高效液相层析 (denaturing HPLC, DHPLC) 用于 DNA 和单核苷酸分析。随之邓大君等将 DHPLC 改进后与 PCR 联用, 从而建立了一种检测甲基化程度的分析方法, 其原理是用重亚硫酸盐处理 DNA, 然后对其进行差异性扩增, 由于原甲基化的胞嘧啶经重亚硫酸盐处理被保留, 因此在随后的 PCR 扩增时, 其变性温度也相应上升, 使最终的 PCR 产物在色谱柱中的保留时间明显延长, 从而判定出 PCR 产物的 DNA 序列甲基化状况。

(七) 亚硫酸盐结合锁式探针法

Nilsson 等首次报道了锁式探针 (padlock probe) 及其应用, 其原理是在没有相应靶 DNA

存在时锁式探针不被连接酶连接, 仅以线性的形式存在, 只有当相应的检测靶 DNA 和锁式探针完全互补配对且同时存在于连接体系中时, 线型锁式探针才能在连接酶的作用下被有效地连接成环型。接着线性形式存在的锁式探针在核酸外切酶的作用下被消化水解, 而连接成环状的锁式探针就可以在接下来的扩增反应中得到扩增。我们对扩增信号或扩增产物进行分析, 就可以判断连接体系中是否存在所需检测的靶 DNA。锁型探针以前是用来专门针对获取外显子并进行测序的。亚硫酸盐结合锁式探针法的程序是: 首先合成长约 150nt 的寡核苷酸链, 酶切裂解转换锁式探针; 然后, 锁式探针文库与经重亚硫酸盐处理的 DNA 退火, 延伸, 再与 5' 端连接环化, 并酶切除线状 DNA 分子; 最后用一对常用引物对所有的环状锁型探针进行 PCR 扩增, 再用鸟枪法测序。这种方法有两个主要的难点: 第一, 本法需要合成的探针数量较多, 工作量庞大, 价格相对昂贵; 第二, 亚硫酸盐处理 DNA 使所有未甲基化的胞嘧啶转换成了尿嘧啶, 序列复杂性显著减少, 从而使亚硫酸盐处理的 DNA 中获得特异性的序列相对于天然 DNA 显得困难。这种方法缺点是获取偏差高、灵敏度低及等位基因随机缺失。因此, 用目前存在的方法获得精确高效的 DNA 甲基化分析是不可能的, 尤其是无法解决等位基因的缺失问题。

全基因组测序单碱基分析的设想在结合新一代测序技术的高通量单碱基解析新一代测序技术的出现后可能会实现。其可以通过针对高通量测序对比检测基因组 DNA 甲基化的位点及程度。此方法的优势在于能够对全基因组甲基化进行精确的分析。最近, 研究人员利用重亚硫酸盐结合的新一代测序技术顺利完成了拟南芥约 120Mb 基因组 DNA 甲基化图谱分析。

三、甲基化 DNA 免疫共沉淀

甲基化 DNA 免疫共沉淀 (MeDIP) 方法是用含有甲基结合结构域蛋白或者用 5-甲基胞嘧啶特异性抗体通过免疫沉淀方法来富集基因组未甲基化或甲基化片段, 其基本程序是: ①提取全基因组 DNA, 经超声波打断成长度为 400~500bp 的片段; ②加热变性, 将变性的单链 DNA 样品平均分为两份; ③其中一份作为 Total input DNA 样品, 另一份加入甲基化 DNA 特异性抗体; ④使用亲和层析分离上一步样品中的 DNA 甲基化片段的抗体复合物, 则样品中其余的非甲基化 DNA 片段被洗脱, 纯化得到甲基化 DNA 片段 (MeD-IPDNA)。最后得到纯化的甲基化 DNA 片段后, 可以结合已有的芯片技术检测基因甲基化, 也可结合荧光定量 PCR 技术定量检测 DNA 甲基化。

(一) MeDIP 结合实时荧光定量 PCR 技术 (MeDIP-qPCR)

之前的实时荧光定量 PCR 技术是在 PCR 反应体系中加入荧光试剂, 利用荧光信号实时检测 PCR 的进程, 最后通过标准曲线对未知模板进行定量分析的方法。实时荧光定量 PCR 能够灵敏、快速、专一、高重复地精确定量起始模板的浓度。MeDIP 技术与实时定量 PCR 技术结合, 先通过 MeDIP 技术富集甲基化的 DNA 片段, 再利用实时荧光 PCR 技术进行检测, 经过数据分析处理得到检测的生物样品中特定定位点的甲基化水平的精确数值 (图 1-2-4)。此方法具有以下优点: 不需要标记探针; 高特异性; 灵敏度高; 精确度高, 重复性好等优势。