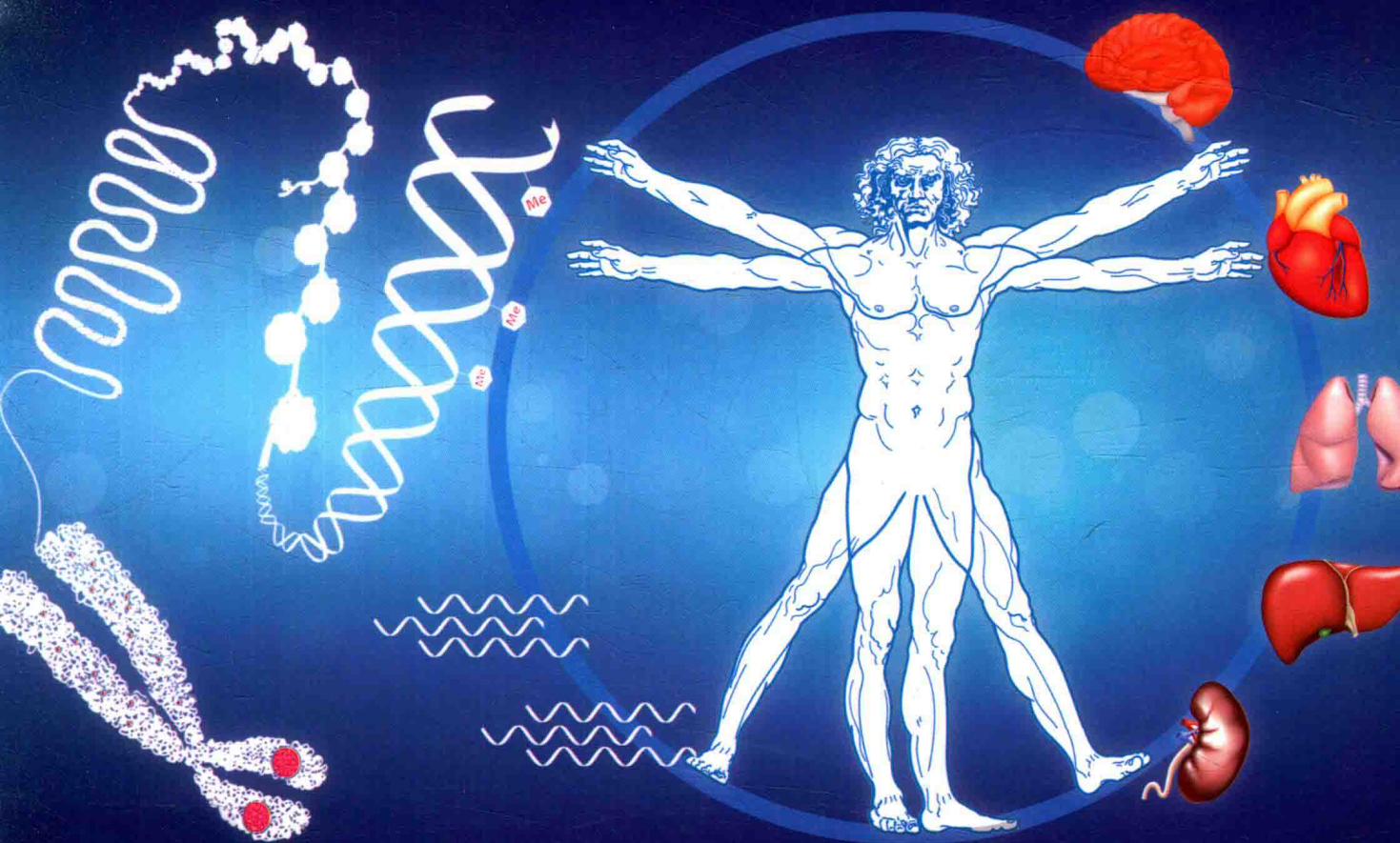


Epigenetics and Complex Diseases

表观遗传学与 复杂性疾病

主 编 陆前进 于文强 吕 红



北京大学医学出版社

表观遗传学与复杂性疾病

Epigenetics and Complex Diseases

主 编 陆前进 于文强 吕 红

北京大学医学出版社

BIAOGUAN YICHUANXUE YU FUZAXING JIBING

图书在版编目 (CIP) 数据

表观遗传学与复杂性疾病 / 陆前进, 于文强, 吕红主编.
—北京: 北京大学医学出版社, 2016.7
ISBN 978-7-5659-1133-0

I. ①表… II. ①陆… ②于… ③吕… III. ①发育遗传学—研究
IV. ①Q344

中国版本图书馆 CIP 数据核字 (2015) 第 122367 号

表观遗传学与复杂性疾病

主 编: 陆前进 于文强 吕 红

出版发行: 北京大学医学出版社

地 址: (100191) 北京市海淀区学院路 38 号 北京大学医学部院内

电 话: 发行部 010-82802230; 图书邮购 010-82802495

网 址: <http://www.pumpress.com.cn>

E - mail: booksale@bjmu.edu.cn

印 刷: 北京大学印刷厂

经 销: 新华书店

策划编辑: 王智敏

责任编辑: 畅晓燕 责任校对: 金彤文 责任印制: 李 啸

开 本: 889 mm × 1194 mm 1/16 印张: 20.5 字数: 648 千字

版 次: 2016 年 7 月第 1 版 2016 年 7 月第 1 次印刷

书 号: ISBN 978-7-5659-1133-0

定 价: 98.00 元

版权所有, 违者必究

(凡属质量问题请与本社发行部联系退换)

作者简介



陆前进，教授，博士生导师，一级主任医师，首届湘雅名医，卫生部有突出贡献中青年专家，中南大学皮肤性病研究所所长，医学表观基因组学湖南省重点实验室主任，中华医学会皮肤性病学分会候任主任委员。长期从事表观遗传学与免疫调控及其在自身免疫病发病机制中的作用及临床应用研究。先后主持国家自然科学基金重大国际合作项目、国家自然科学基金重点项目、“973”计划课题等。以第一作者或通信作者在 *Lancet*、*JAMA*、*Blood*、*J Immunol* 等杂志发表论文 130 篇。作为第一完成人获得 2012 年湖南省科技进步一等奖及 2015 年湖南省自然科学一等奖，2014 年获国际皮肤科联盟杰出贡献奖。



于文强，复旦大学教授，博士生导师，医学博士。瑞典乌普萨拉大学、美国约翰·霍普金斯大学医学院博士后，美国哥伦比亚大学医学院 Associate Research Scientist。目前为复旦大学生物医学研究院高级 PI (principle investigator)，“长江学者”特聘教授，973 项目首席科学家。长期从事器官发育表观遗传学机制和肿瘤抑制基因沉默的表观遗传学机制研究，承担 973、国家自然科学基金、上海市重点基础研究及浦江人才计划等项目。成果先后发表在 *Nature*、*Nature Genetics* 等顶尖杂志上，被 *Cell* 和 *Nature genetics* 等杂志进行重点评论和推介，对其成果给予了很高评价。



吕红，教授，博士生导师，复旦大学生命科学学院，遗传工程国家重点实验室 PI，上海市工业菌株工程技术研究中心主任，中国遗传学会微生物遗传专业委员会委员。长期从事酵母分子遗传学与表观遗传学在基因表达调控中的机制及应用研究。先后承担了国家 863 项目(项目首席)、973 课题(主持)、国家自然科学基金重大研究计划、教育部博士点优先发展领域等科研项目，作为通信作者在 *Journal of Cell Science*、*Cancer Research*、*BMC Genomics*、*Journal of Biological Chemistry* 等杂志上发表论文，授权发明专利 10 项，作为第一完成人获得上海市技术发明二等奖、教育部技术发明二等奖等奖项。

编者名单

主编 陆前进 于文强 吕 红

编者 (按姓名汉语拼音排序)

中南大学湘雅二医院

丁 澎 侯 粲 李亚萍 廉晓日 梁功平 梁云生
廖洁月 刘素芳 龙 海 陆前进 罗双艳 罗湘杭
罗鸯鸯 邱湘宁 苏玉文 谭怡忻 王瑶瑶 肖 嵘
尹 恒 张 鹏 张 庆 张广森 张轶群 赵 明

复旦大学生物医学研究院

陈 坤 董世华 李 晋 李 信 李 岩 彭丽娜
王 翔 肖 敏 肖 瑶 于文强 张 兰 张如奎
赵丽萍 赵欣之 邹清平

复旦大学生命科学院

吕 红 余 垚

中南大学药学院

余聂芳

序言（一）

表观遗传学是研究在 DNA 序列没有改变的情况下，基因功能可逆的、可遗传改变的一门生物学学科。1939 年，生物学家 C. H. Waddington 首先在《现代遗传学导论》(*An Introduction to Modern Genetics*) 中提出 epigenetics 这一术语。1942 年，他在研究细胞发育命运时，再次提出表观遗传学这一概念，并把表观遗传学描述为一个控制从基因型到表现型的机制。随后表观遗传学在研究与经典孟德尔遗传法则不相符的生命现象过程中逐渐发展起来，并成为后基因组时代的重要研究领域。

近十余年来，随着分子生物学技术的迅猛发展，表观遗传学与人类疾病的关系逐步被揭示。大量研究表明，表观遗传修饰异常在人类复杂性疾病如恶性肿瘤、自身免疫性疾病的发病机制中起重要作用。因为表观遗传学基因功能改变的可逆性，从表观遗传学修饰角度研究复杂性疾病的预防、诊断和治疗成为生物医学领域关注的重点。目前，国内虽有介绍表观遗传学的相关著作，但尚无集中阐述与人类复杂性疾病相关的表观遗传机制的专著。因此，我们迫切地需要一本全面、系统阐述表观遗传修饰与复杂性疾病相关研究的参考书和工具书，从而推动我国表观遗传的基础与临床研究，加速研发人类复

杂性疾病的表观遗传诊断方法和治疗药物。

中南大学湘雅二医院陆前进教授、复旦大学生物医学研究院于文强教授及复旦大学遗传工程国家重点实验室吕红教授多年来致力于表观遗传学的基础与应用研究，经过多年的研究与临床实践，取得了一系列原创性研究成果，在国际上具有较高影响力。他们对迄今为止国内外相关研究成果进行系统汇编与梳理，融合自身宝贵经验，联手共同主编出版了《表观遗传学与复杂性疾病》一书，在此表示祝贺！

本书从表观遗传学基础理论着手，结合其相关研究进展进行了系统的论述，从恶性肿瘤、自身免疫性疾病、代谢性疾病、心血管系统疾病及神经精神疾病方面重点阐述了表观遗传修饰在人类复杂性疾病发生、发展过程中的作用及其分子机制，并详细介绍了表观遗传学研究方法、表观遗传修饰的干预药物及其研究进展。全书内容专业权威，紧跟科学前沿，论述深入浅出，将表观遗传学基础研究与临床应用融会贯通。我相信此书的出版必定能极大地推动我国复杂性疾病的表观遗传机制及其临床应用研究进程，故而欣然为之作序，并极力推荐此书！



夏家辉

中国工程院院士
2015 年 12 月

序言（二）

解开生命奥秘、征服疾病是生命科学及医学领域的主要命题。经典遗传学理论难以阐明携带相同基因组的同卵双生子在疾病易感性等方面显示出的极大差异性，同时也解释不了基因转录的时空特异性。表观遗传学解释了很多经典遗传学不能解释的谜团。1939年，生物学家 Waddington 首次提出“表观遗传学”这个概念，为解开生命的奥秘打开了至关重要的一扇门。表观遗传学的研究确实也成为近十多年来最主要的热点。诺贝尔奖获得者、DNA 双螺旋解码者 Watson 说：“你可以继承 DNA 序列之外的一些东西，这正是现在遗传学中让我激动的地方。”表观遗传学是研究在基因的核苷酸序列不发生改变的情况下，基因的表达发生了可遗传的变化的一门学科。其主要研究内容包括 DNA 甲基化、组蛋白修饰、非编码 RNA 调控等。表观遗传的改变可以遗传给子代细胞，且这种改变具有可逆性。

复杂性疾病是在遗传背景的基础上，环境因素通过表观遗传调控机制使疾病相关基因的表达和功能发生可遗传的改变，最终导致其临床表型的出现。

因此，表观遗传在复杂性疾病的发生、发展中起十分重要的作用。利用表观遗传学理论和研究手段，有助于我们深入揭示复杂性疾病的发病机制，同时也为复杂性疾病的预防、新药的研发提供了乐观的前景。

从表观遗传学的提出到现在，人类对复杂性疾病的表观遗传机制有了较深入的认识，及时总结梳理相关知识并出版《表观遗传学与复杂性疾病》一书具有重要的意义。本书以基础研究为着眼点、临床应用为导向，分上、下两篇论述，上篇详细介绍了表观遗传学的基础理论，下篇则阐述了复杂性疾病如肿瘤、自身免疫性疾病、代谢性疾病、心血管系统疾病、神经精神疾病的表观遗传学研究进展，最后还特别介绍了复杂性疾病的表观遗传学研究方法与干预药物。编者团队近年来在表观遗传调控机制及其临床应用研究领域取得了许多原创性的研究成果，积累了丰富的经验。全书内容权威、丰富，紧跟科学前沿发展，论述深入浅出。深信本书的出版必将对复杂性疾病的表观遗传学基础与临床研究起到极大的推动和促进作用。



尚永丰

中国科学院院士

2015年12月

前言

20世纪50年代初，Watson和Crick共同建立了DNA分子双螺旋结构模型，成为生命科学领域具有划时代意义的突破性发现。自此，遗传学迅速兴起，成为现代医学研究领域中的一个重要分支，人类认识到基因缺陷和基因突变可以导致疾病的发生，例如红绿色盲、血友病、珠蛋白生成障碍性贫血（地中海贫血）等。然而，基因并不能决定一切生命过程，特别是对于糖尿病、心血管疾病、自身免疫性疾病、神经精神性疾病、恶性肿瘤等慢性复杂性疾病及衰老等病理生理改变，环境因素在这些生命过程中同样具有重要影响，这在大量同卵双生子的研究中得到证实。事实上，基因的表达活性在不同组织细胞和不同生理、病理状态下存在可变性，而环境的影响可能改变基因表达的调控；有研究表明，这种影响甚至在出生以前就已产生，而且其在胚胎时期所产生的影响可能伴随人的一生。然而，环境因素如何影响基因表达并参与各种生理及病理生理过程尚不清楚。

在这一背景下，表观遗传学应运而生，其主要研究在DNA核苷酸序列不发生改变的基础上而基因表达发生了可遗传的改变。DNA甲基化、组蛋白修饰和microRNA是目前研究最广泛和最深入的表观遗传分子机制。这些表观遗传修饰通过直接或间接地对染色质空间结构及压缩状态进行动态调节，使得基因组中特定功能的基因在特定的场所和特定的时间条件下准确表达或关闭表达，从而影响生长、发育、凋亡等生命过程。换言之，表观遗传修饰可在环境因素的影响下参与基因表达调控，从而影响表型，这一过程不依赖于DNA遗传序列的改变。

事实上，表观遗传修饰异常在人类复杂性疾病的发病机制中起重要作用，这一观点通过近年来大量相关研究报道引起广泛认同。随着人类基因组计划的完成，新兴的高通量测序技术和生物信息技术逐步得以普及，使得我们从全基因组水平进行表观遗传学研究成为可能。这一技术和研究理念革命

性的飞跃使得人类表观基因组计划得以实施，针对恶性肿瘤、自身免疫性疾病等重大复杂性疾病的表观基因组学研究得到迅猛发展。令人瞩目的是，表观遗传修饰本身具有可逆性，这为我们治疗和逆转复杂性疾病带来了希望之光。随着对复杂性疾病表观遗传分子调控机制的深入研究，以及表观遗传药物及干预技术的开发，未来医学有望实现在早期利用表观遗传分子标记对患病风险进行筛查和预警，甚至可能通过特定的饮食、药物或环境干预手段，逆转疾病相关的异常表观遗传修饰状态。

近十余年来，表观遗传学研究已成为国际、国内生物医学领域的热点和前沿，然而国内目前介绍表观遗传学的著作和译著为数不多，而集中阐述与人类复杂性疾病相关的表观遗传机制的著作尚未面世。本书的编写将尝试填补这一空白，通过对迄今为止国内外相关研究成果的汇编与梳理，为表观遗传学研究人员提供一部较为全面、系统和便捷的参考资料和工具书。

本书的上篇将系统阐述表观遗传学的基本概念、内涵、外延及相关研究进展，使读者对DNA甲基化、组蛋白修饰、microRNA、X染色体失活、染色质重塑等重要概念形成比较全面、清晰的认识。下篇将以自身免疫性疾病、代谢性疾病及恶性肿瘤等疾病作为典型代表，重点阐述表观遗传修饰在人类复杂性疾病发生、发展过程中的作用及其分子机制，并介绍表观遗传学研究方法，包括近年来开始兴起的高通量测序方法在表观遗传研究中的应用，最后就表观遗传修饰的干预药物及其研究进展进行综述。

在本书的撰写过程中，我们尽最大努力为读者呈现准确、翔实和前沿的知识，但限于作者学识水平和经验，加之科研工作的繁重和时间紧迫，本书的疏漏之处在所难免，恳请同行专家和广大读者惠予指正。

本书从策划到出版历时3年多，在此过程中全体编写人员在繁重的科研、教学与临床之余付出

了辛勤的劳动，借此机会表示诚挚的谢意！医学遗传学国家重点实验室夏家辉院士及天津医科大学尚永丰院士在百忙之中欣然为此书作序，在此深表谢

意！另外，此书出版得到了北京大学医学出版社的大力支持，在此表示感谢。

中南大学湘雅二医院 陆前进
复旦大学生物医学研究院 于文强
复旦大学遗传工程国家重点实验室 吕红
2015年12月

目 录

上篇 表观遗传学基础理论

第一章 染色质的结构与功能..... 3	第五章 表观遗传修饰调控网络..... 71
第一节 概述 3	第一节 概述 71
第二节 染色质的基本结构单位——核 小体 5	第二节 组蛋白密码 71
第三节 染色质的高级结构 10	第三节 组蛋白修饰与 DNA 甲基化之间的 关联 75
第四节 染色质形成的特殊结构 14	第四节 组蛋白修饰与非编码 RNA 77
第二章 DNA 甲基化与基因表达调控 19	第五节 DNA 甲基化与非编码 RNA 79
第一节 DNA 甲基化 19	第六章 干细胞表观遗传学 81
第二节 真核生物 DNA 甲基转移酶 23	第一节 胚胎干细胞表观遗传学 81
第三节 甲基化 DNA 结合蛋白 28	第二节 成体干细胞表观遗传学 85
第四节 DNA 去甲基化 32	第三节 肿瘤干细胞表观遗传学 88
第五节 DNA 羟甲基化 34	第七章 基因组印记 94
第三章 组蛋白与表观遗传调控 42	第一节 基因组印记概述 94
第一节 组蛋白的翻译后修饰 43	第二节 基因组印记与胚胎发育 95
第二节 组蛋白变体 58	第三节 遗传印记的表达调控 96
第三节 染色质重构 60	第四节 印记相关疾病 97
第四章 非编码 RNA 与基因表达调控 64	第五节 基因印记与 X 染色体失活 98
第一节 非编码 RNA 的发现 64	第八章 X 染色体失活 101
第二节 基因组功能分区 64	第一节 概念 101
第三节 非编码 RNA 分类 66	第二节 分子机制 101
	第三节 X 染色体失活逃逸 104

下篇 表观遗传修饰与复杂性疾病

第九章 概论 108	第二节 表观遗传修饰与肝癌 116
第一节 复杂性疾病概述 108	第三节 表观遗传修饰与肺癌 122
第二节 表观遗传修饰与复杂性疾病 111	第四节 表观遗传修饰与结肠癌 127
第三节 表观遗传修饰与复杂性疾病治疗 ... 112	第五节 表观遗传修饰与前列腺癌 135
第十章 表观遗传修饰与肿瘤 115	第六节 表观遗传修饰与脑瘤 142
第一节 概论 115	第七节 表观遗传修饰与食管癌 147
	第八节 表观遗传修饰与鼻咽癌 153

第九节 表观遗传修饰与卵巢癌	158	第十四章 表观遗传修饰与神经精神疾病.....	252
第十节 表观遗传修饰与皮肤黑色素瘤	164	第一节 概论	252
第十一节 表观遗传修饰与血液系统肿瘤 ...	168	第二节 表观遗传修饰与神经退行性疾病 ...	253
第十一章 表观遗传修饰与自身免疫性疾病.....	184	第三节 表观遗传修饰与常见精神性疾病 ...	258
第一节 表观遗传修饰与系统性红斑狼疮 ...	184	第十五章 表观遗传学研究方法.....	270
第二节 表观遗传修饰与类风湿关节炎	200	第一节 概论	270
第三节 表观遗传修饰与银屑病	205	第二节 DNA 甲基化研究方法	271
第四节 表观遗传修饰与系统性硬皮病	207	第三节 组蛋白修饰研究方法	277
第十二章 表观遗传修饰与代谢性疾病.....	219	第四节 miRNA 研究方法	281
第一节 概论	219	第五节 高通量测序方法在	
第二节 表观遗传修饰与糖尿病	220	表观遗传研究中的应用	285
第三节 表观遗传修饰与骨质疏松	226	第十六章 表观遗传修饰干预药物.....	292
第十三章 表观遗传修饰与心血管系统疾病.....	239	第一节 概论	292
第一节 概论	239	第二节 DNA 甲基化干预药物	296
第二节 表观遗传修饰与高血压	239	第三节 组蛋白去乙酰化酶抑制剂	302
第三节 表观遗传修饰与动脉粥样硬化	244	第四节 展望	307
		索引.....	

上 篇

表观遗传学基础理论

第一章 染色质的结构与功能

第一节 概述

19 世纪末至 20 世纪中叶, 遗传学 (genetics) 取得了巨大的成就, 成为生物学中发展最快的学科之一。1865 年, G. J. Mendel 提出了遗传学的两大定律, 即分离定律和自由组合定律^[1]。1903 年, W. Sutton 和 T. Boveri 将遗传与染色体联系在了一起^[2]。1910—1913 年, T. H. Morgan 及其学生发现了基因的伴性遗传规律和连锁交换定律^[3-4]。1944 年, O. Avery 证明了遗传物质是脱氧核糖核酸 (DNA), 而不是蛋白质^[5]。1953 年, J. Watson 和 F. Crick 提出了 DNA 的双螺旋模型, 自此生物学进入了分子时代^[6]。1958 年, F. Crick 提出了著名的“中心法则”, 概述了遗传物质的传递和翻译过程^[7]。该法则认为 DNA 是遗传物质的核心载体, 它可以进行自我复制; DNA 可以通过 RNA 将遗传信息传递给蛋白质, 但蛋白质不能将信息传递给 DNA。“中心法则”可以简单地用下列公式来表示: $\text{DNA} \rightleftharpoons \text{RNA} \rightarrow \text{蛋白质}$ 。时至今日, 该法则已经过 50 多年的验证和发展, 成为了生物研究的重要基石。

DNA 编码的遗传信息包含在由 A、T、C、G 四种核苷酸组成的序列中。个体所具有的特 DNA 序列被称为基因型 (genotype)。个体表现出的可被检测到的特征, 被称作表型 (phenotype)。经典遗传学认为, 表型是由基因型和不可遗传的环境因素共同决定的。然而, 越来越多的研究结果表明, 即使在相同的基因型和环境条件之下, 生物体仍然可以呈现出截然不同的表型。例如, 单细胞的酵母可以在母代和子代之间转换结合型 (性别), 同卵双胞胎在性格和外貌上的差异, 果蝇含有红白色斑驳复眼细胞, 同一株拟南芥的种子萌发成形状不同的花朵, 克隆猫与其遗传母亲肤色不同等。作为最常见的例子, 在多细胞生物发育过程中, 含有完全相同 DNA 序列的细胞可以分化为具有不同功能的细胞。这些表型的变化是

由基因表达的差异造成的, 但是这些差异并不涉及 DNA 序列的变化。在某些情况下, 类似的表达差异可以通过有丝分裂或减数分裂被稳定地遗传到子代细胞中, 这说明存在一种不同于 DNA 的遗传因子在发挥作用。对于这类“异常”遗传因子的研究几乎与经典遗传学同时兴起, 经过近 50 多年的积累, 发展成为了一门新的学科“表观遗传学 (epigenetics)”。

“表观遗传学”的命名源于一个古老的名词“后生 (epigenesis)”。早在 2000 多年前, 古希腊伟大的哲学家亚里士多德就在其著作中描述并支持了后生理论。该理论认为发育是一个有序展开的过程, 特别是指动物的卵细胞或者植物的种子需要经过一系列的器官形成和细胞分化才能发育成熟。1942 年, C. H. Waddington 将“epigenesis”和“genetics”融合, 创造出了一个新的名词“epigenetics”^[8]。Waddington 认为基因与其产物之间的相互作用决定了表型, 而表观遗传学则是研究这种相互作用的学科。1990 年, R. Holliday 提出表观遗传学研究的是“在复杂生物发育过程中, 在时间和空间上对基因活性进行调节的机制”, 明确了表观遗传学是一门研究基因表达调控的学科^[9]。20 世纪 90 年代后, 表观遗传学得到了飞速的发展, 其定义也随之变得更为精细, 其主要的变化是将表观遗传限定为“非 DNA”的遗传。根据目前的观点, 表观遗传学研究的对象是可遗传的基因表达的变化, 而这些变化并不涉及 DNA 序列的改变。

目前, 表观遗传学研究的三大对象或者说表观遗传的三大调控机制包括: DNA 甲基化、组蛋白修饰和非编码 RNA 调控。这三者都是通过影响染色质 (chromatin) 的结构和功能来调节基因的表达水平, 因此染色质构成了表观遗传调控的基础。染色质, 是指真核生物细胞核中由 DNA 与其紧密结合的蛋白质所构成的复合物。19 世纪 70 年代, W. Flemming

在对细胞进行染色时，发现分布在细胞核中的一些物质能够大量吸收碱性的苯胺染料。因此，他将这种物质命名为染色质（在希腊语中，chromatin 是颜色的意思）。当细胞进入分裂期后，染色质就会凝聚为线状，这种结构的物质后来被命名为染色体（chromosome，意为“带颜色的物体”）^[10]（图 1-1）。1888 年，Kossel 发现了构成染色质的主要蛋白质——组蛋白（histone）^[11]。与遗传学研究的热火景象不同，对于染色质功能的研究在随后的 70 多年里变得相对沉寂。1964 年，V. Allfrey 推测组蛋白修饰（酰基化和甲基化）与染色质转录相关，激发了研究组蛋白修饰与基因表达的热度^[12]。1974 年，R. Kornberg 提出了染色质的基本单位——核小体的结构模型^[13]。随后，T. Richmond 等人对核小体的

结构进行解析，为染色质的功能研究奠定了分子基础^[14-15]。1996，C. D. Allis 发现了四膜虫的组蛋白酰基转移酶与酵母中转录激活因子 Gcn5 同源，从而提供了组蛋白修饰调控基因表达的直接证据^[16]。以此为契机，以染色质研究为基础的表现遗传学进入了一个飞速发展的阶段。DNA 甲基化长久以来一直被认为与基因沉默相关^[17]。1998 年，研究发现 DNA 甲基化能够征集组蛋白去酰基化酶和甲基化酶，揭示了 DNA 甲基化通过改变染色质结构来抑制基因表达的作用机制^[18]。关于 RNA 的研究是 20 世纪 90 年代后期蓬勃兴起的领域之一。2000 年前后，一系列研究表明非编码 RNA 直接参与了异染色质的建立和维持，使得非编码 RNA 的研究成为了表现遗传学中的重要组成部分^[19]。

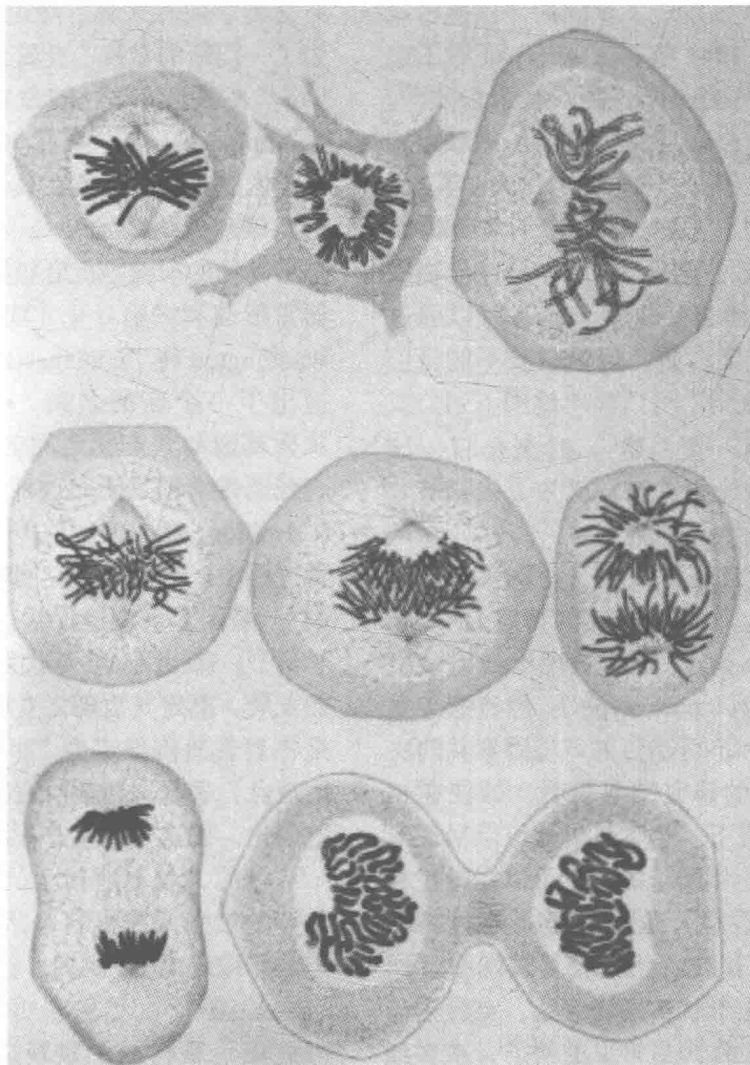


图 1-1 W. Flemming 所描绘的细胞染色体和有丝分裂过程

引自 Flemming W. Zellsubstanz, Kern und Zelltheilung. Leipzig: FCW Vogel, 1882.

染色质作为 DNA 在核内的高级结构，控制了激活及抑制因子与 DNA 的接触和分离，从而在转录水平上调控了基因的表达。更为重要的是，染色质对于基因表达的调控有记忆性，这就是所谓的“表观遗传记忆”。经典的转录模型认为基因的表达是由激活因子所控制的，一旦激活因子消失，转录也就消失了。当转录发生后，染色质的结构发生了改变，这种改变可以是激活性的，也可以是抑制性的，一些染色质结构的改变并不稳定，可能在一个细胞周期内就消失，或者不能被遗传到子细胞中去，这种结构的改变必须要通过转录激活因子的再次征集才可能被重建。而某些结构的改变，通过 DNA 甲基化、组蛋白修饰、非编码 RNA 调控等手段加以稳固，可以被遗传到子细胞中去，从而实现了不依赖 DNA 序列变化的表观遗传。当然，表观遗传不是稳定的，它可以被逆转，这是它与 DNA 遗传之间最主要的区别，而染色质结构的稳定性影响了表观遗传记忆的长短。表观遗传的信号（特定的染色质结构）是如何在有丝分裂和减数分裂中由母细胞遗传到子细胞中去的，是该领域研究

的重点。此外，表观遗传信号是否能够像 DNA 遗传一样，可以通过生殖细胞被传递到下一代中去，仍然是一个存在争议的问题。如果确实存在类似的机制，这将揭示一条蛋白质 → 染色质 → 蛋白质的全新的遗传途径（图 1-2）。关于这个问题的研究将深刻地影响到中心法则、达尔文进化论等生物科学的基石理论，将成为新世纪表观遗传学的热点。

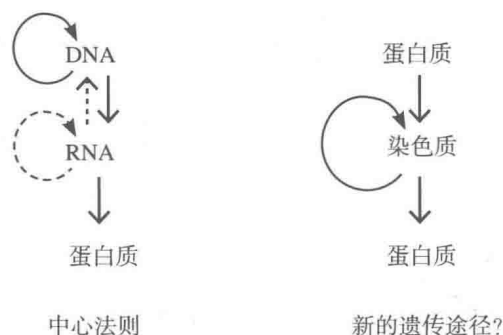


图 1-2 表观遗传学能否证实一条新的遗传途径

第二节 染色质的基本结构单位——核小体

一、核小体的发现

染色质需要经过高度的折叠和压缩才能被容纳在细胞核中。举例来说，人类细胞 DNA 的总长度约为 2m，而细胞核直径仅为 10 μ m 左右，凝聚的过程好比是将长达 8km 的细线塞进一个乒乓球中，这其中还没有考虑与 DNA 紧密结合的大量蛋白质。染色质的凝聚并不是随意的：有些区域需要高度凝聚，以沉默基因的表达或者形成特定的染色质结构，而在某些区域需要疏松的结构，使 DNA 能够接触到其他的生物分子，进行转录、复制、修复等过程。此外，染色质的凝聚程度也随着细胞周期而改变，染色质在分裂期，特别是分裂期中期的凝聚程度最高，而在分裂间期染色质则处于相对疏松的状态，不同时期之间的凝聚程度可以相差 10 倍以上。因此，染色质的凝聚是一个在时空上受到精密调控的过程。

染色质第一级的凝聚，是通过“核小体 (nucleosome)”的结构来实现的。1972 年，A. Olins 和 D. Olins 采用低盐溶液抽提了鸡血红细胞中的细

胞核。在显微镜下，他们观察到了类似“念珠”一样的结构。“念珠”上的珠子就是核小体，其直径约为 10nm，将珠子串联起来的细线是核小体之间的 DNA。1974 年，R. Kornberg 发现了组蛋白之间能够形成复合物，他综合了核酸酶消化染色质能够获得 200bp DNA 片段的数据，提出了核小体的结构模型。他认为核小体是由 200bp 的 DNA 包裹着组蛋白核心所构成的染色质基本结构单位。1984 年，T. Richmond 解析了核小体的结构，分辨率为 7 \AA 。1997 年，T. Richmond 又将解析度提高到了 2.8 \AA ，在接近原子的水平提供了这个重要结构的细节。

二、核小体的结构

1. 核小体的结构 在常规的核小体中，共含有 4 种组蛋白，H3、H4、H2A 和 H2B。每种组蛋白的 2 个拷贝结合在一起，形成了一个组蛋白八聚体；约 146bp 的 DNA 片段按照左手螺旋的方式在这个八聚体外缠绕 1.65 圈，就构成了一个核小体的核心颗粒

(有时习惯把核心颗粒就称作“核小体”)。核小体核心颗粒的直径约为 10nm, 146bp 的 DNA 片段长度为 57nm, 因此 DNA 通过核心颗粒在长度上压缩了 6 倍。核小体还包括了与核心颗粒之间的连接 DNA, 其长度从数个 bp 到 80bp。平均来说, 每个核小体总共含有 200bp 的 DNA。以核小体为单位的染色质将 DNA 的长度缩短到了近 1/3, 这是 DNA 在核内第一级的压缩率^[14,20]。

常规组蛋白, 包括 H3、H4、H2A 和 H2B, 是一类长度约为 100 个氨基酸的小型蛋白质。它们具有非常相似的结构。组蛋白的中部含有名为“组蛋白折叠”(histone fold) 的基序。该基序由 70 个左右的氨基酸组成, 具有“ α 螺旋-环- α 螺旋-环- α 螺旋”的结构特征(图 1-3)。核小体组装时, 组蛋白首先利用该基序发生相互结合, 形成了 H3-H4 的二聚体和 H2A-H2B 的二聚体。随后两个 H3-H4 的二聚体相互结合构成一个四聚体, 该四聚体与 2 个 H2A-H2B 的二聚体结合, 形成了一个完整的组蛋白八聚体。八聚体的核心是一个球状结构域, DNA 在这个结构域外进行缠绕, 就完成了核小体的组装。

2. 核小体与 DNA 之间的结合 每隔 10 个 bp, DNA 双螺旋中的小沟就会靠近组蛋白八聚体, 与之发生结合。组蛋白八聚体主要含有两种 DNA 结合位点, 它们都是由“组蛋白折叠”基序中的序列组成。第 1 种结合位点“ $\alpha 1 \alpha 1$ ”是由 2 个相邻组蛋白所含有的第 1 个 α 螺旋构成, 第 2 种结合位点“L1L2”是由第 1 个和第 2 个环(loop) 构成。八聚体与 DNA

之间的结合包括了多达 142 对的氢键, 其中 50% 发生在组蛋白肽链的骨架与 DNA 的磷酸二酯骨架之间。此外, 带电基团之间的静电作用和非极性基团之间的疏水作用对于结合也贡献颇多。以静电作用为例, 组蛋白富含赖氨酸、精氨酸, 这类氨基酸的侧链带有正电荷侧链, 因此球状核心的表面分布了大量正电荷, 这些正电荷与 DNA 的磷酸二酯骨架上的负电荷相互吸引, 稳定了 DNA 与组蛋白八聚体之间的结合。值得注意的是, DNA 的碱基并没有直接与组蛋白结合, 这说明组蛋白对于 DNA 的高亲和力并不依赖于 DNA 序列, 因此, 组蛋白八聚体几乎能够装配各种序列的 DNA^[21]。

3. 组蛋白的尾巴 从组蛋白的 N 端到“组蛋白折叠”的基序之间, 是一段长度为 10~30bp 的序列, 通常被称作组蛋白的“尾巴”。在核小体的晶体结构中, “尾巴”是看不见的, 说明这段序列十分柔软, 并不具有特定的结构。组蛋白的“尾巴”穿过 DNA “线圈”之间的缝隙, 从八聚体的球状结构域中伸展出去, 暴露在外。与被包裹在球状结构域中的部分不同, 组蛋白尾巴并不直接参与核小体的结构组成, 但它通过与其他蛋白质发生相互作用, 从而对核小体的结构和功能进行调节。举例来说, 如果将酵母组蛋白 H4 的球状结构域的序列进行少量缺失, 就会导致酵母细胞的死亡; 不同的是, 如果将 H4 的尾巴全部去除, 细胞仍然能够存活, 但是会表现出异常的基因沉默^[22]。组蛋白尾巴的主要调节方式是其残基的翻译后修饰。常见的修饰包括了酰基化、甲基

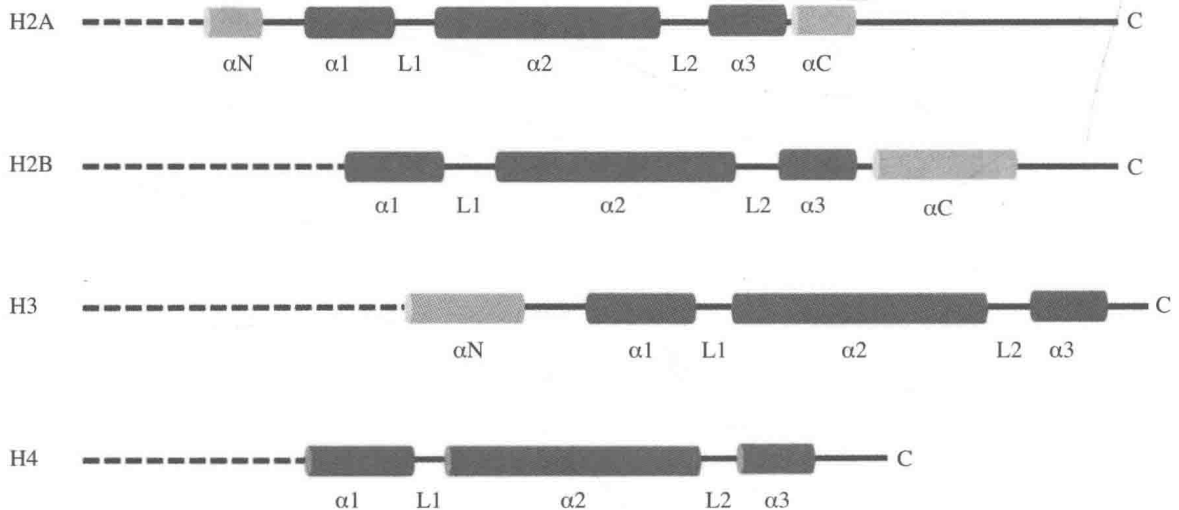


图 1-3 组蛋白都含有名为“组蛋白折叠”的基序(图中“组蛋白折叠”的基序用深灰色表示, 虚线表示 N 端尾巴)