

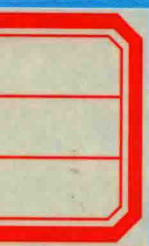
全国普通高等教育临床医学专业“5+3”十二五规划教材

Medical Molecular Biology

医学分子生物学

供临床医学、预防医学、口腔医学
医学影像学、医学检验学等专业用

主编 武军驻



江苏科学技术出版社

中国普通高等教育临床医学专业“5+3”十二五规划教材

Medical Molecular Biology

医学分子生物学

供临床医学、预防医学、口腔医学
医学影像学、医学检验学等专业用

主 编 武军驻

副 主 编 王海生

编 委 武军驻 (武汉大学基础医学院)

王海生 (内蒙古医科大学)

田余祥 (大连医科大学)

汪军梅 (长治医学院)

栗学清 (长治医学院)

于 慧 (包头医学院)

邓秀玲 (内蒙古医科大学)

李俏俏 (山东万杰医学院)

图书在版编目(CIP)数据

医学分子生物学 / 武军驻主编. —南京: 江苏科学技术出版社, 2014. 1

全国普通高等教育临床医学专业“5+3”十二五规划教材

ISBN 978-7-5537-0903-1

I. ①医… II. ①武… III. ①医学—分子生物学—高等学校—教材 IV. ①Q7

中国版本图书馆 CIP 数据核字(2013)第 036590 号

医学分子生物学

主 编 武军驻
责任编辑 庞啸虎
责任校对 郝慧华
责任监制 曹叶平 方 晨

出版发行 凤凰出版传媒股份有限公司
江苏科学技术出版社
出版社地址 南京市湖南路 1 号 A 楼, 邮编: 210009
出版社网址 <http://www.pspress.cn>
经 销 凤凰出版传媒股份有限公司
排 版 南京展望文化发展有限公司
印 刷 江苏凤凰数码印务有限公司

开 本 889mm×1194mm 1/16
印 张 12
字 数 330 000
版 次 2014 年 1 月第 1 版
印 次 2014 年 1 月第 1 次印刷

标准书号 ISBN 978-7-5537-0903-1
定 价 35.00 元

图书如有印装质量问题, 可随时向我社出版科调换。

出版说明

为了全面提高我国普通高等教育医药卫生类专业人才的培养质量,深入落实《国家中长期教育改革和发展规划纲要(2010~2020)》以及服务于医疗教育体系的改革,深入贯彻教育部、卫生部2011年12月联合召开的“全国医学教育改革工作会议”精神,通过全面实施以“5+3”为重点的临床医学教育综合改革方案,进一步深化和推进医学教育深层次改革和发展,通过全面推进临床医学专业课程体系及教育体系的改革和创新,推动临床医学教育内容及教学方法改革和创新,进一步更好地服务教学、指导教学、规范教学,实现临床医学教学质量全面提高,培养高层次、高水平、应用型的卓越医学人才,从而适应我国医疗卫生体制改革和发展的需要,凤凰出版传媒集团江苏科学技术出版社作为长期从事教育出版的国家一级出版社,于2012年1月组织全国50多家高等医学院校开发了国内第一套临床医学专业“5+3”十二五规划教材。

该套教材包括基础课程、专业课程46种,部分教材还编写了相应的配套教材。其编写特点如下:

1. 突出“5+3”临床医学专业教材特色 这套教材紧扣“5+3”临床医学专业的培养目标和专业认证标准,根据“四证”(本科毕业证、执业医师资格证、住院医师规范化培训证和硕士研究生毕业证)考核要求,紧密结合教、学、临床实践工作编写,由浅入深、知识全面、结构合理、系统完整。全套教材充分突出了“5+3”临床医学专业知识体系,渗透了“5+3”临床医学专业人文精神,注重体现素质教育和创新能力与实践能力的培养,反映了“5+3”临床医学专业教学核心思想和特点。

2. 体现教材的延续性 本套教材仍然坚持“三基”(基础理论、基本知识、基本技能)、“五性”(思想性、科学性、先进性、启发性、适用性)、“三特定”(特定的对象、特定的要求、特定的限制)的原则要求。同时强调内容的合理安排,深浅适宜,适应“5+3”本科教学的需求。

3. 体现当代临床医学先进发展成果的开放性 这套教材汲取了国内外最新版本相关经典教材的新内容,借鉴了国际先进教材的优点,结合了我国现行临床实践的实际情况和要求,并加以创造性地利用,反映了当今医学科学发展的新成果。

4. 强调临床应用性 为加快专业学位教育与住院医师规范化培训的紧密衔接,教材加强了基础与临床的联系,深化学生对所学知识的理解,实现早临床、多临床、反复临床的理念。

5. 强调了全套教材的整体优化 本套教材不仅追求单本教材的系统性和全面,更是强调了全套教材的整体优化,注意到了不同教材内容的联系和衔接,避免遗漏和重复。

6. 兼顾教学内容的包容性 本套教材的编者来自全国几乎所有省份,教材的编写,兼顾了不同类型学校和地区的教学要求,内容涵盖了临床执业医师资格考试的基本理论大纲的知识点,可供全国不同地区不同层次的学校使用。

7. 突出教材个性 本套教材在保证整体优化的前提下,强调了个教材的个性,技能性课程突出了技能培训;人文课程增加了知识拓展;专业课程则增加了案例导入和案例分析。

8. 各科均根据学校的实际教学时数编写,文字精炼,利于学生对重要知识点的掌握。

9. 在不增加学生负担的前提下,根据学科需要,部分教材采用彩色印刷,以提高教材的成书品质和内容的可读性。

这套教材的编写出版,得到了广大医学院校的大力支持,作者均来自各学科教学一线,具有丰富的临床、教学、科研和写作经验。相信本套教材的出版,必将对我国当下临床医学专业“5+3”教学改革和专业人才培养起到积极的推动作用。

全国普通高等教育临床医学专业“5+3”十二五规划教材

医学导论	睦建	主编	医学影像学	李坤成	主编	
基础化学	杨金香	主编	临床麻醉学	晁储璋	主编	
有机化学	周健民	黄祖良	全科医学概论	谢波	主编	
生物化学	黄忠仕	翟静	内科学	雷寒	王庸晋	主编
医学分子生物学	武军驻	主编	外科学	康骅	薛昊罡	主编
医学细胞生物学	苗聪秀	主编	妇产科学	段涛	胡丽娜	主编
医学物理学	甘平	主编	儿科学	于洁	主编	
医学伦理学	陈颢	主编	中医学	黄岑汉	主编	
医学心理学	杜玉凤	主编	皮肤性病学	何黎	金哲虎	主编
生理学	白波	杜友爱	康复医学	李雪斌	陈翔	主编
组织学与胚胎学	苏衍萍	王春艳	神经病学	沈霞	主编	
病理生理学	商战平	王万铁	精神病学	王克勤	主编	
病理学	盖晓东	李伟	眼科学	吕帆	主编	
药理学	董志	毛新民	口腔医学	邓锋	主编	
人体寄生虫学	李士根	主编	耳鼻咽喉头颈外科学	龚树生	主编	
医学微生物学	于爱莲	吕厚东	传染病学	周智	主编	
医学免疫学	宋文刚	主编	临床流行病学	冯向先	主编	
临床药理学	许小林	主编	急诊与灾难医学	廖品琥	主编	
核医学	段炼	主编	局部解剖学实践指导及习题集	黄秀峰	吴洪海	主编
医学统计学	景学安	主编	人体寄生虫学学习指导	李士根	主编	
卫生法学	徐晨	蒲川	医学物理学学习指导	甘平	主编	
流行病学	毛淑芳	主编	医学物理学实验	张翼	罗亚梅	主编
预防医学	喻荣彬	主编	眼科学学习指导	吕帆	主编	
法医学	邓世雄	主编	有机化学学习指导	周健民	黄祖良	主编
系统解剖学	李富德	朱永泽	基础化学学习指导	黄锁义	主编	
局部解剖学	吴洪海	黄秀峰	医院感染学	郑文芳	邢玉斌	主编
诊断学	魏武	郑文芝				

我国的临床医学教育中,五年制医学高等教育没有专门开设《医学分子生物学》这门课程,只是在《生物化学》的相关章节中简要介绍了一些相关的基本概念,缺乏方法学的介绍。为了提高我国普通高等医学教育医药卫生类专业人才的培养,深入落实国家中长期(2010~2020)教育改革和发展规划纲要,以及服务医疗教育体系的改革,我们在为长学制(七年制与八年制)开设《医学分子生物学》(2001~2012)的讲授实践,以及自2009年主持五年制武汉大学医学部与芝加哥大学医学院教改实验班整合《医学分子生物学》教学内容进入整合课程的经验基础上,编写了这本《医学分子生物学》教材。

分子生物学的主要研究对象是核酸和蛋白质,因此第一章与第二章分别介绍基因与基因组、蛋白质与蛋白质组的基本概念。而学习医学分子生物学的目的是为临床服务,因此最后两章(第十三章与第十四章)分别介绍基因诊断与基因治疗。以“中心法则”为主线,强调“学以致用”的原则,两章为一个单位分别撰写其余部分。第三章为“DNA的生物合成”,第四章则介绍DNA体外合成即PCR与测序,强调第四章应用过程中涉及的理论部分就是第三章的重点内容,在实际教学中指导学生理解重点之所以成为重点的理由。以此类推:第六章是第五章的展开;第八章是第七章的补充;第九章与第十章简要介绍基因的一般操作;第十一与第十二章阐述体外基因操作如何参与体内的基因表达调控。通过本课程的学习,使学生理解基因信息传递的基本过程、基因表达调控的概念,理解体外进行基因操作的一般原理及它们在医学上的应用。

总之,这本教材是本着创新的教学理念与避免传统教材“大而全”弊端的一次尝试,无论内容与形式都可能存在很多不足,欢迎广大使用者在使用的过程中提出宝贵的指导意见。

武军驻

目 录

第一章 基因、基因组与基因组学	1
第一节 基因	1
一、基因概念的提出与发展	1
二、基因的化学本质	2
三、DNA 的分子结构	3
四、基因的组构	5
第二节 基因组	7
一、真核生物染色体基因组的特点	8
二、原核生物染色体基因组的特点	8
三、病毒基因组的特点	9
第三节 基因组学	9
一、结构基因组学	10
二、功能基因组学	11
三、比较基因组学	12
第二章 蛋白质、蛋白质组和蛋白质组学	13
第一节 蛋白质	13
一、蛋白质的一级结构	13
二、蛋白质的空间结构	14
三、蛋白质结构与功能的关系	17
第二节 蛋白质组与蛋白质组学	18
一、蛋白质组	18
二、蛋白质组学	19
三、蛋白质组学研究的技术方法	20
四、生物信息学在蛋白质学研究中的应用	24
第三章 DNA 的生物合成	25
第一节 遗传学的中心法则和 DNA 复制的特点	25
一、遗传学的中心法则	25
二、DNA 复制的特点	25
第二节 DNA 复制的酶学	27
一、DNA 聚合酶	27
二、DNA 的超螺旋结构及拓扑异构酶	28
第三节 DNA 生物合成过程	30
一、DNA 复制的起始	30
二、DNA 链的延长	31
三、DNA 复制的终止	31

四、其他环状 DNA 分子的复制	32
第四节 DNA 的损伤与修复	33
一、DNA 的损伤	33
二、DNA 损伤的修复	34
第四章 DNA 的体外扩增与测序	36
第一节 聚合酶链反应	36
一、PCR 基本原理	36
二、PCR 反应体系	37
三、PCR 反应条件	39
四、PCR 扩增产物的分析	40
五、PCR 技术应用进展	41
第二节 DNA 测序技术	44
一、双脱氧链末端终止法	44
二、化学降解法	47
三、DNA 自动化测序	48
四、DNA 测序技术的应用	49
第五章 基因表达	51
第一节 RNA 的生物合成	51
一、转录体系	51
二、转录过程	55
三、转录后的加工	57
第二节 蛋白质的生物合成	62
一、蛋白质合成体系	62
二、蛋白质的生物合成过程	65
三、真核生物与原核生物蛋白质合成的异同	73
四、蛋白质生物合成后的加工和运输	74
第六章 体外转录与翻译体系的建立	79
第一节 基因的体外转录	79
一、基因体外转录的基本原理	79
二、基因体外转录的意义及应用	80
第二节 基因的体外翻译	81
一、基因体外翻译的基本原理	81
二、基因体外翻译的应用	84
第七章 基因表达调控	86
第一节 基因表达调控基本概念	86
一、基因表达的规律	86
二、基因表达的方式	87
三、与基因表达调控相关的基本概念	88

四、基因表达的调控环节	88
五、基因表达调控的意义	89
第二节 原核生物基因表达调控	89
一、原核生物基因表达调控的特点	89
二、转录水平调控	90
三、原核基因在翻译水平的调控	94
第三节 真核生物基因表达的调控	95
一、真核生物基因表达调控的特点	95
二、DNA 水平的调控	96
三、转录水平的调控	96
四、转录后水平的调控	99
五、翻译水平的调控	100
六、翻译后水平调控	100
第八章 表观遗传学	101
第一节 概述	101
一、表观遗传学基本概念	101
二、表观遗传学的研究内容	102
三、表观遗传学研究的意义	102
第二节 表观遗传学作用的机制	102
一、DNA 甲基化	103
二、组蛋白修饰	104
三、染色质重塑	106
四、非编码 RNA 的调控	107
五、其他表观遗传机制	107
第三节 表观遗传学与人类疾病	108
一、表观遗传与癌症	108
二、表观遗传学与代谢综合征	108
三、表观遗传与自身免疫性疾病	108
四、表观遗传与心血管疾病	108
第九章 基因工程	110
第一节 基因工程的工具酶	110
一、限制性核酸内切酶	110
二、其他常用的工具酶	112
第二节 基因工程的载体	113
一、常用的克隆载体	114
二、常用的表达载体	118
第三节 基因工程的基本过程	119
一、目的基因的制备及常用分离方法	119
二、载体的选择和构建	120
三、目的基因与载体 DNA 的体外连接	120

四、重组 DNA 导入受体细胞	122
五、重组体的筛选与鉴定	123
六、外源基因的表达	124
第四节 基因工程的应用	126
一、定点诱变技术	126
二、基因诊断与基因治疗	126
三、遗传病的预防	126
四、基因工程药物、疫苗及抗体	127
五、基因靶向技术和转基因技术	127
第十章 转基因技术	128
第一节 转基因动物	128
一、转基因动物的基本原理与方法	128
二、转基因动物的检测	131
三、转基因动物的应用	131
第二节 转基因植物	132
一、基本原理及方法	132
二、转基因植物的应用	133
三、转基因植物的安全性评价	133
第十一章 SiRNA 与 MiRNA	135
第一节 RNAi 及 siRNA	135
一、RNA 干涉的基本过程和共同特点	136
二、RNA 干涉的功能及应用	139
三、RNAi 技术的基本程序	141
第二节 MicroRNA 和基因沉默	143
一、miRNA	143
二、循环 miRNA	147
三、miRNA 的生理功能	148
四、问题和展望	149
第十二章 基因靶向技术	151
第一节 基因靶向技术基础	151
一、同源重组	152
二、胚胎干细胞分离和体外培养	152
第二节 基因靶向技术的原理和方法	153
一、基因靶向技术的原理	153
二、常用于基因敲除的分子生物学方法	153
三、利用同源重组构建基因敲除小鼠模型的步骤	155
四、条件性基因敲除与诱导性基因敲除	158
第三节 基因靶向技术应用及前景	159

第十三章 基因诊断	162
第一节 基因诊断的基本技术	162
一、核酸分子杂交	162
二、聚合酶链式反应	165
三、DNA 序列测定	167
第二节 基因诊断的基本策略	167
一、点突变的检测方法	167
二、基因缺失或插入的检测方法	167
三、基因重排的检测方法	167
四、基因扩增的检测方法	168
五、基因表达异常的检测方法	168
六、DNA 多态性的检测方法	168
第三节 基因诊断的应用	168
一、在遗传性疾病中的应用	169
二、在恶性肿瘤中的应用	169
三、在感染性疾病中的应用	170
四、在法医学中的应用	170
第十四章 基因治疗	172
第一节 基因治疗的基本方法和程序	172
一、基因治疗的基本方法	172
二、基因治疗的基本程序	174
第二节 基因治疗的应用研究	178
一、在遗传性疾病中的应用研究	178
二、在恶性肿瘤中的应用研究	179
三、在心脑血管疾病中的应用研究	179
四、在病毒感染性疾病中的应用研究	180

第一章

基因、基因组与基因组学

生物性状和生命特征为什么能代代相传？相传的生物性状为什么可以改变？控制生物性状的物质基础又是什么？这些都是生物学研究不得不触及的生命本质问题。基因(gene)控制着生物体的遗传性状,但对基因化学本质及功能的真正认识是20世纪40年代以后。基因组(genome)是一套完整单倍体遗传物质的总和,它储存着生物体整套的遗传信息。20世纪末诞生了旨在阐明基因组结构、功能,以及基因之间相互作用的一门学科——基因组学(genomics)。

基因是遗传的物质基础,是DNA或RNA分子上具有遗传信息的特定核苷酸序列。基因组通过复制把遗传信息传递给下一代,使后代出现与亲代相似的性状,并储存着生命孕育、生长、分化、衰老、死亡的全部信息,是决定生物体健康的内在因素。基因的变异不仅是生物进化的分子基础,也是疾病发生的分子基础,更是疾病治疗的分子基础。基因组学研究已经对生物学、遗传学、医药学乃至人类社会产生了深远的影响。

第一节 基 因

基因是核酸分子中编码RNA或多肽链的功能区段。基因信息遵照生物学遗传中心法则传递,即经过转录过程产生各种RNA分子,这样就将基因蕴藏的遗传信息抄录到RNA分子中,其中mRNA作为多肽链合成的模板,经过翻译过程指导多肽链中氨基酸的排列顺序。

一、基因概念的提出与发展

1865年,现代遗传学的奠基人孟德尔(G. J. Mendel)根据他的豌豆杂交试验结果提出了遗传因子学说,并对遗传因子的基本性质做了最早的论述:认为生物体内有某种遗传颗粒或遗传单位,它能传代,并控制生物体的性状。孟德尔通过人工培植这些豌豆,对不同子代豌豆的性状和数目进行细致入微的观察、计数和分析,认为遗传性状是由成对的遗传因子决定的,在生殖细胞形成过程中,成对的遗传因子分开,并分别进入2个生殖细胞中去,这就是遗传学的Mendel第一定律(或称分离律)。他还认为在生殖细胞形成过程中,不同对的遗传因子可以自由组合,此即Mendel第二定律(或称自由组合律)。孟德尔的遗传因子学说为基因概念的提出奠定了基础。

1903年,萨顿(W. S. Sutton)和博韦里(T. Boveri)根据各自的研究,分别注意到孟德尔的“遗传因子”与生殖细胞形成和受精过程中的染色体行为具有平行性,同时提出了遗传的染色体学说:认为遗传因子位于染色体上,第一次把遗传物质和染色体联系起来。这个从细胞学研究得出的结论,圆满地解释了孟德尔遗传现象。1909年,丹麦遗传学家约翰逊(W. L. Johannsen)在《精密遗传学原理》一书中提出“基因”的概念,以此来替代孟德尔的“遗传因子”。从此,“基因”一词一直伴随着遗传学发展至今。基因一词来自希腊语,意思为“生”。约翰逊还提出了提出基因型(genotype)和表型(phenotype)的概念。基因型是逐代传递的成对遗传因子的集合,表型是容易区分的个体特征的总和。基因型和表型概念的提出,初步阐明了基因与性状的关系,不过此时的基因仍然是一个未经证实的、仅靠逻辑推理得出的概念。

1910年,现代实验生物学奠基人美国生物学家摩尔根(T. H. Morgan)以果蝇(*Drosophila*)为研

究材料,发现基因的确存在于生殖细胞的染色体上,基因在每条染色体上是直线排列的,在进行减数分裂形成配子(即生殖细胞)时,排在一条染色体上的基因是不能自由组合的,此即为基因的“连锁”;同源染色体的断离与组合,产生了基因的互相交换,上述发现即是遗传学第三定律——连锁与互换律。摩尔根的实验观察既证实了性染色体在性别决定中的作用,同时也证实了基因是由染色体携带的学说。1926年,摩尔根巨著《基因论》的出版,创立了著名的基因学说,提出:① 基因是携带遗传信息的基本单位;② 基因又是控制特定性状的功能单位;③ 基因也是突变和交换的单位。至此,人们对基因概念的理解更加具体和丰富了。由于在染色体遗传理论方面的开创性工作,摩尔根获得了1933年诺贝尔生理学或医学奖。但基因的化学本质到底是什么?这在当时还是个谜。

二、基因的化学本质

核酸(nucleic acid)是基因的载体,基因的化学本质是核酸分子中编码 RNA 或多肽链的核苷酸序列。核酸的基本组成单位是核苷酸(nucleotide),DNA 由 4 种脱氧核糖核苷酸(dAMP、dGMP、dCMP 和 dTMP)组成,RNA 由 4 种核糖核苷酸(AMP、GMP、CMP 和 UMP)组成。核苷酸由碱基、戊糖和磷酸三部分构成。核酸分子中存在嘌呤碱基和嘧啶碱基,DNA 分子和 RNA 分子均含有腺嘌呤(A)、鸟嘌呤(G)和胞嘧啶(C),而尿嘧啶(U)仅存在于 RNA 分子,胸腺嘧啶(T)仅存在于 DNA 分子。DNA 分子中含有 β -D-2-脱氧核糖, RNA 分子中含有 β -D-核糖。碱基与戊糖通过 β -N-糖苷键共价相连形成核苷(nucleoside)。核苷的戊糖羟基与磷酸之间通过磷酸酯键相连形成核苷酸,其中 5'-核苷酸为核酸分子的基本组成单位(图 1-1)。

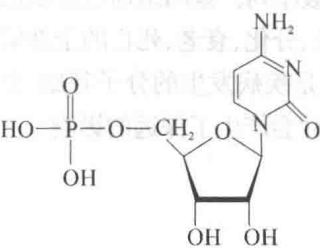


图 1-1 5'-胞嘧啶核苷酸结构

核酸分子中核苷酸的连接方式是 3',5'-磷酸二酯键,即前一个核苷酸的 3'-羟基与后一个核苷酸的 5'-磷酸基之间缩合形成酯键,由此许许多多核苷酸可连接成一个不分支的链状分子。这样的链状分子有一个 5'-磷酸基末端和一个 3'-羟基末端,其极性方向是从 5'→3'(图 1-2)。

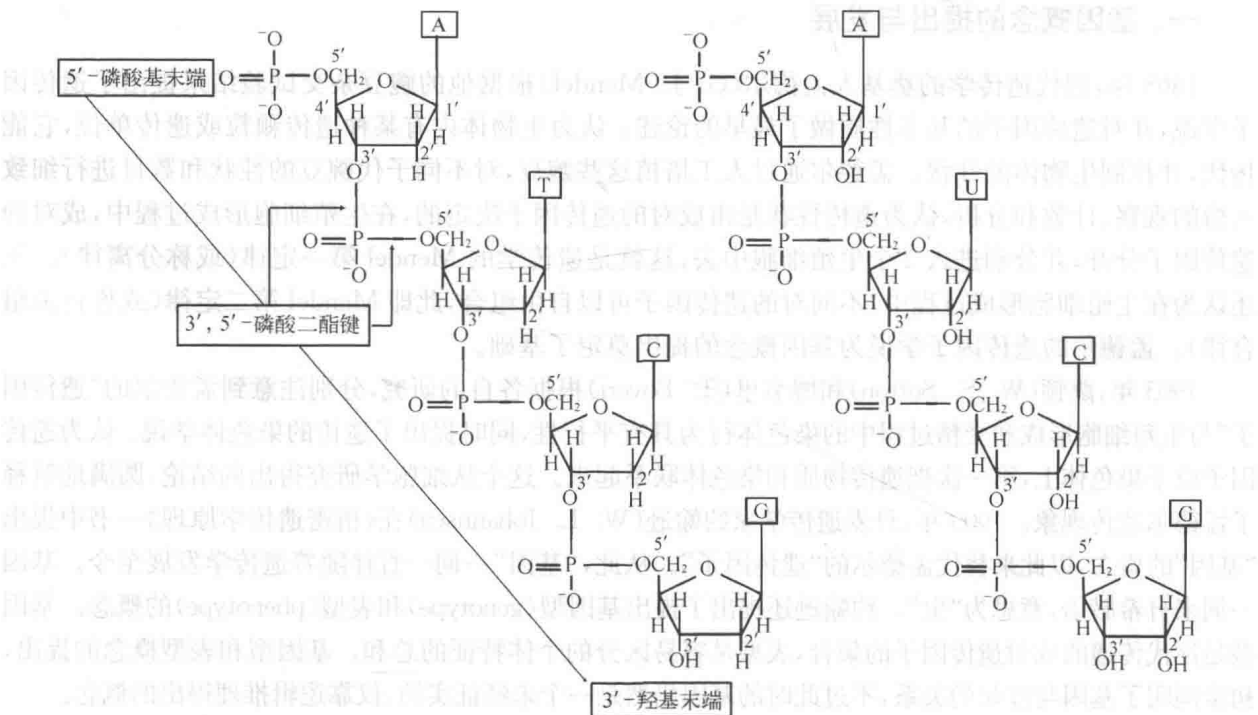


图 1-2 核酸的一级结构及核苷酸的连接方式(左为 DNA,右为 RNA)

核酸的一级结构即是指其分子中核苷酸的排列顺序。各种核苷酸之间的主要差别是所含碱基的不同,因此核酸分子中核苷酸的排列顺序即是碱基的排列顺序,基因信息就蕴藏在碱基的不同排列顺序中。

DNA 是遗传物质的证明

1928 年,英国细菌学家 Fred Griffith 进行的肺炎球菌转化实验在揭示 DNA 作为遗传信息携带者的研究中具有极为重要的作用。他将高温杀死的 S 型细菌和活的 R 型细菌一起注入小鼠体内,结果不仅有许多小鼠死于败血症,而且从死鼠血液中还发现了活的 S 型细菌,由此推测 S 型细菌有某种物质或转化因子进入了 R 型细菌,引起 R 型细菌转化为 S 型细菌。1944 年,美籍加拿大裔细菌学家艾弗里(Oswald Avery)与他的同事麦克劳德(MacLeod)和麦卡蒂(McCarthy)提供了 DNA 是遗传物质的第一个直接实验证据。他们将提取的 S 型细菌的多糖荚膜、蛋白质、DNA 分别与 R 型活菌混合培养,发现是 DNA 引起 R 型菌的转化,证明 DNA 是遗传物质。1952 年,赫尔希(Alfred Hershey)与蔡斯(Martha Chase)用放射性同位素研究噬菌体对细菌的感染,发现感染后噬菌体 DNA 存在于菌体内,此研究结果有力地证明了遗传物质是 DNA 而不是蛋白质。

三、DNA 的分子结构

DNA 的分子结构可分为一级、二级、三级,以至于更高级结构来描述。DNA 的一级结构即是指 4 种脱氧核苷酸的排列顺序(图 1-2),基因对遗传信息的携带和传递就是靠核苷酸排列顺序来实现的。

(一) DNA 的二级结构

真核生物染色体 DNA 的二级结构是双螺旋结构(图 1-3),主要的特点有:

1. DNA 分子是由 2 条反向平行的多核苷酸链围绕同一中心轴旋转而成的右手双螺旋结构。双螺旋结构的表面出现大沟和小沟是调节蛋白与 DNA 相互作用的位点,与基因表达的调控有关。

2. 磷酸-戊糖骨架位于双螺旋的外侧,戊糖平面与中心轴近于平行。碱基位于双螺旋的内侧,碱基平面与中心轴相垂直,每一螺距包含 10 对碱基。

3. 2 条 DNA 链之间 A 与 T、G 与 C 配对(即碱基互补原则)。A 与 T 之间形成 2 个氢键,G 与 C 形成 3 个氢键,配对碱基之间的氢键维系双螺旋的横向稳定。碱基堆砌力维系双螺旋的纵向稳定,是维系 DNA 双螺旋稳定的主要因素。

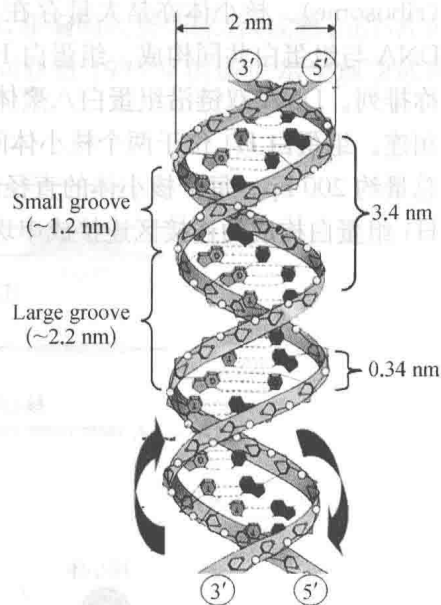


图 1-3 DNA 双螺旋结构

DNA 双螺旋结构模型的建立

1950 年前后,美国生物化学家 Erwin Chargaff 等人采用纸层析和紫外吸收等方法,对不同生物来源的 DNA 的碱基组成进行了测定分析。总结出如下规律:① 腺嘌呤的摩尔数等于胸腺嘧啶的摩尔数,鸟嘌呤的摩尔数等于胞嘧啶的摩尔数,即嘌呤碱的摩尔数等于嘧啶碱的摩尔数;② DNA 的碱基

组成有种属特异性;③ 同一生物个体的 DNA 碱基组成没有组织器官特异性。以上规律被称之为 Chargaff 规则。该规则已预示 A 与 T、G 与 C 配对的可能性。与此同时,英国物理学家 Maurice Wilkins 等人用 X 线衍射技术研究 DNA 的分子结构,女物理学家 Rosalind Franklin 拍摄到了一张十分清晰的 DNA 的 X 线衍射照片,显示 DNA 是一种双链螺旋结构。美国科学家 James Watson 和英国科学家 Francis Crick 在上述研究的基础上,于 1953 年提出了 DNA 双螺旋结构模型(图 1-3)。这个划时代的发现是生物学研究的里程碑,从此揭开了现代分子生物学的序幕。为此,J. Watson, F. Crick 和 M. Wilkins 分享了 1962 年的诺贝尔生理学或医学奖。

(二) DNA 的高级结构

DNA 的三级结构是指 DNA 双螺旋进一步盘曲所形成的构象。原核生物闭环双链 DNA 很容易缠绕盘曲成超螺旋结构(图 1-4)。若双螺旋 DNA 沿右手方向缠绕,即与双螺旋的方向相同时,造成双螺旋的过旋而形成正超螺旋(positive supercoil);若沿左手方向缠绕,即与双螺旋的方向相反时,造成双螺旋的欠旋而形成负超螺旋(negative supercoil)。真核生物的线粒体和叶绿体 DNA 也是环状双链超螺旋结构。真核生物细胞核染色体 DNA 的二级结构为线性双螺旋,当其 2 个末端被蛋白质结合固定时,也可产生超螺旋结构。细胞内 DNA 的超螺旋结构总是处于动态的变化之中。生理条件下,DNA 一般都以负超螺旋构象存在。

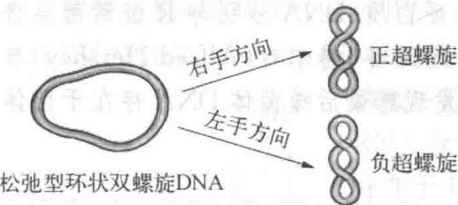


图 1-4 环状及超螺旋 DNA 示意图

真核细胞 DNA 与组蛋白组成更高级的结构——染色质。染色质的基本组成单位是核小体(ribosome)。核小体亦是大量存在的 DNA 三级结构的一种形式(图 1-5),核小体由线性双螺旋 DNA 与组蛋白共同构成。组蛋白 H2A、H2B、H3 和 H4 各两分子构成核小体八聚体核心,呈两层对称排列。DNA 双链沿组蛋白八聚体核心缠绕 1.75 圈,约含 146 bp,其中有 80 bp 紧密与八聚体核心相连。组蛋白 H1 位于两个核小体间的连接区 DNA 上,连接区 DNA 链长约 60 bp,核小体上的 DNA 总量约 200 bp。每个核小体的直径约为 11 nm,高约 5.5 nm。由此,许许多多核小体经由 DNA 和 H1 组蛋白构成的连接区连接成串珠样结构。

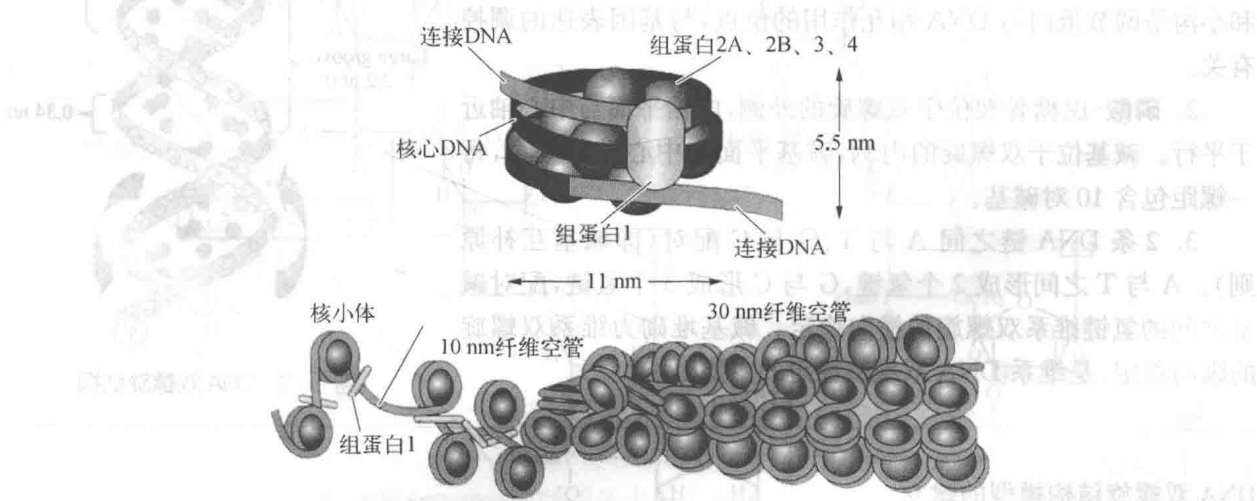


图 1-5 核小体结构及纤维空管结构示意图

核小体形成的串珠样结构再盘曲成直径为 30 nm 的中空纤维状结构——螺线管,每周含 6 个核小体。螺线管再经进一步的盘曲折叠形成直径为 200 nm 的超螺线管和襞状结构,最终形成染色质,

双链 DNA 总长度被压缩约 8 400 倍(图 1-6)。每条染色体由一条双链 DNA 和组蛋白构成。

四、基因的组构

单个基因的组成结构以及个体内的基因组组织排列方式统称为基因组 (gene organization)。基因组成结构不仅包含结构基因区序列,还包括对结构基因表达起调控作用的序列,即调控序列。基因中编码 RNA 或多肽链的 DNA 序列称为结构基因(structure gene);基因两侧或内部通常含有调控序列,如启动子(promoter)及上游启动子元件、增强子(enhancer)、沉默子(silencer)、绝缘子(insulator)、反应元件(response element)、poly(A)加尾信号、终止子(terminator)等。

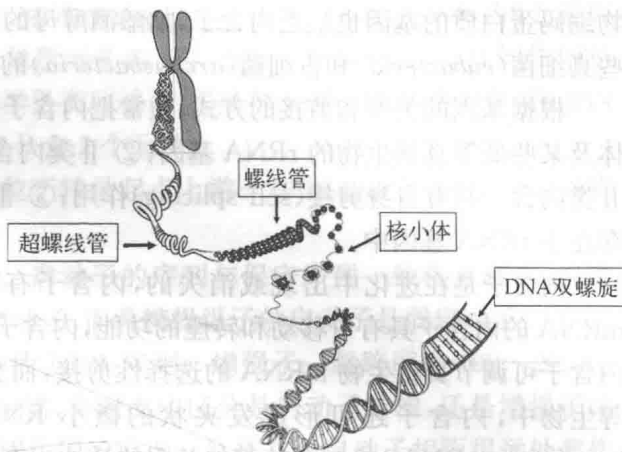


图 1-6 染色质纤维缠绕、折叠示意图

(一) 结构基因中的外显子与内含子

与原核生物相比,真核生物结构基因最突出的特点是编码序列常被一些非编码序列所间隔。20 世纪 70 年代,美国生物学家夏普(Pilip A. sharp)和英国科学家罗伯茨(Richarl J. Roberts)在利用核酸分子杂交实验研究腺病毒时发现:成熟 mRNA 与模板链 DNA 杂交,出现部分的配对(双链区段)和中间不配对(单链区段)现象。夏普认为,hnRNA 中的非编码区片段被切除,而编码区片段被拼接起来。由此,他们提出断裂基因(split gene)的概念:真核生物结构基因是由若干个编码区和非编码区相间排列但又相互连接而成,这些基因被称为断裂基因。第一个被详细研究的断裂基因是鸡的卵清蛋白基因,其全长为 7.7 kb,8 个编码区被 7 个非编码区所间隔,非编码区被切除后,成熟 mRNA 的长度仅为 1.2 kb(图 1-7)。夏普和罗伯茨因发现断裂基因而共同获得 1993 年的诺贝尔生理学或医学奖。

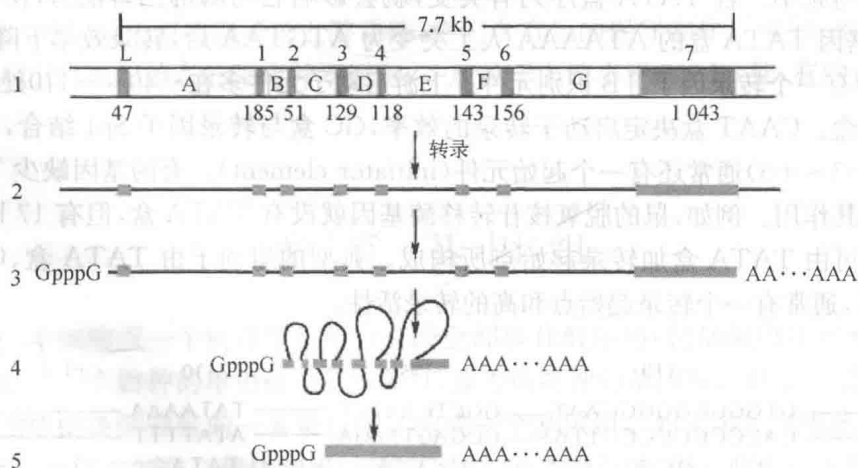


图 1-7 鸡的卵清蛋白基因及转录后加工示意图

通常,把断裂基因中的编码序列称为外显子(exon),而把非编码序列称为内含子(intron)。外显子和内含子均被转录而出现在初级转录产物上。外显子被拼接后成为成熟 mRNA 的核苷酸序列,内含子是隔断基因线性表达而在剪接过程中被除去的核苷酸序列。真核生物基因在内含子和外显子的交界处有两个相当短的保守序列:5'端为 GT,3'端为 AG,此称为 GT-AG 规则。绝大多数脊椎动物编码蛋白质的基因都含有内含子,只有为数不多的基因没有内含子,如组蛋白基因。某些低等真核生

物编码蛋白质的基因也缺乏内含子,如酿酒酵母的许多基因。有些原核生物的基因也有内含子,如一些真细菌(*eubacteria*)和古细菌(*archaebacteria*)的基因。内含子的长度可从 $50\sim 2\times 10^4$ nt。

根据基因的类型和剪接的方式,通常把内含子分为四类:① I类内含子主要存在于线粒体、叶绿体及某些低等真核生物的 rRNA 基因;② II类内含子也发现于线粒体、叶绿体,有相当一部分 I类和 II类内含子具有自身剪接(self splicing)作用;③ III类内含子见于大多数 mRNA 基因;④ IV类内含子存在于 tRNA 基因中。

内含子是在进化中出现或消失的,内含子有利于物种的进化选择。目前已有证据表明,前体 mRNA 的内含子具有可移动和转座的功能,内含子很可能参与了 RNA 介导的细胞调节功能。此外,内含子可调节真核生物 mRNA 的选择性剪接,而且还可产生有功能活性的 RNA。最近还发现在高等生物中,内含子还可形成发夹状的微小 RNA (microRNA, miRNA),利用 RNA 干扰(RNA interference, RNAi)机制调节其他基因的活性。有些内含子具有编码功能,如酵母细胞色素 b 基因的编码内含子,酵母细胞色素 b 基因含有 3 个外显子和 2 个内含子,转录成 mRNA 后首先切除了内含子 1,形成未成熟的 mRNA。此 mRNA 被翻译产生 423 个氨基酸残基的一种叫做成熟酶(maturase)的蛋白,此酶又可将未成熟的 mRNA 中的内含子 2 切除,产生成熟的 mRNA,最终以此 mRNA 为模板翻译成 279 个氨基酸残基的细胞色素 b,其 N 端的 144 个氨基酸与成熟酶 N 端的 144 氨基酸一致。另外,某些遗传性疾病的基因变异是发生在内含子而不在外显子。

(二) 结构基因的调控序列

1. 启动子和上游启动子元件 启动子是控制转录的关键结构,是转录启动时供 RNA 聚合酶识别与结合、并启动转录的一段特殊的 DNA 序列。启动子具有方向性,它决定着转录的方向。不同基因的启动子在序列上具有保守性。

根据 RNA 聚合酶的类别,真核生物启动子分为 I、II 和 III 类启动子。RNA 聚合酶 II 识别的启动子属于 II 类启动子,通常位于转录起始点上游,本身并不被转录,它包括启动子和上游启动子元件(upstream promoter element)等近端调控序列。真核生物的 TATA 盒位于 -30 区域,又称为 Hogness 盒,通常认为这是启动子的核心序列,是转录因子 II D 和 RNA 聚合酶 II 结合的区域,控制转录起始的精确性与频率。若 TATA 盒序列有突变,就会影响它与酶的结合能力,使转录效率下降。例如,兔珠蛋白基因 TATA 盒的 ATAAAA 人工突变为 ATGTAA 后,转录效率下降 80%。在靠近 TATA 盒的上游有一个转录因子 II B 识别元件。上游启动子元件多在 -40~-110 处,比较常见的是 CAAT 盒和 GC 盒。CAAT 盒决定启动子转录的效率,GC 盒与转录因子 Sp1 结合,促进转录过程。在起始点周围(-3~+5)通常还有一个起始元件(initiator element)。有的基因缺少 TATA 盒,此时起始元件可代替其作用。例如,鼠的脱氧核苷转移酶基因就没有 TATA 盒,但有 17 bp 的起始元件。最简单的启动子可由 TATA 盒加转录起始点所构成。典型的启动子由 TATA 盒、CAAT 盒和 GC 盒组成(图 1-8),通常有一个转录起始点和高的转录活性。

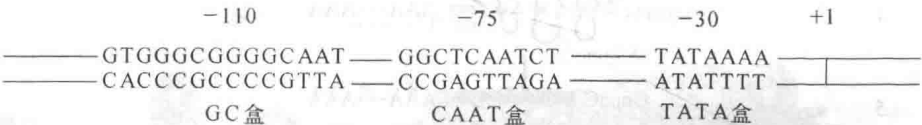


图 1-8 真核生物典型启动子的序列

不典型的启动子不含 TATA 盒,但富含 GC 序列,含多个转录起始点,一般含有数个分离的转录起始点,多见于持家基因。例如,SV40(Simian virus 40,猿猴病毒 40)早期基因没有 TATA 盒,而有多数 GC 盒串联;还有些基因的启动子既不含 TATA 盒,又没有 GC 富含区,可有一个或多个转录起始点,低或无转录活性。