

借

研究生教学用书

Medical
Molecular Biology

医学分子生物学

主编 马文丽



高等教育出版社
Higher Education Press

医学分子生物学

Medical Molecular Biology

主 编 马文丽

副主编 刘兴汉 德 伟

编 者(以姓氏笔画为序)

万福生(南昌大学医学院)

马文丽(南方医科大学)

王艾琳(北华大学医学检验学院)

王海杰(复旦大学上海医学院)

白 怀(四川大学华西第二医院)

刘 誉(暨南大学医学院)

刘兴汉(哈尔滨医科大学)

刘新光(广东医学院)

朱振宇(中山大学基础医学院)

李恩民(汕头大学医学院)

杨光彩(南方医科大学)

林雪松(哈尔滨医科大学)

林德馨(福建医科大学)

费 嘉(暨南大学医学院)

赵 红(中国医学科学院 中国协和医科大学)

梁 爽(南方医科大学)

谭玉珍(复旦大学上海医学院)

德 伟(南京医科大学)



高等教育出版社

Higher Education Press

内容提要

分子生物学的发展日新月异,其新理论、新技术正在不断渗透到生命科学的各个领域。医学分子生物学作为分子生物学的重要分支,与基础医学和临床医学紧密结合,从分子水平研究人体生物大分子的结构、功能、相互作用及其同疾病发生、发展的关系。

作为新编研究生教材,除了介绍分子生物学的基本理论之外,更要体现“更高”、“更新”、“更深”的特点。不但要深入、系统地介绍分子生物学的前沿领域和最新进展,还要注重分子生物学在生命科学领域的应用以及与基础、临床医学的结合。本书编写在组织上有系统性,在内容上有适用性和启发性,不但便于师生教与学,还力争培养研究生的科学思维能力和科研工作能力。

本教材共分三篇 15 章。第一篇为分子生物学基本原理,包括基因组与基因组学, RNA、RNA 组与 RNA 组学,蛋白质、蛋白质组和蛋白质组学,代谢组,基因表达调控和细胞信号传递 6 章分子生物学的基本理论;第二篇为分子生物学应用基础,包括分子生物学常用技术、基因工程、基因诊断与基因治疗和生物信息学在医学分子生物学中应用 4 章内容;第三篇为临床分子生物学,包括遗传性疾病的分子生物学、肿瘤分子生物学、感染性疾病的分子生物学、心血管性疾病的分子生物学和遗传性代谢病的分子生物学 5 章。

本教材主要供基础医学与临床医学研究生教学使用,也可作为研究生入学考试的参考书。

图书在版编目(CIP)数据

医学分子生物学/马文丽主编. —北京:高等教育出版社,2008.2

ISBN 978-7-04-023192-2

I. 医… II. 马… III. 医药学:分子生物学-研究生-教材 IV. Q7

中国版本图书馆 CIP 数据核字(2008)第 004725 号

策划编辑 安琪 责任编辑 孟丽 封面设计 张楠 责任绘图 尹莉
版式设计 王艳红 责任校对 刘莉 责任印制 宋克学

出版发行 高等教育出版社

社址 北京市西城区德外大街 4 号

邮政编码 100011

总机 010-58581000

经销 蓝色畅想图书发行有限公司

印刷 北京人卫印刷厂

开本 889×1194 1/16

印张 27.25

字数 850 000

购书热线 010-58581118

免费咨询 800-810-0598

网址 <http://www.hep.edu.cn>

<http://www.hep.com.cn>

网上订购 <http://www.landaco.com>

<http://www.landaco.com.cn>

畅想教育 <http://www.widedu.com>

版次 2008 年 2 月第 1 版

印次 2008 年 2 月第 1 次印刷

定价 57.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 23192-00

序 言

分子生物学是在分子水平上研究生命本质、生命活动和生命现象的科学,其理论和技术已在医学领域得到广泛应用,是医学院校研究生的一门必修课。医学院校研究生学习分子生物学既要掌握分子生物学的基础理论和基本技术,也要了解分子生物学在医学领域的应用和研究进展。

读过马文丽教授主编的《医学分子生物学》,觉得是一部比较适用的研究生教材。

首先,教材以分子生物学的基础理论为主线,对 DNA、RNA 和蛋白质等生物大分子的结构和功能,遗传信息的传递及调控等经典理论做了深入浅出的详细论述,便于研究生纵览分子生物学基本理论的全貌。

其次,教材在保持基本理论系统性、完整性的基础上,尽可能地引入了基因组学、蛋白质组学、代谢组学和生物信息学等新概念、新发展,有助于研究生掌握分子生物学发展的趋势和走向,分析分子生物学研究的新热点和新焦点,为选择研究课题提供前瞻性的线索。

第三,教材集中了分子杂交、DNA 测序、PCR、RNA 干扰、酵母双杂交等分子生物学的常用技术,强调这些实验技术的原理、用途和应用条件,为研究生在课题研究过程中更准确地选用相关技术提供理论指导。

第四,重点介绍了临床常见病如肿瘤、心血管病、遗传病、代谢病、感染性疾病的分子生物学基础,帮助研究生从分子生物学的角度加深对上述疾病发病机制的认识,指导临床诊断和治疗实践。

总之,这是一部比较适合研究生教学需要的教材。相信这部教材的出版,会在一定程度上缓解研究生教学对教材的需求,为研究生学习分子生物学理论,开展科学研究提供帮助。因此很高兴能为这部教材作序并向同行和研究生推荐。

中国科学院院士
2007. 10. 27.

施 肇 基

前 言

分子生物学(molecular biology)是在分子水平上研究生命现象的科学。医学分子生物学作为分子生物学的分支,运用分子生物学的理论、技术与方法研究人体在正常与疾病状态下的生命现象和活动的规律。

在 Astbury W 于 1950 年首次提出 Molecular Biology 后的 50 余年中,分子生物学取得了长足进步,成为生命科学中发展最快、影响最大的学科。1953 年 Watson J D 和 Crick F H C 发现并提出 DNA 双螺旋结构,1958 年 Crick F H C 提出生物学的中心法则,奠定了分子生物学的经典理论基础。1973 年 Cohen S N 等建立 DNA 重组技术,1985 年 Mullis K 发明 PCR 技术。这些在分子生物学理论指导下陆续诞生的生物技术反过来又极大地推动了分子生物学理论的发展。2003 年人类基因组序列图的提前绘制完成更将分子生物学推进到后基因组时代,功能基因组学、蛋白质组学的研究成果结合生物信息学分析实现了大规模、高通量地筛查、检验致病基因,改变了传统的思维方法和固有的研究模式。分子生物学发展的历程和取得的成就使人们深信,分子生物学已经成为 21 世纪生命科学中的带头学科,生物技术已经成为医学研究中的主导技术,医学分子生物学已经成为医学院校研究生的必修课。

目前国内出版的《医学分子生物学》研究生教材,或因强调“系统”与“完整”,随着新进展的不断涌现,越编越厚,成为“鸿篇巨制”,受教学时数限制,大部分内容不讲授,学生也没有精力通读;或因突出“配套”和“衔接”,将完整的分子生物学内容分散在不同的配套教材当中,给教学活动和研究生学习带来不便。编写一部简明地阐述分子生物学基本理论、基础知识,又尽量反映前沿领域研究进展,篇幅适当,一册在手可以尽览分子生物学全貌,适合多数医学院校研究生教学需要的《医学分子生物学》教材,是教学的迫切需要,也是师生的共同愿望。有鉴于此,高等教育出版社组织国内 13 家高等医学院校的教师尝试编写了这本《医学分子生物学》研究生用教材。

本教材共分三篇 15 章。

第一篇为分子生物学基本理论,介绍基因组学、RNA 组学、蛋白质组学、代谢组学、基因表达调控和细胞信号传递。本教材尝试将生物大分子结构、功能和生物合成结合叙述,以压缩章节和篇幅;在突出基本理论、基本知识、基本概念的前提下,适当引进基因组学、蛋白质组学、代谢组学等新发现、新进展,让研究生知道这些知识的生长点,为研究课题提供前瞻性线索。

第二篇为分子生物学的应用基础,介绍分子生物学常用技术、基因工程、基因诊断与基因治疗和生物信息学。重点讲述这些技术的原理、需要的条件、主要用途和存在的不足。同时适当地介绍相关技术的发展、改进,让研究生从中体会发明者、改进者的科研思路,培养研究生的科学思维方法和解决问题的能力。

第三篇为临床分子生物学,重点介绍遗传性疾病、肿瘤、感染性疾病、心血管性疾病和代谢性疾病的分子生物学知识,引导研究生用分子生物学的理论认识这些疾病的发病机制,指导临床预防与治疗实践。

本教材主要供基础医学与临床医学研究生教学使用,也可作为研究生入学考试的参考书。

本教材在编写过程中得到了高等教育出版社领导的关怀和南方医科大学各级领导的大力支持,各位编委通力合作,南方医科大学的朱丽娜老师参与了大量的编务组织工作,在此一并致以衷心的感谢。

医学分子生物学涉及的知识面广,发展迅速,不断涌现新概念、新成果,编写一部能被国内大多数院校共同接受的研究生教材,是一种尝试和挑战。编者虽努力想使本教材在内容的“广度”、“新度”和“深度”上都能符合多数院校研究生教学需求,但受编者知识结构、专业水平和写作能力的限制,仍会有内容取舍失当和疏漏、错误之处,务请读者批评指正,以便在今后的再版中改进。

编者

2007 年 5 月

郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 58581897/58581896/58581879

传 真：(010) 82086060

E - mail: dd@hep.com.cn

通信地址：北京市西城区德外大街4号

高等教育出版社打击盗版办公室

邮 编：100011

购书请拨打电话：(010)58581118

(警 告)



目 录

第一篇 分子生物学基本原理

第一章 基因、基因组与基因组学	3	思考题	56
第一节 基因的结构与功能	3	主要参考文献	57
一、基因概念的发展	3	第三章 蛋白质、蛋白质组与蛋白质组学	58
二、基因的现代概念	4	第一节 蛋白质的结构与功能	58
第二节 基因组的结构与功能	6	一、蛋白质分子的化学组成	58
一、原核生物基因组	6	二、蛋白质分子的结构	61
二、真核生物基因组	7	三、蛋白质分子结构与功能的关系	65
三、病毒基因组	11	四、蛋白质的理化性质	66
四、逆转录病毒基因组	12	五、蛋白质的分离纯化	67
第三节 基因组学	13	六、蛋白质分子序列及空间结构分析	69
一、基因组学研究内容	13	第二节 蛋白质组与蛋白质组学	69
二、人类基因组计划及其进展	15	一、概述	69
第四节 基因组的复制	16	二、差异蛋白质组学的研究方法及应用	70
一、各种基因组复制的共同机制和不同特点	16	三、蛋白质组学的前景	72
二、原核生物基因组 DNA 复制的特点	17	第三节 蛋白质的生物合成及其干扰	72
三、真核生物基因组 DNA 复制的特点	17	一、参与蛋白质生物合成的物质	72
思考题	19	二、蛋白质的生物合成过程	75
主要参考文献	19	三、蛋白质生物合成的干扰	81
第二章 RNA、RNA 组与 RNA 组学	20	第四节 蛋白质合成后加工	82
第一节 RNA 的结构与功能	20	一、蛋白质分子一级结构的修饰	82
一、RNA 组与 RNA 组学	20	二、蛋白质分子的正确折叠	83
二、RNA 的种类及其功能	20	三、蛋白质分子的化学修饰	84
三、mRNA 的结构与功能	23	四、蛋白质分子的靶向输送	84
四、tRNA 的结构与功能	24	五、蛋白质在细胞内的降解	85
五、rRNA 的结构与功能	27	思考题	86
六、其他小分子 RNA	28	主要参考文献	86
第二节 RNA 的生物合成(转录)	29	第四章 代谢组与代谢组学	87
一、转录的模板和酶	29	第一节 代谢组学的概念	87
二、转录过程	34	一、历史背景	87
第三节 真核生物 RNA 的转录后加工	46	二、代谢组学的概念	87
一、真核生物 mRNA 前体的转录后加工	46	第二节 系统生物学时代的代谢组学	87
二、真核生物 tRNA 前体的转录后加工	52	一、代谢组学与基因组学、转录组学和蛋白质组学的区别	88
三、真核生物 rRNA 前体的转录后加工	54		
四、核酶	54		

二、金属组学	89
第三节 代谢组学的研究方法和分析技术	89
一、代谢组学的研究方法	89
二、代谢组学的分析技术	90
第四节 代谢组学在医药研究中的应用	92
一、在功能基因组研究中的应用	92
二、在疾病诊断研究中的应用	93
三、在药物研发中的应用	93
四、在中医药研究中的应用	93
思考题	94
主要参考文献	94
第五章 基因表达调控	96
第一节 基因表达调控的基本原理	96
一、基因表达调控的分子基础	96
二、基因表达调控的基本方式	97
三、基因表达调控的基本规律	97
第二节 原核生物的基因表达调控	98
一、细菌操纵子的基因表达调控	98
二、Lambda 噬菌体的基因表达调控	103
三、翻译水平的基因表达调控	105
第三节 真核生物的基因表达调控	107
一、染色质水平基因表达调控	108
二、转录水平基因表达调控	109
三、转录后基因表达调控	113
思考题	115
主要参考文献	116
第六章 细胞信号传递	117
第一节 细胞间的信息分子	117
一、细胞间联系的方式	117
二、细胞间的信息分子	117

三、化学信息分子的性质及其与靶细胞的 作用	118
四、细胞信号系统	118
第二节 受体——细胞识别与结合信息分子的蛋 白质	119
一、受体的概念及分类	119
二、受体作用的特点	120
第三节 G 蛋白——细胞跨膜信息传递中关键性膜 结合蛋白	120
一、G 蛋白的结构、特性及分类	120
二、G 蛋白跨膜传递信息的机制	121
三、G 蛋白在信息传递中的作用	122
四、G 蛋白信号传递调节蛋白	123
五、小分子 G 蛋白	124
第四节 细胞内主要信号传递途径及传递的分子 机制	124
一、cAMP 信号系统	124
二、cGMP 信号系统	126
三、IP ₃ 、DG 信号系统	128
四、Ca ²⁺ 信号系统	129
五、受体蛋白酪氨酸激酶信号传递	132
六、细胞因子受体的信号传递	138
七、核受体信号传递	139
八、NF- κ B 信号传递	142
九、整联蛋白介导的细胞信号传递	144
第五节 细胞信号传递与疾病	147
一、G 蛋白异常与疾病	147
二、核受体病	147
思考题	149
主要参考文献	150

第二篇 分子生物学应用基础

第七章 分子生物学常用技术	153
第一节 核酸分子杂交	153
一、核酸探针的种类	153
二、核酸标记物及其选择	154
三、核酸探针的标记	155
四、杂交信号的检测	158
五、核酸分子杂交的类型	158
六、核酸分子杂交实验因素的优化	161
第二节 PCR 技术	163
一、PCR 的基本原理	163

二、PCR 技术的主要特点	163
三、PCR 的反应体系	164
四、PCR 引物设计	166
五、PCR 的反应过程及条件优化	167
六、PCR 衍生技术	168
七、PCR 技术的主要应用	172
第三节 DNA 序列测定	173
一、DNA 测序的基本原理	173
二、DNA 自动测序	176
第四节 生物芯片技术	178

一、生物芯片的概念	178	一、疾病相关基因的分析	224
二、基因芯片技术的基本原理与方法	179	二、对特定基因进行定点诱变	225
三、基因芯片技术的应用	181	三、基因工程医药产品	225
第五节 基因沉默、RNA 干扰与基因剔除技术	183	四、基因诊断与基因治疗	225
一、基因沉默	183	思考题	226
二、RNA 干扰技术	186	主要参考文献	227
三、基因剔除技术	188	第九章 基因诊断与基因治疗	228
第六节 酵母双杂交技术	190	第一节 基因诊断	228
一、酵母双杂交系统的原理	190	一、基因诊断是检测基因结构和表达的异常及病原体基因	228
二、酵母双杂交系统的操作程序	191	二、经典遗传学理论和疾病的分子机制是基因诊断的理论依据	229
三、酵母双杂交系统的应用	191	三、分子生物学技术的进步为基因诊断提供了技术平台	233
四、酵母双杂交系统的局限与发展	192	四、依据检测目的和被测基因性质选择基因诊断方法	237
第七节 免疫印迹技术	193	五、基因诊断在遗传病、肿瘤和感染性疾病检测中的应用	240
一、免疫印迹的基本原理	193	六、DNA 指纹用于法医和亲子鉴定	243
二、蛋白质印迹的基本方式	194	七、基因诊断中生物医学信息网络资源的利用	243
三、抗体探针的标记	195	第二节 基因治疗	244
四、免疫印迹技术的主要用途及发展	195	一、基因治疗是在核酸水平上开展疾病治疗	244
第八节 蛋白质双向凝胶电泳技术	196	二、目的基因导入靶细胞表达治疗疾病	245
一、双向凝胶电泳的原理	196	三、基因干预技术抑制有害基因表达或失控基因过表达治疗疾病	249
二、双向凝胶电泳技术操作的基本过程	197	四、基因治疗用于单基因遗传病、肿瘤和病毒性疾病	250
三、双向凝胶电泳过程中需要注意的问题	197	思考题	254
四、双向凝胶电泳技术在医学中的应用	199	主要参考文献	254
思考题	200	第十章 生物信息学在医学分子生物学中的应用	256
主要参考文献	200	第一节 生物信息学简介	256
第八章 基因工程	202	一、历史背景	256
第一节 基因工程的基本原理	202	二、研究范围	258
一、基因工程技术的诞生及意义	202	三、生物数据库	262
二、基因工程技术的原理	203	第二节 生物序列分析	267
第二节 基因工程的物质基础	203	一、序列分析	267
一、工具酶	203	二、多序列比对	273
二、基因载体	207	三、分子系统树	276
第三节 基因克隆的基本过程	214	四、比较基因组	279
一、目的基因的获取	214	五、多态性分析	282
二、基因载体的选择与制备	215	六、基因功能注释	283
三、DNA 重组体的构建	215		
四、重组 DNA 导入受体细胞	217		
五、筛选与鉴定	218		
第四节 克隆基因的表达	219		
一、克隆基因在原核细胞中的表达	219		
二、克隆基因在真核细胞中的表达	221		
三、克隆基因在其他表达系统中的表达	223		
四、表达产物的分离纯化	224		
第五节 基因工程技术与医学的发展	224		

第三节 基因表达分析	284
一、生物芯片	284
二、基因表达谱分析	284
三、EST 分析	286
四、转录调控研究	289
五、RNA 剪切	290
六、小 RNA 调控	291
第四节 蛋白质结构分析	294
一、结构分析与预测	294
二、蛋白质-蛋白质相互作用(PPI)	295

三、蛋白质相似性与进化	296
四、蛋白质三级结构预测	297
第五节 生物信息与医药应用	298
一、生物靶标	299
二、通路和网络	300
三、生物信息综合分析应用	301
四、药物开发	302
思考题	303
主要参考文献	304

第三篇 临床分子生物学

第十一章 遗传性疾病的分子生物学	309
第一节 遗传性疾病	309
一、遗传性疾病的定义与特征	309
二、遗传性疾病与环境	310
三、遗传性疾病的分类	311
第二节 单基因遗传病	311
一、单基因病的遗传方式	311
二、地中海贫血	313
第三节 多基因遗传病	315
一、多基因病的遗传方式	315
二、阿尔茨海默病	316
第四节 染色体病	317
一、染色体数目畸变	317
二、染色体结构畸变	317
三、几种染色体病	318
思考题	321
主要参考文献	321

相关基因之间的关系	343
第三节 肿瘤细胞侵袭与转移的分子生物学	344
一、肿瘤细胞侵袭与转移的分子机制	345
二、肿瘤细胞侵袭与转移的基因调控	347
三、肿瘤侵袭与转移的分子免疫学基础	349
第四节 肿瘤血管与淋巴管新生的分子生物学	349
一、肿瘤血管新生	350
二、肿瘤淋巴管新生	351
思考题	353
主要参考文献	353

第十二章 肿瘤分子生物学	322
第一节 癌基因与抑癌基因	322
一、癌基因的发现与概念演变	322
二、癌基因的分类	323
三、抑癌基因的发现	325
四、抑癌基因的分类	326
五、癌基因与抑癌基因的功能	326
六、癌基因激活的分子机制	336
七、抑癌基因失活的分子机制	339
第二节 肿瘤细胞增殖与凋亡的分子生物学	340
一、肿瘤细胞增殖相关基因	340
二、肿瘤细胞凋亡相关基因	341
三、肿瘤细胞增殖相关基因与肿瘤细胞凋亡	

第十三章 感染性疾病的分子生物学	355
第一节 病毒致病的分子生物学机制	355
一、病毒的定义与分类	355
二、病毒的基因组	357
三、病毒侵入宿主细胞的分子生物学机制	359
四、病毒基因组进入宿主细胞核的分子生物学机制	365
五、病毒致瘤的分子生物学机制	366
第二节 细菌致病的分子生物学机制	371
一、胞外菌感染的分子生物学机制	372
二、胞内菌感染的分子生物学机制	374
第三节 真菌致病的分子生物学机制	377
一、宿主上皮细胞或细胞外基质黏附分子	377
二、白假丝酵母菌的表面黏附结构	377
三、白假丝酵母菌黏附相关基因	377
四、胞外酶	378
第四节 感染性疾病的分子生物学诊断技术	378
一、细菌 DNA(G+C)摩尔分数测定	378
二、细菌质粒指纹图谱分析	378
三、核酸杂交技术	379

四、PCR 技术	379
五、生物芯片技术	379
思考题	379
主要参考文献	380
第十四章 心血管疾病的分子生物学	381
第一节 原发性高血压的分子生物学	381
一、血压的调节机制及其影响因素	381
二、原发性高血压的遗传学基础	382
三、原发性高血压发生与发展的分子机制	387
四、原发性高血压的基因治疗	388
第二节 动脉粥样硬化的分子生物学	389
一、脂质代谢异常学说	389
二、受体缺陷学说	390
三、慢性炎症学说	391
四、致突变学说(单克隆学说)	392
五、平滑肌在动脉粥样硬化中的作用	392
第三节 扩张型心肌病的分子生物学	393
一、分子遗传学机制	393
二、自身免疫机制	395
三、凋亡机制	396
四、分子生物学治疗	397
第四节 心力衰竭的分子生物学	397
一、心力衰竭的基本概念	397
二、心力衰竭的经典信号传导通路	398
三、心力衰竭的心肌细胞内在分子机制	399

四、心力衰竭的基因治疗	400
思考题	402
主要参考文献	402
第十五章 遗传性代谢病的分子生物学	404
第一节 代谢病的概念	404
第二节 代谢病的分类	404
一、遗传性和获得性代谢病	404
二、小分子代谢病和大分子代谢病	406
第三节 遗传性代谢病的分子机制	407
一、代谢物转运缺陷	407
二、物质代谢中的反应通路缺陷	408
三、代谢调节缺陷	409
第四节 糖尿病	410
一、1 型糖尿病的分子机制	411
二、2 型糖尿病的分子机制	414
三、年轻起病成年型糖尿病的分子机制	416
第五节 溶酶体贮积症	418
一、溶酶体贮积症分类	418
二、糖原贮积症 II 型的分子机制	420
三、唾液酸贮积症的分子机制	421
四、黏多糖病的分子机制	422
思考题	424
主要参考文献	424

第一篇

分子生物学基本原理

第一章 基因、基因组与基因组学

基因的概念是在 19 世纪由遗传学家提出来的,对其化学本质及功能的真正了解是在 20 世纪 40 年代以后。基因是指携带有遗传信息的 DNA 或 RNA 序列,也称为遗传因子,是控制性状的基本遗传单位。基因通过指导蛋白质的合成来表达自己所携带的遗传信息,从而控制生物个体的性状表现。基因有两个特点,一是能忠实地复制自己,以保持生物的基本特征;二是能够突变,绝大多数突变会导致疾病,一小部分突变不致病。非致病突变给自然选择带来了原始材料,使生物可以在自然选择中被选择出最适合自然的个体。绝大部分 RNA 和蛋白质的结构信息都是以基因的形式贮存在 DNA 中的,但除此之外,DNA 中还有大量的并不编码 RNA 或蛋白质的序列,这些序列中同样存在着大量的重要信息。含有一种生物的一整套遗传信息的遗传物质,称为基因组。但是人们经过不懈的努力,渴望解开生命之谜这个多年的愿望并未向前推进多少,以往研究的失败教训使人们开始清醒地认识到,仅在单一学科如细胞学、肿瘤学、人类遗传学或分子生物学领域内研究,都太局限,难以完成人类对自身的认识和保护。在绕了一大段弯路后,人们回过头来决定开始进行人的所有基因即基因组的研究,全面探讨这个人体奥秘,由此形成了基因组学(genomics)和人类基因组计划(human genome project,HGP),其最终目的是对生命进行系统和科学地解码,以此达到了了解和认识生命的起源、种间和个体间存在差异的起因、疾病产生的机制以及长寿与衰老等生命现象。

第一节 基因的结构与功能

基因的概念随着遗传学、分子生物学、生物化学等领域的发展而不断完善。从分子生物学角度看,基因是负载特定生物遗传信息的 DNA 分子片段(有的生物为 RNA),在一定条件下能够表达这种遗传信息,产生特定的生理功能。基因包括合成一个有功能的多肽或 RNA 分子所必需的整个核苷酸序列,即除了蛋白质或 RNA 的编码区外,还包括为获得一个特定转录产物所必需的其他 DNA 序列,如转录控制区,3'端剪切信号和多聚腺苷酸以及初始 RNA 转录物剪接有关的非编码序列。人类只有 5% 的 DNA 序列编码蛋白质、功能 RNA 和调节单位。剩余序列都是间隔 DNA(spacer DNA)和内含子(intron)。人类每个染色体包含单个的分子 DNA 都长达 280 Mb,由组蛋白和非组蛋白逐级压缩。较小的 DNA 分子集中在线粒体和叶绿体。如图 1-1 所示。

一、基因概念的发展

“基因”(gene)一词,是 1909 年丹麦生物学家 Johannsen W 根据希腊文“给予生命”之意创造的,并用“基因”一词代替了 1866 年 Mendel 发表的“植物杂交实验”一文中所用的“遗传因子”这个术语。从 20 世纪一开始,基因研究就一直成为遗传学发展的主线。只是当时的“基因”是遗传性状的符号,并未涉及基因的物质概念。1926 年 Morgan 发表了“基因论”,指出基因是在特定的染色体上,而且是呈直线排列在染色体上的遗传颗粒,位于同源染色体的同一位置的相对基因叫做等位基因。基因是世代相传的,基因决定遗传性状表达。

20 世纪 40 年代,Bendle 和 Tatum 用 X 射线照射链霉菌使其产生不同的突变株,发现突变可以影响代谢反应,故认为突变可致催化代谢的酶发生缺陷,而提出了“一个基因,一个酶”的学说。20 世纪 50 年代初,Benzer 用“顺反子”进一步阐明了基因的概念。Benzer 在研究 T4 噬菌体的一群紧密连锁的突变群时发现了顺反效应,从而导出了“顺反子”概念。两个突变点位于两条染色体上,称反式排列;而两个突变点位于同一

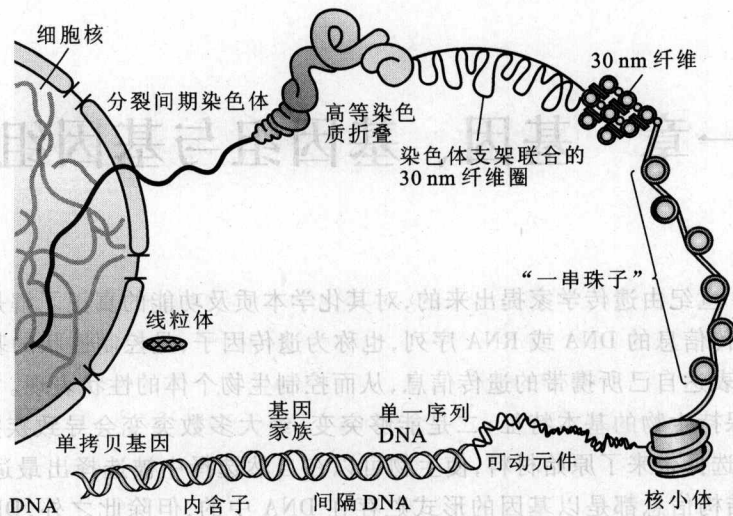


图 1-1 DNA 与染色体

染色体上,这种排列方式称为顺式排列,进而提出了“一个顺反子,一条多肽链”的概念。故现在很多分子生物学的书中仍常用“单顺反子(monocistron)”、“多顺反子(polycistron)”这种名词。但两者不能完全等同,生物学基因的“顺反子”概念只是解释了很多生物现象。随着分子生物学的进展,特别是针对真核生物的基因特征,基因还应有分子生物学的概念。20世纪60年代,遗传密码的破译使人们对基因表达的机理有了更多的了解,将基因的定义修改为:基因是基因组中的1个区域或1段DNA序列,其转录产物编码1条多肽链或者1个结构RNA分子(tRNA或rRNA)。

1994年Alberts提出,基因是一段DNA序列,包括完整的功能单位(编码序列、调节序列和内含子等);基因可以作为1个转录单位,其表达产物通常是1条多肽链或1个RNA分子,但有时编码1组相关的蛋白异形体,有些与蛋白异形体的产生和特殊的转录后加工(如RNA编辑)或者翻译水平的再编码(如核糖体跳跃)有关。

二、基因的现代概念

基因的生物学概念:基因是世代相传的,基因决定了遗传性状的表达,基因的颗粒性主要表现在世代相传的行为和功能表达上具有相对的独立性,基因呈直线排列在染色体上。

基因的分子生物学概念:合成有功能的蛋白质或RNA所必需的全部DNA(除部分病毒RNA外),即一个基因不仅包括编码蛋白质或RNA的核酸序列,还应包括为保证转录所必需的调控序列。

(一) 基因的分类

基因按其功能可分为结构基因和调控基因。

(1) **结构基因** 可被转录形成mRNA,并转译成多肽链,构成各种结构蛋白质、催化各种生化反应的酶和激素等。

(2) **调控基因** 指某些可调节控制结构基因表达的基因。其突变可影响一个或多个结构基因的功能,或导致一个或多个蛋白质(或酶)量的改变。

此外还有一些只转录而不翻译的基因,如核糖体RNA基因和小分子RNA。

(二) 基因的结构

在DNA链上,由蛋白质合成的起始密码子开始,到终止密码子为止的一个连续编码序列,叫做一个可译框(open reading frame, ORF)。人类基因的结构包括编码区、前导区和调节区3个区域。

(1) **编码区** 包括外显子与内含子。

(2) **前导区** 位于编码区上游,相当于mRNA 5'端非编码区(非翻译区)。

(3) 调节区 包括启动子和增强子等基因编码区的两侧序列也称为侧翼序列。

大量实验证明,从病毒到高等生物的细胞核均共用一套遗传密码,从分子水平上为生物同一起源的进化理论提供了有力的证据,也为基因工程提供了可能性。但近年来对人、牛、酵母的线粒体和植物叶绿体基因序列结构的研究发现,其密码子有异于细胞核的遗传密码;AUU、AUG、AUA 都为起始密码子,AUA 也可作为甲硫氨酸起始密码子,UGA 为色氨酸密码子,AGG、AGA 为终止密码子,CUA 为苏氨酸密码子。对于线粒体和叶绿体的这种遗传密码子例外的生物学意义目前尚不清楚。

DNA 的两条核苷酸链都可以作为氨基酸的密码子,不同的蛋白质利用不同的链和不同的 DNA 片段作为它们的基因。但任一链上的 ORF 方向总是由 5'→3',与肽链的 N 端到 C 端的方向是一致的。此外,按照 3 个核苷酸决定一个氨基酸的理论,在序列中的同一个核苷酸有编码 3 种 ORF 的可能性,即密码子有可能重复利用。这种重复性会使 ORF 的信息容量和利用率变得更高。这种重复利用 ORF 主要见于病毒等小型 DNA 中,使不大的 DNA 分子能编码更多的蛋白质。在高等生物中则很少见到重复利用 ORF 的情况。有时一段互补的双链可分别作为 2 个基因的编码序列,如 *ear-1* 和 *ear-7* 基因都位于 17 号染色体,拥有重叠的外显子,但分别从互补链反向转录,其他一些人类基因也可能存在较长的反义重叠 ORF。可能在生物进化的早期,DNA 双链均具有编码能力,而现在只能观察到这种现象的遗迹。

基因是核酸分子中贮存遗传信息的遗传单位,是指贮存有功能的蛋白质或 RNA 序列信息及表达这些信息所必需的全部核苷酸序列。按照这个定义,一个基因不仅仅包括编码蛋白质或 RNA 的核酸序列,还包括保证转录所必需的调控序列、编码区 5' 端上游和 3' 端下游的非编码序列、内含子。从简单的病毒到复杂的高等动植物细胞,RNA 和蛋白质的结构信息都是以基因的形式贮存在 DNA(部分病毒是 RNA)中的。

(三) 基因的功能

基因的功能包括:生物学功能,如作为蛋白质激酶对特异蛋白质进行磷酸化修饰;细胞学功能,如参与细胞间和细胞内信号传递。但以往技术,如 Western 印迹、酶联免疫吸附测定(ELISA)不能对基因进行全面系

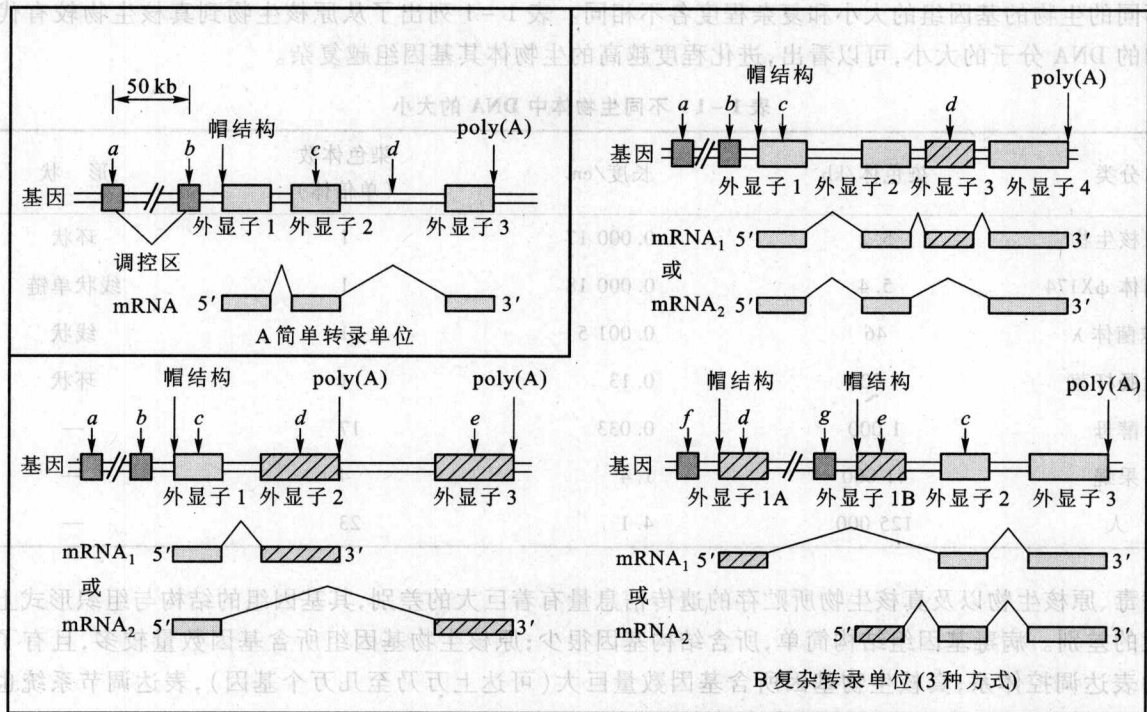


图 1-2 简单的转录单位包括编码蛋白质的区域(从 5' 帽子结构到 3' 聚腺苷酸尾)及其调控区。内含子位于外显子当中,在初级转录产物中被切除(虚线)。转录控制区域(a、b)的突变可以减少或者阻止转录,外显子(c)的突变会产生低功能的异常蛋白。内含子(d)的突变产生了新的剪接体。复杂的转录单位产生初级转录物,可以被多种形式加工处理。如图 B 所示,如果初级转录物包含了两种剪接位点,或者初级转录物包含两个聚腺苷酸尾等,就会产生两种以上转录本。

统的分析。新的技术应运而生,包括基因表达的系统分析技术、cDNA 微阵列、基因芯片等。鉴定基因功能最有效的方法是观察基因表达被阻断或增加后在细胞和整体水平所产生的表型变异,因此需要建立模式生物。

细胞核中的基因在细胞的一生中并非始终处于活性状态,它们有的处于转录状态,即活性状态,这时基因打开;有的处于非转录状态,即基因关闭。在生物体的不同发育期,基因的活性是不同的,而且基因的活性有严格的程序。基因活性的严格程序是生命周期稳定的基础。在真核生物中存在转录单位,转录单位由编码区和调控区组成,可分为简单的和复杂的转录单位,如图 1-2 所示。

生物体的一切遗传性状都受基因控制,但是基因并不等于性状,从基因型到表型(性状)要经过一系列的发育过程。基因控制生物的性状主要通过两条途径,一是通过控制酶的合成来控制生物的性状。这是因为由基因控制的生物性状要表现出来,必须经过一系列的代谢过程,而代谢过程的每一步都离不开酶的催化,所以基因是通过控制酶的合成来控制代谢过程,从而控制生物个体性状的表现的。另一条途径是基因通过控制结构蛋白的成分直接控制生物的形状。蛋白质多肽链上氨基酸序列都受基因的控制,如果控制蛋白质的基因中 DNA 的碱基发生变化,则可引起 mRNA 上相应的碱基的变化,从而导致蛋白质的结构变异。

此外,遗传性状的表现,不但要受到内部基因的控制,还受到外部条件的制约。因此,不同基因型的个体在不同的环境条件下可以产生不同的表型,即使同一基因型的个体,在不同环境条件下,也可以产生不同的表型。也就是说,表型是基因型与环境共同作用的结果。

第二节 基因组的结构与功能

细胞或生物体中,一套完整单倍体的遗传物质的总和称为基因组(genome),如人类基因组包含 23 条(单倍体)染色体上的遗传物质。基因组的结构主要指不同的基因功能区域在核酸分子中的分布和排列情况,基因组的功能是贮存和表达遗传信息。

不同的生物的基因组的大小和复杂程度各不相同。表 1-1 列出了从原核生物到真核生物较有代表性的生物的 DNA 分子的大小,可以看出,进化程度越高的生物体其基因组越复杂。

表 1-1 不同生物体中 DNA 的大小

分类	染色体/kb	长度/cm	染色体数 (单倍体)	形状
原核生物	5.2	0.000 17	1	环状
噬菌体 ϕ X174	5.4	0.000 18	1	线状单链
噬菌体 λ	46	0.001 5	1	线状
大肠杆菌	4 000	0.13	1	环状
酵母	1 000	0.033	17	—
果蝇	41 000	1.4	4	—
人	125 000	4.1	23	—

病毒、原核生物以及真核生物所贮存的遗传信息量有着巨大的差别,其基因组的结构与组织形式上也有着巨大的差别。病毒基因组结构简单,所含结构基因很少;原核生物基因组所含基因数量较多,且有了较为完善的表达调控体系;真核生物基因所含基因数量巨大(可达上万乃至几万个基因),表达调节系统也更为精细。不同的基因组虽然差别巨大,却仍有相似之处。

一、原核生物基因组

原核生物基因组的结构与功能的特点简介如下。