

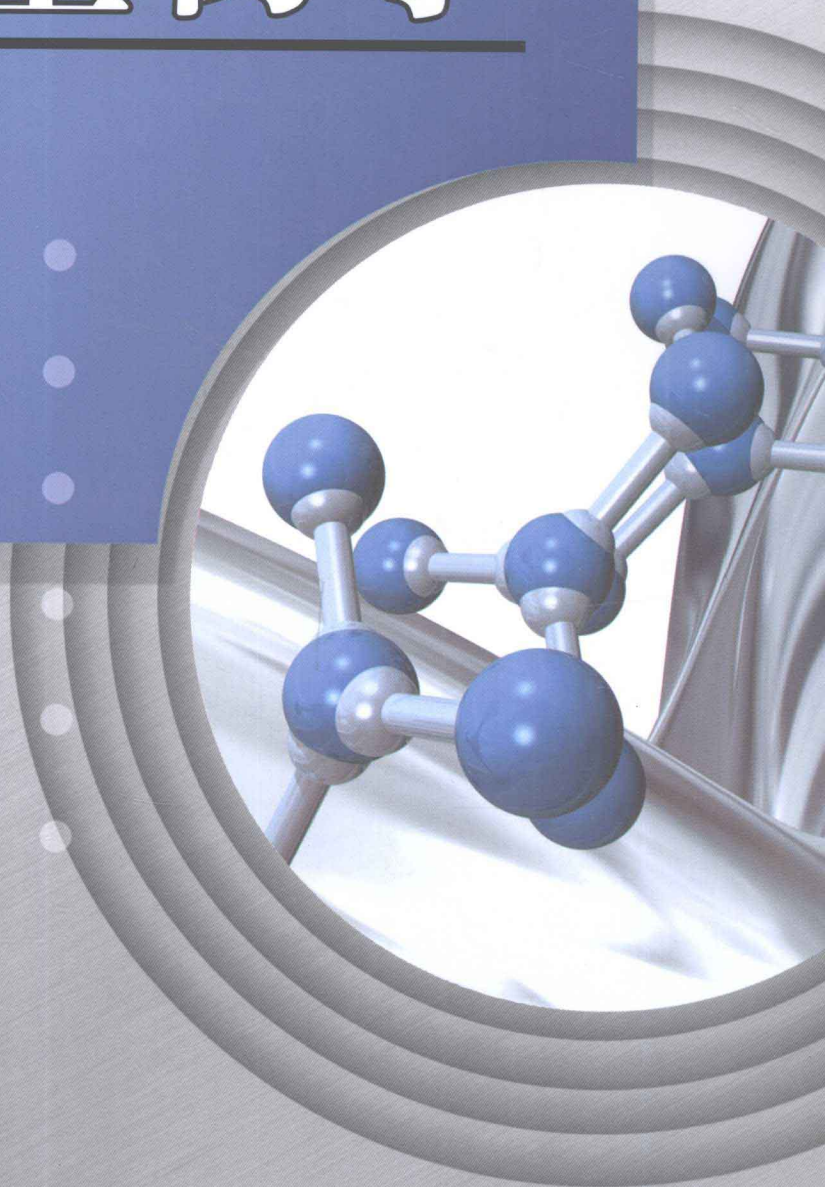


生物科学专业 **6+x** 简明教程系列

MOLECULAR BIOLOGY

分子生物学

蒋继志 王金胜◎主编



科学出版社

生物科学专业“6+X”简明教程系列

分子生物学

蒋继志 王金胜 主编

科学出版社

北 京

内 容 简 介

分子生物学是揭示生物大分子结构与功能的学科,既是当前生命科学领域中最重要基础学科和分支学科,也是该领域中属于最前沿的领军学科,正在强有力地推动着生物学及其相关领域许多学科的飞速发展。本书从构成生物体的大分子开始,沿着遗传信息流向的中心法则,阐述生物大分子及其在复制、修复、重组、转录、翻译、信号转导、基因表达与调控等主要环节中的相互作用与功能。编写过程中着重以简洁精确的方式阐述分子生物学的基本概念和基本理论知识,又注重反映分子生物学的发展趋势。全书共分12章,主要包括生物大分子,基因与基因组,DNA的复制,DNA的损伤、修复与突变,转录,翻译,原核生物基因的表达调控,真核生物基因的表达调控,基因重组、基因工程和分子生物学常用的研究技术。通过本书的学习,力求使学生能了解分子生物学的概况和基本思路,掌握其中的基本概念和 basic 研究方法与技术。

本书可作为普通高等院校生物学类专业包括师范类、普通生物类、医学生物类及农林相关专业的本科生的分子生物学教材,亦可作为生物学类专业研究生和教师的分子生物学或相关课程的教学及科研参考书。

图书在版编目(CIP)数据

分子生物学/蒋继志,王金胜主编. —北京:科学出版社,2011
生物科学专业“6+X”简明教程系列
ISBN 978-7-03-032472-6

I. ①分… II. ①蒋…②王… III. ①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(2011)第 201723 号

责任编辑:席 慧 王国栋 李晶晶 / 责任校对:李 影
责任印制:张克忠 / 封面设计:迷底书装

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>

北京天时彩色印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2011年12月第 一 版 开本:787×1092 1/16

2011年12月第一次印刷 印张:17 1/4

印数:1—4 000

字数:430 000

定价:40.00 元

(如有印装质量问题,我社负责调换)

前 言

分子生物学是研究生物大分子的结构特征和功能及其规律的科学,即在分子水平上揭开生物世界神秘的面纱,阐述生物体在生长、发育、分化过程中各种生物大分子的相互作用、细胞信息传导、基因表达及其调控的机制,从而使人类由被动地响应自然界逐步转向主动地顺应、改造和利用自然界的科学。自然界中从极其微小的病毒、细菌到最高等的动物,其生命活动形形色色、千差万别,但从分子生物学的视角来看,这些生命活动却显示出了十分惊人的相似性,似乎有着共同的规律,那就是一切生命活动都是由生物大分子来实现的。但生命活动是生物体协调各细胞、各组成部分活动的集中体现,并非组成生物体的大分子或各部分的简单叠加,是由这些大分子以及由大分子组成的生物体的各部分通过相互作用来实现的。其中构成生命活动最重要的生物大分子无疑是蛋白质和核酸,对这两类生物大分子的研究最早,也最深入,已分别形成了“基因组与基因组学”、“蛋白质组学”和“后基因组学”;而对另一类生物大分子如糖类的研究开始得较晚,但近年来正在兴起的“糖生物学”,向人们展示了糖及其复合物在生命活动中所担负的极为重要的信息功能,对于构成生物体的细胞和细胞之间的识别、细胞与底物之间的识别,正是依靠比核酸链、多肽链所含信息量大几百倍、几千倍的糖链来完成的;至于对其他生物大分子如脂类及复合物的研究,也正在如火如荼地进行着。除此之外,还有许多未知的领域等待着我们去探索。

迄今,分子生物学已深入和覆盖到生命科学领域的各个分支,特别是生理学、病理学、发育生物学、神经生物学、免疫学等学科,并极大地推动着这些学科的发展。它的原理和研究方法已广泛应用于医药、农林、食品和工程各个领域,并取得了许多令人瞩目的成就。分子生物学自身的迅速发展及其对其他相关学科的强劲推动力有可能完全改变生命科学现有的面貌,也将深刻影响着人类的生活和社会发展。

分子生物学作为一门重要的专业基础课程,是生物学及相关专业知识结构的重要组成部分,在生物学的教学中占有重要的地位。有关分子生物学的教材已有不少,但在内容多寡、难易程度等方面比较适合初学者的教材较为少见。正是在这种情况下,作者在多年讲授本科生及研究生的分子生物学课程的教学实践基础上,参考国内外最新出版的诸多优秀教材编写完成了本书。本书尽可能用言简意赅的语言和简明清晰的图表,阐明重要的理论或实际问题,既易于理解,又便于记忆。全书分为12章。第1章绪论由范永山编写,第2章生物大分子由郭生金编写,第3章基因与基因组由张正光编写,第4章DNA的复制由邢继红编写,第5章DNA的损伤、修复与突变由崔东亚编写,第6章转录由郭小建编写,第7章翻译由王润梅编写,第8章原核生物基因的表达调控由抗艳红编写,第9章真核生物基因的表达调控由周春江编写,第10章基因重组由郭改娥编写,第11章基因工程由闫洪波编写,第12章分子生物学常用的研究技术由许冬梅编写。每一章均有学习目的、本章小结、延伸阅读及思考题,供学生进一步深入研读自测之用。本书初稿完成后,各编写人员进行了交互审阅,其中王立安和郝志敏在分别审阅本书大部分章节的基础上对本书的主要内容进行了充分研讨、补充和更正,最后由主编蒋继志和王金胜教授进行统稿完善。

目 录

前 言

1 绪论	1
1.1 引言	1
1.1.1 分子生物学的概念	1
1.1.2 分子生物学与其他学科间的关系	1
1.2 分子生物学的研究内容	7
1.2.1 结构分子生物学	7
1.2.2 遗传信息的传递规律	8
1.2.3 基因、基因组和蛋白质组	9
1.2.4 基因的表达和调控	10
1.2.5 基因工程	11
1.3 分子生物学的发展历程	12
1.3.1 准备和酝酿阶段	12
1.3.2 现代分子生物学的建立和发展阶段	12
1.3.3 初步认识生命本质并开始改造生命的深入发展阶段	13
本章小结	15
延伸阅读	15
思考题	16
2 生物大分子	17
2.1 生物大分子概述	17
2.1.1 生物大分子的概念	17
2.1.2 生物大分子的化学结构	18
2.1.3 生物大分子的高级结构	20
2.1.4 生物大分子的自组装	22
2.2 核酸	23
2.2.1 DNA 的结构与功能	23
2.2.2 RNA 的结构与功能	32
2.2.3 核酸的变性、复性和分子杂交	36
2.3 蛋白质	38
2.3.1 蛋白质的结构	38
2.3.2 蛋白质的结构与功能	41
2.4 其他生物大分子	42
2.4.1 脂类	42

2.4.2 多糖类大分子	43
2.4.3 核蛋白	44
本章小结	45
延伸阅读	45
思考题	46
3 基因与基因组	47
3.1 基因的概念及历史演变	47
3.2 基因的类型	49
3.2.1 基因的命名	49
3.2.2 断裂基因	50
3.2.3 重叠基因	51
3.2.4 移动基因	51
3.2.5 假基因	52
3.2.6 基因家族	52
3.3 基因组	53
3.3.1 基因组的概念	53
3.3.2 病毒的基因组	54
3.3.3 原核生物的基因组	55
3.4 真核生物的基因组	56
3.4.1 真核生物基因组的特点	56
3.4.2 真核生物基因组的重复序列	56
3.4.3 细胞器基因组	58
3.5 结构基因组学	59
3.5.1 遗传图谱	60
3.5.2 物理图谱	60
3.5.3 基因组图谱的应用	61
3.6 功能基因组学	61
3.6.1 转录组学	61
3.6.2 蛋白质组学	62
3.6.3 生物信息学	62
3.6.4 模式生物体基因组研究	63
3.7 人类基因组计划	63
本章小结	64
延伸阅读	65
思考题	65

4 DNA 的复制	66	5.5.2 定向诱变的应用	101
4.1 DNA 复制的机理与体系	66	本章小结	102
4.1.1 DNA 复制的机理	66	延伸阅读	102
4.1.2 DNA 复制的体系	67	思考题	103
4.2 原核生物 DNA 复制的过程	71	6 转录	104
4.2.1 复制的起始	71	6.1 转录的概述	104
4.2.2 复制的延伸	73	6.1.1 转录的一般特征	104
4.2.3 复制的终止	75	6.1.2 转录的酶学	105
4.3 真核生物 DNA 复制的特点	77	6.2 原核生物的转录	110
4.3.1 染色质复制	77	6.2.1 转录的起始	110
4.3.2 复制起点及复制速度	78	6.2.2 转录的延伸	112
4.3.3 DNA 聚合酶	78	6.2.3 转录的终止	114
4.3.4 RNA 引物和冈崎片段	78	6.3 真核生物的转录	115
4.3.5 DNA 复制与细胞周期	78	6.3.1 真核生物基因转录概述	115
4.3.6 末端复制与端粒酶	78	6.3.2 真核生物转录的机理	117
4.4 DNA 复制的忠实性及复制调控	80	6.4 RNA 转录后的加工	123
4.4.1 DNA 复制的忠实性	80	6.4.1 原核生物 RNA 转录后的加工	123
4.4.2 DNA 复制的调控	81	6.4.2 真核生物 RNA 转录后的加工	125
本章小结	82	6.5 转录的抑制	129
延伸阅读	83	6.5.1 碱基类似物	129
思考题	83	6.5.2 DNA 模板功能抑制物	130
5 DNA 的损伤、修复与突变	84	6.5.3 RNA pol 抑制剂	131
5.1 DNA 的损伤	84	6.6 RNA 的降解	131
5.1.1 DNA 损伤的因素	84	6.6.1 酶促降解	131
5.1.2 DNA 损伤的类型	87	6.6.2 其他方式	132
5.2 DNA 损伤的修复	88	本章小结	133
5.2.1 错配修复	89	延伸阅读	133
5.2.2 切除修复	89	思考题	134
5.2.3 直接修复	91	7 翻译	135
5.2.4 重组修复	92	7.1 蛋白质生物合成概述	135
5.2.5 SOS 反应	94	7.1.1 遗传密码	135
5.3 DNA 的突变	95	7.1.2 蛋白质生物合成的五个阶段	137
5.3.1 DNA 突变的概念	95	7.2 蛋白质生物合成的分子基础	138
5.3.2 DNA 突变的类型	95	7.2.1 mRNA 是蛋白质生物合成的模板	138
5.4 DNA 突变的回复	96	7.2.2 tRNA 转运活化的氨基酸	139
5.4.1 回复突变	96	7.2.3 核糖体是蛋白质生物合成的部位	140
5.4.2 基因内校正与基因间校正	97	7.3 翻译的过程	142
5.5 定向诱变	99	7.3.1 原核生物的蛋白质合成	142
5.5.1 定向诱变的方法	99		

7.3.2 真核生物的蛋白质合成	147	变化相关	179
7.4 蛋白质合成的抑制剂	148	9.2.2 基因扩增	180
7.5 蛋白质合成后的加工及运输	150	9.2.3 基因重排与变换	181
7.5.1 蛋白质合成后的加工	150	9.3 转录水平的调控	182
7.5.2 蛋白质合成后的运输	153	9.3.1 顺式作用元件	182
本章小结	158	9.3.2 反式作用因子	184
延伸阅读	158	9.4 转录后水平的调控	187
思考题	158	9.4.1 mRNA 前体加工的分子机制	187
8 原核生物基因的表达调控	159	9.4.2 转录后加工水平上的调控	189
8.1 原核生物基因表达调控概述	159	9.4.3 RNAi	196
8.1.1 原核生物基因表达调控的几个概念	160	9.5 翻译水平的调控	198
8.1.2 原核生物基因表达调控的两种	161	9.5.1 真核生物 mRNA 的“扫描模式”	199
机制——正调控和负调控	161	与蛋白质合成的起始	199
8.1.3 翻译水平上的调控	163	9.5.2 mRNA 的稳定性与基因表达调控	199
8.2 操纵子学说	164	9.5.3 可溶性蛋白质因子的修饰与翻译	200
8.3 乳糖操纵子学说	165	起始调控	201
8.3.1 组成与结构	165	9.5.4 翻译后调控	201
8.3.2 乳糖操纵子的负调控诱导模型	165	本章小结	201
8.3.3 分解代谢产物阻遏启动子的正调控	166	延伸阅读	201
系统	166	思考题	201
8.4 色氨酸操纵子学说	169	10 基因重组	203
8.4.1 色氨酸操纵子的负控制阻遏模型	170	10.1 基因重组概述	203
8.4.2 色氨酸操纵子的弱化系统	170	10.2 同源重组	204
8.4.3 阻遏与弱化作用的协调	172	10.2.1 同源重组的分子机制	204
8.5 阿拉伯糖操纵子	173	10.2.2 同源重组的酶学	206
8.5.1 阿拉伯糖操纵子结构和功能	173	10.3 细菌的基因转移与重组	207
8.5.2 AraC 蛋白的正、负调节作用	173	10.3.1 接合作用	208
8.6 组氨酸操纵子	175	10.3.2 转化作用	209
8.6.1 组氨酸合成代谢操纵子	175	10.3.3 转导作用	209
8.6.2 组氨酸利用操纵子	175	10.3.4 细菌的细胞融合	209
本章小结	176	10.4 位点特异性重组	210
延伸阅读	176	10.4.1 λ 噬菌体对 <i>E. coli</i> 的整合与切除	210
思考题	176	10.4.2 细菌的特异位点重组	213
9 真核生物基因的表达调控	177	10.5 转座重组	213
9.1 真核生物基因表达调控的概述	177	10.5.1 原核生物的转座子	214
9.2 DNA 及染色体水平的调控	178	10.5.2 真核生物的转座子	216
9.2.1 真核基因的转录与染色质的结构	178	10.5.3 转座的分子机制	219
		10.6 重组 DNA 的应用	221

10.6.1	转基因动物和植物	221
10.6.2	转基因技术应用于疾病诊断、治疗和预防	221
10.6.3	基因打靶	221
	本章小结	222
	延伸阅读	222
	思考题	222
11	基因工程	223
11.1	基因工程概述	223
11.1.1	基因工程的基本过程	223
11.1.2	基因工程的研究意义	224
11.2	基因工程中的工具酶	224
11.2.1	限制性内切核酸酶	224
11.2.2	连接酶	226
11.2.3	DNA 聚合酶	227
11.2.4	逆转录酶	227
11.2.5	其他用于 DNA 重组的工具酶	227
11.3	基因工程的常用载体	228
11.3.1	质粒载体	229
11.3.2	λ 噬菌体载体	230
11.3.3	黏粒载体(柯斯质粒载体)	231
11.3.4	真核生物载体	232
11.3.5	人工染色体载体	234
11.4	基因文库的构建	235
11.4.1	基因文库的类别	235
11.4.2	基因文库的完备性	236
11.5	重组体的构建、转化及鉴定	237
11.5.1	重组体的构建	237
11.5.2	重组体导入宿主细胞	238
11.5.3	重组子的鉴定	239
11.6	基因工程的应用及展望	240
11.6.1	基因工程药物	240
11.6.2	基因治疗	241
11.6.3	基因工程在农业上的应用	241
11.6.4	基因工程在其他领域中的应用	242
11.6.5	基因工程研究展望	242

	本章小结	242
	延伸阅读	243
	思考题	243
12	分子生物学常用的研究技术	244
12.1	概述	244
12.2	核酸的分离提取和电泳检测	244
12.2.1	核酸的分离提取	244
12.2.2	核酸的纯化	247
12.2.3	核酸的定量和保存	247
12.2.4	核酸的电泳检测	248
12.3	聚合酶链反应	248
12.3.1	PCR 反应的原理	248
12.3.2	PCR 反应的体系和步骤	249
12.3.3	PCR 技术的发展及应用	250
12.4	分子标记	251
12.4.1	分子标记的种类及其检测的技术方法	251
12.4.2	分子标记的应用	254
12.5	核酸分子杂交	254
12.5.1	印迹杂交	254
12.5.2	核酸原位杂交	255
12.6	生物芯片技术	257
12.6.1	生物芯片的种类	257
12.6.2	生物芯片的制备	257
12.6.3	生物芯片的应用	258
12.7	基因敲除与基因沉默	259
12.7.1	利用基因同源重组进行基因敲除	259
12.7.2	RNAi 引起的基因沉默	261
12.8	蛋白质组学研究技术	261
12.8.1	蛋白质分离	262
12.8.2	蛋白质分析	262
12.8.3	蛋白质的相互作用	263
	本章小结	264
	延伸阅读	264
	思考题	264

1 绪 论

1.1 引 言

1.1.1 分子生物学的概念

分子生物学(molecular biology)是从分子水平研究生物大分子的结构与功能从而阐明生命现象本质的科学。自 20 世纪 50 年代以来,分子生物学一直是生物学的前沿与生长点。

1.1.2 分子生物学与其他学科间的关系

现代化学和物理学理论、技术和方法在生命本质和生物遗传研究方面的应用催生了分子生物学,分子生物学理论和技术的发展又促进了化学、物理学、生物学、遗传学等相关学科的进步。这些学科之间既相对独立,又有交叉。不同学科之间的交叉又催生了许多边缘学科。

分子生物学与生物化学、细胞生物学和生物物理学的关系非常密切。分子生物学更着重强调的是:①从分子水平进行研究。生物化学、细胞生物学和生物物理学则更强调细胞水平、整体水平和群体水平的研究。②重点研究生物大分子,即蛋白质(protein)、酶(enzyme)、核酸(nucleic acid)、脂质(lipid)体系和部分多糖(polysaccharide)及其复合体系。而一些小分子物质在生物体内的转化属于生物化学的范围。③研究生命活动的普遍规律,即整个生物界所共同具有的基本特征。而研究某一特定生物体或某一种生物体内的某一特定器官的物理、化学现象或变化,则属于生物化学、细胞生物学或生物物理学的范畴。

分子生物学已经渗透到生物学的几乎所有领域,成为生命科学领域的带头学科。

1.1.2.1 生物学、生物分类和生物多样性

生物学(biology)是研究生命现象和生物活动规律的科学,根据研究对象分为动物学、植物学、微生物学等;根据研究内容分为分类学、解剖学、生理学、遗传学、生态学等;根据研究方法分为实验生物学与系统生物学等。20 世纪 40 年代以来,生物学吸收了数学、物理学和化学研究方面的成就,逐渐发展成一门精确的、定量的、深入到分子层次的科学。人们已经认识到生命是物质的一种运动形态,是由蛋白质、核酸、糖类、脂质等生物大分子组成的物质系统,生命现象是物质、能量和信息综合运动与传递的表现。

生物分类学(biological taxonomy)是生物学研究的基础,主要研究生物分类的方法和原理。近年来生物分类学的发展面临着许多问题,无论是二界说(Carl von Linné,1735)、三界说(Hogg,1860;E. N. Haeckel,1866)、四界系统(H. F. Copeland,1938)、五界系统(R. H. Whittaker,1969)、三总界六界系统(陈世骧,1979),还是根据生物的系统发育关系发展起来的支序系统学派(W. Hennig 1950, 1966),利用特殊的、可用于物种鉴定的 DNA 序列——DNA 条形码(DNA bar code)建立的分子系统学(molecular systematics),都尝试根据生物的表型、细胞结构、发育方式、进化关系、生理生化特征、蛋白质和 DNA 结构等对生物进行分类,但由于生命世界的复杂性,目前人们对生物的分界还没有统一的意见。但从 30 亿年前古生物的化石记录到地球上现存生物的情况,都揭示了生物从非细胞结构到细胞结构、从原核生物到真核生

物、从简单到复杂、从低等到高等的进化方向。

非细胞生物主要包括病毒(virus)和亚病毒(subviruses),它们是最原始的生命体。病毒是比细菌还小、没有细胞结构、由一条或多条核酸长链和蛋白质外壳构成、只能在活细胞中增殖的微生物。寄生于细菌的病毒称为噬菌体(bacteriophage)。病毒颗粒很小、以纳米为测量单位,多数要用电子显微镜才能观察到。病毒基因同其他生物的基因一样,也可以发生突变和重组。近年发现了比病毒还要简单的亚病毒,包括只含有具有单独侵染性的较小型的核糖核酸(RNA)分子的一类病毒(viroid),只含有不具备侵染性的RNA的拟病毒(virusoid)和卫星RNA(satellite RNA),以及没有核酸但具有感染性蛋白质颗粒的朊病毒(virino)或蛋白侵染因子(prion, proteinaceous infectious agent)。

具有细胞形态的生物可分为原核生物、古核生物和真核生物。原核细胞(procaryotic 或 prokaryotic cell),是一类没有核膜且不进行有丝分裂(mitosis)、减数分裂(meiosis)、无丝分裂(amitosis)的细胞,在细胞内不发生原生质流动,没有叶绿体(chloroplast)、线粒体(mitochondrion)等细胞器(organelles)的分化,只有核糖体。由这种细胞构成的生物,称为原核生物,包括所有的细菌和蓝藻类。注意支原体、立克次氏体和衣原体均属细菌。它们没有真正的细胞核(nucleus),只有原核或拟核,所含的一个环状双股单一顺序的脱氧核糖核酸分子(circular DNA),没有组蛋白(histone)与之结合,转录和转译(transcription and translation)同时进行,DNA复制后,细胞随即分裂为二。古核生物也称古细菌(archaebacteria),具有原核生物的某些特征,如无核膜及内膜系统;也有某些真核生物的特征,如以甲硫氨酸起始的蛋白质合成、核糖体对氯霉素不敏感、RNA聚合酶和真核细胞的相似、DNA具有内含子并结合组蛋白;还具有既不同于原核细胞也不同于真核细胞的特征,如细胞膜中的脂类是不可皂化的,细胞壁不含肽聚糖等。古核生物多生活在极端的生态环境中。和原核细胞相比,真核细胞的结构更为复杂,它有线粒体等各种膜细胞器,有围以双层膜的细胞核,把位于核内的遗传物质与细胞质分开,DNA与组蛋白及其他蛋白结合而成染色体。真核细胞的分裂分为有丝分裂和减数分裂,分裂的结果使复制的染色体均等地分配到子细胞中去。原生生物是最原始的真核生物,单细胞或多细胞,不分化成组织,可自养、异养或混合营养。植物是以光合自养为主要营养方式的真核生物。典型的植物细胞都含有液泡和以纤维素为主要成分的细胞壁。细胞质中有进行光合作用的叶绿体。真菌是以吸收为主要营养方式的真核生物。真菌的细胞有细胞壁,多含几丁质,也有含纤维素的。真菌细胞没有质体和光合色素。少数真菌是单细胞的,如酵母菌。多细胞真菌的基本构造是分枝或不分枝的菌丝。一整团菌丝称为菌丝体。有的菌丝以横隔分成多个细胞,每个细胞有一个或多个核,有的菌丝无横隔而成为多核体。黏菌是一种特殊的真菌。它的生活史中有一段是真菌性的,而另一段则是动物性的,其结构、行为和取食方法与变形虫相似。黏菌被认为是介于真菌和动物之间的生物。动物是以吞食为营养方式的真核生物。吞食异养包括捕获、吞食、消化和吸收等一系列复杂的过程。单细胞动物吞入食物后形成食物泡,食物在食物泡中被消化。多细胞动物具有复杂的感觉器官和消化系统、排泄系统、呼吸系统、神经系统、内分泌系统和运动系统等。

从类病毒、病毒到微生物、植物和动物,一定范围内不同的生物有规律地构成稳定的生态综合体,这就是生物多样性(biological diversity),既包括物种的多样性,生态系统的多样性,又包括物种的遗传与变异的多样性。由于自然资源的合理利用和生态环境的保护是人类实现可持续发展的基础,因此生物多样性的研究和保护已经成为世界各国普遍重视的一个问题。1992年,在巴西的里约热内卢举行的联合国环境与发展大会上,通过了《生物多样性公约》(Convention on Biological Diversity),这标志着世界范围内的自然保护工作进入了一个新的

阶段。

在生物多样性研究过程中,能够全面地反映基因组变异性的DNA指纹图谱(DNA finger print)的应用越来越广泛。DNA指纹图谱中的谱带能够稳定地从上一代遗传给下一代,其高特异性和体细胞稳定性可用于鉴定个体,在法医学上鉴别犯罪分子和确定个体间的血缘关系具非常重要的价值,其多位点性可用于检测目标基因组在病变及治疗过程中的改变情况。DNA指纹已成为研究生物群体遗传结构、生态与进化、分类等非常有价值的遗传标记(genetic marker)。

1.1.2.2 化学、生物化学和化学生物学

化学(chemistry)是一门在原子、分子层次上研究物质的组成、结构、性质及其变化规律的科学。化学是人类认识和改造物质世界的一种重要方法和手段。作为重要的基础科学之一,化学在与物理学、生物学、自然地理学、天文学等学科的相互渗透中,得到了迅速的发展,也推动了其他学科和技术的发展。例如,核酸化学的研究成果使生物学的研究从细胞水平提高到分子水平,建立了分子生物学。

进入20世纪以后,应用当代科学的理论、技术和方法,化学在认识物质的组成、结构、合成和测试等方面都有了长足的进展,而且在理论方面取得了许多重要成果。

在结构化学方面,应用量子力学研究分子结构,产生了量子化学;从氢分子结构的研究开始,逐步揭示了化学键的本质,先后创立了价键理论、分子轨道理论和配位场理论;应用X射线衍射、电子衍射和中子衍射等方法作为研究物质结构的新分析手段,洞察物质的晶体化学结构;研究物质结构的谱学方法也由可见光谱、紫外光谱、红外光谱扩展到核磁共振谱、电子自旋共振谱、光电子能谱、射线共振光谱、穆斯堡尔谱等;与计算机或电子显微镜联用后,利用积累的大量物质结构与性能相关的资料,人们可以预测或直接观察分子的结构。在放射化学和核化学研究方面,从放射性衰变理论的创立、同位素的发现到人工核反应和核裂变的实现,人类的认识深入到亚原子层次,并创立了同位素地质学、同位素宇宙化学等交叉学科。在化学反应理论方面,随着对分子结构和化学键的认识的提高,以及分子束、激光和等离子技术的应用,使得对不稳定化学物种的检测和研究成为现实。在化学分析方法方面,物质的化学成分和组成分析方法不断改进,分析灵敏度从常量发展到微量、超微量、痕量,深入进行结构分析、构象测定、同位素示踪,直接测定自由基、离子基、卡宾、氮宾、卡拜等各种活泼中间体。在物质的分离技术方面,离子交换、膜分离、色谱法等技术的发展不但可以分离某种特殊物质,而且还可以进行提炼和精制。在化学合成方面,氨、超导材料和纳米物质的无机合成工业和酚醛树脂等高分子材料、蛋白质和核酸等生命基础物质、有复杂结构的天然有机化合物和有特效的药物的合成,为现代工农业、交通运输、医疗卫生、军事技术,以及人们衣食住行各方面,提供了多种性能优异而成本较低的重要材料,成为现代物质文明的重要标志。

生物化学(biochemistry)是研究生命的化学组成及其在生命活动中变化规律的一门学科。生物化学的主要任务是从化学组成角度理解生物大分子和生物代谢,与分子生物学的关系最为密切。

生物体(living organism)是由生物大分子所组成,具有系列特殊功能的系列复杂机构,它具有生长、代谢、感应、繁殖等生命的基本特征。细胞(cell)是生物体的基本构成单位,生物体通过细胞的活动进行各种生命活动。原生质(protoplasm)泛指生物细胞内的全部生命物质,主要包括蛋白质、核酸、多糖、脂类4种生物大分子作为生物体骨架和提供能量,维生素、激素、有机酸等次生物质催化、调节新陈代谢,以及无机盐、水等。

新陈代谢(metabolism)是生物体与外界环境进行物质交换与能量交换的过程。机体内的代谢反应相互联系、协同制约组成许多代谢途径和网络,在严密精巧的调控下,有条不紊地进行。代谢调控是近代生物化学研究的一个重要方面,主要有三种途径:①通过代谢物的诱导或阻遏作用控制酶的合成。这是在转录(transcription)水平的调控,如乳糖诱导乳糖操纵子(operon)合成有关的酶。②通过激素与靶细胞的作用,引发一系列生化过程,如环腺苷酸激活的蛋白激酶通过磷酸化反应对糖代谢的调控。③效应物通过别构效应直接影响酶的活性,如终点产物对代谢途径第一个酶的反馈抑制。生物体内绝大多数调节过程是通过别构效应实现的。

生物大分子的多种多样功能与它们特定的结构有密切关系。蛋白质的主要功能有催化、运输和储存、机械支持、运动、免疫防护、接受和传递信息、调节代谢和基因表达等。酶的作用具有催化效率高、专一性强等特点,酶与蛋白质、核酸等生物大分子的相互作用和在工农业生产、国防和医学上的应用是酶学研究的重点方向。核酸是遗传信息的储存和传递者,主要包括脱氧核糖核酸(DNA)和核糖核酸(RNA),核糖核酸还可分为信使核糖核酸(mRNA)、转移核糖核酸(tRNA)、核蛋白体核糖核酸(rRNA)和其他一些小分子核糖核酸(sRNA),它们在蛋白质生物合成和基因表达调控中起着重要作用。生物体的糖类物质包括单糖、寡糖和多糖,单糖是生物体能量的主要来源,寡糖和蛋白质或脂质可以形成糖蛋白、蛋白聚糖和糖脂,多糖是生物体的结构物质和营养物质。蛋白质、核酸、酶和糖类是生物化学的四大研究对象。

使用小分子作为工具解决生物学的问题或通过干扰(调节)正常过程了解蛋白质的功能,是化学生物学(chemical biology)或化学遗传学(chemical genetics)的研究范畴。目前,化学生物学在国际上已得到广泛认可并正迅速发展,如 *Nature* 出版集团出版了 *Nature Chemical Biology*,美国化学会出版了 *ACS Chemical Biology* 等化学生物学方面的专业杂志。

1.1.2.3 物理学、生物物理学和物理生物学

物理学(physics)是研究物质的结构、相互作用及其运动规律的学科,它研究物质世界的一些最简单、最基本、最普遍的运动形式,如机械运动、电磁运动、热运动和微观粒子的运动等,以及由这些简单运动构成的更高级运动形式,如生命、遗传、思维等。物理学研究的是物质世界普遍而基本的规律,这些规律对有机界和无机界同样适用。

生物物理学(biophysics)是物理学与生物学相结合的一门边缘学科,它应用物理学的概念和方法研究生物各层次结构与功能的关系,分析生命活动的物理过程和物质在生命活动过程中表现的物理特性。生物物理学又存在分子生物物理学、放射生物物理学、生理生物物理学和数学或理论生物物理学等不同分支。分子生物物理学利用电子显微镜、超速离心机、X射线衍射技术等物理手段研究在生物中起重要作用的大分子及类似大小的粒子,放射生物物理学研究生物体对电离辐射和紫外线的反应,生理生物物理学利用物理学原理解释生物体的部分功能和物理环境变化时生物体正常功能的变化,数学或理论生物物理学则以数学和物理学的理论来解释生物的行为。研究生命物质的物理规律,不仅能进一步阐明生物的本质,更重要的是能使人们对自然界整个物质运动规律的认识达到新的高度。

利用新近发展起来的物理学先进概念和技术,精确地测量和描述生物体系的结构、功能和行为,把生物学建立在定量的物理学基础之上,是物理生物学(physical biology)的重点研究内容。物理生物学的出现是现代学科交叉发展到一定阶段的必然趋势。物理生物学的概念一经提出,立即得到了国际学术界的高度重视。例如,哈佛大学设立了“物理生物学”方向和专业,霍普金斯大学制订了“物理与工程生物学”研究生计划,英国 Institute of Physics (IOP)出版社发行了 *Physical Biology* 杂志。

1.1.2.4 遗传学和分子遗传学

遗传学(genetics)是研究生物遗传和变异的科学。遗传(heredity)是指亲代与子代、子代与子代之间相似的现象。“种瓜得瓜,种豆得豆”,“龙生龙,凤生凤”等通俗用语都是对遗传现象的简单说明。变异(variation)是指同种生物个体间的差异。生物界没有绝对相同的两个个体,“龙生九子,子子不同”就是指的这个道理。遗传与变异是对立统一的关系。遗传是相当稳定的但这种稳定性是相对的,暂时的;而变异则是绝对的,永恒的。生物通过不断地发生变异,然后通过遗传把某些变异在后代中巩固下来。因此遗传和变异是生物进化和物种形成的内在依据。

遗传学研究的核心是遗传物质的本质、遗传物质的传递和遗传信息的实现。遗传物质的本质包括它的化学本质、它所包含的遗传信息、它的结构、组织和变化等;遗传物质的传递包括遗传物质的复制、染色体的行为、遗传规律和基因在群体中的数量变迁等;遗传信息的实现包括基因的原初功能、基因的相互作用、基因作用的调控及个体发育中的基因的作用机制等。细胞分裂、染色体复制和基因功能的研究是奠定遗传学基本理论的基础。正是分子生物学在基因本质和基因结构与功能研究方面的一个又一个重大突破,使遗传学得到迅速发展,衍生出分子遗传学、发育遗传学和遗传工程学三个方向,并形成了细胞遗传学、发育遗传学、行为遗传学和免疫遗传学等新的分支学科。随着分子生物学实验技术,如限制性内切酶降解(restriction enzyme digestion)、聚合酶链反应(polymerase chain reaction, PCR)、核苷酸序列分析、DNA重组(DNA recombination)等技术的应用和发展,现代遗传学已从对现有生物进行改良进入到定向改造生物的遗传结构的新领域。转座子(transposon)、操纵子(operon)、基因表达调控(gene expression regulation)的发现,使遗传学开始与生产实践相结合,经过基因改良的生物(gene modified organism, GMO)正在悄悄地改变着这个世界。

分子遗传学(molecular genetics)是在分子水平上研究生物遗传和变异机制的遗传学分支学科。经典遗传学主要研究基因在亲代和子代之间的传递,而分子遗传学重点研究基因的本质、基因的功能及基因的变化。分子遗传学是在微生物遗传学和生物化学的基础上发展起来的,它的早期研究工作都以微生物,特别是以大肠杆菌和它的噬菌体作为研究材料完成的。

分子遗传学的研究方法是得到一系列使某一种生命活动不能完成的突变型(mutant),如不能合成某一种氨基酸、不能进行DNA复制、不能进行细胞分裂或不能完成某些发育过程,然后对它们进行遗传互补测验、生物大分子和细胞的精细结构分析、核酸和蛋白质的顺序分析,了解这些突变型代表几个基因和各个基因在染色体上的位置,判断蛋白质的哪一部分和特定的功能有关等。分子遗传学研究的方法已经成为许多遗传学分支学科的重要研究方法。分子遗传学的发展,使人类对DNA、蛋白质、遗传密码及遗传机制的演变有了更深入的认识,通过这些方面的研究,对生物进化过程将会有更加本质性的了解。

1.1.2.5 细胞生物学、细胞遗传学和干细胞的发育生物学

细胞生物学(cell biology)是在显微、亚显微和分子水平三个层次上,研究细胞的结构、功能和各种生命规律的一门科学。

自从1665年英国人R. Hooke用自己设计与制造的显微镜(放大倍数为40~140倍)观察软木(栎树皮)细胞的构造,并首次用“cell”(小室)称呼他所观察到的封闭状小室以后,人们对大量动植物进行了观察。1839年德国人施莱登(M. J. Schleiden)和施旺(T. Schwann)提出了“细胞学说”(cell theory):一切植物、动物都是由细胞组成的,细胞是一切动植物的基本单

位。1880年, J. von Hanstein 又提出“原生质体”(protoplast)的概念,并发展成为原生质理论:动、植物有机体细胞的组织单位是一团原生质,分化为细胞核和细胞质。1879~1886年, W. Flemming、E. Strasburger、van Beneden 等相继发现了无丝分裂(amitosis)、有丝分裂和减数分裂现象。1883~1958年,电镜技术的应用揭示出细胞的亚显微世界,发现了一些光镜下看不见的细胞器及其结构,如中心体、线粒体、高尔基体、内质网、核蛋白体、叶绿体等。20世纪60年代以后,细胞结构与功能开始和生化生理相结合,尤其是基因重组技术出现以后,细胞生物学将研究细胞的分子结构及其在生命活动中的作用作为主要任务,细胞周期调控(cell cycle control)、细胞衰老(cellular senescence)与凋亡(cell apoptosis)、细胞信号转导(signal transduction)、真核细胞基因表达与调控(regulation of gene expression)、细胞起源与进化等成为研究热点。

细胞遗传学(cytogenetics)是从细胞学的角度,特别是从染色体的结构和功能,以及染色体和其他细胞器的关系来研究遗传现象,阐明遗传和变异的机制。研究对象主要是真核生物,特别是包括人类在内的高等动植物。

早期的细胞遗传学着重研究分离、重组、连锁、交换等遗传现象的染色体基础及染色体畸变和倍性变化等染色体行为的遗传学效应,并涉及各种生殖方式(如无融合生殖、单性生殖及减数分裂驱动等方面)的遗传学和细胞学基础。以后又衍生出一些分支学科,研究内容进一步扩大,如体细胞遗传学——主要研究体细胞,特别是离体培养的高等生物体细胞的遗传规律;分子细胞遗传学——主要研究染色体的亚显微结构和基因活动的关系;进化细胞遗传学——主要研究染色体结构和倍性改变与物种形成之间的关系;细胞器遗传学——主要研究细胞器,如叶绿体、线粒体等的遗传结构;医学细胞遗传学——主要研究染色体畸变与遗传病的关系等,对遗传咨询和产前诊断具有重要意义。

干细胞(stem cell, SC)是一类具有自我复制能力(self-renewing)的多潜能细胞,在一定条件下,它可以分化成多种功能细胞。根据干细胞所处的发育阶段分为胚胎干细胞(embryonic stem cell, ES 细胞)和成体干细胞(somatic stem cell)。根据干细胞的发育潜能分为全能干细胞(totipotent stem cell, TSC)、多能干细胞(pluripotent stem cell)和单能干细胞(unipotent stem cell)。胚胎干细胞的发育等级较高,是全能干细胞,而成体干细胞的发育等级较低,是多能或单能干细胞。

当受精卵分裂发育成囊胚时,内层细胞团(inner cell mass)的细胞即为胚胎干细胞。胚胎干细胞具有全能性,可以自我更新并具有分化为体内所有组织的能力。早在1970年 M. J. Evans 已从小鼠中分离出胚胎干细胞并在体外进行培养。而人的胚胎干细胞的体外培养直到最近才获得成功。但人类 ES 细胞的研究工作在全世界范围内存在很大争议,出于社会伦理学方面的原因,有些国家甚至明令禁止进行人类 ES 细胞研究。

成年动物的许多组织和器官,如表皮和造血系统,具有修复和再生的能力。成体干细胞在其中起着关键的作用。在特定条件下,成体干细胞或者产生新的干细胞,或者按一定的程序分化,形成新的功能细胞,从而使组织和器官保持生长和衰退的动态平衡。造血干细胞是体内各种血细胞的唯一来源,它主要存在于骨髓、外周血、脐带血中。造血干细胞的移植是治疗血液系统疾病、先天性遗传疾病及多发性转移性恶性肿瘤疾病的最有效方法。关于神经干细胞研究起步较晚,理论上讲,任何一种中枢神经系统疾病都可归结为神经干细胞功能的紊乱。

干细胞在细胞发育控制方面和医学治疗方面的潜在的巨大理论价值和应用价值,使其迅速成为当前的研究热点。对干细胞的发育生物学和分子生物学研究,使人类不但可以控

制细胞的发育方向,实现细胞的定向分化,并将其应用于细胞移植治疗等实用研究,而且能够发现一些有直接治疗功能或提供药物治疗重要靶点的基因。利用胚胎干细胞可以检测基因在人类胚胎发育中的作用和新药功能,从而避免基因敲除(gene knock-off)技术和新药用于人类临床治疗的一些伦理及法律上的障碍。

1.2 分子生物学的研究内容

分子生物学的研究对象主要是核酸和蛋白质这两种生物大分子。目前普遍认为,在所有生物中,①构成生物大分子的单体都是相同的,它们具有共同的核酸语言和共同的蛋白质语言;②生物体内一切有机大分子的建成都遵循共同的规则;③生物大分子单体(核苷酸、氨基酸)组成和排列方式的不同是产生功能差异的基础,不同的生物大分子之间的互作是造成物种特性差异的根本原因。

核酸和蛋白质的结构分析及遗传物质和遗传信息传递规律的研究对分子生物学学科的发展起到了巨大的推动作用,在这些研究的基础上,分子生物学在基因组学和功能基因组学、基因的表达和调控、蛋白质组学(proteome)和基因工程(DNA重组技术)等方面取得了非常卓越的成就。

1.2.1 结构分子生物学

生物大分子,特别是蛋白质和核酸结构和功能的研究,是分子生物学在分子水平上研究生命现象本质的基础。所谓的分子水平,指的是那些携带遗传信息的核酸和在遗传信息传递及细胞内、细胞间通讯过程中发挥着重要作用的蛋白质等生物大分子。这些生物大分子均具有较大的分子质量,由简单的小分子核苷酸或氨基酸排列组合以蕴藏各种信息,并且具有复杂的空间结构以形成精确的相互作用系统,由此构成生物的多样化和生物个体精确的生长发育和代谢调节控制系统。阐明这些复杂的结构及结构与功能的关系是分子生物学的主要任务。

要了解一种生物大分子的功能,通常要先研究其结构。例如,对DNA的结构的研究使认识基因突变(gene mutation)、染色体复制(chromosome replication)和遗传重组(genetic recombination)成为可能;tRNA分子的三维结构的研究解释了DNA储存的遗传信息如何翻译到蛋白质的氨基酸顺序中;对初始转录本mRNA和成熟mRNA排列顺序的比较让我们认识到RNA加工在基因表达过程中的重要性;对DNA结合蛋白的激活结构域和DNA结合结构域的诠释展现了生物大分子间相互作用的主要方式。

结构分子生物学的任务是通过阐明生物大分子的三维结构来解释细胞的生理功能。在蛋白质结构分析方面,1951年L. C. Pauling等提出的 α 螺旋结构描述了蛋白质分子中肽链的一种构象;1955年F. Sanger完成了胰岛素的氨基酸序列的测定;1957年和1959年J. C. Kendrew和M. F. Perutz在X射线分析中分别应用重原子同晶置换技术和计算机技术阐明了鲸肌红蛋白和马血红蛋白的立体结构;1965年中国科学家合成了有生物活性的胰岛素,首先实现了蛋白质的人工合成。在核酸结构分析方面,1944年O. T. Avery等研究细菌中的转化现象,证明了DNA是遗传物质;1953年J. D. Watson和F. H. C. Crick提出了DNA的双螺旋结构,开辟了分子生物学研究的新纪元;1961年F. Jacob和J. L. Monod提出了操纵子的概念,解释了原核基因表达的调控。到20世纪60年代中期,关于DNA自我复制、RNA转录和蛋白质合成的一般性质已基本清楚,遗传和变异的规律也随之逐渐明朗。

1.2.2 遗传信息的传递规律

遗传物质可以是 DNA,也可以是 RNA。细胞的遗传物质都是 DNA,只有一些病毒和亚病毒的遗传物质是 RNA。以 DNA 为模板按照碱基互补配对的原则合成 DNA,以及以 RNA 为模板合成 RNA,都称为复制(replication)。以 DNA 为模板按照碱基互补配对的原则合成 RNA,称为转录(transcription)。转录产生的并不是成熟的 mRNA,而只是成熟 mRNA 的前体,或称为初始转录本 mRNA,它们需要进行一定的剪切和连接才能成为成熟的 mRNA,这个过程称为 RNA 加工(RNA processing)。以成熟的 mRNA 为模板在核糖体上进行蛋白质多肽链的合成,称为翻译(translation)。翻译产生的只是蛋白质前体,它需要加工、修饰、折叠和组装后,转运到适当的位置后才能发挥作用。这种从 DNA 到 RNA,再到蛋白质的遗传信息传递方向,叫做直线形中心法则(central dogma)(图 1-1A)。

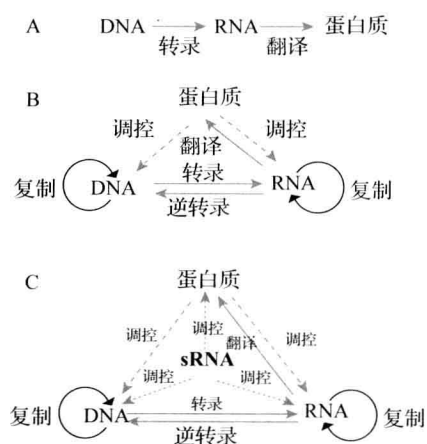


图 1-1 遗传信息的传递

A. 直线形中心法则; B. 三角形中心法则; C. 圆锥形中心法则

以 RNA 为遗传物质的病毒称为逆转录病毒(reverse virus),在这种病毒的感染周期中,单链的 RNA 分子在逆转录酶(reverse transcriptase)的作用下,可以逆转录成单链的 DNA,然后再以单链的 DNA 为模板生成双链 DNA。在逆转录酶催化下, RNA 分子产生与其序列互补的 DNA 分子,称为互补 DNA (complementary DNA, cDNA),这个过程即为逆转录(reverse transcription)。由此可见,遗传信息并不一定是从 DNA 单向地流向 RNA, RNA 携带的遗传信息同样也可以流向 DNA。但是 DNA 和 RNA 中包含的遗传信息只是单向地流向蛋白质,迄今为止还没有发现蛋白质的信息逆向地流向核酸。另外, DNA 和 RNA 在复制、转录和翻译过程中都需要蛋白质的参与和调节,因此,蛋白质在遗传信息传递过程中起着非常重要的调控作用。这种 DNA、mRNA 和蛋白质间复杂的相互作用类似一个三角形,它是对直线形中心法则的必要补充,可形象地称之为三角形中心法则(图 1-1B)。

随着对 RNA 种类的不断发现,现代分子生物学家对 RNA,尤其是小分子 RNA (small RNA, sRNA)或非编码 RNA(non-coding RNA, ncRNA)的功能有了更深入的认识。许多小分子细胞核 RNA (snRNA)与 RNA 加工有关,一些小分子细胞核仁 RNA (snoRNA)参与 rRNA 的合成,19~22nt 的微小 RNA (miRNA)参与 DNA 和 RNA 的修饰、mRNA 的稳定性、蛋白质的合成。因此,有的科学家认为在遗传信息传递过程中,小分子 RNA 起着中心的调控作用,它们对 DNA、RNA 和蛋白质的结构与功能都有着非常重要的影响。因此,这种 DNA、mRNA、蛋白质和 sRNA 间的复杂关系,可称为圆锥形中心法则(图 1-1C)。

虽然中心法则对遗传信息的传递方向进行了系统的概括,但还是存在一些特别现象曾对中心法则提出严重的挑战,如朊病毒的发现。朊病毒是一种蛋白质传染颗粒(proteinaceous infectious particle),它是羊瘙痒病、疯牛病和人类的库鲁病(Kuru disease)和克-杰氏综合征(Creutzfeldt-Jacob disease, CJD)的病原体,能在寄主中传播,并在受感染的宿主细胞内产生与自身相同的分子,且实现相同的生物学功能,这意味着这种蛋白质分子也是负载和传递遗传信息的物质。但更深入的研究表明,朊病毒只是由基因编码产生的一种正常蛋白质的异构体,它进入宿主细胞后并不是自我复制,而是将细胞内基因编码产生的 PrP 蛋白(prion related pro-

tein)由正常的 PrP^c 异构体转变成致病的 PrP^{sc} 异构体。因此,朊病毒并不是遗传物质。当然,不依赖核糖体的非核糖体肽合成酶(NRPS)和 RNA 编辑(RNA editing)的发现,使人们认识到遗传信息的传递规律还有待进一步的完善和发展。

1.2.3 基因、基因组和蛋白质组

基因(gene)是 DNA 分子中含有特定遗传信息的一段核苷酸序列,它包含合成一种功能蛋白或 RNA 分子所必需的全部 DNA 序列。根据其是否具有转录和翻译功能,基因可分为三类:①编码蛋白质的基因,具有转录和翻译功能,包括编码酶和结构蛋白的结构基因及编码阻遏蛋白的调节基因;②只有转录功能而没有翻译功能的基因,包括 tRNA 基因和 rRNA 基因;③不转录的基因,它对基因表达起调节控制作用,包括启动基因和操纵基因,启动基因和操纵基因有时被统称为控制基因。随着分子生物学研究的深入,人们又发现在基因结构中存在有“断裂基因”、“重叠基因”、“假基因”、“移动基因”等。移动基因,又称跳跃基因(jumping gene)或转座子(transposon)。这些结构的发现,使人们对基因的功能有了更深入的理解。

基因位于染色体上,并在染色体上呈线性排列。基因不仅可以通过复制把遗传信息传递给下一代,还可以使遗传信息得到表达。不同个体之间在形态、发育和功能等方面的不同,都是基因差异所致。基因是表现生物体遗传性状的物质基础。

单倍体细胞中的全套染色体或病毒粒子所含的全部 DNA 分子或 RNA 分子,称为该生物体的基因组(genome)。基因组中既含有编码序列,也含有非编码序列。基因组的大小用全部 DNA 的碱基对(base pair, bp)总数表示。

1986 年美国科学家 T. Roderick 提出了基因组学(genomics),指对所有基因进行基因组作图、核苷酸序列分析、基因定位和基因功能分析的一门科学。基因组研究主要包括以全基因组测序为目标的结构基因组学(structural genomics)、以基因功能鉴定为目标的功能基因组学(functional genomics)或后基因组(postgenome)、研究和利用模式生物基因组测序产生的大量基因组信息进行基因结构和功能分析的比较基因组学(comparative genomics)。

结构基因组学是一门通过基因作图、核苷酸序列分析确定基因组成、基因定位的科学。遗传信息在染色体上,但染色体不能直接用来测序,必须将基因组这一巨大的研究对象进行分解,使之成为较易操作的小的结构区域,这个过程就是基因作图(gene mapping)。根据使用的标志和手段不同,基因作图主要有三种类型:①遗传连锁图——通过遗传重组所得到的基因在具体染色体上的线性排列图谱。利用遗传标志之间的重组频率,确定它们的相对距离,一般用厘摩(cM,即每次减数分裂的重组频率为 1%)来表示,遗传标志有 RFLP(限制性酶切片段长度多态性)、RAPD(随机引物扩增多态性 DNA)、AFLP(扩增片段长度多态性)、STR(短串联重复序列,又称微卫星)和 SNP(单个核苷酸的多态性)等。②物理图谱——利用限制性内切核酸酶处理染色体,根据重叠序列确定片段间连接顺序的图谱。利用遗传标志之间物理距离[碱基对(bp)、千碱基(kb)或兆碱基(Mb)]确定图距,遗传标志主要采用序列标签位点(sequence tag site, STS),染色体定位明确且可用 PCR 扩增的单拷贝序列。③转录图谱——利用表达序列标签(expressed sequence tag, EST)作为标记所构建的分子遗传图谱。mRNA 或 cDNA 的 5' 或 3' 端序列称为表达序列标签,一般长 300~500bp。

功能基因组学是利用结构基因组所提供的信息和产物,发展和应用新的实验手段,通过在基因组或系统水平上全面分析基因的功能,使得生物学研究从对单一基因或蛋白质的研究转向多个基因或蛋白质同时进行系统研究的学科。研究内容包括基因功能发现、基因表达分析及突变检测。采用的技术包括减法杂交、差示筛选、cDNA 代表差异分析、mRNA 差异显示等