

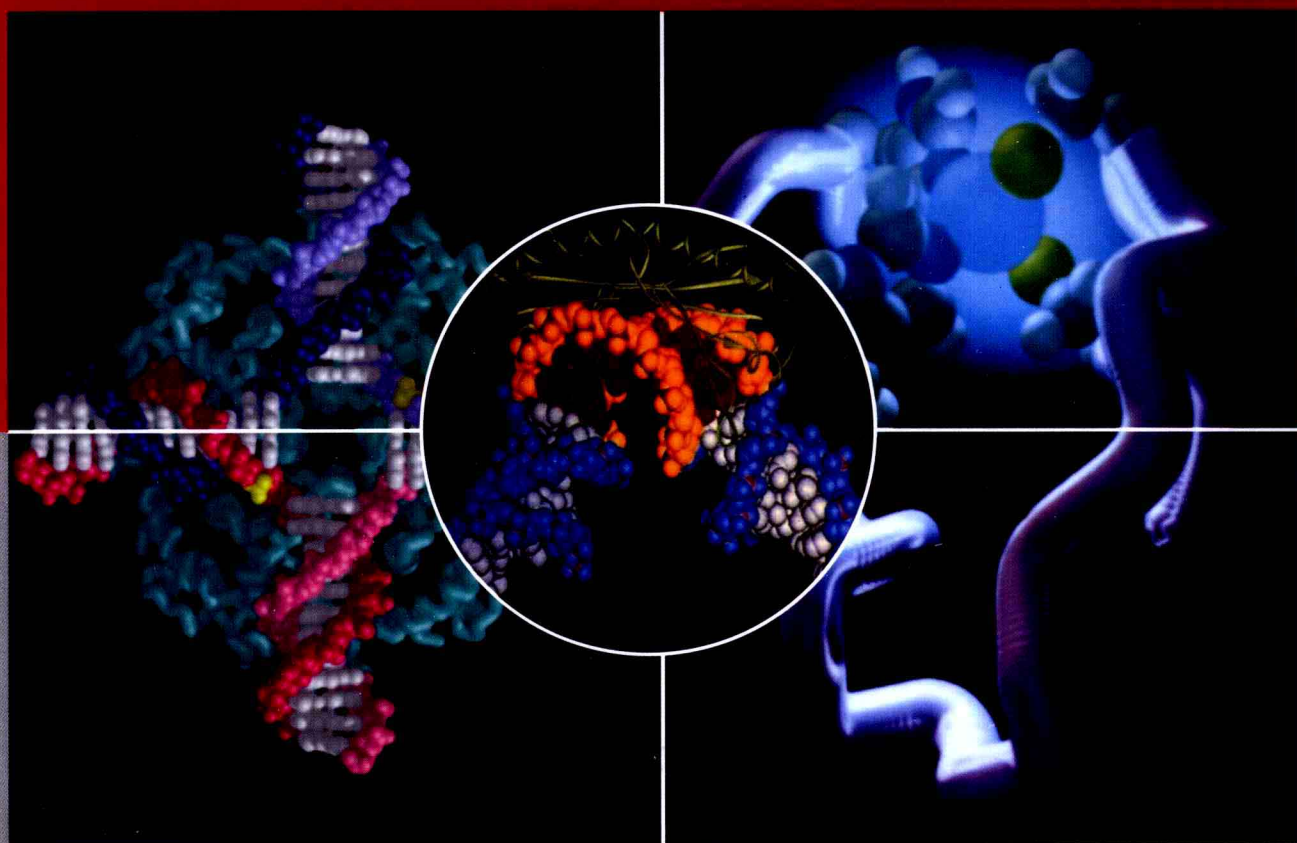


普通高等教育“十一五”国家级规划教材

分子生物学教程 (第三版)

赵亚华 编著

MOLECULAR BIOLOGY COURSE (THIRD EDITION)



科学出版社

普通高等教育“十一五”国家级规划教材

分子生物学教程

Molecular Biology Course

(第三版)

赵亚华 编著

科学出版社

北 京

内 容 简 介

本书从分子生物学的定义出发,以 DNA 和 RNA 这两类生物大分子为主线,由浅入深地讲述了这些大分子的结构与功能及其基因的复制、转录、转录后加工、翻译、原核基因与真核基因的表达调控等过程。本书共分 12 章,以较简明的形式概括了分子生物学的核心内容,既全面地阐述了分子生物学的基本理论,又突出介绍了学科发展的前沿研究。在本课程重要的知识点后都附有小结,每章末尾也做了重点归纳,以便读者能够快速搜索查找到相关的重点内容,使教师在教学中能掌握要点,学生在复习考试中能抓住重点。

本书可作为综合性大学、医科大学、师范院校和农林学科院校生命科学本科生、研究生的分子生物学的教材,也可作为生命科学类的研究人员、教师等的参考书。

图书在版编目(CIP)数据

分子生物学教程=Molecular Biology Course/赵亚华编著. —3 版.

—北京:科学出版社,2011

(普通高等教育“十一五”国家级规划教材)

ISBN 978-7-03-031689-9

I. ①分… II. ①赵… 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 118236 号

责任编辑:王国栋 王 玥/责任校对:钟 洋

责任印制:张克忠/封面设计:科地亚盟图文设计有限公司

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>

北京天竺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

*

2004 年 6 月第 一 版 开本:A4(890×1240)

2006 年 9 月第 二 版 印张:23 1/2 插页:1

2011 年 7 月第 三 版 字数:728 000

2011 年 7 月第一次印刷 印数:1—5 000

定价:49.00 元

(如有印装质量问题,我社负责调换)

第三版前言

在生命科学教学、研究与人才培养的全部过程中,分子生物学是十分重要的专业基础课之一。随着相关学科理论与技术研究的不断深入,分子生物学的研究领域不断扩大,学科进展也非常迅速。因此,全面系统地掌握分子生物学的知识非一日之功可成,而配备一本好的教科书,使教师和学生能在有限的教学时间内较充分地学习这门课程的核心知识,无疑是非常重要的。本书以 DNA 和 RNA 这两类生物大分子为主线,深入浅出地描述了这些大分子物质的结构与功能及其复制、转录、翻译和表达调控的全部过程。在绪论部分择要介绍了分子生物学的发展历程,内容涉及广泛,从宏观的生命科学发展整体水平上力求使学生认识到该学科在生命科学中的重要地位,并了解生命科学发展的前沿。

本书作者在 2004 年编写了《分子生物学教程》第一版,但由于这门学科发展十分迅速,原书的内容已不能满足教学和参考的需要。2006 年,作者在第一版编写基础上,参考了国内外的一些优秀教科书,博采众长,并根据平时的教学积累和学生用书需求,重新全面修订撰写了第二版《基础分子生物学教程》。第二版出版后,得到了广大教师和学生的广泛认同,在经历了 6 年教学实践的基础上,作者继续参阅近年来国内外同类教材和书籍,仔细比较,汲取精华,增加了较多的量化描述,枚举了很多国际上经典的实例,对许多图表做了细致明确的标示和修改,以增强自明性。同时,在多届学生中做了书面用书调查,听取意见和建议,在此基础上,全面修订改版形成了第三版。与第二版相比,第三版删去了生物大分子结构与功能和 DNA 重组技术这两章,同时对原来的章节做了许多必要的充实并加入了大量实例,如增加了对非编码 RNA 的重要性认识,核糖体开关、RNAi 干扰和 microRNA 的研究进展,两种新的碱基切除修复机制,几个选择性剪接的重要例子,转录及转录激活因子的调控等内容。在叙述体裁上,全书系统地在重要的知识点后都附有简要小结,每章末尾的本章重点归纳也进行了相应升级。新版既有利于读者能够快速查找相关的重点概念和内容,也便于教师在教学中能够掌握要点,并使学生在复习过程中能够撷取重点。第三版的修订比例约占第二版的 35%。

分子生物学的内容十分丰富而繁杂,本书的特点是将最基本而重要的科学内容汇集于一体,条理清晰,语言精练,取材新颖,概念陈述准确、知识点全面、信息量大、实例丰富,具有很强的可读性、可参考性和适用性,不啻为指导学生快速掌握分子生物学基础知识的理想教材。由于分子生物学发展迅速,资料浩瀚,且国内外从事这方面工作的专家学者很多,关注的侧重点各有不同,本书难以面面俱到,疏漏和不足之处在所难免,敬请广大读者批评指正,不胜感激。

本书的再版被列入高等教育“十一五”国家级规划教材和华南农业大学“十一五”教材建设的规划项目,得到了科学出版社,华南农业大学教务处、生命科学学院及许多同行、同仁及高年级学生的大力支持,作者在此一并表示衷心的感谢!

作 者

2011 年 3 月于广州五山

第一版前言

近半个世纪以来,生命科学领域取得了举世瞩目的重大成就和进步,而分子生物学是生命科学中发展最迅速,取得成果最多的学科之一。分子生物学是研究核酸、蛋白质等生物大分子的结构与功能,并从分子水平上阐述蛋白质与核酸、蛋白质与蛋白质之间相互作用的关系及其基因表达调控机理的科学。人类对生物学的研究最早从研究动物、植物的形态、解剖和分类开始,到以后对细胞学、遗传学、微生物学、生理学、生物化学的研究进入了细胞水平。从20世纪50年代以来,以生物大分子为研究目标的分子生物学开始逐步形成了独立的学科,并迅速成为现代生物学领域中最具活力的科学。随着相关学科不断发展,生物学与其他学科相互之间的渗透越来越深入,物理学和化学的理论、术语和方法不断地用于生物学的研究。目前,科学家已经建立了一整套分子生物学研究的方法、系统和一般的逻辑推理原则,使分子生物学的研究能迅速地向深度和广度发展。

目前,分子生物学已经深入到生物学科各个领域之中,并正在产生一系列新的分子科学,改变了或正在改变着整个生物学的面貌,其研究成果已在工业、农业、医学以及生物制药等领域得到了广泛的应用。在生命科学的教学活动中,分子生物学是生命科学领域十分重要的专业基础课之一。随着相关学科研究的不断深入,分子生物学的内容愈来愈多,因此,真正全面系统地掌握分子生物学的知识并不容易,而具备一本好的教科书,使教师和学生能在有限的教与学的时间内较好地掌握这门课程的内容,无疑是非常重要的,作者综合了国内外一些最新教科书,博采众长,根据平时的教学积累和目前学生用书现状,编写了这本教程。

本书以DNA和RNA两类生物大分子为主线,深入浅出地介绍了这些大分子的结构与功能及其复制、转录、翻译和表达调控。由于病毒分子生物学的内容十分丰富,而DNA重组技术是分子生物学领域的重要部分,考虑到全书的系统性与先进性,本书将这两章的内容也做了简单介绍。在绪论中,简要介绍了分子生物学的发展历程,内容涉及广泛,从宏观的生命科学发展整体水平上,力求使学生认识到分子生物学在生命科学中的地位,并了解生命科学的前沿。本书共分13章,以简明的形式概括了分子生物学的核心内容,既全面重点地阐述了分子生物学的基本理论,又突出介绍了学科发展的前沿动态。

分子生物学的内容十分丰富而繁杂,本书的特点是将最基本而核心的内容汇集一体,取材新颖、简明扼要、条理清晰、语言精练、基本概念准确、知识点全面、信息量大,具有一定的可读性和适用性,是指导学生在有限的时间掌握分子生物学基础知识的理想教材。由于分子生物学发展迅速,资料浩瀚,专家学者们关注的侧重点往往不同,因此本书难以面面俱到,疏漏之处在所难免,敬请广大读者批评指正,不胜感激。

本书的编写与出版被列入华南农业大学“十五”教材建设的规划项目,得到了教务处、生命科学学院以及教研室同仁的大力支持,在此表示衷心感谢!

作 者

2003年8月于广州五山

目 录

第三版前言

第一版前言

第 1 章 绪论	1
1.1 分子生物学的概念	1
1.2 分子生物学研究的主要内容	1
1.2.1 基因与基因组的结构与功能	1
1.2.2 DNA 的复制、转录和翻译	1
1.2.3 基因表达调控的研究	2
1.2.4 DNA 重组技术	2
1.2.5 结构分子生物学	2
1.3 分子生物学与生物化学之间的关系	2
1.4 分子生物学发展的历程	3
1.4.1 人类对 DNA 和遗传信息传递的认识阶段	3
1.4.2 重组 DNA 技术的建立和发展阶段	3
1.4.3 重组 DNA 技术的应用和分子生物学的迅猛发展阶段	4
1.5 21 世纪分子生物学发展的趋向	5
1.5.1 功能基因组学	5
1.5.2 蛋白质组学	6
1.5.3 生物信息学	6
【本章重点归纳】	7
【思考题】	7
第 2 章 核酸的结构与功能	8
2.1 细胞内的遗传物质	8
2.1.1 DNA 是主要的遗传物质	8
2.1.2 RNA 也是遗传物质	9
2.2 核酸的化学组成与共价结构	9
2.2.1 核酸的化学组成	9
2.2.2 多聚核苷酸的结构	11
2.3 DNA 的高级结构与功能	11
2.3.1 双螺旋模型特征	11
2.3.2 DNA 高级结构的其他形式	12
2.3.3 DNA 结构的动态性与精细结构	16
2.3.4 DNA 的超螺旋结构与拓扑学性质	18
2.4 真核生物的染色体及其组装	20
2.4.1 真核生物的染色体	20
2.4.2 染色体中的蛋白质	21

2.4.3 核小体的形成	22
2.4.4 染色质的高级结构	22
2.5 RNA 的结构与功能	23
2.5.1 RNA 的结构特点及与 DNA 的区别	24
2.5.2 RNA 在细胞中的分布	24
2.5.3 细胞中 RNA 分类概述	25
2.6 核酸的变性、复性与分子杂交	28
2.6.1 核酸的变性	28
2.6.2 核酸的复性与分子杂交	30
【本章重点归纳】	32
【思考题】	33
第 3 章 基因与基因组的结构与功能	34
3.1 基因的概念	34
3.1.1 基因与 DNA 的关系	34
3.1.2 基因与多肽链的关系	35
3.2 基因的命名	36
3.3 基因组	37
3.3.1 基因组的概念	37
3.3.2 基因及基因组的大小与 C 值矛盾	37
3.4 病毒及其基因组	39
3.4.1 病毒基因组一般特点	39
3.4.2 病毒的核酸	39
3.4.3 噬菌体基因组	40
3.4.4 几种病毒的基因组	42
3.5 细菌基因组	44
3.5.1 细菌基因组的一般特点	44
3.5.2 细菌的染色体基因组	45
3.6 真核生物基因组	46
3.6.1 真核生物基因组的特点	46
3.6.2 真核生物基因组的结构	46
3.6.3 线粒体基因与基因组的结构	59
3.6.4 叶绿体基因与基因组的结构与功能	60
3.6.5 人类基因组简介	61
【本章重点归纳】	72
【思考题】	75
第 4 章 DNA 的复制	76
4.1 DNA 复制概述	76
4.1.1 DNA 复制的一些概念	77

4.1.2 复制方向	79	6.3.1 转座子的概念	125
4.1.3 复制方式	80	6.3.2 转座子的分类	125
4.1.4 DNA 复制的酶体系	82	6.3.3 转座子的转座机制	126
4.1.5 DNA 的半不连续复制	86	6.3.4 转座子转座的基本特征	128
4.1.6 DNA 合成的保真性	87	6.3.5 DNA 转座引起的遗传学效应	129
4.1.7 DNA 拓扑异构酶	87	6.3.6 真核生物的转座子	129
4.1.8 单链 DNA 结合蛋白	89	6.4 逆转录转座子	132
4.2 细菌 DNA 复制的机制	89	【本章重点归纳】	136
4.2.1 大肠杆菌复制的起始	89	【思考题】	137
4.2.2 真核生物 DNA 复制的引发	91	第 7 章 RNA 的转录合成	138
4.2.3 大肠杆菌复制的延伸	91	7.1 RNA 转录概述	138
4.2.4 复制的终止	94	7.1.1 RNA 转录的一般特点	138
4.3 真核生物 DNA 的复制	95	7.1.2 原核生物和真核生物基因转录的差异	139
4.3.1 真核生物的 DNA 聚合酶	95	7.2 启动子的结构与功能	139
4.3.2 真核生物染色体端粒的复制	96	7.2.1 启动子的结构	139
4.4 原核细胞 DNA 复制的调控	98	7.2.2 启动子的功能	140
4.4.1 大肠杆菌染色体 DNA 的复制调控	98	7.3 细菌的 RNA 聚合酶	142
4.4.2 ColEI 质粒 DNA 的复制调控	98	7.3.1 RNA 聚合酶概述	142
4.4.3 R6K 质粒 DNA 的复制调控	99	7.3.2 大肠杆菌的 RNA 聚合酶	142
4.4.4 单链 DNA 噬菌体的复制调控	99	7.3.3 T7 RNA 聚合酶	143
4.4.5 λ 噬菌体 DNA 的复制调控	99	7.3.4 σ 因子的结构与功能	143
4.5 真核生物 DNA 复制调控简述	100	7.3.5 核心聚合酶的结构与功能	145
4.5.1 病毒 SV40 DNA 的复制调控	100	7.3.6 RNA 聚合酶全酶的结构与功能	146
4.5.2 腺病毒 DNA 的复制调控	101	7.3.7 原核生物 RNA 的转录过程	150
4.5.3 酵母染色体 DNA 的复制调控	101	7.4 真核生物的 RNA 聚合酶及其转录	155
【本章重点归纳】	101	7.4.1 真核生物基因转录概述	155
【思考题】	103	7.4.2 真核生物基因转录的 RNA 聚合酶	156
第 5 章 DNA 的损伤、修复和基因突变	104	7.5 真核基因转录的启动子	159
5.1 DNA 的损伤	104	7.6 类型 II 基因转录的转录因子	163
5.1.1 DNA 分子的自发性损伤	104	7.7 类型 II 基因转录起始复合物的装配	167
5.1.2 物理因素引起的 DNA 损伤	105	7.8 类型 I 和 III 的转录因子	169
5.1.3 化学因素引起的 DNA 损伤	105	7.9 RNA 转录的抑制	173
5.2 DNA 的修复	106	【本章重点归纳】	174
5.3 基因突变	111	【思考题】	177
【本章重点归纳】	114	第 8 章 RNA 转录的剪接与加工	178
【思考题】	114	8.1 原核生物 RNA 的转录后加工	178
第 6 章 DNA 的重组与转座	115	8.1.1 原核生物 rRNA 前体的加工	178
6.1 同源重组	115	8.1.2 原核生物 tRNA 前体的加工	179
6.1.1 同源重组的分子模型	115	8.1.3 原核生物 mRNA 前体的加工	180
6.1.2 同源重组的酶学分子机制	117	8.2 真核生物 RNA 的加工	181
6.1.3 酵母的减数分裂重组	119	8.2.1 真核生物 tRNA 前体的转录后加工	181
6.1.4 异源双链与基因转换	120	8.2.2 真核生物 rRNA 前体的转录加工	182
6.1.5 细菌的基因转移与 DNA 重组	121	8.2.3 细胞核 mRNA 前体剪接概述	184
6.2 特异位点重组	122		
6.3 DNA 的转座	125		

8.2.4 细胞核 mRNA 前体剪接机制和过程	185	10.11 重叠基因的调控作用	279
8.2.5 真核生物 mRNA 前体的选择性剪接	193	10.12 细菌中 DNA-蛋白质的相互作用	279
8.2.6 RNA 的自我剪接	195	10.13 综合实例——噬菌体基因的表达调控	284
8.2.7 核酶	199	10.13.1 噬菌体的生活周期	284
8.3 反式剪接	200	10.13.2 噬菌体裂解过程中基因表达调控是级联反应	285
8.4 mRNA 5'端加帽	201	10.13.3 噬菌体 SP01-替换 σ 亚基改变宿主的转录对象	286
8.5 mRNA 3'端的多聚腺苷酸化	203	10.13.4 T4 噬菌体-修饰核心酶并替换 σ 亚基改变宿主的转录对象	286
8.6 多聚腺苷酸化作用的机制	204	10.13.5 T7 噬菌体——RNA 聚合酶的代换	286
8.7 前体 mRNA 的剪切和多聚腺苷酸化	205	10.13.6 λ 噬菌体基因组的表达调控	287
8.8 mRNA 加工事件的协同运作	206	【本章重点归纳】	293
8.9 RNA 的编辑	210	【思考题】	294
8.9.1 RNA 编辑的机制	210	第 11 章 真核生物的基因表达调控	295
8.9.2 RNA 编辑的类型	212	11.1 真核基因表达调控的特点	295
8.9.3 RNA 编辑的生物学意义	213	11.2 真核细胞基因表达调控的不同层次	296
8.10 RNA 的再编码	213	11.3 DNA 染色体水平的调控	297
8.11 RNA 干扰	214	11.4 DNA 水平上的调控	307
【本章重点归纳】	218	11.5 真核基因转录水平的调控	308
【思考题】	220	11.6 基因表达的转录后水平的调控	318
第 9 章 遗传密码与蛋白质的生物合成	221	11.7 转录因子对基因表达的调控	320
9.1 遗传密码的破译	221	11.7.1 细胞对转录因子的调控	320
9.2 遗传密码的基本特性	222	11.7.2 转录激活因子的类型与结构	325
9.3 蛋白质的生物合成	226	11.7.3 转录激活因子结合 DNA 的结构基序	327
9.3.1 概述	226	11.7.4 与 DNA 相互作用的其他蛋白质因子	332
9.3.2 蛋白质生物合成的分子基础	226	11.7.5 调控蛋白对特异 DNA 序列的识别	334
9.3.3 蛋白质生物合成的过程	230	11.8 甾体激素对基因转录的调控	334
9.3.4 蛋白质合成的抑制	244	11.8.1 甾体激素的应答元件	334
9.3.5 蛋白质合成的调节	244	11.8.2 甾体激素对基因转录的调控	335
9.4 蛋白质合成后的运输	249	11.9 RNA 结合蛋白对基因表达的调控	335
9.5 蛋白质前体的共价修饰	253	【本章重点归纳】	337
9.6 蛋白质的折叠	254	【思考题】	338
【本章重点归纳】	255	第 12 章 病毒的分子生物学简介	339
【思考题】	256	12.1 病毒基因组的一般结构	339
第 10 章 原核生物基因表达调控	257	12.2 病毒基因组的复制	339
10.1 基因表达调控概述	257	12.3 逆转录病毒	342
10.2 原核基因表达调控的若干概念	258	12.4 逆转录的生物学意义	345
10.3 乳糖操纵子的调控	260	12.5 腺病毒	346
10.4 阿拉伯糖操纵子的调控	267	12.6 丙型肝炎病毒	348
10.5 色氨酸操纵子的调控	270		
10.6 受双启动子调控的半乳糖操纵子	274		
10.7 组氨酸操纵子的调控	275		
10.8 细菌的应急反应	275		
10.9 正调控系统和负调控系统	276		
10.10 受多重启动子调控的操纵子	278		

12.7 艾滋病与 HIV	349	【本章重点归纳】	356
12.8 病毒对宿主细胞的影响	353	【思考题】	356
12.9 病毒与肿瘤发生	354	参考文献	357
12.10 病毒的基因工程疫苗及病毒载体	354	索引	359
12.11 亚病毒	354	彩图	



1.1 分子生物学的概念

分子生物学(molecular biology)是研究核酸、蛋白质等生物大分子的结构与功能,并从分子水平上阐述这些大分子之间相互作用的关系及其基因表达调控机理的科学,是人类从分子水平上真正揭开生物世界的奥秘,由被动地适应自然界转向主动地改造和重组自然界的基础学科。

人类对生物学的研究经历了相当漫长的历程。最早从研究动物和植物的形态解剖和分类开始,到以后对细胞学、遗传学、微生物学、生理学以及生物化学的研究,逐步进入了细胞水平。从 20 世纪 50 年代以来,以生物大分子为研究目标的分子生物学开始逐步形成独立的学科,并迅速成为现代生物学领域中最具活力的科学。随着相关学科不断发展,生物学与其他学科之间的相互渗透越来越深入,物理学、化学和电子计算机的理论、术语和方法不断地用于生物学的研究。目前,科学家已经建立了一整套分子生物学研究的方法、系统和一般的逻辑推理原则,以及数目十分庞大的分子生物信息数据库等,使分子生物学的研究迅速地向着纵深发展。

广义上讲的分子生物学包括对蛋白质和核酸等生物大分子结构与功能的研究,以及从分子水平上阐明生命的现象和生物学规律。例如,蛋白质的结构、运动和功能、酶的作用机理和动力学、膜蛋白结构与功能和跨膜运输,等等。从这个视角看,分子生物学几乎包括了生物学领域的很多方面,但实际上这些内容随着其研究的深入已逐步发展形成了各自独立的学科。由此通常采用狭义的概念,将分子生物学的定义偏重于对核酸(基因)的分子生物学范畴,主要研究基因或 DNA 结构与功能、复制、转录、表达和调节控制等的分子过程,其中也涉及与这些过程相关的蛋白质和酶的结构与功能的研究。

小结:分子生物学是研究核酸和蛋白质等生物大分子的结构与功能、并从分子水平上阐述它们间相互作用关系及基因表达调控机理的科学,是人类由被动适应自然界转向主动改造和重组自然界的学科。

1.2 分子生物学研究的主要内容

分子生物学是研究所有生物学现象的分子基础。从这个意义上说,分子生物学包括了地球上所有的生物学。因此,对分子生物学研究内容的界定是困难的。随着科学技术突飞猛进地向纵深发展,各学科之间的划分越来越细,在长期的科学研究与实践中,对生物学中的许多分支与科目,分子生物学家并未将它列为分子生物学研究的范畴,例如,某些生物学反应就像一个标准的化学反应一样通过酶和产物浓度被调节,对这些反应调节的研究就属于生物化学的范围。但如果一个酶催化的反应是通过酶基因或酶分子结构的改变而被调节,则属于分子生物学的内容。这类似于把细胞内的化学成分的排列和结构的研究称为细胞生物学。但当人们分离到了昆虫和某些原虫的行为突变体后,就要对其进行分子生物学分析。因此,分子生物学与生物学其他各分支之间的界限越来越不明显了。尽管分子生物学涉猎的范围十分广泛,研究内容也包罗万象,但是按照狭义分子生物学定义,我们可将现代分子生物学的主要研究内容概括为以下几个大的方面。

1.2.1 基因与基因组的结构与功能

基因的研究一直是影响整个分子生物学发展的主线。在不同的历史时期对基因的研究有不同的内容,20 世纪 50 年代以前,主要从细胞染色体水平上进行研究,是基因的染色体遗传学内容;50 年代之后,主要从 DNA 大分子水平上进行研究,属于基因的分子生物学阶段。近 20 多年来,由于重组 DNA 技术的不断完善和应用,人们已经改变了从表型到基因型的传统研究基因的途径,而能够直接从克隆目的基因出发,研究基因的功能及其与表型的关系,使基因的研究进入了反向生物学阶段。在这个历程中,对基因与基因组的微细及高级结构与功能的研究始终是分子生物学研究内容最基础最重要的部分。

1.2.2 DNA 的复制、转录和翻译

这一方面研究的重点是 DNA 或基因怎样在各系

统相关的酶与蛋白等因子作用下,按照中心法则进行自我复制、转录、反转录和翻译。同时,对 mRNA 分子进行各种剪接、加工修饰、编辑以及对新生多肽链折叠成有功能的空间结构的分子机理研究。

1.2.3 基因表达调控的研究

基因表达的实质是遗传信息的转录和翻译。在生物个体的生长、发育和繁殖过程中,遗传信息的表达按照一定的时序发生变化(时序调节的表达);并且,随着内外环境的变化而不断地加以修正(环境调控表达)。

基因表达的调控主要发生在转录水平和翻译水平上。原核生物基因组和染色体结构都比真核生物简单,转录和翻译在同一时空内发生,基因表达调控主要发生在转录水平。真核生物有细胞核结构,转录和翻译过程在时间和空间上都被分隔开,且在转录和翻译后都有复杂的分子信息加工过程,其基因表达的调控发生在各种不同的水平,主要表现在对上游调控序列、信号传导、转录因子以及 RNA 剪辑等多个方面。

1.2.4 DNA 重组技术

分子生物学的核心是遗传信息的结构、传递和控制,在这个过程中 DNA 重组技术是不可缺少的手段之一。DNA 重组技术是 20 世纪 70 年代初兴起的一门科学技术。应用此技术能将不同的 DNA 片段进行人为的重组和定向连接,并指定在特定的受体细胞中与载体同时复制和表达,产生影响受体细胞的新的遗传性状。严格地说,DNA 重组技术并不完全等于基因工程,因为后者还包括其他能使生物细胞基因组结构发生改变的体系。DNA 重组技术是核酸化学、遗传学、细胞学、病毒学、蛋白质化学、酶工程以及微生物学等长期深入研究的结果,反过来,这些学科的发展又以 DNA 重组技术作为重要手段而进行。在这个过程中,限制性内切核酸酶、DNA 连接酶及其他工具酶的发现与应用是这一技术得以建立的关键。

作为分子生物学研究的内容之一,DNA 重组技术的主要目的是:①用于大量生产某些在正常细胞代谢中产量很低的多肽,如激素、抗生素、酶类及抗体等,提高产量,降低成本,使许多有价值的多肽类物质得到广泛的应用。例如,用于治疗艾滋病的基因工程白介素 12 可有效地阻止病情发展,恢复 HIV 病毒携带者的免疫系统和功能;②用于定向改造某些生物的基因组结构,使它们所具备的特殊功能更符合人类生活的需要,其经济价值能成百上千倍地提高。例如,一种含有分解各种石油成分的重组 DNA 超级细菌能快速分解石油,可用来恢复被石油污染的海域和土壤;③DNA 重

组技术用于进行基础研究,已经成为研究分子生物学领域一切基础性问题的技术方法和常规武器。

1.2.5 结构分子生物学

任何一个生物大分子当它在发挥生物学功能时都必须具备两个前提,一是必须拥有特定的空间结构(三维结构);二是在它发挥生物学功能的过程中必定存在着结构和构象的变化。结构分子生物学的发展就是研究生物大分子特定的空间结构以及结构的运动变化与其生物学功能关系的科学。它包括结构的测定、结构运动变化规律的探索和结构与功能相互关系三个方向的研究。近年来科学家不断研究发现了大量新的生物大分子静态结构,并逐步深化对其动态功能的认识。已经能够研究运动时间最短达 $10^{-15} \sim 10^{-12}$ s,运动幅度最小为 0.1pm 的分子运动。由单一分子研究深入到对复合乃至多亚基、多分子复合体研究,使对多种生物大分子联合起来的复杂结构与相互作用功能的认识与研究成为可能。对核糖体的结构与功能的研究就是这方面的一个例子。目前,研究生物大分子空间结构及其运动规律的手段主要是 X 射线衍射晶体技术、二维和多维核磁共振成像、电子衍射、中子衍射、电镜三维重组以及各种波谱学的方法。

结构分子生物学在今后仍然是生命科学发展的基础学科。在这一领域中,仍需要有生物学家、生物化学家、物理学家、化学家以及计算机和工程学的专家的共同努力。

小结:分子生物学主要研究内容包括基因与基因组的结构与功能;遗传信息的传递与表达;基因表达调控;DNA 重组技术;结构分子生物学等。

1.3 分子生物学与生物化学之间的关系

当代生命科学的一大特点是几乎所有关于生命的分支学科均已被分子生物学渗透,由此涉及了许多难以穷尽的方方面面难以界定。实际上,分子生物学与生物化学之间的关系是非常紧密而难以区分的。但随着这两门学科的研究向纵深发展,内容越来越多,科学家们不得不将它们做相对的划分。分子生物学的定义如前所述,它从分子水平上研究生命现象,生物化学是从分子水平上研究生命现象的化学本质;从学科范畴讲,分子生物学包括了生物化学;但从研究的基本内容上,例如,在遗传信息流从 DNA→mRNA→蛋白质的代谢传递过程中,许多内容又属于生物化学的范围,等等。因此,分子生物学与生物化学这两门学科是“你中有我”、“我中有你”,而不能截然分开。但这两门学科的研究方向和研究方法与手段也显示出了明显的

区别。

在研究方向上,分子生物学主要是研究蛋白质、核酸和其他大分子的结构与功能,以及它们之间的相互作用,着重解决细胞中的信息传递和代谢调节等问题。而生物化学主要研究大分子物质的组成、性质结构与功能及其在生命活动中的代谢转化等动态的过程,包含大量有机小分子的参与。因此分子生物学与生物化学虽然在研究内容上有相同之处,但在研究方向上,分子生物学的着重点是大分子的结构与功能,而生物化学则以生物分子的动态代谢转化为主。

在研究方法上,分子生物学是以化学和物理学的方法研究大分子结构,采用生物化学与遗传学相结合的方法探索其功能,解决大分子结构与功能及其代谢调节的关系。而生物化学主要采用生物化学与化学以及生理学的方法,探索生命的化学过程,解决分子转化与能量转换的问题。所以,分子生物学与生物化学,在分离、纯化生物分子时,或许采用同样的方法,而在分别探索其研究的问题时,却采用了许多不同的手段。

1.4 分子生物学发展的历程

分子生物学在人类文明史上的光辉成就,以前所未有的速度推动着生物学的发展,使整个生物学的面貌发生了巨大的变化。由于无数分子生物学家的不懈追求与刻苦研究,使我们现在不但能从分子水平上认识了核酸的结构与功能以及复制、转录、翻译、剪接、加工、修饰等的详细过程,而且已经测知了许多重要生物的基因组及其结构与功能,真正从分子水平上对这些基因控制的生长、发育和变异等一系列生物学问题有了更深入的了解,获得了令人振奋的结果。为了从学习和认识上的条理更加清晰,在此将其发展过程简单地概括为3个阶段。

1.4.1 人类对DNA和遗传信息传递的认识阶段

1928年,F Griffith做的肺炎双球菌的转化试验奠定了DNA是遗传物质的基础。

1944年,Oswald Avery等用生物化学和物理化学手段对F Griffith的肺炎双球菌转化试验进一步做了分析,证实DNA是生物的遗传物质。这一重大发现打破了长期以来许多生物学家认为的只有像蛋白质那样的大分子才能作为细胞遗传物质的观点,在遗传学上树立了DNA是遗传信息载体的理论。

1950年,E Chargaff提出了DNA碱基组成的等比例规律。与此同时R Hotchkiss对Avery的转化物做了纯化,进一步证实了高纯度的DNA是遗传物质。

1952年,Hershey和Martha Chase用同位素示踪

技术,将T2噬菌体侵染大肠杆菌细胞,证实了主要是核酸进入细菌体内,而病毒外壳蛋白留在细胞外,且进入菌体的DNA能利用细菌的生命过程合成噬菌体自身的DNA和蛋白质,并能自我组装成与亲代完全相同的子代噬菌体。烟草花叶病毒的重建实验也证明,病毒蛋白质的特性由RNA决定,即遗传物质是核酸而不是蛋白质。至此,DNA作为遗传物质才被普遍地接受。

1953年,是开创生命科学新时代具有里程碑意义的一年,Watson和Crick发表了“脱氧核糖核酸的结构”的著名论文,提出了DNA双螺旋结构模型,为人类充分揭示遗传信息的传递规律奠定了坚实的理论基础。同年,Sanger历经8年,完成了第一个蛋白质——胰岛素的氨基酸全序列分析。

1954年,Crick在前人研究工作基础上,提出了中心法则理论;Gamnow从理论上研究了遗传密码的编码规律;1961年M N Johann和H Matthaai做了一个突破性试验,破译了第一批遗传密码,对正在兴起的分子生物学研究起了重要推动作用,将永载史册。

1957年,A Kornberg在大肠杆菌中发现了DNA聚合酶I,这是能在试管中合成DNA的第一种核酸酶,这种酶已被广泛用于制备标记的DNA探针。

1958年,M Meselson和F Stahl通过经典的同位素¹⁵N标记DNA的CsCl密度梯度超速离心试验对DNA复制的3种可能的机制进行了辨别。

1940~1965年,F Jacob和J Monod等历经大量试验验证,提出了著名的原核基因表达操纵子学说。

1.4.2 重组DNA技术的建立和发展阶段

1967年,Gellert发现了DNA连接酶。

1970年,Smith和Wilcox等分离到第一种限制性内切核酸酶。同年Temin和Baltimore在RNA肿瘤病毒中发现了逆转录酶,证实了Temin 1964年提出的“前病毒假说”。逆转录酶已成为目前分子生物学研究中的一个重要工具。

1972~1973年,重组DNA时代到来。H Boyer和PBerg等发展了重组DNA技术,并完成了第一个细菌基因的克隆,开创了基因工程的新纪元。

1975年,Southern发明了DNA片段印迹法;Gruenstein和Hogness建立了克隆特定基因方法;O'Farrell发明了双向电泳的蛋白质分析方法。Blohel等提出了信号肽假说。

1975~1977年,Sanger,Maxam和Gilbert发明了DNA序列测定技术。

1977年,第一个全长5387bp的噬菌体ΦX174基

基因组序列测定完成。

1.4.3 重组 DNA 技术的应用和分子生物学的迅猛发展阶段

1979 年, Solomon 和 Bodmer 最先提出至少 200 个限制性片段长度多态性(RFLP)可作为连接人类基因组图谱的基础。

1980 年, Wigler 等把非选择性基因导入哺乳动物细胞; Cohen 和 Boyer 获得一项克隆技术的美国专利。

1981 年, Cech 等发现四膜虫 26S rRNA 前体的自我剪接作用, 具有催化作用 RNA(核酶)的发现, 促进了 RNA 研究的飞速发展; 同年, Palmiter 等获得转基因小鼠; Spradling 等培育出转基因果蝇。

1982 年, Prusiner 等在感染瘙痒病的仓鼠脑中发现了朊病毒(prion); 同年, 第一个由基因工程生产的药物——胰岛素在美国和英国获准使用; Sanger 及其合作者完成 λ 噬菌体全长 48 502bp 的基因组 DNA 全序列测定。

1985 年, Saiki 等发明了聚合酶链式反应(PCR); Sinsheimer 首先提出人类基因组图谱制作计划设想; Smith 等报导了 DNA 测序中应用荧光标记取代同位素标记的方法; Miller 等发现 DNA 结合蛋白的锌指结构。

1989 年, Greider 等在纤毛原生动物中发现了端粒酶是以内源性 RNA 为模板的逆转录酶; Hiatt 等首次报道了在植物中亦可产生单克隆抗体。同年, 美国专利商标局宣布将接受基因工程植物和基因工程动物方面的专利申请。第一个具有专利权的用于医药研究的动物——杜邦肿瘤鼠诞生。

1990~1992 年, 转基因玉米及转基因小麦诞生, 谷物基因工程开始变为现实。

1992 年, 欧洲共同体各国 35 家实验室发表第一个真核生物染色体(酵母 3 号染色体)DNA 全序列共 315kb。

1990 年, 人类基因组计划全面正式启动; Simpson 等发现了对 mRNA 前体编辑起指导作用的小分子 RNA; Sinclair 等在人类 Y 染色体上发现了新的性别决定基因。

1991 年, 由欧洲共同体组织 17 个国家 35 个实验室的 147 位科学家, 手工测序完成了第一条完整染色体的测序工作。

1994 年, 日本科学家在 *Nature Genetics* 上发表了水稻基因组遗传图; Wilson 等完成了线虫 3 号染色体连续的 2.2Mb 的测定, 预示着百万碱基规模的 DNA 测序时代的到来。

1995 年, S Zimmerly 等发现 II 类内含子的结构及剪接机制。

1996 年, W Dietrich 等绘制了小鼠基因组的完整遗传图谱。

1997 年, Wilmut 等首次不经过受精, 用成年母羊的体细胞遗传物质成功获得克隆羊——Clone Sheep Dolly; Willard 等首次构建了人染色体(HACs); Salishury 等发现 DNA 一种新的结构形式——四显性组合是基因交换期间 DNA 联结的一种方式。

1998 年, 第一个多细胞真核生物线虫基因组被揭示; 美国科学家 James T 等用 ES 和 EG 细胞建立了胚胎干细胞系, 为研究胚胎干细胞的发育和利用胚胎干细胞治疗疾病提供了全新的途径; D Doyle 等勾画出了清晰的第一个钾离子通道结构, 这是分子生物学上的一个重要进展。

1999 年 J Cate 等第一次绘制完整核糖体晶体结构, 并揭示了很多细节; 国际人类基因组计划(HGP)研究小组完成了人第 22 号染色体测序工作。

2000 年, 国际 HGP 联合研究小组完成了人类第 21 号染色体的测序, 从原定预计的 2003 年 6 月提前到 2001 年 6 月; 果蝇和拟南芥的基因组测序完成; 由 Alan Coleman 领导的研究小组完成了世界首例克隆猪在苏格兰诞生; 中国超级杂交水稻基因组计划”正式启动; 2000 年 3 月, 塞莱拉公司宣布完成果蝇基因组测序; 2000 年 12 月, 英国、美国等国科学家宣布绘出拟南芥基因组的完整图谱, 这是人类首次全部破译一种植物的基因序列。此专题有多篇重要文章发表。

2001 年, *Nature* 和 *Science* 同时发表了 HGP 全序列。美国科学家 Hartwell、英国科学家 Hunt 和 Nurse 因对细胞周期调控因子的研究而分享诺贝尔生理学奖。

2002 年 4 月 5 日, 以杨焕明为首的中国科学家在 *Science* 上发表了水稻(籼稻)全基因组框架序列图。

2003 年, 美国、日本、英国、法国、德国、俄罗斯和中国科学家共同宣布 HGP 序列图绘制成功, HGP 目标基本完成; 中国内地、台湾、香港科学家宣布联手启动“中华人类基因组单体型图”计划; 中国、美国分别测定出非典型肺炎病毒的基因图谱。

2004 年, 以色列学者 Aaron Ciechanover 和 Avram Hershtko 和美国 Irwin Rose 发现泛素蛋白调节的蛋白质降解机制。

2005 年, 由美国、中国、日本等国的 200 多位学者参加的“国际人类基因组单体型图计划”取得重要成果, 公布了人类基因组“差异图”, 发现了 100 多万个常见 SNP 位点, 标定了单体型“模块”在 DNA 链上的“边

界”,并划分了基因组上包含最常见 DNA 变异的 10 个区域。利用这份“差异图”,在糖尿病、早老性痴呆症和癌症等疾病的研究中,将患者与健康人全基因组的 SNP 进行比较,更高效地寻找与疾病相关的基因变异。

2006 年,美国 Andrew Z Fire 和 Craig C Mello 发现了 RNAi 干扰机制,目前已被广泛用作研究基因功能的一种手段,有助于诊断和治疗遗传病;英国 The Wellcome Trust Sanger Institute and colleagues 和美国杜克大学人类遗传学中心等研究机构的研究人员联合在 *Nature* 上发表人类最后一个染色体——1 号染色体基因测序。在人体全部 22 对常染色体中,1 号染色体包含基因数量最多,破译难度最大。

2007 年,美国犹他大学人类遗传学研究所 Mario R Capecchi 和北卡罗来纳州大学 Oliver Smithies 与英国 Martin J Evans 对胚胎干细胞的研究获得诺贝尔生理学奖。

2008 年,日本、美国和美籍华裔科学家研究绿色荧光蛋白作出特殊贡献而获奖;*Science* 杂志评出的 2008 年十大科学进展中,细胞重编程被评为第一位,这项研究几乎在一夜之间开启了一个生物学的新领域,它有望成就挽救生命的医学上的重大进步。细胞重编程质疑细胞单向发育的理论,认为已分化的细胞在特定条件下能被逆转,恢复到全能性状态,或形成胚胎干细胞系,或进一步发育成一个新个体。通过插入回拨细胞发育时钟的基因,科学家正在深入了解疾病和研究细胞如何决定其命运的生物学。

2009 年, Venkatraman Ramakrishnan、Thomas A Steitz 及以色列 Ada E Yonath 3 位科学家对核糖体的结构和功能的研究而获得诺贝尔化学奖;3 位美国学者 Elizabeth H Blackburn、Carol W Greider 及 Jack W Szostak 被授予诺贝尔生理学奖,表彰他们发现端粒和端粒酶对染色体的保护作用。

现代分子生物学的研究成就还包括以下几个方面的重要突破:①对人类基因组计划的实施取得了多项重大成果,建立了多种不同类型和层次的文库;各类分子探针如限制性内切核酸酶片段长度多态性(RFLP)、扩增片段长度多态性(AFLP)、随机扩增片段长度多态性(RAPD)等的应用;聚合酶链式反应的普及应用和日益高度自动化并由机器人操作的核酸自动测序仪,以及大容量先进的计算机分析存储技术,等等。②在互联网上注册公布的有详细注释的许多生物如噬菌体、大肠杆菌、酵母、果蝇、线虫、小鼠、水稻、拟南芥等的全基因组序列,以及大量在过去认为根本无法得到的蛋白质序列,为研究者提供了共享生物分子信息资源的绝好平台。③对肿瘤、艾滋病、心血管疾病、高血

压以及糖尿病等的防治,从分子机理上已取得了可喜的成果。④农作物抗病虫害、动植物品种和品质改良、抗病毒病等方面已逐步走向实用阶段。

综合以上不难看出,20 世纪以核酸研究为核心,带动着分子生物学向纵深发展。50 年代的双螺旋结构,60 年代的操纵子学说,70 年代的 DNA 重组,80 年代的 PCR 技术,90 年代的 DNA 测序都是分子生物学发展的里程碑,将生命科学带向一个由宏观到微观,再到宏观的时代。

1.5 21 世纪分子生物学发展的趋向

人类基因组 DNA 测序计划已经提前完成了,人类基因组研究的重点正在由序列结构向基因功能转移,进入了以基因组功能研究为主要内容的“后基因组”(post-genomics)时代。主要任务是研究细胞内全部基因的表达图式和全部蛋白图式,或者说是“从基因组到蛋白质组”。由此,分子生物学的重点又回到了蛋白质上来,使生物信息学应运而生,生命科学正在进入一个全新的时代。

1.5.1 功能基因组学

功能基因组学(functional genomics)依赖于对 DNA 序列的认识。应用基因组学的知识和工具,人们能够了解和认识影响整个生命过程的特定序列表达谱。例如,酿酒酵母(*S. cerevisiae*)16 条染色体全部序列的测序工作已于 1996 年完成,基因组全长 13 040kb,含有 5885 个可能编码蛋白质的基因,140 个编码 rRNA 基因,40 个编码 snRNA 基因和 275 个 tRNA 基因,共计 6340 个基因。功能基因组学就是进一步研究这些基因在一定条件下,如在孢子形成期,同时有多少基因协同表达才能完成这一发育过程?这就需要研究这一时期的全套基因表达谱(gene expression pattern)。解决如此复杂的问题必须在方法学上有重大突破,创造出高效、快速地同时测定基因组成千上万个基因活动的方法。

功能基因组学也包括了在测序后对基因功能的研究。例如,酵母有许多功能重复的基因,分布在染色体两端,当处于丰富培养基条件时,这些基因似乎是多余的,但环境改变时就显示出其功能。**基因冗余**现象实际上是对环境的适应,冗余基因的存在为进化适应提供了可选择的余地。在基因组全序列中还保留了基因组进化的遗迹,提示基因重复常发生在近中心粒区和染色体臂中段。目前,研究者已把**酵母基因组**作为研究真核生物基因组功能的模式,计划建立酵母基因组 6000 多个基因的**单突变体文库**(single mutant library),并

用于其他高等真核生物基因组的“基因功能作图”。

总之,功能基因组学的任务是对成千上万的基因表达进行分析比较,从基因组整体水平上阐述其活动规律。核心问题是基因组的多样性和其进化规律,基因组的表达及其调控,模式生物基因组研究,等等。这门新学科的形成,是在后基因组时代生物学家的研究重点从揭示生命的所有遗传信息转移到在整体水平上对生物功能研究的重要标志。

1.5.2 蛋白质组学

1994年Wilkins等首先提出了**蛋白质组**(proteome)的概念,随后得到国际生物学界的广泛承认。蛋白质组的定义是指一个基因组所表达的全部蛋白质的总和(proteome indicates the proteins expressed by a genome);“proteome”由蛋白质一词的前几个字母“prote”和基因组一词的后几个字母“ome”拼接而成。

蛋白质组学是以蛋白质组为研究对象,研究细胞内所有蛋白质及其动态变化规律的科学。蛋白质组与基因组不同,一种生物体的基因组基本上是固定不变的,即同一生物不同细胞中基因组基本上相同,人类基因的总数为6万~10万个。单从DNA序列的信息并不能解释某个基因的表达时间、表达量、表达产物蛋白质翻译后加工和修饰等情况,以及它们在亚细胞群中的分布等。这些问题有望在蛋白质组学研究中找到答案。因为细胞内的蛋白质组是动态的,有它的时序性、可调节性、相关性,进而能够在细胞和生物体的整体水平上阐明生命现象的本质和活动规律。蛋白质组研究的数据库与基因组数据库的整合,将对功能基因组的研究发挥重要作用。

随着研究的深入和需要,科学家将蛋白质组由原来的定义,即一个基因组所表达的蛋白质,改为细胞内基因组所表达的全部蛋白质。但要获得如此完整的蛋白质组,在实践中是极其困难的。因为蛋白质的种类和形态总是处在细胞新陈代谢的动态过程之中难以准确测定。所以,1997年,Cordwell和Humphrey-Smith提出了**功能蛋白质组**(functional proteome)的概念,指在特定时间、特定环境和实验条件下基因组活跃表达的蛋白质。与此同时,中国生物学家提出了**功能蛋白质组学**(functional proteomics)新概念,研究定位在细胞内与某种功能有关或在某种环境条件下的一群或一套蛋白质。

功能蛋白质组只是总蛋白质组的一部分,通过对功能蛋白质组的研究,既能阐明某一群体蛋白质的功能,也能丰富总蛋白质数据库,它是从生物大分子水平到细胞水平研究的重要桥梁。无论是蛋白质组学还是

功能蛋白质组学,首先都要求分离亚细胞结构、细胞或组织等不同生命结构层次的蛋白质,获得总蛋白质谱。为了尽可能分辨细胞或组织内所有蛋白质,目前一般采用高分辨率的双向凝胶电泳。一种正常细胞的双向电泳图谱通过扫描仪扫描并数字化,运用二维分析软件可对数字化的图谱进行各种图像分析,包括被分离蛋白在图谱上的定位、信息计数、图谱间蛋白质差异表达的检测,等等。一种细胞或组织优质的蛋白质组双向电泳图,可以得到几千甚至上万种蛋白质。为了适应这种大规模的蛋白质组分析,质谱技术已成为蛋白质鉴定的重要技术。从质谱技术测得完整蛋白质的相对分子质量、肽质谱以及部分肽序列等数据,通过相应数据库的搜寻,从而比对鉴定目标蛋白质。最后,再对蛋白质翻译后修饰的类型和程度进行分析,在蛋白质组定性和定量分析的基础之上,建立蛋白质组数据库。

从1997年P E Hodes等构建成第一个完整的蛋白质组数据库——酵母蛋白质数据库(yeast protein database, YPD),到目前蛋白质组的研究进展速度极快。随着新思路和新技术的不断涌现,这项新技术将会不断完善,发展成为后基因组时代的重点技术手段。

1.5.3 生物信息学

HGP和其他重要生物基因组大量序列信息和功能信息的积累和在互联网上的注册与详细注释等网络资源,催生了生物信息学(bioinformatics)这门新学科。对浩如烟海的DNA和蛋白质分子中各种类型信息进行识别、存储、分析、比对、注释、模拟和传输是生物信息学的主要内容,它由各类数据库、计算机网络和应用软件三大部分组成。主要在基因的染色体定位、寻找感兴趣基因的同源序列的快速搜索、基因识别和文献查寻等方面提供帮助,已成为分子生物学家必不可少的重要工具。

国际上现有大型生物信息中心,主要是美国国家生物信息学中心(NCBI)、基因组序列数据库(GS-DB)、欧洲分子生物学研究所(EMBL)和日本DNA数据库(DDBJ)。这些中心和全球的基因组研究实验室通过网站、电子邮件或者直接与服务器和数据库联系而获得的搜寻系统,使研究者可以在多种不同的分析系统中对序列数据进行查询、利用和共享巨大的生物信息资源。

随着大规模DNA自动测序的迅速发展,序列数据呈爆炸性地增长,国际上许多著名实验室曾在HGP启动的同时,就与信息科学和数据库技术同步发展,收集、存储、处理了庞大的数据,使生物信息学逐步走向成熟,在各基因组计划中发挥了无以取代的重要作用。

所建立的核苷酸序列数据库,已存有大量生物的 cDNA 和基因组 DNA 序列的信息。在已应用的软件中,有 DNA 分析、基因图谱构建、RNA 分析、多序列比较、同源序列检索、三维结构观察与演示、进化树生成与分析,等等功能。

在蛋白质组研究中,由于蛋白质组是随着生物个体发育阶段和所处的内外环境而变化的,mRNA 丰度与蛋白质的丰度非显著相关,以及需要经历翻译后的加工修饰等,因而对蛋白质的生物信息学研究,在内容上有许多特殊之处。当前通用软件所建立的数据,主要有蛋白质分子质量、序列信息、结构域、二维电泳、三维结构、特殊性质、模拟酶解、翻译后修饰、代谢及蛋白质分子相互作用,等等。

利用生物信息学资源研究基因产物——蛋白质的性质与功能已成为生命科学的重要组成部分。传统的基因组分析是得到连续的 DNA 序列信息,而蛋白质组连续系则是基于多重分子质量、等电点范围、空间结构和分子构象等参数,构建活细胞内全部蛋白质表达的图谱,使人们在研究不同条件下细胞和组织乃至整个生命体的活动成为可能。以基因组学、转录组学、蛋白质组学及代谢组学等不同信息资源层次的“组学”研究为基础的系统生物学,人类已经能够研究生物系统中所有组成成分的变化规律及在特定遗传或环境条件下相互关系。

分子生物学已经渗透到生物学科各个领域之中,并正在产生一系列新的分子科学,改变了或正在改变着整个生物学的面貌,其研究成果已在工业、农业、医学以及生物制药等领域得到广泛的应用。对分子生物学的深入研究将使整个生物学在分子水平上统一起来,即统一的生物学,愈来愈多的研究成果说明生命的本质具有高度的有序性和一致性,这就是所谓生长、发育与进化的统一理论。

【本章重点归纳】

分子生物学有广义与狭义之分。广义上的分子生物学指对蛋白质和核酸等生物大分子结构与功能的研

究,以及从分子水平上阐明生命现象和生物学规律的学科;狭义上指的是主要研究基因或 DNA 结构与功能、复制、转录、表达和调节控制等过程,也涉及与这些过程相关的蛋白质和酶的结构与功能的研究。目前一般采用的是狭义上的概念。

分子生物学是在多个学科的基础上发展起来的,因此与多个学科都存在着内容上的交叉,特别是生物化学。分子生物学是从分子水平上研究生命现象,而生物化学是从分子水平上研究生命现象的化学本质;在研究方向和方法上两者着差异。归纳分子生物学的研究内容,可概括为 5 个方面:基因与基因组的结构与功能;DNA 复制、转录和翻译;基因表达调控研究;DNA 重组技术和结构分子生物学。分子生物学的发展速度是以往任何学科都无法比拟的。发展过程概括为 3 个阶段:人类对 DNA 和遗传信息传递的认识阶段、重组 DNA 技术的建立和发展阶段、重组 DNA 技术的应用和分子生物学的迅猛发展阶段。随着人类基因组计划的完成,人类已经掌握了许多模式生物的基因组信息。但基因组信息资源并不能直接阐明基因的功能,更未能预测某基因所编码蛋白质的功能与活性,所以并不能指导人们充分正确地利用这些基因的产物。于是出现了“后基因时代”,旨在快速、高效、大规模鉴定基因的产物和功能。基因组研究的重点由结构向功能转移,这成为今后发展的主流方向。相关学说和理论相应诞生,如功能基因组学、蛋白组学和生物信息学,等等。生命科学正在进入一个崭新的时代。

【思考题】

1. 从广义和狭义上给出分子生物学的定义。
2. 现代分子生物学研究的主要内容有哪几个方面?什么是反向生物学?后基因组时代?
3. 写出 3 个分子生物学发展史中的大事件(发明年代、发明者、简要内容)。
4. 21 世纪分子生物学的发展趋势将是怎样的?



2.1 细胞内的遗传物质

2.1.1 DNA 是主要的遗传物质

生命的基本特性之一是生物体内具有物质和能量以及信息的变化(或转化)。生物体信息的变化包括两个基本方面,一是从亲代到子代的信息传递,二是生物个体内遗传信息的表达,同时决定该个体的性状特征。遗传学早期的基因学说认为,基因作为遗传因子决定着生物的性状,并能自我复制,稳定地传递给后代。随着学科的发展,至今人类对基因的认识仍处在发展之中。人类对DNA是遗传物质的认识经历了三个阶段。

第一阶段是1928年Frederick Griffith做的肺炎双球菌转化试验。研究者从有荚膜、菌落光滑的S型肺炎球菌(*Pneumococcus*)细胞中提取DNA,加入到无荚膜、菌落粗糙的R型细菌培养物中,发现DNA能使部分R型细胞获得合成S型细胞特有的多糖荚膜的能力。若将DNA预先用脱氧核糖核酸酶降解,则失去转化能力。已经转化了的细菌,其后代仍保留合成S型荚膜的能力,说明此性状可以遗传给后代。

第二阶段是1944年,Oswald Avery等用生物化学和物理化学手段对肺炎双球菌转化试验进一步分析证明DNA是遗传物质。1952年,Hershey和Martha chase用 ^{35}S 和 ^{32}P 同位素示踪技术,将T2噬菌体侵染*E. coli*细胞,证明了主要是核酸进入细菌体内,而病毒外壳蛋白留在细胞外,且进入菌体的DNA能利用细菌的生命过程合成噬菌体自身的DNA和蛋白质,并能自我组装成与亲代完全相同的子代噬菌体。烟草花叶病毒的重建实验也证明,病毒蛋白质的特性由RNA决定,即遗传物质是核酸而不是蛋白质。

第三阶段是1953年,Watson和Crick提出了DNA双螺旋结构模型。这一理论的建立是基于1950年Erwin Chargaff等测定各种生物的DNA碱基组成后,发现不同生物的DNA碱基组成有着严格的种族特异性,嘌呤碱基总是与嘧啶碱基进行氢键配对,A+T与G+C的摩尔数相等并互补的数据的支持。双螺旋结构的这一最基本特性,决定了DNA分子具有自我复制和亲代向子代传递遗传信息的能力。至此,DNA作为遗传物

质被普遍接受。DNA双螺旋结构模型为人类充分揭示遗传信息的传递规律奠定了坚实的理论基础。这一重大发现打破了长期以来许多生物学家认为的只有像蛋白质那样的大分子才能作为细胞遗传物质的观点,在遗传学上树立了DNA是遗传信息载体的理论。

生物体内DNA分子上的遗传信息通过表达产生各种蛋白质和RNA实现其功能,DNA在执行其作为遗传物质功能的同时,也具有一定的稳定性和灵活性。DNA是通过碱基互补配对形成的双链分子,碱基互补是其复制、转录、表达遗传信息的基础。复制过程中,通过碱基互补配对的机制,准确地把遗传信息传递给子代。DNA分子在生理状态下性质稳定,适于作为遗传物质;而作为生物进化的分子基础,DNA也会少量发生突变(mutation),且这种突变可稳定地遗传。DNA的转化可用于细菌、动物和植物各类细胞,现已成为分子生物学实验室常用的方法。实际上,供体细胞DNA进入受体细胞而引入新的遗传形状是一个自然过程。一些生物体DNA的含量见表2.1。

表 2.1 一些细胞和病毒的DNA含量

生物种类*	DNA量(pg/细胞或病毒)	碱基对数目/bp
哺乳类*	6	5.5×10^9
鸟类*	2	2×10^9
爬虫类*	5	4.5×10^9
两栖类*	1.6~160	$1.5 \times 10^9 \sim 1.5 \times 10^{11}$
鱼类*	0.6~21	$6 \times 10^8 \sim 2 \times 10^{10}$
昆虫类*	0.2~10	$2 \times 10^8 \sim 1 \times 10^{10}$
线虫*	0.08	8×10^7
细菌	0.002~0.06	$2 \times 10^6 \sim 6 \times 10^7$
T4噬菌体	0.00024	1.7×10^5

注:pg= 10×10^{-12} ; * 真核细胞指体细胞的数值。

小结:生物体内DNA分子上的遗传信息通过表达产生各种蛋白质和RNA实现其功能。DNA是通过碱基互补配对形成的双链分子,碱基互补是其复制、转录、表达遗传信息的基础。DNA作为遗传物质主要有以下特性:① 贮存遗传信息;② 将遗传信