

FENZI SHENGWU

FANYING YUANLI YU FANGFA YANJIU

分子生物

反应原理与方法研究

于天飞 万秀清, 汪军梅 编著



中国水利水电出版社
www.waterpub.com.cn

014040848

Q7
72

FENZI SHENGWU

FANYING YUANLI YU FANGFA YANJIU

分子生物

反应原理与方法研究

于天飞 万秀清 汪军梅 编著



中国水利水电出版社
www.waterpub.com.cn



北航

C1728097

Q7
72

内 容 提 要

分子生物学是生命科学发展过程中诞生的一门新兴学科。全书以基本原理为基础,包括绪论,DNA 的生物合成反应,DNA 的损伤、修复与突变,DNA 重组与克隆,RNA 的生物合成,蛋白质的生物合成,肽链合成后的加工和输送,原核生物基因表达的调控,真核生物基因表达的调控,分子生物学方法等内容,通过对基本原理的分析而逐步展开对分子生物反应原理与方法的研究。本书论述逻辑严密、条理分明,内容丰富、详略得当,文字通俗流畅,叙述由浅入深,是一本关于分子生物不错的参考书,可供相关专业的人员阅读、参考。

图书在版编目(CIP)数据

分子生物反应原理与方法研究/于天飞,万秀清,
汪军梅编著. — 北京:中国水利水电出版社,2013.12
ISBN 978-7-5170-1503-1

I. ①分… II. ①于… ②万… ③汪… III. ①分子生
物学—研究 IV. ①Q7

中国版本图书馆 CIP 数据核字(2013)第 298789 号

策划编辑:杨庆川 责任编辑:杨元泓 封面设计:崔 蕾

书 名	分子生物反应原理与方法研究
作 者	于天飞 万秀清 汪军梅 编著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址:www. waterpub. com. cn E-mail:mchannel@263. net(万水) sales@waterpub. com. cn
经 售	电话:(010)68367658(发行部)、82562819(万水) 北京科水图书销售中心(零售) 电话:(010)88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京厚诚则铭印刷科技有限公司
印 刷	三河市天润建兴印务有限公司
规 格	184mm×260mm 16 开本 17.5 印张 426 千字
版 次	2014 年 4 月第 1 版 2014 年 4 月第 1 次印刷
印 数	0001—3000 册
定 价	62.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社发行部负责调换

版权所有·侵权必究

前 言

分子生物学是在分子水平上研究基因结构和功能的科学,是生命科学发展过程中诞生的一门新兴学科。近半个多世纪以来,随着生物技术的突飞猛进,分子生物学的理论也得以迅速发展与不断完善,并以极为强势的姿态,渗透和影响多个相关学科和产业领域。

作者在结合多年来的实践和经验、咨询了相关专业的学者和同事、吸取了近年来相关图书的优点的基础之上,萌生了编撰此书的想法,确定本书为《分子生物反应原理与方法研究》。全书以基本原理为基础,包括绪论,DNA的生物合成反应,DNA的损伤、修复与突变,DNA重组与克隆,RNA的生物合成,蛋白质的生物合成,肽链合成后的加工和输送,原核生物基因表达的调控,真核生物基因表达的调控,分子生物学方法等共10章的内容,通过对基本原理的分析逐步展开对分子生物反应原理与方法的研究。

本书在写作过程中,力求做到概念明确,逻辑严密、条理分明,书中内容详尽、详略得当、图文并茂,便于阅读,深度和广度适宜,文字通俗流畅,语言简练,注重理论联系实际,极力贯彻基础性、系统性、科学性、实用性等原则,保证逻辑性和系统性。本书是关于分子生物的一本很好的参考书,可供相关专业的人员阅读、参考。本书的目标:一是帮助读者获得简明而正确的概念和理解,准确地掌握分子生物学的基本理论和学科前沿,二是使读者了解开展分子生物学理论研究的方法,而适当的联系生活实际应用,以希望能够增强阅读者的学习兴趣,并因此获得思维、逻辑与分析的启迪和创新能力的提升。

全书由于天飞、万秀清、汪军梅撰写,具体分工如下:

第4章~第6章:于天飞(齐齐哈尔大学);

第1章、第2章、第8章、第9章:万秀清(牡丹江烟草科学研究所);

第3章、第7章、第10章:汪军梅(长治医学院)。

该书能顺利出版,首先要感谢学校领导的大力支持和关怀以及众多同事的帮助。其次,本书在撰写过程中得到了许多专家的指导和意见,同时参阅了许多学者的著作和文献资料,在此对相关的专家、学者一并表示感谢。此外,出版社的工作人员为本书稿的修订做了许多工作,感谢你们为本书顺利问世所作的努力。

由于能力有限、经验不足、时间仓促,本书在内容的组织上难免存在错漏和不妥之处,竭诚希望各同行专家以及广大读者提出宝贵的意见和建议,以使本书不断完善。

作 者

2013年11月

目 录

第 1 章 绪论	1
1.1 分子生物学的概念	1
1.2 分子生物学研究内容	1
1.3 分子生物学的发展与展望	3
第 2 章 DNA 的生物合成反应	13
2.1 DNA 复制的一般特征	13
2.2 DNA 的复制	23
2.3 DNA 的逆转录	35
第 3 章 DNA 的损伤、修复与突变	40
3.1 DNA 损伤	40
3.2 DNA 损伤的修复	46
3.3 DNA 突变与回复	57
3.4 定向诱变	70
第 4 章 DNA 重组与克隆	75
4.1 同源重组	75
4.2 位点特异性重组	87
4.3 转座与逆转座	92
4.4 DNA 克隆	111
4.5 克隆基因的表达	123
第 5 章 RNA 的生物合成	132
5.1 原核生物的转录	132
5.2 真核生物的转录	145
5.3 转录校对	153
5.4 转录过程的选择性抑制	153
第 6 章 蛋白质的生物合成	157
6.1 蛋白质构象	157
6.2 遗传密码	159
6.3 tRNA	163
6.4 氨酰-tRNA 的合成	165
6.5 原核生物的蛋白质合成	167
6.6 真核生物的蛋白质合成	172
6.7 蛋白质生物合成的抑制剂	178

第 7 章	肽链合成后的加工和输送	179
7.1	肽链合成后加工	179
7.2	肽链合成后的输送	187
第 8 章	原核生物基因表达的调控	196
8.1	NDA 水平的调控	196
8.2	转录水平的调控	201
8.3	翻译水平的调控	204
8.4	操纵子对基因表达的调控	210
8.5	核开关	217
第 9 章	真核生物基因表达的调控	221
9.1	真核生物基因调控的特点	221
9.2	DNA 及染色体水平的调控	222
9.3	转录后水平的调控	230
9.4	翻译水平的调控	233
9.5	翻译后水平的调控	236
9.6	真核生物发育水平的调控	237
9.7	脂蛋白基因的调控	241
第 10 章	分子生物学方法	252
10.1	DNA 测序和基因组测序	252
10.2	核酸杂交法	265
10.3	LPL 内含子突变检测技术	269
参考文献		274

第 1 章 绪 论

分子生物学是随着遗传学、生物化学和细胞生物学等学科的发展兴起的。1953 年 DNA 双螺旋结构模型的提出,是分子生物学兴起的标志。50 多年来,关于基因的复制、转录和翻译的过程及其基因表达调控的研究,取得很大成绩,基因工程和基因组学得到蓬勃发展。随着研究技术的不断进步,分子生物学会不断加快发展速度,使人们可以更加深入的认识基因的结构和功能,以及基因表达的过程和调控机制,这些研究成果逐渐走向应用,可能为农、林、牧业提供新的优良品种,在解决粮食问题、能源问题、环境问题等方面发挥作用。还可能为疾病的预防和治疗提供新方法,提高人类的生活质量。

1.1 分子生物学的概念

分子生物学是指对生物大分子结构与功能的研究,是从分子水平研究生命本质的科学。

20 世纪初期,现代科学技术迅速发展起来。当时的研究者们已经在细胞学、遗传学、生物化学和物理学等领域取得了大量的研究成果。到了 20 世纪中期,科学家们不仅研究豌豆、果蝇、玉米等材料,也对一些简单的生物,如细菌、噬菌体等做了深入的研究。所谓的简单只是与高等生物相比较而言,这些细菌和噬菌体其实也是相当复杂的。科学家们发现:在大多数生物中,DNA 是主要的遗传信息载体;DNA 的结构使它的复制与修复近乎完美;DNA 的线状结构编码了蛋白质的三维结构。所有这些研究成果表明:那些控制简单生物的基本生物学原则对那些复杂的生物同样适用。从此,分子生物学真正发展起来。

从广义上来说,蛋白质及核酸等生物大分子的结构与功能的研究都属于分子生物学的范畴,也就是从分子水平阐明生命现象和生物学规律,如蛋白质的结构、功能和运动;酶的作用机理和动力学;膜蛋白的结构、功能和跨膜运输等,都属于分子生物学的内容。但是有些题目一般不属于分子生物学内容,如代谢中的某些反应,如果这些反应由反应物和产物的浓度来调节,一般就认为是典型的生物化学反应。另外,细胞结构与各种细胞成分的组织则属于细胞生物学。

从狭义上来说,分子生物学偏重于核酸的分子生物学,主要研究基因或 DNA 的复制、转录、表达及调节控制等过程,其中也涉及与这些过程有关的蛋白质和酶的结构与功能的研究。

本书的内容也侧重于狭义分子生物学的概念。

1.2 分子生物学研究内容

半个世纪以来,分子生物学发展迅速,涉及的范围也很大,取得了丰硕的成果。概括起来,分子生物学的研究内容大致可以分为以下三类。

1. 核酸的结构

核酸的一级结构是指 DNA 链中脱氧核苷酸的排列顺序(或 RNA 链中核苷酸的排列顺序),它包含了核酸的全部信息,是遗传信息的结构基础。

核酸还可以形成高级结构。自从 B-DNA 的双螺旋结构被阐明以后,其他 DNA 的结构也被揭示出来。例如三链和四链 DNA、A-DNA 和 Z-DNA,真核生物的线粒体和叶绿体 DNA,原核生物中的质粒 DNA。随着 RNA 的发现, RNA 的结构,包括 tRNA、rRNA 和 mRNA 以及一些小分子 RNA 的结构得到了深入的研究。

基因和基因组的结构是人们非常关注的内容。真核生物具有内含子和外显子交替排列的割裂基因,真核生物的基因可能有许多拷贝,这些拷贝可能成簇排列,可能串联起来,可能位于基因组的不同染色体上。真核基因组中有许多序列并不编码蛋白质,还有一些多次重复的小片段。

现在,已经测定出很多生物的完整基因组序列,包括人类基因组的序列。人类基因组 DNA 的全序列为 3×10^9 碱基对,约含有 10 万个基因。但是要彻底了解这些基因的产物的功能,基因的表达调控机理,要理解 80% 以上非编码序列的作用等,都还要进行长期的艰苦研究。

生物大分子,包括核酸在进行各种生物功能时,空间结构会发生变化。因此,生物大分子的空间结构、空间结构的变化和生物大分子功能之间的关系也是分子生物学的研究内容之一。

2. 核酸的功能

说到核酸的功能,只是习惯的说法,并不十分准确。例如:DNA 的作用就是携带遗传信息,并不执行其他的功能。另外,只靠 DNA 分子本身其实什么也做不成,所有的 DNA 事务都是由蛋白质处理的。所谓“有功能的”或“有活性的”基因,是说一个基因具有完整的结构的、加工的和调控的信息等,可以精确地指导所要进行的生命过程。所有这些过程都包括了核酸与蛋白质间的识别与相互作用,蛋白质与蛋白质之间的相互作用。

核酸的功能包括:

DNA 复制的机理。这个部分的研究还包括了很多与复制有关的酶和蛋白质因子以及它们的结构与作用机理。DNA 突变和修复的分子机制以及 DNA 的重组机制也受到人们的关注。

DNA 转录成 RNA,与转录有关的酶,以及 RNA 前体的转录后加工是分子生物学的重要研究领域。从 RNA 反转录产生 DNA 是对中心法则的重要补充,现在已经成为分子生物学的重要技术之一。

由 DNA 转录出的 mRNA 是蛋白质合成的模板,mRNA 上的三联体密码子由携带了特定氨基酸的 tRNA 识别。64 个密码子中,61 个编码蛋白质中的氨基酸,3 个是肽链合成终止的信号。遗传密码的知识使我们从 DNA 的序列就能预测蛋白质的序列。

中心法则是分子生物学的基本理论体系。但是,中心法则的运行并不是机械的,运行的速度并不是一成不变的。生命体中所有的过程都要受到严格的甚至多重的调控,使细胞的代谢发生变化,以应对外界环境的改变。基因表达调控是分子生物学的重要领域之一。

3. 综合利用上述的研究成果

人们综合利用上述研究成果,衍生出了分子遗传学、基因工程、蛋白质工程等。为医学、农业、工业和环境保护等打开了新局面。

新技术的不断涌现促进了分子生物学的不断进步。例如核酸的化学合成,限制性内切酶、连接酶等重要工具酶的发现,DNA序列的快速测定,聚合酶链式反应(PCR)技术等,都为分子生物学向更广阔的领域发展提供了动力。现在,打破种属界限,将不同来源的DNA在体外重组,从而大量获得目标蛋白已经成为一种常规的操作。

转基因动植物的研究方兴未艾。用转基因动物能够获取治疗人类疾病的重要蛋白质,世界上有几百种基因工程药物及其他基因工程产品在研制中。基因诊断与基因治疗是基因工程在医学领域发展的一个重要方面。在转基因植物方面,转基因玉米、转基因大豆已经相继投入商品生产。

当然,克隆技术还有其他一些问题需要研究。例如:转基因生物是否会引发伦理学问题?在我国,转基因的作物是否具有自主知识产权?那些转基因农作物的商业化是否会增加农民的种植成本?转基因的农作物商业化是否会影响我国的粮食安全?这些问题也需要给予关注。

1.3 分子生物学的发展与展望

科学领域中任何一门学科的形成和发展,一般很难准确地说明它是何时、何人创始的。分子生物学的产生和发展,同其他学科一样,经历了漫长而艰辛的过程,逐步走向成熟而迅速发展的道路。

1871年,Lankester就提出,生物不同种属间的化学和分子差异的发现和ación对确定系统发生的关系要比总体形态学的比较研究更为重要。后来,随着德国、美国生理化学实验室的建立和生物化学杂志的创办,促进了生物化学的发展。当生物化学深入到研究生物大分子时,1938年Weaver在写给洛克菲勒基金会的报告中,首次使用了分子生物学一词。他写道:“在基金会给予支持的研究中,有一系列属于比较新的领域,可称之为分子生物学……”。一年以后,研究蛋白质结构的Astbury使用了这个名词,以后它变得越来越普遍。特别是在1953年,Watson和Crick发表了著名论文“脱氧核糖核酸的结构”以后,DNA双螺旋结构的发现促进了遗传学、生物化学和生物物理学的结合,推动了分子生物学的形成和迅速发展,使生命科学全面地进入分子水平研究的时代,这是生物科学发展史上的重大里程碑。1956年剑桥医学研究委员会率先建立了分子生物学实验室,1959年《分子生物学》杂志创刊,1963年成立了欧洲分子生物学国际组织,分子生物学从而成为崭新的独立学科,带动着生命科学迅猛发展,成为现代自然科学研究中的重要领域。

1.3.1 分子生物学的发展基础

分子生物学是一门交叉学科,与其他学科互相促进,互相渗透。回顾分子生物学的发展历史,许多学科的成果都成为分子生物学发展的基础。分子生物学发展的历史充满着科学家的艰苦付出,也有很多逸闻趣事。

1. 遗传学基础

1859 年,达尔文发表了《物种起源》,提出了适者生存的进化理论。他认为,动植物在长期的生命过程中会发生一些微小的变化,其中一些动植物积累了这些变化,对环境更加适应,从而得到更好的生存与繁衍。

1856—1864 年,孟德尔(Mendel)的著名实验开创了现代遗传学,他提出了遗传的分离定律和独立分配定律,还提出了遗传因子的概念。

1903 年,Sutton 自己的工作解释了孟德尔的实验,并推测遗传因子是细胞中染色体的一个部分。尽管 Sutton 的试验并没有直接证明遗传的染色体理论,但非常重要,因为 Sutton 第一次把遗传现象与细胞学联系在一起。

1910 年以后,摩尔根(Morgan)等人用果蝇作试验时,发现有些基因在染色体上的距离近些,有些则远些。他们根据大量的试验构建了遗传图谱,并提出了连锁遗传规律。

1915 年,摩尔根证实了遗传的染色体基础。

不同的基因可能会位于同一条染色体上产生遗传的连锁现象,然而连锁通常并不完全。Janssens 提出了染色体交换理论,染色体在减数分裂过程的联会阶段,发生断裂,然后交叉连接起来,造成了不完全连锁。

1931 年,Barbara McClintock 用玉米为材料证实了染色体的断裂重接。

一旦遗传的规律被阐明,就可以解释生物的变异现象和进化理论。但是,每个基因上发生的变化都很小,这些小的变化足以产生新的物种吗? Wright 等人认为:由于地球的年龄很大,又由于选择的压力很温和,一些小的有利的性状足以积累起来形成新的物种。到 20 世纪 40 年代,生物学家 Huxley、遗传学家 Dobzhansky、古生物学家 Simpson 和鸟类学家 Mayr 从各自的研究结果出发,都证实经典遗传学与进化论确实是一致的。

2. 物理学与生物化学基础

几乎就在孟德尔的遗传定律刚刚发现以后,遗传学家就开始思考基因的化学结构,以及基因是如何工作的。但是在很长时间内都没有实质性的进展。因为当时核酸和蛋白质的结构都不清楚。

1927 年,Muller 和 Stadler 分别独立地发现了 X 射线可以诱导突变,由突变的频率可以估计一个基因的大小。以后的时间里,很多科学家都发现基因的突变可以影响到细胞中的蛋白质,由此发现基因与蛋白质之间具有一定的关系。此时,Beadle 和 Edward 基于对红色面包霉的研究,提出了一个基因一个酶的假说。

但是,蛋白质的结构是什么?基因的结构是什么?基因是如何工作的?决定遗传信息的究竟是蛋白质还是基因?经典遗传学无法提供有效的研究手段来研究基因的化学本质,因此需要借助其他学科的方法。

1928 年,Griffith 证明灭活的致病性肺炎链球菌中的某些成分可以转化非致病性的肺炎链球菌。1944 年,Avery 鉴定出了这种成分的化学本质,这是来自于致病性肺炎链球菌的 DNA。Avery 由此证明了 DNA 是遗传信息的载体。

1952 年,Hershey 和 Chase 用噬菌体也证明了 DNA 是遗传信息的载体。

一个多世纪以前,很多科学家认为研究生物学一定要用完整的活细胞,因为他们相信细胞

具有某种“活力”，一旦细胞破碎，“活力”就会丧失，研究就无法进行。一旦科学家决定打开一个活细胞，分子生物学就诞生了。

1949年，生物化学家 Chargaff 研究了 DNA 的化学组成，提出了 Chargaff 规则。当蛋白质的结构开始用 X 射线分析时，一些科学家也在用 X 射线分析 DNA 的结构。

到了 20 世纪 50 年代，Wilkins 和 Franklin 获得了 DNA 纤维结构高质量的 X 射线衍射图谱。这个图谱显示 DNA 是螺旋结构，而且可能由 2 条或 3 条多核苷酸链组成。1952 年，有机化学家发现 DNA 的多核苷酸链间的连键是 $3' \rightarrow 5'$ 磷酸二酯键。

1953 年，Watson 和 Crick 推导出了 DNA 的双螺旋结构，并推测了 DNA 分子自我复制的机制。DNA 双螺旋结构的发现是一场伟大的革命，从此 DNA 不再神秘，它与丙酮酸、甘油等一样，都是化学分子，都可以在实验室进行研究。

当 Watson 和 Crick 推导出了 DNA 的双螺旋结构的时候，Kornberg 正在研究嘧啶和嘌呤核苷酸的合成。Kornberg 和同事们发现了嘌呤和嘧啶核苷酸的从头合成途径以及这些途径中的很多酶。

1956 年，Kornberg 等人发现 DNA 合成的前体是四种脱氧三磷酸核苷，随后又发现了合成 DNA 的酶（现在我们知道这是 DNA 聚合酶 I），并用纯化的 DNA 聚合酶合成了具有侵染性的噬菌体 Φ X174 的 DNA。

1958 年，Meselson 和 Stahl 发现 DNA 在复制的过程中两条链要分开，子代 DNA 分子中只有一条链是新合成的，一条链保留在子代分子中，就是半保留复制。

1968 年，冈崎 (Okazaki) 提出 DNA 半不连续复制的模型。

双螺旋的发现彻底终结了 DNA 是否是遗传物质的争论。但是，DNA 不可能是蛋白质合成的直接模板。因为在细胞中，活跃地合成蛋白质的位置上不存在 DNA。况且真核生物的 DNA 位于细胞核中，而蛋白质合成是在细胞质中。那么，一定还有另外一种分子，它从 DNA 那里得到了遗传信息，然后移动到细胞质中作为合成蛋白质的模板。这种分子很可能就是第二类核酸——RNA。Caspersson 和 Brachet 已经在细胞质中发现了大量的 RNA。

但是，RNA 的核苷酸顺序如何转变成蛋白质中的氨基酸顺序呢？一开始人们想象可能是 RNA 折叠起来形成一个疏水的空穴，空穴的形状恰好与一种特殊的氨基酸契合。但是 Crick 认为这根本不可能。因为从化学性质上看，RNA 中碱基上的基团更可能与水溶性的基团反应。其次，即使某些 RNA 片段能够形成疏水的空穴，这样的结构也无法区分 Gly 与 Ala，或者 Val 与 Ile。因为这 2 对氨基酸的侧链非常相似，各自只差一个甲基。在 1955 年，Crick 推测一定存在一种接头分子 (adaptor)，这种分子既可以识别核酸，又能够连接氨基酸，它很可能也是一种 RNA。Crick 在 1956 年又提出了中心法则：遗传信息可以自我复制，可以从 DNA 流向 RNA，然后再流向蛋白质，这就构成了中心法则的基本内容。

1953 年，Zamecnik 等发展了在体外用无细胞系统合成蛋白质的体系。几年以后，他们又发现在合成蛋白质以前，氨基酸首先要连接到 tRNA 上；同时还发现了催化这个反应的是氨酰-tRNA 合成酶。tRNA 就是 Crick 预言的接头分子，tRNA 的一端用反密码子与蛋白质合成的模板识别，另一端连接氨基酸。

蛋白质合成的模板是什么？最初有人认为是 rRNA。但是经过仔细研究之后发现这是不可能的。第一，核糖体由大、小两个亚基组成，每个亚基都含有 rRNA；第二，几乎在所有

的细菌、植物和动物中,所有小亚基中的 rRNA 大小都非常相似;第三,尽管相应的 DNA 中的 AT/GC 比例不同,大、小 rRNA 的碱基组成几乎相同。这种相似性怎么可能指导合成大量的蛋白质呢?所以,rRNA 不可能是蛋白质合成的模板。现在我们知道 rRNA 是核糖体的组成成分之一。

1960 年,Brenner 及 Gross 等人发现,用噬菌体 T2 感染大肠杆菌后,细菌不再合成自己的 RNA,只从 T2 的 DNA 上转录噬菌体的 RNA。转录出来的 T2 RNA 不是与核蛋白结合形成核糖体,而是附着在核糖体上,然后沿核糖体的表面移动到可以结合氨酰-tRNA 的位置上去合成蛋白质。因为 T2 RNA 从 DNA 上接受遗传信息,又转移到核糖体上合成蛋白质,所以把它叫做信使 RNA(mRNA)。从此发现,mRNA 才是蛋白质合成的模板。

1958 年,Weiss 和 Hurwitz 的实验室分别独立地发现了依赖于 DNA 的 RNA 聚合酶。遗传密码的破译是分子生物学的伟大成就之一。1961 年,Nirenberg 用体外蛋白质合成系统,加入人工合成 poly U 和各种同位素标记的氨基酸,得到的是多聚苯丙氨酸的肽链。这个实验证明了 UUU 是苯丙氨酸的密码子,还证明可以用合成的 RNA 链作为蛋白质合成的模板。后来他用同样的方法证明了 AAA 编码 Lys,CCC 编码 Pro。但 GGG 容易形成三股螺旋,后来使用另外的方法才证明 GGG 编码 Gly。

Nirenberg 多核苷酸磷酸化酶合成 RNA 链。这个酶合成 RNA 时不需要模板,当使用混合的核苷酸为底物时,模板的序列是随机的。虽然可以根据加入反应体系中的 NTP 的比例来估计核苷酸在密码子中的比例,但这个方法不能确定密码子的碱基顺序。

同时,有机化学家 Khorana 合成了具有确切碱基顺序的多聚核糖核苷酸。进一步证实每个密码子由三个核苷酸组成,但由于不能确定确切的起始核苷酸,还是不能确认到底哪一个密码子对应于哪一个氨基酸。

1964 年,Nirenberg 发现以三核苷酸为模板就足以使相应的氨酰-tRNA 结合到核糖体上,但是三核苷酸还是随机的片段。Khorana 将自己的方法与 Nirenberg 的方法结合,很快解读了约 50 个密码子。后来又破译了其他的密码子,并发现 3 个密码子是终止密码子,AUG 既是 Met 的密码子,也是起始密码子。这些发现也包括其他实验室的工作。

上述重要发现共同建立了以中心法则为基础的分子遗传学基本理论体系。1970 年 Temin 和 Baltimore 又同时从鸡肉瘤病毒中发现了以 RNA 为模板合成 DNA 的逆转录酶,又进一步补充和完善了遗传信息传递的中心法则。

操纵子学说使人们开始了解基因表达调控。最初发现了原核生物基因表达调控的一些规律,后来研究逐渐扩展到真核基因表达调控机理,真核生物发育与细胞分化的机理。目前的研究已经使我们了解到真核生物基因表达调控机理在于:基因的顺式调控元件与反式作用因子的相互识别与作用,核酸与蛋白质之间的相互识别与作用,蛋白质与蛋白质之间的相互识别与作用。

在研究基因表达调控的过程中,细胞核内及其他小分子 RNA 具有特殊功能。还发现了具有催化活性的 RNA,即核酶。1995 年,Guo 等发现一些短的 RNA 片段可以造成同源的 mRNA 降解,从而调节基因表达,叫做 RNA 沉默(RNA silence)。

3. 技术方法的基础

1967—1970 年,R. Yuan 和 H. O. Smith 等发现的限制性核酸内切酶为基因工程提供了

有力的工具。

1975—1977年, Sanger、Maxam 和 Gilbert 先后发明了 DNA 序列的快速测定法。

1985年, Mullis 等发明聚合酶链式反应(polymerase chain reaction, PCR), 这种特定核酸序列扩增技术以其高灵敏度和特异性被广泛应用, 对分子生物学的发展起到了重大的推动作用。目前分子生物学已经从研究单个基因发展到研究生物整个基因组的结构与功能。

20世纪90年代, 全自动核酸序列测定仪问世。

新技术的不断涌现促进了分子生物学的不断进步。现在打破种属界限, 将不同来源的 DNA 在体外重组, 从而大量获得目标蛋白已经成为一种常规的操作。

1.3.2 分子生物学的发展

在分子生物学的形成和发展过程中, 有许多重大的发现和事件, 具体情况如下。

1864年, Hoope-Seyler 结晶并命名了血红蛋白。

1869年, Miescher 第一次分离了 DNA。

1871年, Lankester 首先提出生物不同种属间的化学和分子差异的发现与分析, 对确定系统发生的关系, 要比总体形态学的比较研究更为重要。

1926年, Sumner 从刀豆的提取物中得到脲酶结晶, 并证明此蛋白质结晶有催化活性。同年, Svedberg 创建了第一台分析用超高速离心机, 并用其测定了血红蛋白的相对分子质量约为 6.8×10^4 。

1931年, Pauling 发表了他的第一篇关于“化学键特性”的论文, 详细说明了共价键联结的规律。后来, 又建立了处理生物分子的量子力学理论。

1934年, Bernal 和 Crowfoot 发表了第一张胃蛋白酶晶体的详尽的 X-射线衍射图谱。

1941年, Astbury 获得了第一张 DNA 的 X-射线衍射图谱。

1944年, Avery 提供了在细菌的转化中, 携带遗传信息的是 DNA, 而不是蛋白质的证据。实验证明, 使无毒的 R 型肺炎双球菌转变成致病的 S 型, DNA 是转化的基本要素。8年后, 1952年, Hershey 和 Chase 又用同位素示踪技术证明 T_2 噬菌体感染大肠杆菌, 主要是核酸进入细菌内, 而病毒外壳蛋白留在细胞外。烟草花叶病毒的重建实验证明, 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。

1950年, Chargaff 以不同来源 DNA 碱基组成的精确数据推翻了四核苷酸论, 提出了 Chargaff 规则, 即 DNA 的碱基组成有一个共同的规律, 胸腺嘧啶的摩尔含量总是等于腺嘌呤的摩尔含量, 胞嘧啶的摩尔含量总是等于鸟嘌呤的摩尔含量, 即 $[A]=[T]$ 和 $[G]=[C]$ 。

1951年, Pauling 和 Corey 应用 X-射线衍射晶体学理论研究了氨基酸和多肽的精细空间结构, 提出了两种有周期规律性的多肽结构学说, 即 α -螺旋和 β -折叠理论。

1953年, 这是开创生命科学新时代的第一年, 具有里程碑意义的是 Watson 和 Crick 发表了“脱氧核糖核酸的结构”的著名论文, 他们在 Franklin 和 Wilkins X-射线衍射研究结果的基础上, 推导出 DNA 双螺旋结构模式, 开创了生物科学的新纪元。同年, Sanger 历经 8 年的研究, 完成了第一个蛋白质——胰岛素的氨基酸全序列分析。

1954年, Gamow 从理论上研究了遗传密码的编码规律。

1956年, Volkin 和 Astrachan 发现了 mRNA(当时尚未用此名)。

1958年, Hoagland等发现了tRNA在蛋白质合成中的作用; Meselson和Stahl应用同位素和超离心法证明DNA的半保留复制; Crick提出遗传信息传递的中心法则。

1960年, Marmur和Dofy发现了DNA的复性作用, 确定了核酸杂交反应的专一性和可靠性; Rich证明DNA-RNA杂交分子与核酸间的信息传递有关, 开创了核酸实际应用的先河。与此同时, 在蛋白质结构研究方面, Kendrew等得到了肌红蛋白0.2nm分辨率的结构, Perutz等得到了血红蛋白0.55nm分辨率的结构。

1961年, Jacob和Monod提出操纵子学说, 发表了蛋白质合成中遗传调节机理的论文, 此论文被誉为是分子生物学中文笔优美的经典论文之一。同年, Brenner等获得mRNA的证据; Hall和Spiegelman证明 T_2 DNA和 T_2 专一性RNA的序列互补; Crick等证明了遗传密码的通用性。

1962年, Arber提出第一个证据, 证明限制性核酸内切酶的存在, 导致以后对该类酶的纯化, 并由Nathans和Smith应用于DNA图谱和序列分析。

1965年, Holley等采用重叠法首先测定了酵母丙氨酰-tRNA的一级结构, 为广泛、深入地研究tRNA的高级结构奠定了基础。

1967年, Gellert发现了DNA连接酶, 该酶将具有相同粘末端或者平末端的DNA片段连接在一起。同年, Philips及其同事确定了溶菌酶0.2nm分辨率的三维结构。

1970年, Temin和Baltimore几乎同时发现了反转录酶, 证实了Temin 1964年提出的“前病毒假说”。在劳氏肉瘤病毒(RSV)感染以后, 首先产生的是含有RNA病毒基因组全部遗传信息的DNA前病毒, 子代病毒的RNA是以前病毒的DNA为模板进行合成的。反转录酶已成为目前分子生物学研究中的一个重要工具。

1972年—1973年, 重组DNA时代到来。Berg、Boyer和Cohen等创建了DNA克隆化技术, 在体外构建成具有生物学功能的细菌质粒, 开创了基因工程新纪元。与此同时, Singer和Nicolson提出生物膜结构的液态镶嵌模型。

1975年, Southern发明了凝胶电泳分离DNA片段的印迹法; Gruenstein和Hogness建立了克隆特定基因的新方法; O'Farrell发明了双向电泳分析蛋白质的方法, 为分子生物学的深入发展创造了重要的技术条件; Blobel等报导了信号肽。

1976年, Bishop和Varmus发现动物肿瘤病毒的癌基因来源于细胞基因(即原癌基因)。

1977年, Berget等发现了“断裂”基因; Sanger、Maxam和Gilbert创立了“酶法”“化学法”测定DNA序列的方法, 标志着分子生物学研究新时代的到来。

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

1980年, Wigler等通过与某个选择性标志物共感染, 从而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, 随后又证明前体中的居间序列(intervening sequence, IVS)有五种酶的活力。几乎在同时, Altman从纯化的RNase P中, 证明催化tRNA前体成熟的催化剂是RNase P中的RNA。具有催化作用RNA(ribozyme)的发现, 促进了RNA研究的飞速发展。

1982年, Prusiner等在感染瘙痒病的仓鼠脑中发现了朊病毒。

1983年,Herrera-Estrella等用Ti质粒作为转基因载体转化植物细胞获得成功。

1984年,McGinnis等发现果蝇、非洲爪蟾等同源异形基因中的同源异形盒的核苷酸序列;Schwartz和Cantor发明了脉冲梯度凝胶电泳法;Simons和Kleckner等发现了反义RNA。

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的设想;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。

1986年,Dryja等发现成视网膜细胞瘤(Rb)基因是一种抑癌基因;Robin等采用X-光晶相学,证实了DNA结合蛋白的螺旋—转角—螺旋结构。

1987年,Mirkin等在酸性溶液的质粒中发现三链DNA;Burke等用酵母人工染色体(YAC)作载体克隆了大片段DNA;Hoffman等确定了Dnchenne肌肉萎缩病灶的蛋白产物是萎缩素;Hooper等和Kuehn等分别用胚基细胞进行哺乳动物胚的转基因操作,取得重大进展。

1988年,Landschalz等在对CyC3(细胞色素C基因调节蛋白)、癌基因产物(MyC、V-jun、V-fos)和CBP(CCAAT盒结合蛋白)的研究过程中,发现了结合区亮氨酸序列的周期性,提出DNA结合蛋白的亮氨酸拉链结构模型;同年,Whyfe等证明癌的发生是癌基因的激活和抑癌基因失活的结果。

1989年,Greider等首先在纤毛原生动物中发现了端粒酶是以内源性RNA为模板的反转录酶;Hiatt等首次报导了在植物中亦可产生单克隆抗体。

1990年,人类基因组计划(HGP)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(guide RNA);Sinclair等在人类Y染色体上发现了新的性别决定基因——SRY基因。

1991年,由欧洲共同体(EC)组织17个国家35个实验室的147位科学家,以手工测序为主要手段,首先完成了第一条完整染色体(酵母3号染色体)的315kb的测序工作;Hake等首次报导在植物中发现含有同源异形盒基因;Blackburn等提出调节聚合序列[通式为 $(T/A)_m G_n, m=124, n=1\sim 8$]的单链DNA可形成分子内或分子间的四螺旋结构,起着稳定染色体的作用。

1993年,Jurnak等在研究果胶酸裂解酶时,发现一种新的蛋白质结构——平行 β 螺旋;Yuan等在哺乳类细胞内发现一种参与调节细胞凋亡并具有剪切作用的蛋白质——IL-1 β 转换酶。

1994年,日本科学家在《Nature Genetics》上发表了水稻基因组遗传图;Wilson等用3年时间完成了线虫(*C. elegans*)3号染色体连续的2.2Mb的测定,预示着百万碱基规模的DNA序列测定时代的到来。

1995年,Cuenoud等发现了具有酶活性的DNA;Tu等在*E. coli*中发现了具有转运与信使双功能的RNA——10Sa RNA。

1996年,Lee等首次报导了酵母转录因子GCN4中的氨基酸片段能自动催化合成自我复制的肽;洪国藩等采用“指纹—锚标”战略构建了高分辨率的水稻基因组物理图谱,DNA片段的长度为120kb;Ooffeau等完成了酵母基因组DNA全序列(1.25×10^7 bp)的测定。

1997年,Wilmot等首次不经过受精,用成年母羊体细胞的遗传物质,成功地获得克隆

羊——多莉(Dolly); Willard 等首次构建了人染色体; Salishury 等发现 DNA 一种新的结构形式——四显性组合, 这可能是基因交换期间 DNA 联结的一种方式。

1998 年, Renard 等用体细胞操作获得克隆牛——Marguerite, 再次证明从体细胞可克隆出遗传上完全相同的哺乳动物; Gene Bank 公布了最新人的“基因图谱 98”, 代表了 30181 条基因定位的信息; Venter 对人类基因组计划提出新的战略——全基因组随机测序, 毛细管电泳测序仪启动。

从以上所述分子生物学的发展中可以看出, 20 世纪是以核酸的研究为核心, 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 测序都具有里程碑的意义, 将生命科学带向一个由宏观到微观再到宏观, 由分析到综合的时代。

1.3.3 分子生物学的展望

分子生物学已经取得巨大成就, 成为生命科学的带头学科之一, 有力地促进了各分支学科的发展。随着新世纪的到来, 生命科学又将进入一个新时代, 下面是分子生物学的一些学科特征和发展趋向。

1. 功能基因组学

遗传学最近的定义是对生物遗传的研究和对基因的研究。功能基因组学是依附于对 DNA 序列的了解, 应用基因组学的知识和工具去了解影响发育和整个生物体的特定序列表达谱。以酿酒酵母(*S. cerevisiae*)为例, 它的 16 条染色体的全部序列已于 1996 年完成, 基因组全长 12086kb, 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。功能基因组学是进一步研究这 6000 多个基因, 在一定条件下, 譬如酵母孢子形成期, 同时有多少基因协同表达才能完成这一发育过程, 这就需要适应这一时期的全套基因表达谱。要解决如此复杂的问题就必须在方法学上有重大的突破, 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。目前用于检测分化细胞基因表达谱的方法, 有基因表达连续分析法、微阵列法、有序差异显示和 DNA 芯片技术等。今后, 随着功能基因组学的深入发展, 将会有更新更好的方法和技术出现。

功能基因组亦包括了在测序后对基因功能的研究。酵母有许多功能重复的基因, 常分布在染色体的两端, 当酵母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。基因丰余现象实际上是对环境的适应, 丰余基因的存在为进化适应提供了可选择的余地。基因组全序列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。

当前, 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, 并可用于其他高等真核生物基因组之“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和比较, 从基因组整体水平上阐述基因活动的规律。核心问题是基因组的多样性和进化规律, 基因组的表达及其调控, 模式生物体基因组研究等。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。

2. 蛋白质组学

蛋白质组对不少人来说,目前还是一个比较陌生的术语。它是在1994年由澳大利亚Macquarie大学的Wilkins等首先提出的,随后得到国际生物学界的广泛承认。他们对蛋白质组的定义为:“蛋白质组指的是一个基因组所表达的全部蛋白质”(proteome indicates the proteins expressed by a genome)。“proteome”是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,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是一样的,人类的基因总数约是6~10万个。单从DNA序列尚不能回答某基因的表达时间、表达量、蛋白质翻译后加工和修饰的情况,以及它们的亚细胞分布等。这些问题有望在蛋白质组研究中找到答案,因为蛋白质组是动态的,有它的时空性、可调节性,进而能够在细胞和生命有机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与基因组数据的整合,亦会对功能基因组的研究发挥重要的作用。

蛋白质组由原定义一个基因组所表达的蛋白质改为细胞内的全部蛋白质,比较更为全面而准确。但是,要获得如此完整的蛋白质组,在实践中是难以办到的。因为蛋白质的种类和形态总是处在一个新陈代谢的动态过程中,随时发生着变化,难以测准。1997年,Cordwell和Humphery-Smith提出了功能蛋白质组(functional proteome)的概念,它指的是在特定时间、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,中国生物科学家提出了功能蛋白质组学(functional proteomics)新概念,把研究定位在细胞内与某种功能有关或在某种条件下的一群蛋白质。

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,亦能丰富总蛋白质数据库,是从生物大分子(蛋白质、基因)水平到细胞水平研究的重要桥梁环节。

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,包括分离蛋白在图谱上的定位,分离蛋白的计数、图谱间蛋白质差异表达的检测等。一种细胞或组织的蛋白质组双向电泳图,可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分析,质谱已成为蛋白质鉴定的核心技术。从质谱技术测得完整蛋白质的相对分子质量、肽质谱(或称肽质量指纹以及部分肽序列等数据,通过相应数据库的搜寻来鉴定蛋白质。此外,尚需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上建立蛋白质组数据库。

从提出蛋白质组的概念到现在短短几年中,已于1997年构建成第一个完整的蛋白质组数据库——酵母蛋白质数据库,进展速度极快,新的思路和技术不断涌现,蛋白质组学这门新兴学科,在今后的实践中将会不断完善,充实壮大,发展成为后基因组时代的带头学科。

3. 生物信息学

HGP大量序列信息的积累,导致了生物信息学这门全新的学科的产生,对DNA和蛋白