



普通高等教育“十一五”国家级规划教材

分子遗传学

(第四版)

李振刚 编著

MOLECULAR
GENETICS
(FOURTH EDITION)



科学出版社

普通高等教育“十一五”国家级规划教材

分子遗传学

(第四版)

李振刚 编著

科学出版社

北 京

内 容 简 介

本书为普通高等教育“十一五”国家级规划教材，是在前三版基础上，根据学科发展和教学改革修订而成。本书以基因、染色质的结构与功能为基础，以真核细胞的基因调控为重点，以核酸与蛋白质的相互作用为线索，对蛋白质遗传、RNA 遗传进行了论述，对分子生物学的基石——中心法则，从分子遗传学的角度进行了探讨，提出了广义中心法则。对癌变、发育等重大生物学问题从分子遗传学的角度进行了论述。全书立论严谨，叙述流畅，观点明确，提出了不少新的见解与论点，跟踪了分子遗传学的发展，把科学知识的学习置于大胆讨论、活跃思维之中。

本书既是高等院校生物、医学、农林等专业的基础教材，又可作为研究生的参考用书，同时，还可供生物、遗传、医学、农林科技工作者参考使用。

图书在版编目(CIP)数据

分子遗传学/李振刚编著. —4 版. —北京: 科学出版社, 2014. 6
普通高等教育“十一五”国家级规划教材
ISBN 978-7-03-041121-1

I. ①分… II. ①李… III. ①分子遗传学-高等学校-教材 IV. ①Q75

中国版本图书馆 CIP 数据核字(2014)第 128988 号

责任编辑: 刘 畅/责任校对: 赵桂芬
责任印制: 阎 磊/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>

北京市文林印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2000 年 1 月第	一	版	开本: 787×1092	1/16
2004 年 9 月第	二	版	印张: 31	
2008 年 3 月第	三	版	字数: 813 000	
2014 年 6 月第	四	版	2014 年 6 月第十七次印刷	

定价: 68.00 元

(如有印装质量问题, 我社负责调换)

前 言

从 20 世纪 50 年代起, DNA 双螺旋的“旋风”高擎着中心法则 (central dogma) 的“大旗”席卷了古老的生物学的每一个角落。伴随分子生物学的热潮, 人们发现, “在现代科学中, 生命与非生命的界限几乎已经消失了”(Asimov, 1960)。B. Commoner 不无讽刺地说: “生物学应当改造成为核酸及其创造物的化学。”(Science, 1961, 133: 3466)

诚然, 从 DNA 的复制到蛋白质分子的合成, DNA 双螺旋模型与中心法则在分子水平上揭开了生物学崭新的一页。然而, 基因是怎样通过蛋白质决定个体性状的? 分子生物学的理论框架已显得过于简单而难以应付。化学诱变的先驱者 C. Auerbach 在第十五届国际遗传学大会的演讲中说: “双螺旋的发现是我们在理解突变方面的巨大飞跃; 但这既不是突变研究的开始, 也不是突变研究的终结……突变是在活细胞内发生的一种过程, 这样的过程不能用单纯的物理或化学模式进行充分描述。”著名的分子生物学家 E. Pennisi (2001) 曾指出: “基因占据生物学的中心舞台已经数十年。但是最近几年的工作提出, 虽然基因经常是挂牌的‘名角’(star billing), 可它们更像一些‘木偶’(puppets)。各种蛋白质, 有时是 RNAs, 在拉着‘线’(strings), 告诉基因在何时何地打开或关闭。”因此, 本书不仅论述 DNA 的遗传, 而且对近年来分子生物学在 RNA 遗传、蛋白质遗传方面的进展, 开辟专章进行论述; 不仅论述 DNA 密码遗传, 而且强调 DNA、RNA 及蛋白质的表观遗传 (epigenetics)。所以, 本书不是 DNA 双螺旋的生物化学, 更不是中心法则的演绎, 而是一切生物大分子遗传信息的分子生物学。

本书认为, 分子遗传学首先是遗传学, 本书坚实的理论基础仍然是 T. H. Morgan 的《基因论》, 而不是 DNA 双螺旋模型与中心法则。虽然中心法则论述了遗传信息从 DNA 到蛋白质的传递以及表达的机制, 受到了生物学家的普遍欢迎, 推动了分子生物学的崛起。但是, 20 世纪 70 年代, 反转录酶的发现使 Crick 重新修订了中心法则 (Nature, 1970, 227: 561)。内含子 (intron) 的发现, 更使 Crick 感到震撼 (Science, 1979, 204: 264)! 之后, 朊病毒 (prion) 及疯牛病使蛋白质遗传重新活跃于生物学界, 表观遗传 (DNA 甲基化、组蛋白密码、RNAi) 的提出更使中心法则面临严峻的挑战! 这说明, 中心法则、DNA 双螺旋学说与传统遗传学之间还有一定的差距, 《基因论》仍然是分子遗传学的坚实的理论基础。

本书特辟第 12 章中心法则导论进行讨论, 显示了本书另一个特点: 大学专业知识的教材不应该是传统教条的灌输, 而应该是带有研讨性质的论述, 使学生向学者过渡, 使他们不仅是本学科的学习者, 而且是本学科的研究者。

自 1980 年起, 我在中国科学技术大学开设分子遗传学课程, 并建立分子遗传学研究室, 从事转基因与染色质遗传的研究, 凡 20 余年。在本科教学以及指导硕士、博士研究生的过程中, 逐步形成了本书的内容。1985 年, 安徽科学技术出版社出版了一本小开本的《分子遗传学》; 1990 年, 中国科学技术大学出版社出版《分子遗传学概论》; 2000 年, 科学出版

社出版了新版本的《分子遗传学》，2004 年出版了第二版，2007 年被审定普通高等教育“十一五”国家级规划教材，2008 年出版第三版，本次修订后推出第四版。

本书自出版以来，得到读者的广泛支持，累计印刷 16 次。为了反映近年分子遗传学领域的一些新进展、新思维，笔者对本书进行了修订。本次修订在保持原有体系的基础上，增加了一些新的内容，比如端粒、非编码 RNA (ncRNA)、衰老的分子遗传学、表观遗传 (epigenetics) 与肿瘤等，同时对部分专业名词的进行了规范，希望通过对本书持续的修订，给读者带来更好的阅读与学习体验。感谢一批批读者们对本书的支持，希望广大的同行老师和同学在使用过程中提出宝贵的意见。

中国科学技术大学

李振刚

2014 年 5 月

zgli@ustc.edu.cn

目 录

前言

第 1 章 引论	1
1.1 分子遗传学的含义	1
1.2 分子遗传学的产生	3
1.3 分子遗传学的展望	8
参考文献	16
第 2 章 基因	18
2.1 基因的分子概念的发展	19
2.2 基因组学时代的基因概念	22
2.3 蛋白质基因概念的提出	26
2.4 组蛋白密码——对基因“唯 DNA”的质疑	27
2.5 新基因的产生	34
2.6 基因的进化	36
2.7 基因与 DNA	39
2.8 重复序列	48
2.9 重复基因	57
2.10 重复序列及重复基因的起源	63
2.11 断裂基因	64
2.12 重叠基因	67
2.13 模糊基因	69
2.14 转座子	70
2.15 新的基因概念的展望	82
参考文献	83
第 3 章 染色质	88
3.1 染色体与染色质——遗传物质的两种存在形式	88
3.2 常染色质与异染色质——染色体的两种功能状态	88
3.3 染色体单线性	89
3.4 染色质的分子组成	90
3.5 核小体的结构	98
3.6 常染色质基因表达的分子基础	104
3.7 异染色质形成的分子机制	108
3.8 染色质的非组蛋白框架	114
3.9 微生物的类染色质	115
3.10 染色质的复制与转录	115
3.11 染色体端粒	119

参考文献	121
第 4 章 基因的复制、转录与表达	127
4.1 中心法则	127
4.2 DNA 复制	128
4.3 转录过程——RNA 合成	140
4.4 mRNA——蛋白质合成的模板	142
4.5 蛋白质合成	150
参考文献	163
第 5 章 基因的调控	166
5.1 基因调控的基本模型	166
5.2 调控序列与调控蛋白	167
5.3 基因的分子调控	171
5.4 原核生物操纵子的特点	174
5.5 σ 因子级联调控模型	177
5.6 真核基因的分子调控——多因子调控	178
5.7 真核基因的染色质调控	185
5.8 转录后的基因调控	189
5.9 真核基因的调控模型——Davidson-Britten 模型	191
5.10 真核基因的多位点协同调控	193
参考文献	197
第 6 章 蛋白质遗传	200
6.1 朊病毒——感染性蛋白质	201
6.2 朊病毒的繁殖	203
6.3 朊病毒是细胞中的非孟德尔遗传因子	205
6.4 朊病毒的遗传标准	209
6.5 朊病毒蛋白质——蛋白质基因	210
6.6 朊病毒蛋白中有一个独立的 prion 决定域	211
6.7 消耗性蛋白质与遗传性蛋白质	212
6.8 作为细胞结构的“蛋白质复合体”的遗传	213
参考文献	222
第 7 章 RNA 遗传	224
7.1 RNA 世界	224
7.2 RNA 干扰	227
7.3 RNAi 对基因表达的作用	232
7.4 RNA 编辑	236
参考文献	242
第 8 章 动物发育的分子生物学	245
8.1 发育分化理论	245
8.2 胚胎极性与背腹的决定——卵皮层的旋转与发育的启动	248
8.3 器官组织的分化——诱导的分子机制	248

8.4	发育程序的分子机制	253
8.5	形态发生的分子机制	257
8.6	非 A-P 型 <i>H</i> 基因: <i>En</i> 、 <i>Pax</i> 、 <i>Eux</i> 等	262
8.7	细胞凋亡的概念	265
8.8	发育基因调控网	269
8.9	衰老的分子遗传学	273
	参考文献	278
第 9 章	癌分子遗传学	282
9.1	癌与癌基因	282
9.2	癌的发生——单克隆起源	283
9.3	癌变的起因	284
9.4	抑癌基因	287
9.5	原癌基因转变为癌基因的途径	289
9.6	单一突变不足以引起癌变——癌变的多阶段性质	291
9.7	细胞癌变多阶段性的分子基础	292
9.8	细胞癌基因与信息传递	293
9.9	与发育相关的癌基因	295
9.10	表观遗传 (epigenetics) 与肿瘤	303
	参考文献	305
第 10 章	突变、修复与重组	307
10.1	基因的突变	307
10.2	自发突变	308
10.3	诱发突变	310
10.4	DNA 突变(损伤)的修复	316
10.5	突变不完全是随机过程	324
10.6	呼救 (SOS) 系统	324
10.7	基因的重组	326
	参考文献	366
第 11 章	植物发育的分子遗传学	369
11.1	植物发育的分子遗传学特点	369
11.2	植物胚胎发育的极性——发育的起点	371
11.3	植物的体型格局的发育	377
11.4	植物的形态发生	384
11.5	植物发育与形态发生中的基因	395
	参考文献	420
第 12 章	中心法则导论	423
12.1	引言	423
12.2	中心法则的提出及修正	424
12.3	对中心法则的挑战	427
12.4	中心法则在生命系统中的地位	448

12.5 中心法则与遗传信息流·····	458
12.6 中心法则的未来·····	461
参考文献·····	467
分子遗传学习题·····	473
习题答案·····	480



第1章 引 论

1.1 分子遗传学的含义

分子遗传学是研究遗传信息大分子的结构与功能的科学，又称为狭义的分子生物学^[1]。它依据物理、化学的原理来解释遗传现象，并在分子水平上研究遗传机制及遗传物质对代谢过程的调控^[2]。

分子遗传学不同于一般的遗传学。传统的遗传学“主要研究遗传单元在各世代的分布情况”^[3]，而分子遗传学则着重研究遗传信息大分子在生命系统中的储存、复制、表达及调控过程。它的研究范畴如图 1.1 所示。

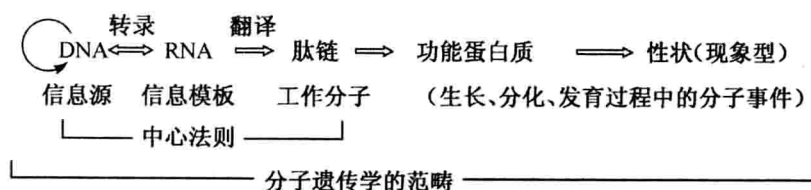


图 1.1 分子遗传学的研究范畴

1.1.1 分子遗传学不等于中心法则的演绎——中心法则可能是一个过于简化的、错误的法则

分子生物学家把分子遗传学理解为：“DNA 与 RNA 的复制与转录以及 RNA 的翻译”^[1]，也就是中心法则的全过程。其实，分子遗传学的研究范畴要比中心法则广泛得多，深刻得多。它首先是遗传学，其坚实的理论基础仍然是摩尔根（T. H. Morgan）的《基因论》^[3]，而不是中心法则。中心法则只是对 DNA 编码序列（基因）及其表达在分子水平上的解释。但是，从中心法则到性状的形成，仍然是一个复杂的、语焉不详的生物学过程；它不是中心法则所能解释清楚的。正如 DNA 突变研究的先驱者 Auerbach^[4]所说：“DNA 双螺旋的发现是我们在理解突变方面的巨大飞跃，但这既不是突变研究的开始，也不是突变研究的终结”，“突变是在活细胞内发生的一种过程，这样的过程不能用 DNA 的简单化学反应来说明。”分子遗传学家把目光紧紧盯在 DNA 上的时间已经太久，他们应该研究活细胞内与遗传变异有关的一切分子事件，并把目光扩展到 RNA、蛋白质甚至糖分子上面。

J. Cotter（国际绿色和平组织科学顾问）和 C. Thea（德国绿色和平组织项目主任）指出：“分子生物学的中心法则是基于 DNA 是活跃的，并可以控制基因表达的基础上提出‘DNA 生成 RNA，RNA 生成蛋白质’。50 年来的研究已经改变了这个认识。50 年来的发现使我们的认识从 DNA 是储存信息、控制生成蛋白质的主宰分子转变为 DNA 只是一个多层

次复杂调控网络中的信息储存库 (store of information)”。所以,“在新的科学语言中,DNA 的角色变得更加被动;DNA 的物理化学性质说明了它的被动性。尽管分子质量很大,DNA 仍然是稳定的分子,表现出不活跃、被动的性质”。他们大胆地指出:“中心法则的语言将被修改,DNA 不再是主动的,而变成了被动的。DNA 不是自我复制,而是由一系列相互作用的酶复制的。DNA 不能复制 DNA,它只是模板”。并最后断言:“回顾这 50 年,许多大大小小的发现动摇了最初的中心法则和 DNA 是恒定的、不可改变的观点。这些新发现构成了现代分子生物学的基础,它们显示了虽然中心法则的核心部分仍然正确,但是很显然是一个过于简化的模型。”(J Cotter, C Thea. 2003. 50 years since the double helix: genetic engineering is crude and old-fashioned. greenpeace org)

笔者认为,如果如 Cotter 和 Thea 上述的“中心法则的核心部分”只是指“DNA 生成 RNA, RNA 生成蛋白质”的话,可以说中心法则是一个过于简化但仍然正确的模型。但是,中心法则的核心部分不会是从 DNA 到蛋白质的一般的生化过程,而是具有遗传学意义的法则。Crick 对中心法则的表述例如:“信息一旦进入蛋白质,它就不可能再输出”,“详细点说,信息的传递从核酸到核酸或从核酸到蛋白质或许是可能的,但从蛋白质到蛋白质或从蛋白质到核酸则是不可能的”(详见本书 12.2.3),在朊病毒能够复制、繁殖的事实面前,“信息一旦进入蛋白质,它就不可能再输出”就是错误的。因此,中心法则可能是一个过于简化的、错误的法则。

1.1.2 分子遗传学不是核酸及其衍生物(蛋白质)的生物化学

分子遗传学研究的对象是分子水平上的生物学过程——遗传及变异的过程^[1]。它研究的是动态的生命过程,而不是在试管里或电泳仪上孤立地研究生物大分子的结构与功能的简单的因果关系。有人认为分子遗传学是生物化学的一部分,是研究核酸及其衍生物(蛋白质)的一个分支。这正说明他(她)对生物学与生命现象认识的肤浅。早在 1992 年, *Nature* 杂志的主编 J. Maddox 曾著文“*Is molecular biology yet a science?*”^[5], 指出:“现在有那么一群叫做分子生物学家的人,他们的文章无视整体的动物、植物,也很少言及它们的生理学。对这些人来说,实验的资料大部分来自所谓‘凝胶’……分子生物学在很大程度上正变成定性的科学……如果事情只是简单地说明某个基因版本与某种遗传病相关,那么分离这种片段(如电泳),然后测序,已足矣。”但是,“以往几年的巨大成就表明,生命过程是由严格控制下进行的一些有序事件所组成”。“如果在生物的发育过程中不同形态的出现是决定于形态因子(molecular morphogens)的分子扩散作用,那么它们要何种浓度才能满足正常的发育?”这就需要研究活细胞内的动态的、整体性的分子事件才能作出比较真实的回答。而那种只管因(DNA/RNA)果(蛋白质/酶),不管过程的定性式的研究是无法回答这类实质性问题的。Maddox^[5]警告说:“在人们长期地为细胞生物学现象寻找定性的解释中,他们将会相信细胞只不过是一个充满了分子开关的‘袋子’,它们作为分子传动器或开或关而出现在预定的事件序列中。这会造成一种印象,按照还原主义者的程序所进行的,对有条理地分割与干燥了的生物体的描述不会远离事物的真相……因此分子生物学家应当修正他们的路线,重振质量作用定律(law of mass action)。”

因此要真正地在分子水平上了解遗传变异的本质,仅仅研究核酸或蛋白质的生物化学是远远不够的。对于那些从活细胞中分离出来的、“干燥”了的生物大分子的化学研究是必要的,但绝不是分子遗传学研究的中心内容,更不是它的全部内容。分子遗传学所研究的应该

是细胞中动态的遗传变异过程以及与此相关的所有的分子事件。很显然, 这些事件决不限于中心法则, 也不限于核酸、蛋白质。

1.2 分子遗传学的产生

分子遗传学的产生标志着分子生物学的崛起。分子遗传学是微生物学、遗传学、化学、物理等学科相互交叉、相互渗透的产物, 究其来源, 错综复杂。这里我们仅将分子遗传学发展的几个主要阶段做一简略叙述。

1.2.1 物理学的渗透——分子遗传学的物理学语言

1945年, 量子力学的创始人之一薛定谔(E. Schrodinger)的《生命是什么》^[6]一书出版。他倡导用物理学与化学的思想和方法研究生命的秘密。他说: “目前的物理和化学虽然还缺乏能力来说明生物体中发生的各种事件, 然而丝毫没有理由怀疑它们是不能用这两门科学来说明的。”

《生命是什么》一书可以说是分子生物学研究开始的标志。薛定谔是第一个用物理和化学语言对生命进行了系统描述的物理学家。薛定谔用物理与化学的语言为我们提供了现代分子生物学完整的理论框架, 把生命、遗传、基因彻底分子化, 并把信息论、量子力学的概念引入生命过程之中。

1.2.1.1 生命之本——负熵

“生命的特征是什么? 一块物质什么时候可以说是活的呢? 那就是当它继续在‘做某些事情’、运动、新陈代谢等, 而且可以指望它比一块无生命物质在相似情况下‘维持生活’的时间要长得多。”

20世纪前期, 人们认为生命现象并不服从热力学定律, 因而不能用物理学定律来解释。根据热力学第二定律, 自然界演化的方向是从有序到无序, 而生命的发生、演化、分化、生长等过程, 显然是从组织程度较低的无序到组织程度较高的有序。这是无生命世界中难以实现的。

薛定谔^[6]说: “我们不必因为物理学的普遍定律难以解释生命而感到沮丧。”“一个有机体能够避免很快地衰退为惰性的‘平衡’态, 似乎成了难解之谜, 以致在人类思想的最早时期, 曾经认为有某种特殊的非物质的力或超自然的力(‘活力’)在有机体里起作用, 现在还有人是这样主张的。”

薛定谔用“负熵”的概念, 使物理学渗透到生物学领域, 开启了分子生物学的崛起!

熵是一个可以计算的物理学的量。在绝对零度时, 任何一种物体的熵等于零。当你以缓慢的、可逆的、微小的变化使物体进入另一种状态时, 熵的增加是这样计算的: 在此步骤中必须供给的每一小部分热量, 除以供给热量时的温度(摄氏), 然后把所有这些求得的商数加起来。例如, 当熔解一种固体时, 熔化热除以熔点温度, 就是它的熵的增加数。因此, 熵的单位是卡/摄氏度。

“ $\text{熵} = k \lg D$, k 是所谓的玻尔兹曼常数 ($k = 3.2983 \times 10^{-24} \text{ cal}^\circ/\text{C}$), D 是有关物质的原子无序状态的数量量度。要用简短的非专业性的术语对 D 这个量做出精确的解释几乎是不

① $1 \text{ cal} = 4.1868 \text{ J}$

可能的。它所表示的无序，一部分是那种热运动的无序，另一部分是存在于随机混合的、不是清楚地分开的各种原子或分子中间的无序。”^[6]

薛定谔认为，既然 D 是无序的度量，它的倒数 $1/D$ 可以作为有序的度量。因为 $1/D$ 的对数正好是 D 的负对数，玻尔兹曼的方程式可以表达为负熵 $= k \lg(1/D)$ 。所以，负熵就是取负号的熵，它本身是有序的一个量度。

“在我们的食物里，究竟含有什么样的宝贵东西能够使我们免于死亡呢？那是很容易回答的。每一个过程、事件、事变——你叫它们什么都可以，一句话，自然界中正在进行着的每一件事，都是意味着它在其中进行的那部分世界的熵的增加。因此，一个生命有机体在不断地增加它的熵（你或者可以说是在增加正熵），并趋于接近最大值的熵的危险状态，那就是死亡。要摆脱死亡，就是说要活着，唯一的办法就是从环境里不断地汲取负熵，我们马上就会明白负熵是十分积极的东西。有机体就是以负熵为生的。或者，更确切地说，新陈代谢中的本质的东西，乃是使有机体成功地消除了当它自身活着的时候不得不产生的全部的熵。”^[6]

因此，生命是以负熵为生。有机体本身吸引了一连串的负熵去抵消它在生命活动中增加的熵，从而使它自身维持在一个稳定的而又很低的熵的水平上，生命因此得以继续。他指出一个开放系统的熵不一定增加，它可以从外界引入“负熵”；生命正是一个开放系统。

$$S = \Delta S_e + \Delta S_i, \Delta S_i \geq 0$$

式中， ΔS_e 是外部引入的熵，它可以是负的。 ΔS_i 是内部产生的熵，不能小于零。因此，系统的总熵 S 可正可负，只要 $\Delta S_i \geq 0$ 就不违反热力学第二定律。

1.2.1.2 基因之本——非周期性晶体分子

薛定谔认为，“一个有机体在它自身集中了‘秩序之流’，从而避免了衰退到原子混乱（从合适的环境中‘吸取秩序’）。这种惊人的天赋似乎同‘非周期性固体’，即染色体分子的存在有关。”^[6]他指出，在非周期性晶体中是靠每个原子和每个自由基在固体里发挥各自的作用。在这种结构里，不必有大量的原子就可产生出几乎是无限可能的排列。

薛定谔据此推导出遗传密码的存在。他说：“一个基因——也许是整个染色体纤丝——是一种非周期性的固体。”“就基因分子的图式来说，微型密码是丝毫不错地对应于一个高度复杂的特定的发育计划，并且包含了使密码发生作用的手段，这一点已经不再是难以想象的了。”^[6]

他还推算了一个基因的大小。他认为基因“在染色体上定位以后，测量那条染色体的长度并除以特定的数目，再乘以染色体的横截面，就得出了我们所需要的估计数”。“果蝇的某些细胞（唾腺细胞），由于某种原因是大大地增大了的，它们的染色体也是如此。在这些染色体上，你可以分辨出纤丝上的深色横纹的密集图案。C. D. Darlington 曾经说过，这些横纹的数目（他当时说是2000个）虽然比较多，但大体上等于用繁育试验得出的、位于染色体上的基因数。他倾向于认为，这些横纹带标明了实际的基因（或基因的分离）。在一个体积正常的细胞里测得的染色体长度，除以横纹的数目（2000），他发现一个基因的体积等于边长为 $300\text{\AA}$ ^① 的一个立方体。”“ $300\text{\AA}$ 大约只有 100 个或 150 个原子距离，所以，基因应该包含的原子，肯定不会超过几百万个。”^[6]

① $1\text{\AA}=0.1\text{nm}$

1.2.1.3 基因的突变——生物学的量子论

“遗传的机制是同量子论的基础密切相关的，确切地讲，是建立在量子论的基础之上的。”^[6]“量子论的最大启示是在‘自然界的圣典’里发现了不连续性的特点，而当时的观点却认为自然界中除了连续性外全都是荒谬的。”

按量子力学来说，一个物体在大范围内连续地改变着它的能量。例如，摆的摆动会由于空气的阻力而逐渐地缓慢下来。但是对于微观系统来说，如在原子这一级上的行为就不同了。它们大多数是不连续地发生变化的，是量子化的，通常称之为量子跃迁。薛定谔^[6]对基因的突变进行了量子力学的描述：“我们将假定一个基因的结构是一个巨大的分子，只能发生不连续的变化，这种变化在于原子的重新排列并导致一种同分异构的分子。这种重新排列也许只影响到基因中的一小部分区域，大量的各种不同的重新排列也许是可能的。从任何可能的同分异构体中，把实际的构型分离出来的阈能一定是很高的（这是同一个原子的平均热能相比），致使这种变化成为一种罕有事件。这种罕有事件我们认为就是自发突变。”“就是说，突变是不出现中间形式的，而是‘跃迁式’的变异。”“跃迁式，并不是说这个变化是相当大的，而是说这是一种不连续的变化，在未变和少许改变之间没有中间形式。德弗里斯称之为突变。”德弗里斯的突变论，不妨称为生物学的量子论。因此，“突变并不是由连续的小剂量辐射相互增强而产生的一种积累效应。突变一定是由辐射期间发生在一条染色体中的单一事件所产生的”。

薛定谔还估算了一个突变点在遗传物质上的大小。他说：“突变的单一事件正是在生殖细胞的某个‘临界’体积内发生的电离作用（或类似的过程）。这种临界体积有多大呢？它可以根据观察到的突变率，按照这样的考虑来做出估计，即如果每立方厘米产生 50 000 个离子的剂量，使得任何一个配子（它们是在照射的区域里的）以那种特定的方式发生突变的机会只是 1 : 1000，那么，我们就可断定那个临界体积，即电离作用要引起突变所必须“击中”的“靶”的体积只有 $1/50\,000\text{cm}^3$ 的 $1/1000$ ，就是说，只有五千万分之一立方厘米。”^[6]

“如果在距离染色体上某个特定的点不超过‘10 个原子距离’的范围内发生了一次电离（或激发），就有产生突变的一次机会。”

综上所述，我们可以看到薛定谔在《生命是什么》这本小册子中，用当时水平的物理与化学语言，给我们构筑了现代分子生物学的框架。如果把他的“非周期性晶体”换成 DNA，那就是一个现代分子生物学的宣言！

1.2.2 微生物学向遗传学的靠拢

虽然摩尔根的“基因论”在 1926 年已经问世，但 20 世纪 30 年代的微生物学家却往往采用拉马克的遗传观点，这是因为他们对微生物的遗传可塑性往往有极深刻的印象。例如，在含有致死药物的培养基上，可以很容易地培育出对各种致死药物有抗性的微生物品系；把不能利用乳糖的微生物放在以乳糖为主要营养来源的培养基上，可以培育出利用乳糖的新品种。似乎人们所期望的微生物的任何变异，都能通过适当的培养而产生出来，这就容易使人相信培养基中的物质可以引起微生物遗传结构的定向变异。20 世纪 40 年代抗生素的大规模使用，发生了病原菌的抗药性问题。许多灵丹妙药经过几个月的使用就会失灵，这使医药界大伤脑筋。面对这种局面，微生物学家必须回答这样一个问题：细菌的抗药性是后天获得的定向变异，还是早已发生的变异而被药物所筛选？只有明确了这个问题，医生们才能确定使

用抗生素的方案：两种以上的抗生素，是同时使用还是交替使用？

实验表明，病原菌的一种抗药性，在没有该药物存在的情况下随机地发生了。利用影印培养 (replica plating)，可以在从来没有接触过该种致死药物（如青霉素）的培养基中分离出抗该药的菌落。这说明抗药性的产生并不是由于微生物在某种药物作用下的后天获得性遗传，而是随机发生的自发突变经过药物的筛选作用，使不具有抗药性的菌体死亡，使具有抗药性的变异菌体大量繁殖起来。拉马克学说在微生物学的最后一道防线崩溃了，人们开始转向摩尔根基因突变理论。

1.2.3 生化遗传学的出现

近代遗传学的基础已稳固建立，人们开始转向研究基因是怎样发生作用的问题。生物化学家很自然地把性状的差别与不同的生化反应联系起来，把支配性状的基因与控制生化反应的酶联系起来。1923 年，英国人加罗德 (A. E. Garrod) 证明人类的黑尿症 (alcaptonuria) 是一种隐性遗传病。当这种纯合隐性基因存在时就不能产生尿黑酸酶，使尿黑酸（蛋白质的代谢产物）不能最终分解为二氧化碳和水而积累于血液中。这样，一部分尿黑酸多聚物沉积于软骨及其他结缔组织中，使患者年老时发生黄褐病 (ochronosis)，患者双颊、鼻、巩膜及耳廓呈灰黑色或褐色，有时并发变性关节炎；一部分尿黑酸则随尿排出，暴露在空气中氧化成黑色素，使尿迅速转为黑色。这一症状自婴儿期即出现（常在尿布上见黑色斑点），终身如此，故称黑尿症。这表明基因通过对酶合成的控制而影响遗传性状的发育。

但是用高等生物来研究各种突变型的生化细节，是一项繁杂艰难的工作，后来由于采用了粗糙脉孢霉为材料才得以突破。粗糙脉孢霉 (*Neurospora crassa*) 的菌丝体由许多菌丝细胞组成，每个菌丝中含有多个单倍体细胞核，菌丝体能通过产生单倍体的分生孢子进行无性繁殖，有性繁殖时则必须在两个不同接合型 (mating type) 的菌丝体之间才能发生 (图 1.2)。一种接合型的分生孢子通过另一相对接合型的受精丝进入子实体中，然后分裂成若干单倍体核并与子实体中的单倍体核融合，形成合子。合子核在狭窄的孢子囊中进行减数分裂，减数分裂后的 4 个细胞再进行一次有丝分裂，最后形成 8 个单倍体的子囊孢子。子囊孢子萌发后，产生新的菌丝体。

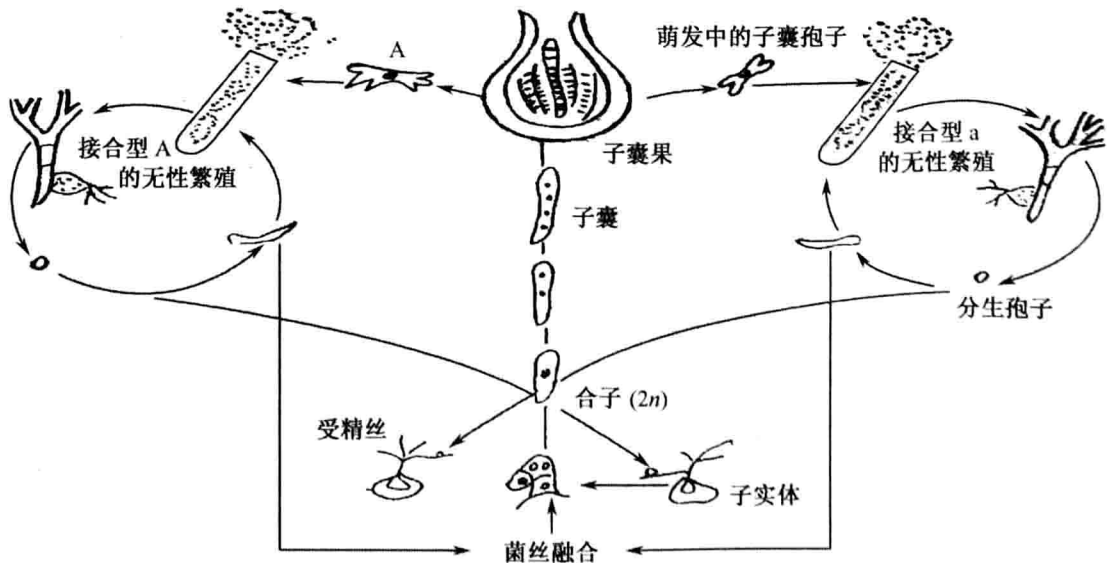


图 1.2 粗糙脉孢霉生活史

粗糙脉孢霉作为研究材料的特点是：① 因为是单倍体，没有显性基因掩盖隐性基因问题；② 孢子能在简单的化学培养基上生长并产生一系列营养缺陷型突变体；③ 一定的基因型表现为一定的性状，基因型的分离与重组必然表现为性状的分离与重组。在粗糙脉孢霉的减数分裂中，如果四分体的非姐妹染色体不发生交换，则 8 个孢子来自两种基因型（两种亲本型），必然发生 4 : 4 的性状分离。如发生交换，则 8 个孢子来自四种基因型，必然发生 2 : 2 : 2 : 2 的性状重组。如果某一性状是由单基因所决定的，则在交换后也就不存在性状重组问题，在后代中总是出现 1 : 1 的性状分离。用诱变剂处理粗糙脉孢霉能得到各种营养缺陷型的突变体。有一种精氨酸营养缺陷型，必须在培养基中加入精氨酸才能正常生长发育，这种突变型的共同特点是最终不能形成精氨酸。但从具体的生化反应步骤上看，这种突变型又可分为三类（表 1.1）。

表 1.1 粗糙脉孢霉精氨酸缺陷型三类突变体的遗传控制

三类突变体生长时培养基中所必须加入的氨基酸			精氨酸形成中的酶反应与基因控制	
突变体 1	突变体 2	突变体 3		
			谷氨酸	$\begin{array}{c} \text{NH}_2 \\ \\ \text{COOHCH}_2\text{CH}_2\text{CHCOOH} \end{array}$
		+	鸟氨酸	$\begin{array}{c} \text{NH}_2 \\ \\ \text{NH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CHCOOH} \end{array}$
	+	+	瓜氨酸	$\begin{array}{c} \text{O} \quad \text{NH}_2 \\ \quad \\ \text{NH}_2\text{—CNHCH}_2\text{CH}_2\text{CH}_2\text{CHCOOH} \end{array}$
+	+	+	精氨酸	$\begin{array}{c} \text{NH} \quad \text{NH}_2 \\ \quad \\ \text{NH}_2\text{CNHCH}_2\text{CH}_2\text{CH}_2\text{CHCOOH} \end{array}$

从表 1.1 可见第一类突变型不能形成精氨酸，可能因为它们不能把瓜氨酸转化为精氨酸；第二类突变型不能形成精氨酸是由于它们不能把鸟氨酸转化为瓜氨酸。第三类突变型缺乏形成精氨酸的能力，是因为它们不能把谷氨酸转化为鸟氨酸。当把第一类突变型与正常亲本杂交后，一个子囊中的 8 个孢子总是表现 4 : 4 的性状分离，其中 4 个孢子能在缺少精氨酸的培养基上生长，另外 4 个则只有在加入精氨酸的培养基上才能生长。这说明第一类突变型是单一基因突变的结果。这个基因只能是表 1.1 中的基因 1，同理可证，第二类、第三类突变型分别是基因 2、基因 3 突变的结果。这意味着一个专一的基因控制一个特异的生化反应，而每一个特异的生化反应都涉及一个特异酶的催化作用。这种基因与酶之间的相关性就产生了“一个基因一个酶”的假说^[7]。它表明一个基因控制着一个酶的形成。酶是蛋白质，它第一次暗示在基因的分子结构与其产物（蛋白质）之间存在着对应关系。

1.2.4 从生化遗传学到分子遗传学

基因与酶（蛋白质）的对应性，使人们想到了基因在遗传信息上与其产物相关。以下的三项重要发现更促成了从生化遗传学向分子遗传学的转变。

(1) 20 世纪 40 年代解决了遗传的物质基础问题。1928 年，格里菲思 (F. Griffith)^[8]把活的 RⅡ 型肺炎球菌（无致病力）与灭活的 SⅢ 型肺炎球菌（活时有致病力）分别注射入小白鼠体后，小鼠仍然健康，但是当用 RⅡ 型活菌与灭活的 SⅢ 型死菌共同注入鼠体后，则小

白鼠被感染死亡，在死鼠体中发现大量活的 SⅢ 型肺炎球菌。这说明 SⅢ 死菌中的遗传物质使 RⅡ 型转化为 SⅢ 型。1944 年艾弗里 (O. T. Avery)^[9] 进一步证明使 RⅡ 型转化为 SⅢ 型的遗传物质正是 SⅢ 菌体的 DNA。后来赫尔希等 (A. Hershey 和 M. Chase)^[10] 用 ^{35}S 与 ^{32}P 分别标记 T2 噬菌体的蛋白质外壳与核心 DNA。发现在感染过程中蛋白质外壳留在菌体外面，只有 DNA 进入菌体。在感染后 25min 左右菌体被溶解，产生出 100~150 个完整的 T2 噬菌体。这个实验令人信服地证明 DNA 是遗传的物质基础，它含有产生整个 T2 噬菌体的遗传信息。

(2) 20 世纪 50 年代确定了分子水平上的遗传机理问题。1953 年沃森 (J. Watson) 和克里克 (F. Crick) 提出的 DNA 分子的双螺旋结构模型 (图 1.3)^[11]，其主要内容是：双螺旋的两条链以氢键相连，碱基的配对原则是 A 与 T，C 与 G，这个模型合理地解释了 DNA 复制和转录过程，解决了 DNA 的自我复制问题，巩固了 DNA 作为遗传物质的地位。

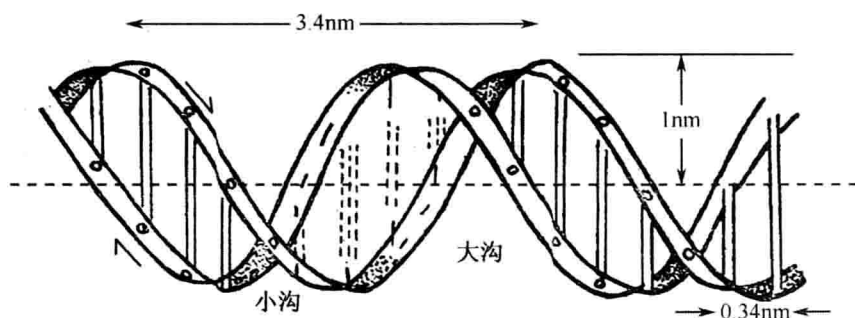


图 1.3 Watson-Crick DNA 双螺旋模型

(3) 20 世纪 60 年代解决了遗传密码问题。1955 年桑格 (F. Sanger) 测定了牛胰岛素中氨基酸残基的准确顺序；1958 年克里克提出中心法则 (centre dogma)^[12]。这些工作鼓舞着人们把核酸与蛋白质的线性结构联系起来，终于导致了 1967 年“遗传密码字典”的问世。至此，分子遗传学有了稳固的基础。

1.3 分子遗传学的展望

分子遗传学的产生，标志着分子生物学已逐步地成熟为一门科学。它的每一项成就，从 DNA 双螺旋结构、遗传密码到基因工程，都使整个科学界为之鼓舞。基因是什么？基因如何复制？基因怎样行动？这些问题在 DNA 的层面上，我们现在似乎都已得到了满意的答案，但是分子遗传学并不就此止步，它正以旺盛的生命力向新的领域推进。

1.3.1 基因的概念

任何一门科学都是以概念为基础的。化学是以原子-分子概念为基础的，而遗传学则是以基因概念为基础的。基因概念的演变，标志着遗传学的发展。

摩尔根在《基因论》中提出遗传粒子理论，整个基因论是以粒子性的基因彼此独立，互不重叠为依据的。他认为基因连接成直线，好像线上的联珠。但是，分子遗传学的发展证明了基因不仅可以重叠^[13]，而且可被分隔^[14]。这在基因概念上是个突破。它不仅使摩尔根的基因概念过时，也使风行一时的“基因是功能单位”的顺反子概念不能成立。为此，吉尔伯特 (Gilbert) 提出“基因是转录单位”的新概念^[15]。这个概念也还不能令人满意，因为有