

基因与基因组分析

Analysis of Genes and Genomes

著 者 Richard J Reece
主 译 刘树林

 WILEY



北京大学医学出版社

基因与基因组分析

著 者 Richard J Reece

主 译 刘树林

副主译 石 爽 郭志荣 赵革新

译 者 (以姓氏拼音为序)

| | | | | |
|-----|-----|-----|-----|-----|
| 艾效曼 | 部小霞 | 范 芸 | 冯 海 | 高 倩 |
| 高志勇 | 龚 俊 | 郝 翰 | 胡代伦 | 黄 鹤 |
| 黄远深 | 李 凡 | 李文德 | 李永华 | 李有泉 |
| 廉文清 | 梁靖瑞 | 林 然 | 刘茜玮 | 刘若琳 |
| 罗凌琪 | 桑 田 | 宋韩明 | 汪业军 | 温娟娟 |
| 解玮琳 | 鄢谢桥 | 尹利民 | 原鹏波 | 岳 睿 |
| 赵亚楠 | 赵宇池 | 郑 鑫 | | |

北京大学医学出版社

Analysis of Genes and Genomes

Richard J Reece

Copyright © 2004 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England

All Rights reserved.

Authorised translation from the English Language edition published by the John Wiley & Sons, Ltd.

Simplified Chinese Translation Copyright © 2006 Peking University Medical Press

北京市版权局著作权合同登记号：图字：01-2005-1900

JIYIN YU JIYINZU FENXI

图书在版编目 (CIP) 数据

基因与基因组分析 / (英) 里斯 (Reece, R. J) 原著; 刘树林译.

—北京: 北京大学医学出版社, 2006.8

书名原文: Analysis of Genes and Genomes

ISBN 7-81071-960-2

I. 基 ... II. ①里 ... ②刘 ... III. 基因—分析—医学院校—教学参考资料
②基因组—分析—医学—教学参考资料 IV. Q343.1

中国版本图书馆 CIP 数据核字 (2006) 第 0445569 号

基因与基因组分析

主 译: 刘树林

出版发行: 北京大学医学出版社 (电话: 010-82802230)

地 址: (100083) 北京市海淀区学院路 38 号 北京大学医学部院内

网 址: <http://www.pumppress.com.cn>

E-mail: booksale@bjmu.edu.cn

印 刷: 莱芜市圣龙印务有限责任公司

经 销: 新华书店

责任编辑: 安林 责任校对: 杜悦 责任印制: 郭桂兰

开 本: 787mm × 1092mm 1/16 印张: 23.5 字数: 575 千字

版 次: 2006 年 10 月第 1 版 2006 年 10 月第 1 次印刷

书 号: ISBN 7-81071-960-2 / R · 960

定 价: 105.00 元

版权所有, 违者必究

(凡属质量问题请与本社发行部联系退换)

作者前言

像“基因工程”和“克隆”这样能激发人情绪的词语实不多见。报纸和电视总是用这些词语来描述不很正确的事物，甚至可能违反了自然规律。基因工程和基因的改变总能令人联想到奇怪的食物和异常的动物。但是，我希望读者在阅读本书时，能意识到基因工程及其赖以存在的分子生物学技术是理解生物如何工作的关键因素。几千年来，人们一直在运用基因，经常是不自觉地，例如通过有选择育种来获得预期提高的某种品质。我们在如何看待基因方面现正处于一个分水岭阶段：迄今的50年里，我们了解了遗传物质的结构；今后要做的是应用我们的研究能力来了解我们体内各基因之间及其与环境条件之间是如何反应的。举例来说，目前尚不可能在生化基础上确定为什么人们对药物治疗存在不同反应，但是我们已经掌握了探讨有关问题的技术。因此，后基因组时代的到来在带给我们兴奋的同时，也使我们有能力去解决问题——不管我们是否真的要去解决。

对基因和基因组的分析很容易被写成一系列用于解决某特定问题的技术。我在写作时刻意避免了这一点，并在适宜之处使用具体的例子来解析问题和提出可能的解决方案。我尽量引用已发表的著作，而且引用的全部是原始文献，以便有兴趣的读者继续探讨。这样做也使我得以将大量的观点和实验结合其历史背景加以描述。有一种广为流行的误解，认为我们关于基因如何工作的了解应该完全归功于 Watson 和 Crick。的确，他们的贡献绝不应被低估，然而其他很多人的工作也不应被忽略。人类全基因组测序，以及同样重要或甚至更为重要的实验模式生物的基因组研究，给未来生物科学的发展提供了特殊的机遇。现在越来越多的实验可以在全基因组水平上进行，而我们现在才刚刚开始认识其结果。

我写这本书时所遇到的问题之一，就是如何把握深度和广度之间的平衡。我特意集中于容易操作的实验系统——即用 *E.coli* 代表原核生物、酵母代表真核生物。此外，我所提到的高等真核生物几乎都是哺乳动物、特别是人，其目的在于为读者传递一些有关当今流行的研究思路和实验方法的信息，同时，也介绍一些与目前实验相关的历史背景。我们不可忽略前人的工作。这一做法使得果蝇及除 *E.coli* 之外的其它原核生物被排除在外，但这并不意味着对这些领域的轻视。我尽量用具体的例子概述本领域的全貌，而不是对方方面面蜻蜓点水或针对少数几个事例深究其错综复杂的细节。我是否确实将这一平衡把握得恰当尚待读者评判。然而，我可以断言的是，生物学研究现在正处于一个空前的令人激动的时期，希望本书能够对此有所反映。

致谢

在编写这本书的过程中，我得到了众多朋友的支持。百密一疏的失误或差错都是我自己的责任，但我在此感谢使错误得以减少到最低可能的 David Timson, Noel Curtis, Cristina Merlotti, Chris Sellick, Carolyn Byrne, Ray Boot-Handford 和 Ged Brady。我的感激之情同样地呈献给 Robert Slater (University of Hertfordshire) 和 Mick Tuite (University of Kent)，感谢他们极具建设性的评论和建议。同时，感谢书中提及的为此书无私地提供数据、特许引用其工作的朋友和同事们。我特别感谢的还有 Jordi Bella，他为我展示了简便易学的分子制图程序。John Wiley 出版公司的 Nicky McGirr 告诉我这是一个好立项，她无比的热诚和鼓励使我得以度过我不自信的一段时光。当然，她是对的。用于检验本书诸多观点的“豚鼠”是曼彻斯特大学基因工程系各年级的学生们。我感谢读过本书部分手稿的学生对我从前的观点提出的挑战。Judith, Daniel 和 Kathryn 对写这本书从一开始到写作的全过程都表现出惊人的耐心。读者如果觉得开卷有益，那么要感谢的应该是他们而不是我。最后，我希望对我的老师 Tony Maxwell 和 Mark Ptashne 表示感谢，他们以其各自不同的方式，使我体味到他们对科学发自内心的热忱和严谨的科研作风。

译者前言

在我们的《医学微生物学》课堂上，学生们总是表现出对生命科学的极大兴趣和高涨的学习热情。每次课后，学生们都通过E-mail踊跃提问，其中很多问题都是围绕着遗传物质的结构和基因组学研究前沿的。然而医学微生物学这门课的重点并不在这里。有鉴于此，我们计划组织有关教师编写一部适合于医学生的基因与基因组学方面的辅助或参考教材。但这是一项浩大的工程，一切都要从头做起。其中最重要的是要保证教材的质量，因此不是短期所能完成的。为此，我们筛选了近年来国际上出版的有关基因与基因组学方面的书籍，以期找到内容的深度和广度适宜且能反映基因组学领域最新进展的教材或专著，供医学专业的教师和学生选用。英国曼彻斯特大学的Richard J Reece所著《Analysis of Genes and Genomes》正好符合我们这些方面的要求。

《Analysis of Genes and Genomes》这本书的最大特点是深入浅出，易于被初涉本领域的学生接受；内容较为广泛而系统，使读者在阅读后对本领域能有个较全面的了解；图文并茂，使很多复杂的内容变得直观和易于理解。考虑到我们推荐此书的主要对象是没有接触过基因组学领域的医学生，他们对基因组学的英文专业词汇不熟悉，于是我们组织了一些对基因和基因组学领域的研究有基本了解和极大学习热情的医学本科生和研究生共同翻译了此书，学生的翻译初稿由博士后研究人员和教师审校。我们的翻译原则是直译以最大限度地忠实于原文。

本书在翻译和出版过程中，曾蒙多方面的支持和帮助。在此，我们衷心感谢北京大学医学出版社对我们给予的鼓励和帮助。由于时间和水平所限，错误和疏漏在所难免，因此我们敬请读者批评指正。

译者

2006年4月

缩写词及缩略语表

| | | |
|------|---|------------------|
| AAT | α_1 -antitrypsin | 抗胰蛋白酶 α_1 |
| AAV | adeno-associated virus | 腺病毒相关病毒 |
| AD | activation domain | 转录激活结构域 |
| BAC | bacterial artificial chromosome | 细菌人工染色体 |
| CaMV | cauliflower mosaic virus | 花椰菜花叶病毒 |
| CAP | catabolite activator protein | 分解代谢产物激活蛋白 |
| CBD | chitin-binding domain | 甲壳素结合结构域 |
| CDK | cyclin-dependent kinase | 细胞周期蛋白依赖性激酶 |
| cDNA | complementary DNA | 互补 DNA |
| CFI | cleavage factor I | 切割因子 I |
| CFII | cleavage factor II | 切割因子 II |
| CHEF | contour-clamped homogeneous electric field | 等压电场 |
| ChIP | chromatin Immunoprecipitation | 染色质免疫沉淀 |
| CMV | cytomegalovirus | 巨细胞病毒 |
| CPSF | cleavage and polyadenylation specificity factor | 切割与多聚腺苷酸化特异因子 |
| CStF | cleavage stimulation factor | 切割刺激因子 |
| CTD | carboxy-terminal repeat domain | 羧基末端重复序列域 |
| DBD | DNA binding domain | DNA 结合域 |
| DEAE | diethylaminoethanol | 二乙氨基乙醇 |
| DHFR | dihydrofolate reductase | 二氢叶酸还原酶 |
| DNA | deoxyribonucleic acid | 脱氧核糖核酸 |
| DTT | dithiothreitol | 二硫苏糖醇 |
| ECM | extra-cellular matrix | 细胞外基质 |
| EMS | ethyl methane sulphonate | 甲基磺酸乙酯 |
| ER | endoplasmic reticulum | 内质网 |
| ES | embryonic stem | 胚胎干 (细胞) |
| EST | expressed sequence tag | 表达序列标签 |
| FIGE | field inversion gel electrophoresis | 倒转电场凝胶电泳 |
| FISH | fluorescent <i>in situ</i> hybridization | 荧光原位杂交 |
| FRET | fluorescence resonance energy transfer | 荧光共振能量传递 |
| GST | glutathione-S-transferase | 谷胱甘肽-S-转移酶 |
| HAC | human artificial chromosome | 人类人工染色体 |

缩写词及缩略语表

| | | |
|--------|--|-------------------|
| HAT | histone acetyltransferase | 组蛋白乙酰转移酶 |
| H-DAC | histone deacetylase | 组蛋白脱乙酰酶 |
| HSV | herpes simplex virus | 单纯疱疹病毒 |
| IMAC | immobilized metal ion affinity chromatography | 固相金属离子亲和层析 |
| IMPACT | intein mediated purification with an affinity chitin binding tag | 内含肽介导的几丁质亲和纯化 |
| ITR | inverted terminal repeats | 反向末端重复序列 |
| LTR | long terminal repeats | 长末端重复序列 |
| MBP | maltose binding protein | 麦芽糖结合蛋白 |
| mRNA | messenger RNA | 信使 RNA |
| MCS | multiple cloning site | 多克隆位点 |
| MLP | major late promoter | 主要晚期启动子 |
| MSV | maize streak virus | 玉米条纹病毒 |
| NLS | nuclear localization signal | 核定位信号 |
| OD | optical density | 光密度 |
| ORF | open reading frame | 开放读码框 |
| PABII | polyA binding protein II | 多聚腺苷酸结合蛋白 II |
| PAC | P1 artificial chromosome | P1 人工染色体 |
| PAP | polyA polymerase | 多聚腺苷酸聚合酶 |
| PCR | polymerase chain reaction | 聚合酶链反应 |
| PFGE | pulsed-field gel electrophoresis | 脉冲场凝胶电泳 |
| RdRp | RNA-dependent RNA polymerase | RNA 依赖性的 RNA 聚合酶 |
| RF | release factor replicative form | 释放因子, 终止因子 复制型 |
| RFLP | restriction fragment length polymorphism | 限制性片段长度多态性 |
| RIP | ribosome inactivating protein | 核糖体失活蛋白 |
| RISC | RNA-induced silencing complex | RNA 诱导沉默复合物 |
| RNAi | RNA interference | RNA 干扰 |
| rRNA | ribosomal RNA | 核糖体 RNA |
| RT | reverse transcription reverse transcriptase | 逆转录 逆转录酶 |
| RT-PCR | reverse transcription-polymerase chain reaction | 逆转录 - 聚合酶链反应 |
| SAM | S-adenosylmethionine | S-腺苷基甲硫氨酸 |
| SDS | sodium dodecyl sulphate | 十二烷基硫酸钠 |
| siRNAs | small inhibiting RNAs | 小干扰 RNA |
| SNP | single-nucleotide polymorphism | 单核苷酸多态性 |
| snRNP | small nuclear ribonucleoprotein | 小核内核糖核蛋白 |
| SRB | suppressor of RNA polymerase | RNA 聚合酶抑制子 |

| | | |
|---------|---------------------------------------|--------------|
| STS | sequence tagged site | 序列标签位点 |
| SV40 | simian virus 40 | 猿病毒 40 |
| TAF | TATA-box binding associated factor | TATA 盒结合相关因子 |
| TBP | TATA-box binding protein | TATA 盒结合蛋白 |
| TdT | terminal deoxynucleotidal transferase | 末端脱氧核糖核酸转移酶 |
| TGMV | tomato golden mosaic virus | 番茄金色花叶病毒 |
| TK | thymidine kinase | 胸腺嘧啶核苷激酶 |
| tRNA | transfer RNA | 转运 RNA |
| VA RNAs | viral associated RNAs | 病毒相关 RNA |
| VNTR | variable number tandem repeats | 可变数目串联重复序列 |
| YAC | yeast artificial chromosome | 酵母人工染色体 |

目 录

| | |
|----------------------------|----|
| 第 1 章 DNA: 结构和功能 | 1 |
| 1.1 核酸是遗传物质 | 1 |
| 1.2 核酸结构 | 4 |
| 1.3 双螺旋结构 | 8 |
| 1.3.1 反向平行螺旋 | 10 |
| 1.3.2 碱基对和碱基堆积 | 11 |
| 1.3.3 不破坏双螺旋而得到其中的信息 | 11 |
| 1.3.4 氢键 | 12 |
| 1.4 DNA 的可逆变性 | 13 |
| 1.5 细胞中 DNA 的结构 | 15 |
| 1.6 真核细胞的核小体 | 17 |
| 1.7 DNA 的复制 | 21 |
| 1.8 DNA 聚合酶 | 23 |
| 1.9 DNA 复制的过程 | 24 |
| 1.10 重组 | 26 |
| 1.11 基因和基因组 | 27 |
| 1.12 基因组中的基因 | 29 |
| 1.13 转录 | 31 |
| 1.13.1 原核生物的转录过程 | 31 |
| 1.13.2 真核基因的转录过程 | 34 |
| 1.14 RNA 的加工 | 39 |
| 1.14.1 RNA 的剪接 | 41 |
| 1.14.2 选择性剪接 | 42 |
| 1.15 翻译 | 43 |
| 第 2 章 基因分析的基本技术 | 47 |
| 2.1 限制酶 | 48 |
| 2.1.1 限制修饰系统的类型 | 50 |
| 2.1.2 其他修饰系统 | 51 |
| 2.1.3 II 型限制酶是如何工作的? | 52 |
| 2.2 DNA 分子的连接 | 54 |
| 2.3 克隆的基础 | 57 |

| | |
|---|-----|
| 2.4 细菌转化 | 60 |
| 2.4.1 化学转化 | 61 |
| 2.4.2 电穿孔 | 62 |
| 2.4.3 基因枪 | 62 |
| 2.5 凝胶电泳 | 63 |
| 2.5.1 聚丙烯酰胺凝胶电泳 | 63 |
| 2.5.2 琼脂糖凝胶电泳 | 63 |
| 2.5.3 脉冲场凝胶电泳 | 67 |
| 2.6 核酸印迹杂交 | 69 |
| 2.6.1 NDA 印迹杂交 (Southern blotting) | 70 |
| 2.6.2 “东西南北”印迹杂交 | 72 |
| 2.7 DNA 纯化 | 74 |
| | |
| 第 3 章 载体 | 78 |
| 3.1 质粒 | 80 |
| 3.1.1 pBR322 | 82 |
| 3.1.2 pUC 质粒 | 83 |
| 3.2 选择性标记 | 87 |
| 3.3 λ 载体 (λ vectors) | 89 |
| 3.4 柯斯质粒载体 (cosmid vectors) | 96 |
| 3.5 M13 噬菌体载体 (M13 vectors) | 96 |
| 3.6 噬菌粒 (phagemids) | 100 |
| 3.7 人工染色体 | 101 |
| 3.7.1 YAC | 101 |
| 3.7.2 PAC | 103 |
| 3.7.3 BAC | 105 |
| 3.7.4 HAC | 106 |
| | |
| 第 4 章 聚合酶链式反应 | 108 |
| 4.1 PCR 反应条件 | 112 |
| 4.2 耐热 DNA 聚合酶 | 114 |
| 4.3 模板 DNA | 115 |
| 4.4 寡核苷酸引物 | 116 |
| 4.4.1 寡核苷酸引物的合成 | 117 |
| 4.5 引物错配 | 119 |
| 4.6 PCR 与遗传病诊断 | 122 |
| 4.7 PCR 产物的克隆 | 124 |
| 4.8 RT-PCR | 124 |

| | |
|---|-----|
| 4.9 实时 PCR | 126 |
| 4.10 PCR 的应用 | 128 |
| 第 5 章 克隆基因 | 129 |
| 5.1 基因组文库 | 130 |
| 5.2 cDNA 文库 | 134 |
| 5.3 定向 cDNA 克隆 | 137 |
| 5.4 应用 PCR 方法构建的文库 | 139 |
| 5.5 消减文库 (Substraction libraries) | 140 |
| 5.6 后基因组时代的文库的构建 | 143 |
| 第 6 章 基因鉴定 | 144 |
| 6.1 核酸杂交筛选 | 144 |
| 6.2 免疫筛选 | 148 |
| 6.3 功能筛选 | 152 |
| 6.4 相互作用筛选 | 153 |
| 6.5 噬菌体展示 | 153 |
| 6.6 双杂交筛选 | 153 |
| 6.6.1 双杂交筛选系统存在的问题以及解决方法 | 158 |
| 6.7 其他相互作用的筛选方法 - 关于一个主题的不同方面 | 160 |
| 6.7.1 单杂交筛选 | 160 |
| 6.7.2 三杂交筛选 | 160 |
| 6.7.3 反相双杂交筛选 | 161 |
| 第 7 章 人工突变 | 162 |
| 7.1 应用引物延伸的定点突变方法获得突变 DNA | 163 |
| 7.2 突变 DNA 单链的选择方法 | 166 |
| 7.2.1 硫代磷酸 DNA 突变链的选择方法 | 166 |
| 7.2.2 <i>dut⁻ ung⁻</i> (或 Kunkel) 链选择 | 167 |
| 7.3 盒式突变 | 168 |
| 7.4 通过 PCR 方法实现突变 | 168 |
| 7.5 QuickChange [®] 突变 | 173 |
| 7.6 对特定基因进行随机突变 | 175 |
| 7.7 蛋白质工程 | 178 |
| 第 8 章 蛋白质表达与纯化 | 179 |
| 8.1 在 <i>E.coli</i> 中表达 | 180 |
| 8.1.1 <i>lac</i> 启动子 | 180 |

| | | |
|---------|--|-----|
| 8.1.2 | tac 启动子 | 180 |
| 8.1.3 | λ P _L 启动子 | 181 |
| 8.1.4 | T7 表达系统 | 181 |
| 8.2 | 酵母表达 | 184 |
| 8.2.1 | 酿酒酵母 (<i>Saccharomyces cerevisiae</i>) | 184 |
| 8.2.1.1 | GAL 系统 | 184 |
| 8.2.1.2 | CPUI 系统 | 185 |
| 8.2.2 | 巴氏毕赤酵母 (<i>Pichia pastoris</i>) | 186 |
| 8.2.3 | 粟酒裂殖糖酵母 (<i>Schizosaccharomyces pombe</i>) | 186 |
| 8.3 | 昆虫细胞表达 | 186 |
| 8.4 | 高等真核细胞表达 | 188 |
| 8.4.1 | Tet-on/Tet-off 系统 | 188 |
| 8.5 | 蛋白质纯化 | 190 |
| 8.5.1 | His 标签 | 191 |
| 8.5.2 | GST 标签 | 193 |
| 8.5.3 | MBP 标签 | 195 |
| 8.5.4 | IMPACT | 196 |
| 8.5.5 | TAP 标签 | 196 |
| 第 9 章 | 基因组测序项目 | 199 |
| 9.1 | 基因组制图 | 199 |
| 9.2 | 遗传图谱 | 200 |
| 9.3 | 物理图谱 | 203 |
| 9.4 | 寡核苷酸测序 | 205 |
| 9.4.1 | 人工 DNA 测序 | 205 |
| 9.4.2 | 自动的 DNA 测序 | 208 |
| 9.5 | 基因组测序 | 210 |
| 9.6 | 人类基因组计划 | 212 |
| 9.7 | 寻找基因 | 213 |
| 9.8 | 基因功能的确定 | 214 |
| 9.9 | 生物信息学 | 215 |
| 第 10 章 | 后基因组分析 | 217 |
| 10.1 | 基因表达的总体变化 | 218 |
| 10.1.1 | 差异显示 | 218 |
| 10.1.2 | 芯片 | 218 |
| 10.1.3 | 染色质免疫沉淀 | 225 |
| 10.2 | 全基因组水平蛋白质功能研究 | 226 |

| | |
|---|------------|
| 10.3 基因敲除分析 | 226 |
| 10.4 反义技术和 RNA 干扰 (RNAi) | 228 |
| 10.5 全基因组双杂交筛选 | 230 |
| 10.6 蛋白质检测芯片 | 232 |
| 10.7 结构基因组学 | 233 |
| 第 11 章 植物基因工程 | 235 |
| 11.1 植物中的克隆 | 235 |
| 11.1.1 根瘤农杆菌 (<i>agrobacterium tumefaciens</i>) | 235 |
| 11.1.2 细胞核直接核转化 | 239 |
| 11.1.3 病毒载体 | 239 |
| 11.1.4 叶绿体转化 | 241 |
| 11.2 植物转基因的商业开发 | 244 |
| 11.2.1 延熟 | 244 |
| 11.2.2 除虫剂抗性 | 245 |
| 11.2.3 除草剂抗性 | 245 |
| 11.2.4 病毒抗性 | 246 |
| 11.2.5 真菌抗性 | 246 |
| 11.2.6 终止子技术 | 247 |
| 11.3 基因工程农作物的伦理学 | 248 |
| 第 12 章 动物细胞工程 | 249 |
| 12.1 细胞培养 | 249 |
| 12.2 动物细胞转染 | 249 |
| 12.2.1 化学法转染 | 250 |
| 12.2.2 电穿孔法 | 250 |
| 12.2.3 脂质体介导的转染法 | 251 |
| 12.2.4 肽法 | 251 |
| 12.2.5 直接 DNA 转移法 | 252 |
| 12.3 病毒载体 | 253 |
| 12.3.1 SV40 | 253 |
| 12.3.2 腺病毒 | 254 |
| 12.3.3 腺病毒相关病毒 (AAV) | 255 |
| 12.3.4 逆转录病毒 | 256 |
| 12.4 动物细胞的可选择标记和基因扩增 | 258 |
| 12.5 动物细胞中的基因表达 | 259 |

目录

| | |
|-----------------------|-----|
| 第 13 章 转基因动物 | 261 |
| 13.1 原核注射 | 261 |
| 13.2 胚胎干细胞 | 264 |
| 13.3 核转移 | 268 |
| 13.4 基因疗法 | 273 |
| 13.5 基因治疗的示例及前景 | 273 |
| 术语表 | 275 |
| 蛋白质 | 282 |
| 诺贝尔奖获得者 | 285 |
| 参考文献 | 288 |
| 索引 | 329 |

第1章 DNA：结构与功能

要点

- 核酸是遗传信息的携带者
- DNA 为反向平行的双链螺旋结构
- DNA 螺旋的两条链通过碱基配对 (A 与 T, G 与 C) 相联结
- DNA 复制是双链解链后两条链分别复制
- 分子生物学中心法则
DNA → RNA → 蛋白质
- 转录是以 DNA 的一条链为模板合成 RNA 的过程
- 翻译是 RNA 分子解码合成蛋白质的过程

每种生物都携带着为自身生长和繁衍所需的信息。遗传的基本概念以及由此导致对基因的认识，可追溯到1865年，是由Gregor Mendel通过他的研究提出的；V. Orel对此有详细的回顾 (Orel, 1995)。孟德尔根据他的豌豆培育试验结果，得出结论：每株豌豆的每个基因都由两个等位基因 (alleles) 组成，但只呈现出一种表型 (phenotype)。尽管孟德尔对此变化之中的分子过程一无所知，他却能够清楚地认识这个复杂的现象，也许这才是孟德尔的伟大之处。大约在同一时期，人们发现遗传是通过精子与卵子的结合实现的。此时Ernst Haeckel注意到精子的主要物质成分是细胞核，于是推测遗传信息可能就存在于细胞核之中。

1.1 核酸是遗传物质

自从“遗传”这一概念诞生以来，人们就相信遗传物质是亲代通过物理方式传给子代的。蛋白质和核酸都曾被认为有可能是遗传物质。在20世纪40年代以前，许多科学家倾向于蛋白质。这主要有两个原因，首先，细胞之中的蛋白质含量丰富，尽管每种蛋白质在不同类型细胞之中的含量差异很大，但对于大部分细胞来讲，蛋白质总量占细胞干重的一半以上。其次，人们推测，遗传性状的表达需要复杂的信息传递过程，而核酸则显得过于简单了。DNA (脱氧核糖核酸) 首先由瑞士化学家Johann Frederick Miescher在1869年分离出来。他先将细胞核从细胞之中分离出来，之后又从提取出的细胞核中分离出了一种他称之为核素 (nuclein) 的酸性物质。Miescher发现，与蛋白质不同，核素之中含有大量的磷而不含硫。Miescher非常有远见，他说：“如果有人想要假设某种单纯的物质导致受孕的话，那么他就毫无疑问地首先应该考虑核素。”

1926年，Levene和Simms根据DNA分子是由大约等量的四种核苷酸所组成的知识，并且在测定了核苷酸之间连接方式的基础之上，假设了一种四核苷酸的结构 (图1.1) 来解释核苷酸在核酸之中的化学排列 (Levene & Simms, 1926)。他们设想了一种非常简单的四核苷酸

单元，这种单元不断重复，最终形成一条长的核酸分子。由于四核苷酸结构相对简单，人们普遍认为核酸无法提供遗传物质所预期的化学多样性。而蛋白质由于包含有20种不同的氨基酸，则可能满足物质多样性的基本要求。

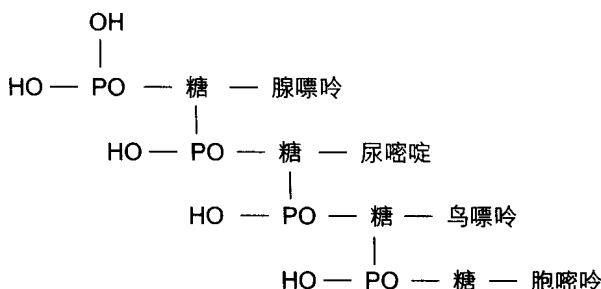


图 1.1 Levene 和 Simms 在 1926 年提出的四核苷酸核酸模型。在提出这个模型的时候，曾考虑到植物和动物的核酸可能不同，而且此时 DNA 和 RNA 的区别也还不是很清楚。

1928 年，Frederick Griffith 用不同型别的肺炎链球菌做了一个实验 (Griffith, 1928)。一些菌株有毒力，能够使人和小鼠患肺炎。另一些菌株无毒力，不致病。两种肺炎链球菌在形态学上的区别在于，有毒株在菌体外有一多糖荚膜，而且在琼脂平板表面上菌落呈光滑型。无毒株则没有荚膜包裹，同时菌落呈粗糙型。光滑型的菌株有毒力，是因为其表面的多糖荚膜使得它们不易被宿主的免疫系统吞噬，因此他们易于繁殖而引起肺炎。粗糙型的菌株没有多糖荚膜的保护作用，于是很容易被宿主的免疫系统吞噬并清除掉。

Griffith 了解到，只有在向小鼠体内注射活的光滑型肺炎链球菌的时候，才能够引起肺炎。加热杀死之后的光滑型肺炎链球菌和粗糙型肺炎链球菌单独注射到小鼠体内时，都不会引起肺炎。Griffith 又做了一个关键的实验 (图 1.2)：虽然粗糙型肺炎链球菌和杀死之后的光滑型肺炎链球菌分别注入到小鼠体内均不致病，但如果将两种混合在一起注射，则小鼠死亡。对死亡小鼠的血液进行分析，可以查到大量能够致病的光滑型肺炎链球菌。于是 Griffith 得出结论，加热杀死的光滑型肺炎链球菌能够把活的粗糙型肺炎链球菌转化成光滑型。他将这种现象称为“转化”(transformation)，同时提出，转化原理可能是多糖荚膜的某些成分，或者是荚膜合成所需的某种物质，尽管他也注意到了荚膜单独是无法引起肺炎的。

1944 年，Oswald Avery, Colin MacLeod 和 Maclyn McCarty 发表了他们关于细菌转化的研究工作，证明引起转化的因素是 DNA (Avery, Macleod & McCarty, 1944)。他们首先培养了大量的光滑型肺炎链球菌，然后收集细菌并将其加热杀死。在匀质化以及经过几轮洗涤剂提取之后，他们得到了一种提取物。将这种提取物和粗糙型肺炎链球菌混合注射，仍然有细菌转化效应。提取物中的蛋白质已经通过氯仿萃取除去，同时通过酶促反应除去多糖。最终，沉淀出少量产物，酒精处理得到一种纤维性物质，这种物质仍然能够引起粗糙型肺炎链球菌的细菌转化。从最初的 75 L 培养的细菌细胞中，最终得到了 10 ~ 25mg 这种“活性因子”。进一步实验无可置疑地证实了这种转化因子是 DNA。实验中提取得到的那种纤维性物质，通过分析其氮/磷比值，结果与 DNA 完全吻合。为了除去最终提取物质中可能存在的污染物，他们使用胰蛋白水解酶和糜蛋白酶对其进行处理，之后又用核糖核酸酶 (一种 RNA 分解酶)