

基因工程技术 在基础与临床医学中的应用

蔡良琬 主编

人民卫生出版社

基因工程技术 在基础与临床医学中的应用

蔡良琬 主编

顾健人 汪 健
沈 岩 王申五
司静懿 张 沛
袁建刚 蔡良琬

编写



人 民 卫 生 出 版 社

内 容 简 介

本书主要介绍了近年来基因工程技术在基础与临床医学中应用的研究进展。全书分三篇：第一篇简要介绍基因工程与基因分析技术；第二篇讨论人的基因克隆、基因定位、癌基因生长因子、多样性免疫蛋白的生成机制及某些病毒的分子生物学研究；第三篇介绍有关技术在遗传病、糖尿病与冠心病以及临床病毒诊断上的应用。可供高等医学院校师生及研究人员、医务工作者参考。

责任编辑 郝巨为

基因工程技术在基础与临床医学中的应用

蔡良琬 主编

人民卫生出版社出版

(北京市崇文区天坛西里10号)

流研所印刷厂印刷

新华书店北京发行所发行

787×1092毫米16开本 13印张 324千字

1990年1月第1版 1989年12月第1版第1次印刷

印数：00,001—3,000

ISBN 7-117-01220-X/R·1221 定价：6.80元

前 言

基因工程技术出现才十几年，因包括有纯化基因与扩增基因技术，故立即为广大生物学与医学研究工作者重视并应用。这项新技术不仅把分子生物学提高到了基因水平，而且也把生物学与医学中的其他学科引上基因研究的道路。其意义和影响是极为深远的。我们编写此书目的是帮助我国医务工作者进一步认识和熟悉这些技术，了解应用它的重要性。书中对基因工程及有关的基因分析只作了简单的介绍，用较多的篇幅阐述了近十年来它们在医学中的应用，如人的基因分子克隆，纯化与扩增的现状，部分技术在基因定位中的应用，癌基因与生长因子对人体细胞的作用，多样性免疫蛋白的生成机制，以及某些与人类密切相关的病毒研究与临床上的应用。临床医学方面除介绍在遗传病诊断上的应用外，重点阐明了有关技术用于研究糖尿病和冠心病中近年来所取得的研究进展。基因工程技术，正在医学领域中得到迅速应用和发展，会成为医务工作者从事研究和临床实践的有力武器，相信在2000年来到之前，会有一些疑难病的本质得到阐明，其发病的奥秘被揭开。

参加本书编写的除我本人外，还有上海肿瘤研究所的顾健人与汪健同志，中国医学科学院基础医学研究所的沈岩、王申五、司静懿、张沛、和袁建刚同志。由于编写者工作繁忙，故又请了基础医学研究所的赵敏顺同志协助整理，家姐蔡良班与基础所绘图室几位同志帮忙抄写与绘图，在此表示感谢。编写人为求本书内容深入浅出，易于为广大读者理解吸收，作出了很大努力；只是由于本人对医学中不少学科理解不够，在编写中肯定会有不少不足之处，希望读者热心给予批评指正。

蔡良班

目 录

第一篇 基因、基因工程技术与基因分析	(1)
第一章 目前基因研究的概况	(1)
一、基因的三种概念	(1)
二、各类生物体的基因与基因组	(1)
(一) 基因与基因组概念	(2)
(二) 基因组的二级与高级结构	(2)
(三) 各种生物体基因组的特点	(3)
(四) 病毒的基因组	(3)
(五) 基因组织研究的重要性	(9)
(六) 细菌类的基因组	(10)
(七) 真核生物的基因组	(11)
三、原核与真核生物体内的基因表达	(15)
(一) 原核生物基因表达的特点	(15)
(二) 真核生物的基因表达	(17)
参考文献	(20)
第二章 基因工程技术的内容与发展	(21)
第一节 基因工程技术的内容	(21)
一、分子克隆技术的内容	(21)
二、用于分子克隆的限制性内切酶	(22)
三、运载体的选择	(23)
第二节 分子克隆技术的发展	(24)
一、文库的建立	(24)
(一) 基因文库	(24)
(二) cDNA库	(25)
二、染色体上的步移	(26)
参考文献	(28)
第三章 基因的几种重要分析方法	(29)
一、几种重要的研究方法	(29)
二、限制性内切酶的图谱分析	(32)
(一) 限制性内切酶的多选择性	(32)
(二) 限制性内切酶酶谱的制作	(33)
三、分子杂交	(34)
(一) 分子杂交的特异性	(34)
(二) 分子杂交的种类	(35)
(三) 分子杂交应用的广泛性	(35)

(四) 分子杂交的局限性·····	(35)
四、核苷酸顺序分析·····	(36)
(一) 从核苷酸顺序了解蛋白质的结构·····	(36)
(二) 对基因表达的分析·····	(37)
(三) 基因之间相似性的比较·····	(37)
(四) 核苷酸顺序直读法的特点与发展·····	(37)
五、基因分析在医学发展上的某些重要作用·····	(38)
参考文献·····	(39)
第二篇 基因工程技术在基础医学中的应用·····	(40)
第四章 人类基因的分子克隆现状·····	(40)
第一节 单一基因的分子克隆·····	(40)
一、人类基因分子克隆的来源·····	(40)
二、人类基因的分子克隆数及增长趋势·····	(41)
(一) 激素类多肽基因的分子克隆·····	(42)
(二) 人体中一些比较重要蛋白质的分子克隆·····	(42)
(三) 与免疫体系相关的多样性蛋白基因的分子克隆·····	(42)
(四) 癌基因的分子克隆·····	(43)
(五) 活性多肽的分子克隆·····	(43)
(六) 各种受体基因的分子克隆·····	(43)
(七) DNA重复顺序及某些特殊的单拷贝顺序·····	(43)
三、某些未知功能的基因片段的利用·····	(44)
(一) 入RC8及有关其他片段的研究·····	(44)
(二) G8及其研究现状·····	(45)
四、人基因分析的现状·····	(45)
第二节 单一染色体基因文库的建立·····	(45)
一、X染色体基因文库的建立·····	(46)
二、各染色体中DNA的含量·····	(46)
三、各染色体基因文库的建立·····	(46)
参考文献·····	(49)
第五章 染色体的研究与基因定位·····	(49)
第一节 染色体研究的现状·····	(49)
一、核小体染色质与染色体·····	(50)
(一) DNA的超盘绕·····	(50)
(二) 核内的骨架结构·····	(51)
(三) 有活性的染色质与高度超盘绕的染色体·····	(51)
二、真核生物染色体的结构特点·····	(52)
(一) 染色体的线状结构·····	(52)
(二) 着丝粒的分子克隆与基因分析·····	(53)
(三) 端粒·····	(53)

三、人工染色体.....	(55)
(一) 一个短线形质粒的构建.....	(55)
(二) 长线状人工染色体的形成.....	(56)
第二节 人基因在染色体上的定位.....	(56)
一、分子克隆技术在基因定位上的应用.....	(56)
(一) 分子克隆基因用于基因定位的优越性.....	(57)
(二) 目前定位研究上的局限性.....	(57)
二、人的基因定位图谱.....	(57)
三、应用分子克隆及基因分析法的基因定位.....	(57)
(一) 应用分子克隆的基因与体细胞杂交相结合的定位.....	(57)
(二) 基因分析与基因定位的结合定位.....	(58)
参考文献.....	(59)
第六章 癌基因、生长因子和抗癌基因.....	(60)
第一节 癌基因.....	(61)
一、癌基因的概念.....	(61)
二、癌基因和人体肿瘤.....	(64)
三、癌基因的激活.....	(65)
(一) 点突变.....	(66)
(二) DNA 重排.....	(66)
(三) 基因放大.....	(67)
(四) 癌基因结构的变化.....	(67)
四、癌基因与癌的发生与演进.....	(67)
(一) 化学致癌物和癌基因.....	(67)
(二) 癌基因的作用同组织类型和细胞分化阶段有关.....	(67)
(三) 癌基因的协同作用与多阶段致癌学说.....	(68)
(四) 癌基因和癌的演进.....	(68)
第二节 生长因子和癌基因表达产物.....	(69)
一、生长因子.....	(69)
二、生长信号和细胞内传递途径.....	(71)
(一) 生长因子和生长因子受体结合.....	(71)
(二) 信号放大系统和第二信使.....	(71)
(三) 细胞效应阶段.....	(72)
(四) 细胞最后的反应.....	(72)
三、癌基因产物在细胞内的作用和癌基因的分类.....	(73)
(一) 生长因子和生长因子受体.....	(74)
(二) 酪氨酸激酶类.....	(74)
(三) ras基因家族.....	(74)
(四) 核蛋白类.....	(74)
(五) 其它.....	(75)

第三节 抗癌基因	(75)
一、抗癌基因的概念	(75)
二、抗癌基因存在的证据	(75)
(一) 染色体稳定性和肿瘤的发生	(75)
(二) 细胞杂交试验	(75)
(三) 遗传性肿瘤中某些基因的丢失	(76)
(四) Rb基因的分离和Wilms瘤的定性	(76)
三、抗癌基因的生物学特性	(77)
(一) 癌基因产物的拮抗物	(77)
(二) 抑素	(77)
(三) 干扰素	(78)
(四) 染色体稳定性	(78)
参考文献	(79)
第七章 淋巴细胞中多样性蛋白的生成机制	(80)
第一节 免疫球蛋白的生成机制	(80)
一、Dreyer-Bennett 假说	(81)
二、免疫球蛋白基因的分子克隆	(81)
三、轻重链基因的活化	(82)
(一) 轻链基因的重排	(82)
(二) 重链基因的重排	(84)
四、多样性抗体的形成	(85)
(一) 淋巴细胞的成熟与抗体的生成	(86)
(二) 不同抗体的转换机制	(87)
(三) 抗体形成的基因转录与加工过程	(88)
(四) 抗体多样性形成的几种可能机制	(89)
第二节 T淋巴细胞上多样性蛋白的生成机制	(90)
一、抗原受体等肽链与其基因的研究	(90)
(一) T细胞特异mRNA的分子克隆	(90)
(二) cDNA及其分子杂交	(91)
(三) β 多肽链基因	(91)
(四) α 多肽链基因	(92)
(五) γ 多肽链基因	(92)
(六) T细胞上其他肽链的分析	(93)
二、T细胞抗原受体蛋白的形成与基因重排	(94)
三、T细胞抗原受体基因在染色体上的定位	(94)
四、T细胞抗原受体基因重排及表达的规律	(94)
第三节 淋巴细胞多样性蛋白基因重排在临床诊断上的应用	(95)
一、B淋巴细胞恶性肿瘤的诊断	(95)
二、T淋巴细胞恶性肿瘤的检测	(96)

参考文献	(98)
第八章 病毒结构与功能的研究现状	(99)
第一节 人类逆转病毒的研究现状	(99)
一、动物的逆转病毒	(99)
(一) 内源性 与 外源性	(100)
(二) 逆转病毒中的几个主要蛋白基因	(100)
(三) 逆转病毒基因组的5' 及3' 末端结构与原病毒的LTR	(101)
(四) 逆转病毒的逆转录环化与整合	(101)
(五) 应用分子克隆方法分析逆转病毒基因组	(102)
二、人类逆转病毒	(104)
(一) HTLV-I 型的结构	(105)
(二) HTLV-I 的转录活化	(106)
(三) HTLV-Ⅱ病毒的结构与功能	(106)
三、LTR的结构与生物功能的关系	(107)
(一) LTR 的结构	(107)
(二) HTLV-I 与HTLV-Ⅱ型病毒LTR的比较	(108)
第二节 乙型肝炎病毒的分子生物学研究	(109)
一、人乙肝病毒基因组的分子克隆	(109)
(一) 表面抗原 (HBsAg) 蛋白基因及其表达	(111)
(二) 核心抗原基因与核心抗原HBcAg及e抗原HBeAg的形成	(111)
(三) DNA聚合酶	(112)
(四) HBx 基因的研究	(112)
二、动物乙型肝炎病毒的研究	(113)
三、乙型肝炎病毒DNA的复制	(114)
四、乙型肝炎病毒DNA的转录	(114)
(一) mRNA 的形成	(115)
(二) 转录调节原件的组织特异性	(115)
五、乙型肝炎病毒DNA的整合	(115)
六、基因工程制备的乙型肝炎疫苗	(116)
参考文献	(116)
第三篇 基因工程技术在临床医学上的应用	(118)
第九章 基因工程技术在遗传病诊断上的应用	(118)
第一节 基因诊断与分析基因缺陷的基本方法	(118)
一、应用基因探针直接分析某些遗传病	(118)
二、应用基因探针间接分析某些遗传病	(119)
三、应用相关的DNA片段做间接的基因分析	(120)
四、分析基因缺陷的基本方法——RFLPs法	(120)
(一) 人基因组DNA 的分离	(121)
(二) 分子杂交探针	(122)

(三) 限制性内切酶图谱分析·····	(123)
第二节 基因工程技术在血友病研究中的应用·····	(124)
一、FⅧ与FⅨ基因的分离及其组织结构分析·····	(125)
(一) FⅧ基因的克隆化分离·····	(125)
(二) FⅧ的基因结构与蛋白质结构分析·····	(126)
(三) FⅨ基因的克隆化分离与结构分析·····	(127)
二、血友病分子病理学·····	(127)
(一) 乙型血友病的基因缺陷·····	(127)
(二) 甲型血友病的基因缺陷·····	(127)
三、血友病的基因诊断·····	(128)
(一) RFLPs连锁分析·····	(128)
(二) FⅨ基因的RFLPs·····	(129)
(三) FⅧ基因的RFLPs·····	(130)
四、基因工程技术在血友病治疗中的应用·····	(131)
(一) 利用培养细胞生产凝血因子的研究·····	(132)
(二) 凝血因子基因在转基因动物体内的表达·····	(132)
第三节 苯丙酮尿症的基因诊断·····	(134)
一、苯丙酮尿症的特点与产生的原因·····	(134)
二、PAHcDNA在苯丙酮尿症研究中的应用·····	(134)
三、苯丙酮尿症的家族分析·····	(135)
四、苯丙酮尿症的产前诊断·····	(137)
五、PAH基因进一步研究的必要性·····	(137)
参考文献·····	(138)
第十章 地中海贫血的基因缺陷·····	(139)
一、正常血红蛋白及其基因控制·····	(139)
(一) 正常人血红蛋白·····	(139)
(二) 正常人珠蛋白基因与基因簇·····	(140)
(三) 珠蛋白基因的结构·····	(141)
(四) 人珠蛋白基因的表达·····	(142)
二、 α 地中海贫血的基因缺陷·····	(143)
(一) 缺失型 α 地中海贫血的基因缺陷·····	(143)
(二) 非缺失型 α 地中海贫血的基因缺陷·····	(144)
(三) α 地中海贫血的产前基因诊断·····	(144)
三、 β 地中海贫血的基因缺陷·····	(148)
(一) 基因突变所致的 β 地中海贫血·····	(149)
(二) 缺失型 β 地中海贫血·····	(149)
(三) β 地中海贫血的基因分析和产前诊断·····	(149)
参考文献·····	(157)
第十一章 某些常见病的遗传性与相关的基因分析·····	(158)

第一节 糖尿病的遗传性质与胰岛素基因的多态性	(158)
一、糖尿病的免疫遗传学基础	(158)
(一) 人的HLA基因系统	(159)
(二) HLA系统的几个重要抗原	(160)
(三) HLA遗传性与I型糖尿病的相关性	(160)
(四) I型糖尿病与HLA系统外基因的相关性	(161)
(五) HLA系统在胰岛破坏过程中可能出现的相关性	(161)
二、胰岛素基因的多态性及它与糖尿病的关系	(161)
(一) 胰岛素基因的多态性	(161)
(二) 5'末端插入顺序与糖尿病之间的关系	(162)
(三) B肽与A肽突变分子的研究	(165)
(四) 前胰岛素原分子的变异	(166)
三、胰岛素受体的研究与糖尿病的关系	(166)
(一) 胰岛素与胰岛素受体的结合作用	(167)
(二) 胰岛素受体	(167)
(三) 胰岛素受体cDNA的研究	(168)
第二节 与冠心病有关的脂类代谢研究	(169)
一、载脂蛋白的结构与功能	(169)
(一) 八种载脂蛋白的基因分析	(169)
(二) 载脂蛋白的功能	(170)
(三) ApoAI的基因分析	(171)
(四) ApoB的分子克隆	(171)
二、脂蛋白受体的研究	(172)
(一) LDL受体的结构	(173)
(二) LDL受体在胆固醇运输中的作用	(174)
(三) LDL受体变异对脂类运输的影响	(175)
三、从家族性高脂血症患者的分析看遗传性缺损	(176)
参考文献	(177)
第十二章 基因工程技术在临床病毒学方面的应用	(178)
第一节 基因工程技术在乳头瘤病毒研究中的应用	(178)
一、乳头瘤病毒的一般特征及分类	(178)
二、乳头瘤病毒基因组组织结构及基因产物的作用	(180)
三、HPV DNA在生殖道癌前和癌中的存在	(182)
四、人类生殖道肿瘤HPV的整合	(184)
五、HPV DNA在宫颈癌细胞中的表达	(185)
六、HPV的转化模型	(186)
第二节 重组HBV DNA探针及杂交技术在乙型肝炎研究中的应用	(187)
一、重组HBV DNA探针制备方法简介	(187)
二、杂交方法的检出水平	(188)

三、在乙型肝炎血清学研究中的应用.....	(189)
(一) 乙型肝炎患者血清的点样及处理.....	(189)
(二) 斑点杂交.....	(189)
(三) 血清免疫学指标与HBV DNA的关系.....	(190)
四、在肝细胞内HBV DNA检测中的应用.....	(191)
(一) 细胞DNA 的 制 备.....	(191)
(二) 印迹杂交.....	(191)
(三) 在肝癌和肝硬化发病机理研究中的应用.....	(191)
(四) HBV DNA探针的纯化及其对杂交结果的影响.....	(192)
参考文献.....	(193)

第一篇 基因、基因工程技术与基因分析

第一章 目前基因研究的概况

一、基因的三种概念

基因有三种概念，即遗传学的概念、分子生物学的概念及生物大分子结构研究的概念。

基因一词是在1865年孟德尔提出遗传的概念下被确定的。它在遗传学中代表任何一种生物体中的遗传物质，现广泛应用于遗传学中。1868年生物化学家发现了核酸，经过50年以上的化学分析，证明它是生物体中一类重要的化学物质。至于它的生物功能，则是经过1928年Griffith与1944年Avery等的转化实验才证实。Griffith利用肺炎球菌做感染实验，在实验中发现当把被杀死的、有荚膜、光滑的Ⅲ型肺炎球菌与不光滑、无荚膜、未经处理的活肺炎球菌一起注入小鼠体内。经过一段时间，在小鼠体内可以找到新的有荚膜、光滑的Ⅲ型肺炎球菌。据此，他指出这种变异肺炎球菌的出现，是由一种有转化作用的物质所引起的。1944年Avery等重复了这一实验，采用了提取出的核酸来代替被杀死的Ⅲ型肺炎球菌，得到转化作用的成功。他们的实验是历史上首次承认核酸是遗传物质。不过直到1953年才用化学结构说明核酸就是遗传物质。那是两个青年科学工作者Watson与Crick总结了当时核酸中DNA的一些物理与化学上的结构分析结果，提出了DNA的双螺旋结构模型。Watson与Crick提出这一结构模型后，很快又用它解释了DNA作为遗传物质的可能性与理论依据。这一学说不仅奠定了分子生物学的基础，也确立了核酸就是基因的分子生物学概念，从而大大促进了核酸的研究，推动了它的分子生物学及大分子结构的研究。

目前以研究核酸（即基因）的功能及其结构的关系为主的工作，则是从基因的分子生物学概念出发，至于应用物理学与化学研究核酸大分子结构，是在基因的大分子结构概念基础上发展进行的。这三方面从不同角度出发的研究，都很重要，不仅都取得了相当大的进展，而且正在彼此补充深化，用以解释生物体的各种生命现象。

本章从核酸大分子概念出发，介绍各种生物体的遗传物质、基因、基因组与基因组织等几个概念，以及某些主要的生物功能与核酸结构的关系。应该指出，近年来基因的研究达到如此深度，很大程度是靠基因工程技术的发展，使人们愈来愈认清基因在生物体中遗传的特性，及其在生命本质中的重要性。

二、各类生物体的基因与基因组

核酸有两大类：DNA与RNA（脱氧核糖核酸与核糖核酸）。前者是由含腺嘌呤（A）、鸟嘌呤（G）、胞嘧啶（C）与胸腺嘧啶（T）四种碱基组成的多聚核苷酸；后者除以上A、G、C外没有T，而有尿嘧啶（U）的核苷酸所组成的多聚核苷酸。DNA与RNA分子都很大，都是链状的多聚体。在任何一种生物体内都有核酸，即使是无细胞的病毒、噬菌体或类病毒体内也都有核酸，只是大部分病毒中只有一种核酸，即DNA或RNA。这类生物体的遗传信息就存在于其所含的DNA或RNA之中。1944年Avery的实验证明DNA是遗传物质，只

不过他用的是细菌。在细菌以上生物体中，不论单细胞与多细胞的生物，其细胞中都含有DNA及RNA，但在这两类中真正的遗传物质则是DNA，当然相应的RNA也很重要，也有其特殊功能。DNA包括全部基因，参与并完成生物生长、传代等许多生物过程，如基因的复制（繁殖），基因的表达（生物表型的出现）等等。不同RNA有不同功能，如核糖体RNA（rRNA）组成核糖体，是蛋白质生物合成的场所。转移RNA（tRNA）的作用是携带与供应各种氨基酸。信使RNA（mRNA）则来自基因DNA，从DNA转录出相应的mRNA后，它被作为模板，翻译出相应的蛋白质。一切生物体都服从这一规律，无细胞的病毒也不例外。这是生物体繁殖与生存的基础。至于生物体内的蛋白质既不是遗传物质，也不是基因，但它们非常重要，因为任何一个生物体在生长与繁殖过程中，如没有其特异蛋白质（如酶）参与基因的活动，基因都不可能起作用，也可以说基因的任一活动，都包括了核酸与蛋白质的相互作用。这是在生物大分子功能研究中极为重要的。至于脂类与多糖类，它们虽也重要，但对微小的生物来说（如病毒），往往不重要，因为有些病毒中并没有这些物质。由此可见，生物体中核酸与蛋白质是生命的基础，而遗传的物质基础是核酸。

（一）基因与基因组概念

基因是生物体的遗传物质，一般指的是生物体与某些性状有关部分的核酸，如遗传中的红眼基因（多种蛋白的结构基因），如分子生物学中的半乳糖苷酶基因（形成此酶蛋白的结构基因），以及负责调节控制基因活动的调控基因，是遗传物质研究的重要对象。它们也就是指单一蛋白或几个相关蛋白的结构与调控基因。核酸分子很大，含基因很多，故经常需要把一个基因与生物体内带有该生物体的全部基因区别开。通常称一个生物体的全部基因为基因组（genome）。如人的基因组，就包括人的全部基因，在任何染色体中的都计算在内（目前在研究中，也有人用基因组来表示一大组基因，而不是生物体的全部基因）。总之，基因组所代表的范围远大于基因。至于我们将介绍的基因工程技术，则是一种分离纯化基因的操作方法，可以从一个基因组中纯化出任一个基因及纯化与扩增某个基因的方法。从理论上说，应用这一技术，任何一个研究工作者都可获得他所要了解的基因，并可得到大量的该基因来从事研究。因此，基因工程技术对了解生物体的基因组与任一基因都是不可缺少的研究手段。

（二）基因组的二级与高级结构

核酸都是由很长的链状多核苷酸组成，不论是单链DNA，双链DNA，单链RNA与双链RNA，在生物体内都保持有一定的结构状态。其中决定二级结构的，主要是两链间的氢键，或单链内部折叠所形成的氢键。从生物体中提取核酸时，往往会破坏这些原有结构（如折断），只有小的环状核酸，由于小、又是闭环的，不易被破坏，故在提取后还能保持其原有的结构。

核酸的高级结构是指螺旋状的长链所形成的空间结构，如DNA的二级结构为双螺旋结构，而在生物体内，这些双螺旋结构又都有序地缠绕成更高层次的超螺旋。因此，超螺旋是基因组存在的高级结构形式。超螺旋结构同样与基因的活性有关。对于一些简单的生物体，经研究已知，一定的超螺旋结构还是保持生物活性的必要形式。不过任何一种生物大分子的高级结构，都是以其一级结构为基础，一级结构确定二级结构再缠绕或折叠才出现高级结构。在这点上核酸与蛋白质是一致的。本书中虽然有时也涉及基因的二级或高级结构，但重点讨

论的是一级结构。

分析一级结构就是分析核酸中的核苷酸排列顺序,因为它决定任一基因或基因组的特点,象蛋白质中的氨基酸排列顺序一样,所以还可用一个蛋白编码基因的一级结构,与密码的关系,推算出应翻译出的蛋白质的一级结构。还可从该基因在生物体中所出现的二级结构(双螺旋、单螺旋、线型等结构)推算出其二级结构变化的可能性(如能否形成局部的DNA左手螺旋、十字架结构等等,有关这些二级及高级结构,将在后面哺乳动物基因组中简单介绍)。总之,一级结构是研究基因结构的基础,即分析任何一种基因时,不只要说明含有多少核苷酸,还应说明其核苷酸排列顺序,才能解释一个基因的功能与结构的关系。

(三)各种生物体基因组的特点

生物体有大小之分,在细胞以上水平的生物体有原核与真核生物之分,这是形态与特征上的区别。在多细胞动物中又有类、属、纲、目、科等之分。从生物进化角度看,还有古老与新生之分。举凡这些差别常常能反映在其基因的特点上,虽然有些尚不易简单区别,但由于一个生物体的基因组带有其全部特有的基因,它应能反映生物体的特性,只不过还需深入,才能找出并应用于区别罢了。

(四)病毒的基因组

病毒是一种极小的生物,如果不谈论类病毒,病毒则是唯一不能单独繁殖的生物体,它必须依赖某一特定细胞而繁殖。这类生物体小,所含的基因组也小,通常只有一种核酸,DNA或RNA。病毒中的核酸常常只含有几千个核苷酸,如为双链DNA则为几千个核苷酸对。因为一个核苷酸正好有一个碱基。而特异性又正好反映了该核苷酸的特点。故为了简单表示,人们常用碱基(base)来代表核苷酸,用碱基对(bp)来表示核苷酸对。一些大的病毒(包括噬菌体在内)也有基因组在40kbp(kilo base pair)以上的,如 λ 噬菌体。从一个基因组的大小、含几条核酸链以及所含核苷酸的特点与一级结构,可以知道该病毒是简单还是复杂。当然只凭基因组的大小并不能知道该病毒的危险性大小。如有些小的病毒,其对生物的危害却很大,甚至难以治服。由于任何一种基因都服从一般基因转录及翻译蛋白质的规律(根据相同密码子的关系),故基因组大时,所含的编码蛋白质会相应多些。目前已发现的蛋白质,以不同亚基而论,都未超过几百个氨基酸,故以平均一个蛋白质含有300个氨基酸计算,可以估计出900到1000个核苷酸的核酸将能编码一个特异蛋白质。用这种方法去计算一个病毒基因组可能编码的蛋白质,虽然不能完全正确,但有一定参考意义。病毒虽只含有一种核酸,但也有不同的高级结构,如常见的有线型、环型、单链及双链等不同二级结构。

本书介绍的几个病毒是有代表性的,属于与人类疾病密切相关的病毒,重点说明的是它们基因组的特点。此外,还将介绍几个与基因工程技术密切相关的病毒,如 ϕ x174、 λ 与M13三种噬菌体,它们都属于细菌病毒,但它们常应用于基因工程与基因分析中。病毒在许多方面与细菌有严格的区别,但从病毒的大小与基因组分析上看又有一定的关系。

上面已提到病毒分子中都只有一种核酸,甚至大到痘苗病毒。痘科病毒中的一些病毒都很大,但也只含有一种双链的DNA;而比它小的衣原体(包括砂眼病毒与鹦鹉热病毒)却含有两种核酸,即有双链DNA与RNA,它们有向单细胞生物过渡的特点。

病毒学中过去就有用带有那一种核酸如DNA或RNA来分类的,如小RNA病毒、肝DNA

病毒等；但目前对大多数病毒还是按形态特征来分类。从所介绍的几个典型病毒，可以简单了解一个生物体的生命过程，如繁殖与生长和基因组的密切关系。

1. 脊髓灰质炎病毒 (Poliovirus) 脊髓灰质炎病毒属于肠道小RNA病毒，此病毒的直径为29~30nm，呈颗粒状，中有一个16nm的致密核，在此核中有一条单链的RNA，就是它的基因组。脊髓灰质炎病毒的基因组实际上是一个RNA，严格说，则是一条mRNA。此基因组分子的5'末端有mRNA应有的帽子结构，其分子3'末端有一多聚A的尾，总基因组中有7.7kb左右个核苷酸，不同型有一些差别。这一基因组的核苷酸排列顺序已基本测出，此RNA中编码有该病毒的全部蛋白质及RNA复制酶等。当此mRNA翻译成蛋白质时，首先生成一个相联在一起的前体蛋白，然后经过几次剪切加工，才形成病毒蛋白VP1、VP2、VP3、VP4等及RNA复制酶等（图1—1）。

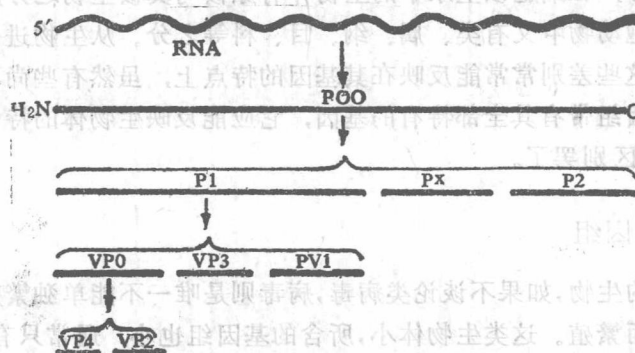


图1-1 脊髓灰质炎病毒基因组及其翻译产物

脊髓灰质炎病毒的繁殖是在它接触宿主细胞（如猴肾细胞）RNA注入细胞后开始，其中很少一部分mRNA翻译成蛋白质，另一部分则复制出相应的mRNA负链，以负链mRNA为模板，合成大量的正链mRNA，为新生的病毒提供新的基因组，达到一定时间后，组成新的病毒，至此达到了繁殖的目的。脊髓灰质炎病毒不大，含有的蛋白质与核酸都有限，它本身并不能完成这一复杂过程，要靠宿主细胞的亚细胞器如核糖体及细胞中的tRNA等其他因子，不然不能完成蛋白质的合成。另外病毒中虽只有一条RNA链，并可翻译出全部病毒的蛋白，但各种蛋白的翻译量、拷贝数并不一样。加上由负链mRNA复制出的基因组的量，也须适应生成新脊髓灰质炎病毒蛋白的比例，才能组成新病毒。这说明在RNA基因组中必定含有调节控制区来调节各种蛋白、酶以及新mRNA的合成，只是脊髓灰质炎病毒中的RNA很简单，不含特殊的调控区。一般认为其3'末端的多聚A的长短有调控作用，另外RNA分子本身折叠所形成的内部双螺旋结构，以及其他特殊结构（如所含的核苷酸顺序和所形成的发卡结构）等都会影响基因组的复制与翻译。从这一病毒分子的繁殖过程，不难看出基因的结构与核苷酸排列顺序及其功能有密切关系。另外了解基因组的结构，对控制其感染，制备疫苗也很有用，如作为脊髓灰质炎病毒的减毒株，实际上就是野型基因组的变异体，在变异后改变了其生物性质，降低了病毒体的毒性（图1—2）。

2. 流行性感冒病毒 此病毒简称流感病毒。它属于粘病毒，呈球型，可分类为甲、乙、丙三型。甲型根据其中含有的血凝素（HA）与神经氨酸（NA）抗原的不同，又分为若干亚型。仅从流感病毒的HA基因的差别上已找到有12种以上，NA有9个亚型，故常用H₂

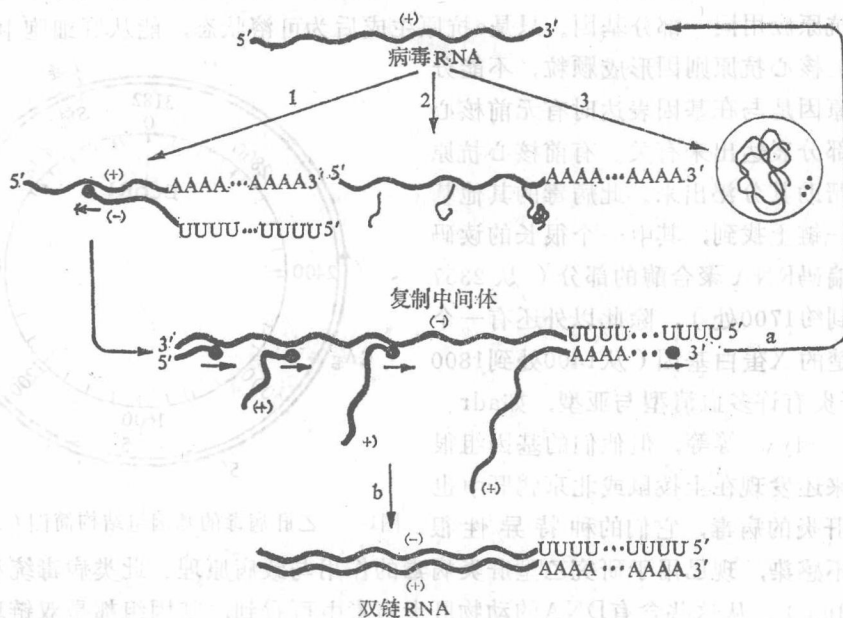


图1-2 脊髓灰质炎病毒RNA的新合成与新病毒的形成

$H_1H_2H_3\cdots$ 及 $N_1N_2N_3\cdots$ 来代表某一亚型流感病毒。有些动物体内也有流感病毒,而且与人的流感病毒能发生基因重组,并引起基因变异,故流行的感冒病毒亚型十分复杂,不易用免疫法使人产生终身的免疫力。流感病毒甲乙型分析的比较清楚,它们都含有8个单链的RNA片段,这8个RNA片段的核苷酸顺序大部分都已测出,它们都是由负链mRNA基因编码的。从N基因所推出的氨基酸不同亚型之间有一定的相似性(图1—3)。

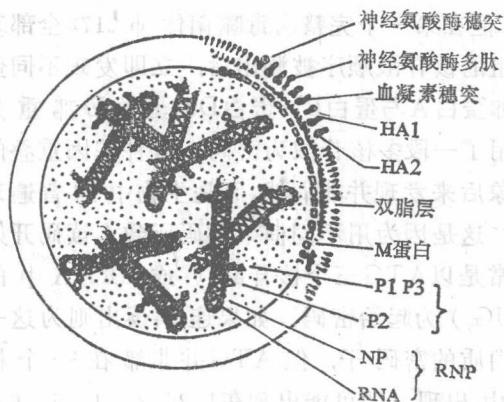


图1-3 流感病毒结构的示意图

流感病毒中的8个负链 mRNA各自编码一种或两种特异的蛋白质(如病毒蛋白质)。其中 P_1 、 P_2 与 P_3 三种蛋白由一个负链mRNA编码。病毒中有一个核蛋白及一个神经氨酸酶,它们也各由一条mRNA负链所编码,其他三条负链mRNA则与血凝素 HA_1+HA_2 ,基质蛋白 M_1+M_2 以及核蛋白中的非结构蛋白 NS_1+NS_2 有关。在 M_1 与 M_2 基因及 NS_1+NS_2 基因中有密码与基因重叠的现象,这种基因重叠的现象在病毒基因组中常出现,其重叠的特点将在病毒的基因组中4介绍。

过去认为只是HA与NA的变异或仅由HA决定此病毒的毒性,现已知并不完全正确,实际上变异与 P_3 (可能是聚合酶)和NP也有关。

3.乙型肝炎病毒 乙型肝炎病毒(hepatitis B Virus HBV)是一种DNA病毒,它的基因组是由一个双链环状DNA组成,但在已找到的病毒分子中,DNA都不是完整的双链环,而是由一个长链及一个短链组成。乙型肝炎病毒DNA比较小,只有约3200bp(以补成完整双链环计算)其结构如图1—4。在一条链环上已找到两个重要蛋白的编码基因,一个