

谈谈遗传工程

庚 镇 城 编
上海科学技术出版社



现代科学技术讲座

谈谈遗传工程

庚 镇 城 编

上海科学技术出版社出版

(上海瑞金二路 450 号)

新华书店上海发行所发行 上海市印刷四厂印刷

开本 787×1092 1/32 印张 2 字数 41,000

1979 年 9 月第 1 版 1979 年 9 月第 1 次印刷

印数 1—30,000

书号: 13119·807 定价: 0.17 元

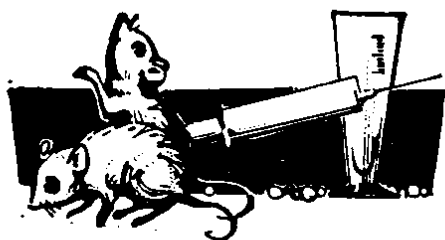
出版说明

为了帮助广大读者初步了解现代科学技术一些重要部门的基本面貌,我们请有关同志分头编写了这组《现代科学技术讲座》,扼要介绍现代农业科学技术、能源科学技术、材料科学技术、电子计算机科学技术、激光科学技术、空间科学技术、高能物理和遗传工程等的概况,及其在四个现代化中的重要意义,供广大干部和各条战线上的同志们参考。

错误和不足之处,欢迎广大读者批评指正。

目 录

一、漫长的过程.....	1
二、分子遗传学的蓬勃发展.....	6
三、改造生命类型的工程——遗传工程.....	21
四、遗传工程的应用前景.....	39
五、潜在的危险.....	48
六、安全措施问题的争论.....	54



一、漫长的过程

“遗传工程”现在几乎是妇孺皆知的名词了。但是，相当多的同志虽然见过它的面，却并不一定清楚它究竟指的怎么一回事。如果从广义上说，遗传工程就是用人工的手段去改造生物的遗传性状。它可以是个体水平的(如杂交育种)，细胞水平的(如细胞融合技术)，染色体水平的或是基因水平的。但现在说的遗传工程，一般是指在基因水平上改变遗传性状，所以也称基因工程。遗传工程是在分子遗传学基础上产生的，它是借助生物化学的手段，将一种生物的遗传物质抽提出来，在体外进行切割和重组，然后引入另一种生物的活细胞内，以改变或创造新的生物品种的遗传学技术。

现代的遗传工程是七十年代才出现的新事物。但是遗传工程的根源却可以追溯到几千年前。人类对遗传物质的认识，经历了漫长的过程。

所谓遗传，就是亲代的性状能够通过遗传物质传给下一代的现 象。它是生命延续和种族繁衍的保证。俗话说“种瓜得瓜，种豆得豆”，就是对遗传现象的生动描述。不过，子女虽然象父母，却又总有些不完全象的地方，“一母生九子，九子各别”。这种生物体因环境的改变而引起自身遗传性改变的现象就叫变异，它是生物进化与发展的根据。遗传与变异都是生物体的重要特征。研究生物遗传及其变异规律的科学，就是遗传学。

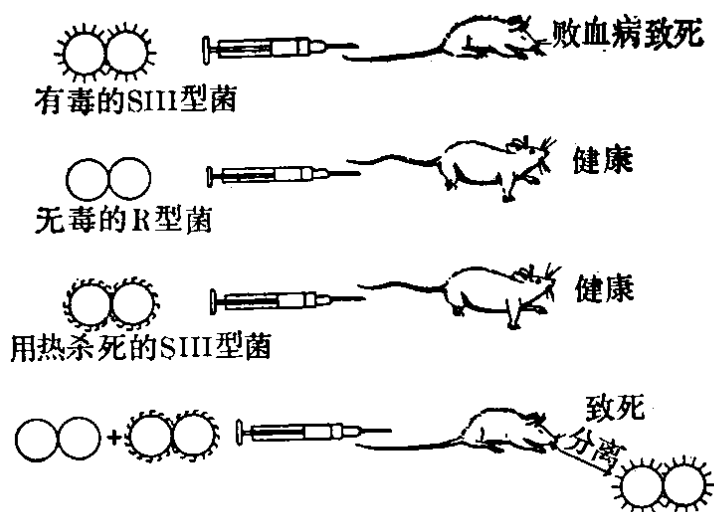
远在史前时期，人类就已经发现了生物的遗传与变异，劳

动人民正是基于对遗传变异现象的初步认识，才开创了驯养动物和栽培植物的农牧业。根据考古学家近年来的研究，早在公元前 7500—6750 年间，西亚的原始居民就已经驯养狗了，狗大概是首先被人类驯化的动物。稍迟，又驯化了猫、羊、牛等动物。人类栽培植物的历史也有几千年了。公元前 6000～前 4000 年间，古巴比伦人和亚述人已经学会在椰枣树的雌株与雄株间进行人工授粉。公元前 4000 年左右，古埃及人已种植大麦、小麦、水稻、大豆、棉花等多种作物。差不多相同时期，我们勤劳智慧的祖先，也已开始从事培育动植物品种的活动了。在新石器时代(1 万年～4 千年前)中期的仰韶文化中，农业和畜牧业已经占了重要的地位。在几千年漫长岁月中，人们通过选择，杂交等方式，驯化和培育出了许许多多动植物品种，人类的活动是有成效的。但是，那时候人们对于遗传现象的本质及其规律性还缺乏深刻的了解，因此，人类改造生命类型的活动，免不了要受许多盲目性、偶然性的支配，人类在大自然面前所能赢得的自由是很有限的。

十八世纪中叶，在资本主义经济发展的基础上，欧洲人开始对生物遗传变异现象进行实验性的研究。一七六〇年俄国彼得堡科学院悬赏征求“关于植物性别”的论文，瑞典著名生物学家林耐获得了奖赏。继之，欧洲许多国家的学者在这方面做了大量的研究工作，到十九世纪六十年代，奥国布隆地方(今捷克斯洛伐克的布尔诺市)的僧侣、植物学家孟德尔以豌豆为材料，进行了多年的杂交试验，终于发现了生物遗传的基本规律。孟德尔可谓是近代遗传学的鼻祖。根据他的理论，生物的各种遗传性状均由细胞内相应的颗粒性遗传因子(廿世纪初改称为基因)所决定。譬如，豌豆的红花性状是由红花因子决定，白花性状由白花因子决定；高植株由高茎因子决

定,矮植株由低茎因子决定;等等。但是,遗传因子存在于细胞的什么部位?它的物质载体是什么?这些问题,在孟德尔的时代,当然是无法回答的。一九一〇年以后,以美国学者摩尔根为代表的细胞遗传学家们,创立了染色体^①遗传理论。这个理论确定,细胞核内的染色体是遗传的物质基础。染色体上面排列着许多基因,它们既有分工又相互协调地控制着细胞的分化和性状的发育。但是,由于当时的遗传学家普遍地忽视生物化学,很少去研究基因的化学本质,所以直到一九四〇年左右,基因的概念还只停留在形式上,而没有确切的内容。由于蛋白质具有巨大的多样性(估计蛋白质在一百万种以上,最简单的细菌至少也包含两千种以上的蛋白质,人体内大约有五万种以上的蛋白质),所以,五十年代以前,遗传学界普遍把蛋白质误认为基因的化学本质。四十年代中期,美国学者艾弗里等人成功地进行了证明基因的本质是脱氧核糖核酸(DNA)的实验,冲破了这种传统的错误观点。艾弗里他们使用的研究材料是肺炎双球菌。肺炎双球菌有外侧包有荚膜的S型和由于突变失去荚膜的裸露的R型。S型菌因为具有荚膜,不被白血球吞食,因而是强毒性的,只要给小白鼠注射很少量,即可使小白鼠感染肺炎而死亡。R型细菌则因为没有荚膜而很容易被白血球吞食,所以属于无毒型。一九二八年,英国的医生格利菲斯在研究肺炎双球菌时,发现了一件令人惊异的事实:他把少量无毒的R型细菌同大量的已被高温

① 细胞在进行有丝分裂时,细胞核的核膜消失,出现易被碱性染料着色的丝状物体,称染色体。染色体在体细胞中成对,在生殖细胞中成单。各种生物染色体的数目与形态是一定的。例如,果蝇染色体数是4对8条,家蚕是28对56条,金鱼是47对94条,人是23对46条,豌豆是7对14条,水稻是12对24条。



杀死的有毒细菌
SIII 型混到一起，
注射到小白鼠体
内，结果小白鼠
死了，后来从小
白鼠的血液中分
离出多量 SIII 型
的活细菌(图1)。

基因的化学

图1 格利菲斯发现的令人惊异的现象

本质是 DNA 这一点，
并不限于艾弗里等人的
实验证明，也为后来美
国研究者赫尔希和蔡斯
(1952)的噬菌体标记实
验所确认。噬菌体属于
病毒类，是一类没有细
胞结构但有遗传、变异、
共生、干扰等生命现象
的极小的微生物，它们
的组织结构极其简单，
基本上只由一个蛋白质
的外壳和壳体内的核酸
组成。赫尔希和蔡斯用
放射性同位素硫³⁵和
磷³²分别标记 T2 噬菌
体的蛋白质外壳和藏在
噬菌体头部内的DNA。

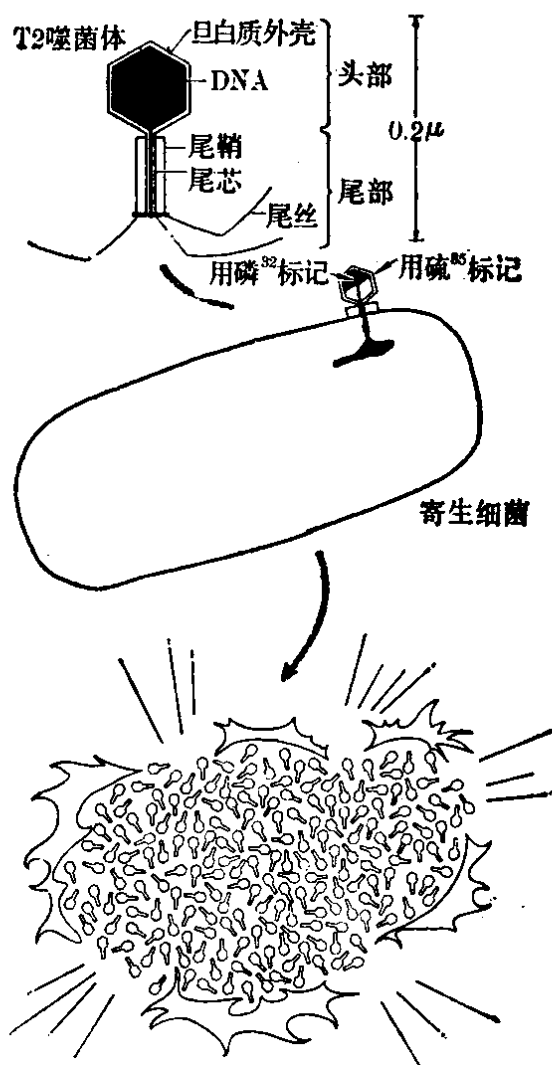
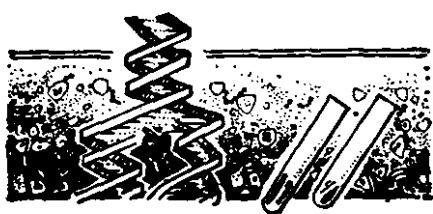


图2 T2噬菌体侵染宿主细菌时的情形

结果发现，噬菌体侵染宿主细菌时，蛋白质外壳留在宿主菌体外，只有噬菌体的 DNA 部分侵入到宿主菌体内；而且在很短时间内（大约廿分钟）DNA 就繁殖出许多同原噬菌体一模一样的、具有完整蛋白质外壳的第二代噬菌体。这说明噬菌体的 DNA 中包含着整个噬菌体的全部遗传信息（图 2）。

基于这些令人信服的实验结果，五十年代初，绝大多数遗传学家已认清基因的化学本质不是蛋白质，而是核酸，特别是脱氧核糖核酸(DNA)。从此遗传学的研究别开生面，迎来了急风暴雨般的发展时期。



二、分子遗传学的蓬勃发展

DNA 分子的双螺旋结构

一九五三年,在遗传学发展史上是一个非常重要的年份。在这一年的四月,英国《自然》杂志上发表了美国学者华生和英国学者克里克合写的一篇题为《核酸的分子结构》的文章。这篇文章很短,不足一千个词,但它却标志着遗传学的一个新

时代的开端。以它为界标,遗传学的研究从细胞水平深入到分子水平。从此,二十五年来,分子遗传学取得了巨大的惊人的进展,一个又一个重大成就接踵而至。

华生和克里克根据X光衍射资料,提出了DNA分子的双螺旋结构模型(图3)。

根据这个模型,DNA分子由两条多核苷酸长链构成,每条核苷酸链由许多核苷酸连接而成,每个核苷酸包括三种成分:脱氧核糖(S)、磷酸(P)和碱基。碱基又分两类:嘌呤和嘧啶。嘌呤中有腺嘌呤(A)和鸟嘌呤(G),嘧啶中有胸腺嘧啶(T)和胞嘧啶(C)。脱氧核糖

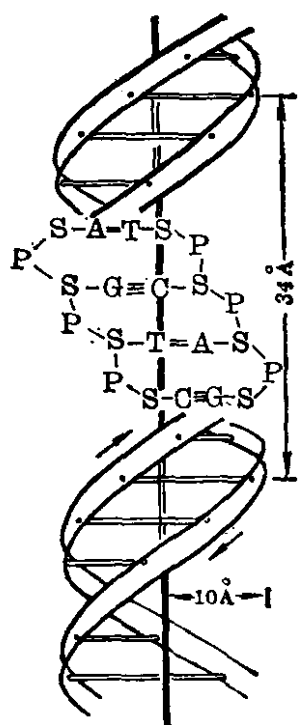


图3 华生-克里克的DNA结构模型

鸟嘌呤(G),嘧啶中有胸腺嘧啶(T)和胞嘧啶(C)。脱氧核糖

和磷酸相间排列,以磷酸二酯键相连,构成螺旋梯的两臂。两条链子围绕一个抽象的中轴向右旋转,每一个螺旋纵距是 $34\text{\AA}$ ($1\text{\AA} = 10^{-10}\text{米}$),臂距中轴的距离为 $10\text{\AA}$ 。螺旋梯的“横档”由一个嘌呤和一个嘧啶组成,它们之间由氢键连接起来,并且存在着严格的配对关系(互补性),即A总是和T, G总是和C配对。华生-克里克的模型后来得到了一系列实验工作的支持,特别是一九六九年加利福尼亚大学的研究人员用电子显微镜拍摄到了放大七百三十万倍的DNA双螺旋结构的照片,从而给予直接的证实^①。

从华生-克里克的模型出发,可以圆满地解释许多问题。生命界形形色色,多种多样,即便是同一物种的个体也千差万别,无一尽同。拿人来说,除一卵双生儿之外,没有两个个体在遗传上是一样的,不论他们是父子、母女,还是兄弟、姊妹。生命的多样性自然是以遗传物质的多样性为基础的。分子遗传学的研究表明,DNA分子完全能满足这种要求。

DNA分子的结构似乎很简单,变化也不多。在核苷酸的三个组成部分中,磷酸和核糖是大家相同的,只有碱基有些变化。但是碱基只有四种,而且是严格配对的,一对核苷酸的两个碱基之间,可能有四种相互组合的方式,即:



不过,我们必须注意到,DNA是由许多核苷酸组成的。核

① 不过近两年,华生-克里克的模型受到了挑战。新西兰的研究人员在重新推敲了DNA的X线衍射结果之后,认为双螺旋结构未必说是正确的,并于1976年提出一个新模型——“并排模型(side-by-side)”。照他们看来,DNA的结构与其说是双螺旋的,勿宁说是象梯子那样并排着的。倘此新模型被证实的话,以华生-克里克模型为基础分子遗传学将会受到很大冲击。

核苷酸对之间虽然有互补关系，相邻两对核苷酸的排列顺序却没有什么限制。这就好比拍电报用的电码一般，虽然电码符号只有“·”和“—”两种，但只要把·或—按照不同的顺序排列起来，就可以表示各种不同的意思了。由于 DNA 都是大分子，一个 DNA 分子的核苷酸至少在 100 对以上，因此就可能有 4^{100} 种以上的排列方法。这是一个十分庞大的数字，它表明 DNA 分子的多样性几乎是没有限制的。现在认为，DNA 的多样性可以跟蛋白质相比，甚至超过蛋白质。DNA 分子碱基的排列顺序，就是生物遗传的电报，其中包含的大量信息完全可以表达生物的不同遗传性状。

另外，生物的种类不同，其核酸分子的大小也不同(见表 1)，越是复杂的生物体，它细胞中的 DNA 分子就越大。据今所

表 1 几种生物核酸的比较

生物种类	核 酸 类 型	道尔顿数①	核苷酸对数	基 因 数
人(2 <i>n</i>)	双链 DNA	4.2×10^{12}	6.5×10^9	—
人的精子(<i>n</i>)	双链 DNA	2.1×10^{12}	3.3×10^9	—
鸡(2 <i>n</i>)	双链 DNA	1.5×10^{12}	2.3×10^9	—
果蝇(2 <i>n</i>)	双链 DNA	1.4×10^{11}	2.2×10^8	—
酵母(2 <i>n</i>)	双链 DNA	2.4×10^{10}	3.7×10^7	—
大肠杆菌(<i>n</i>)	环状双链 DNA	2.8×10^9	4.3×10^6	—
T2 噬菌体	线状双链 DNA	1.3×10^8	2.0×10^5	~100
T7 噬菌体	线状双链 DNA	2.5×10^7	38,000	~30
φX174 噬菌体	环状单链 DNA	1.7×10^6	5500	~8
MS2 噬菌体	线状单链 DNA	1×10^6	3300	3

知，DNA 分子在大分子中可说是最大的。而且，不同的 DNA 分子在适宜的条件下还可以互相连接起来，形成更大的分子，

① 道尔顿：分子重量的单位，1 道尔顿 = 1.65×10^{-24} 克。

这样，它含有的基因就更多了。象人的 DNA 分子含有 6.5×10^9 对核苷酸，其不同的排列方法可能有 $8000^{1,000,000,000}$ 种！包含的基因数以万计。

现在可以很确定地说，生物的性状就是通过遗传信息即碱基的排列顺序传递给下一代的。不同种类的生物，因为核酸不同，遗传信息也就不同，所以表现出不同的性状；同种的生物，也由于每个个体都具有特定的碱基排列顺序，所以表现出的性状也有差异（当然还有环境的因素，但这是内因）；子女为什么会象父母？这个几千年来使人迷惑不解的谜语的答案就是，父母把相应的碱基排列顺序传给了子女，并且在子女身上得到了表现。

基因的一个重要特性是能够进行复制。华生 - 克里克模型说明了亲代 DNA 分子在聚合酶的参与下，如何形成两个与自身相同的子代 DNA 的过程(图 4)。在复制过程中，DNA 的两条多核苷酸长链彼此松开，各以自己为样板，同化或吸收周围已经合成的核苷酸，根据碱基互相配对的规律，按照自己的结构方式组合一条新的长链，跟自己并列盘旋，成为稳定的双螺旋结构。这样就完成了复制过程。DNA 的复制过程是很复杂的，已故的日本学者冈崎对此作出了重要贡献。他提出的不连续复制模型，不论在原核生物中还是真核生物

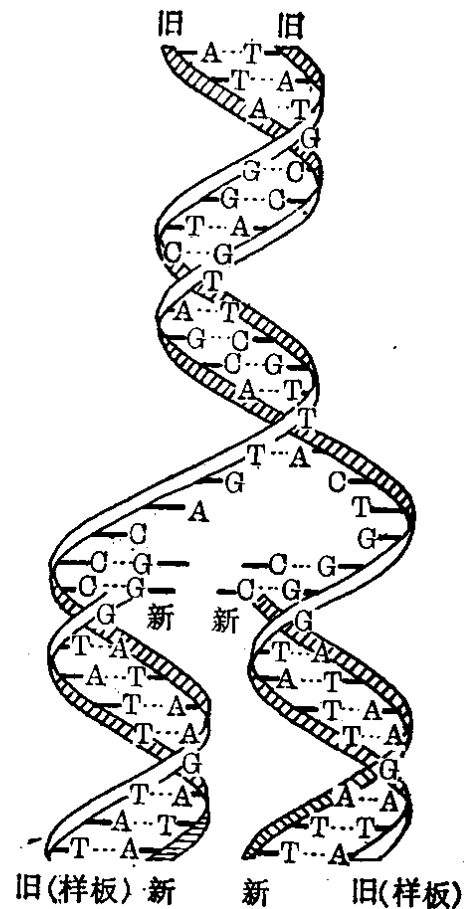


图 4 DNA 复制模型

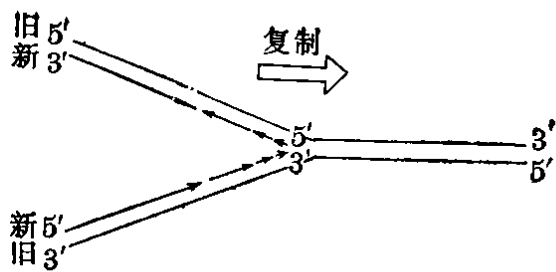


图5 不连续复制

中①，都得到了证实。按照冈崎的模型，DNA的复制，是在聚合酶的参与下，一个片段一个片段分头进行的(图5)。

DNA的另一重要功能
是转录。转录是指以一条

DNA链(信息链)为样板，在依赖DNA的RNA聚合酶参与下，根据碱基配对(A-T, G-C)原理，合成互补的核糖核酸(RNA)的过程。在核糖核酸中没有胸腺嘧啶(T)，而代之尿嘧啶(U)，所以在转录过程中，是尿嘧啶(U)与DNA中的腺嘌呤(A)呈现互补关系(图6)。转录的结果形成信使RNA(mRNA)，转移RNA(tRNA)和核糖

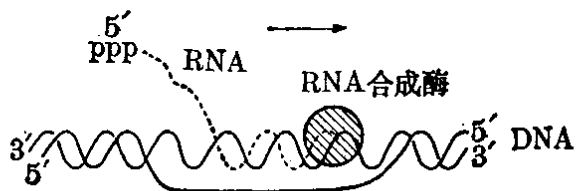


图6 转录示意图

体RNA(rRNA)。它们都参与蛋白质(结构蛋白质与酶)的生物合成，细胞内合成蛋白质的过程称之为翻译(图7)。DNA的转录和翻译关系到细胞分化和性状发育。这与基因通过对于酶合成的控制来控制细胞的生化过程，也是完全一致的。

记录在DNA分子上的遗传信息一方面通过复制而传种接代，另一方面通过转录而指导细胞内蛋白质(包括酶)的合

① 依据染色体的存在形态，生物可分两大类：原核类与真核类。真核类是高等的生物类型，细胞通常有细胞核，在细胞分裂期间，出现用光学显微镜可以观察到的由核酸和蛋白质组成的染色体结构。而原核类则属低等的生命类型如病毒、细菌等，无细胞核，其“染色体”仅由裸露的DNA组成。

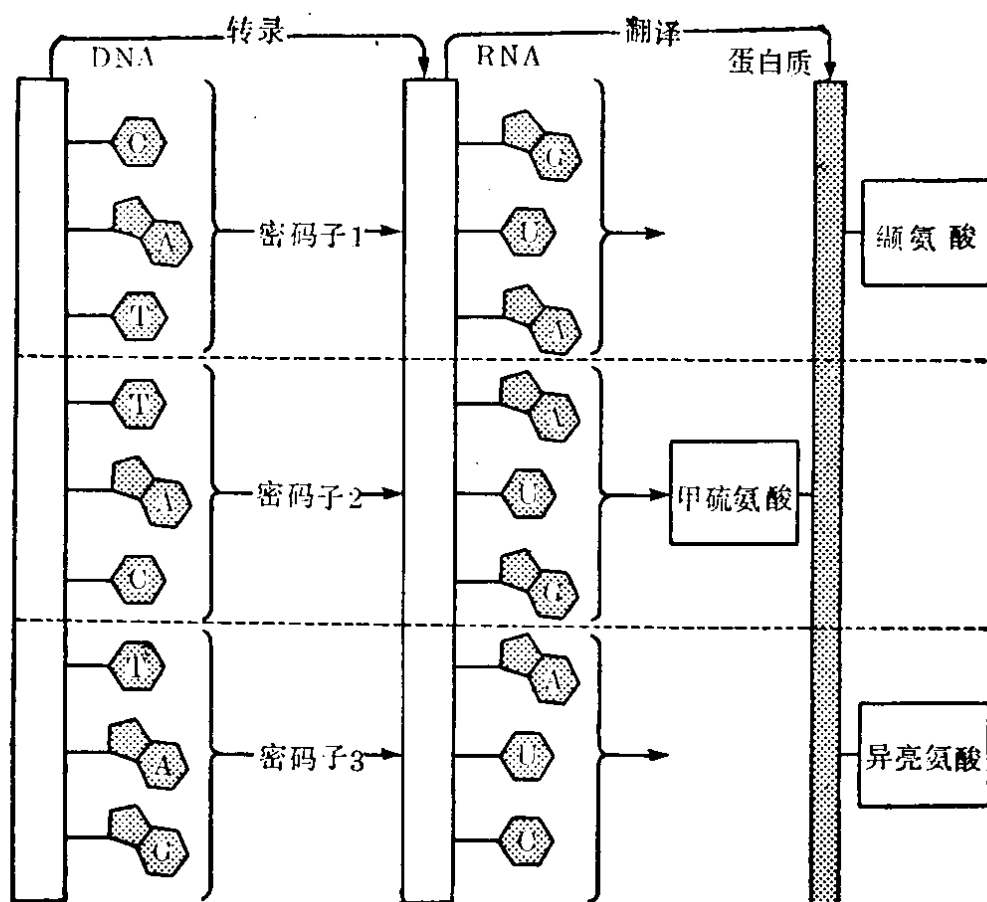


图7 转录、翻译的示意图

成,进而控制细胞的分化和发育。这就是根据华生-克里克的DNA 结构模型对遗传过程作出的解释。

三体密码和遗传密码表

遗传信息是 DNA 分子中碱基的排列顺序,这句话说起来挺轻巧。可是,碱基的排列顺序又是按照什么规律编制成遗传密码的呢?再有,生物体内有成千上万种蛋白质(包括酶),它们或构筑机体,或决定代谢,任何生命活动都离不开它们。可以说,没有蛋白质就没有生命。蛋白质都是由氨基酸组成。尽管蛋白质的种类非常多,但氨基酸却只有二十种,各种蛋白质在结构及功能上的差别,皆因氨基酸的排列次序不

同所致。那么氨基酸的排列次序是怎样决定的呢？遗传信息是怎样“指导”蛋白质合成的呢？美国学者尼伦伯格、奥巧阿、科拉纳等人经过十来年的努力，终于在六十年代中期搞清了在细胞合成蛋白质(包括酶)的过程中，即所谓翻译的过程中，是 DNA 的转录产物——信使 RNA (mRNA) 的碱基次序决定了氨基酸的排列次序。并且确定 mRNA 中的每三个相邻的核苷酸成为一组，规定一种氨基酸，称之为三体密码。四种碱基中任取三个，共有 $4^3 = 64$ 种不同排列方式，这 64 种三体密码就规定着 20 种氨基酸(见表 2)。

表 2 三 体 密 码 表

第一位碱基 第二位碱基	U	C	A	G	第三位碱基
U	苯丙氨酸 苯丙氨酸 亮氨酸 亮氨酸	丝氨酸 丝氨酸 丝氨酸 丝氨酸	酪氨酸 酪氨酸 ~终止~ ~终止~	半胱氨酸 半胱氨酸 ~终止~ 色氨酸	U C A G
C	亮氨酸 亮氨酸 亮氨酸 亮氨酸	脯氨酸 脯氨酸 脯氨酸 脯氨酸	组氨酸 组氨酸 谷氨酰胺 谷氨酰胺	精氨酸 精氨酸 精氨酸 精氨酸	U C A G
A	异亮氨酸 异亮氨酸 异亮氨酸 甲硫氨酸	苏氨酸 苏氨酸 苏氨酸 苏氨酸	天冬酰胺 天冬酰胺 赖氨酸 赖氨酸	丝氨酸 丝氨酸 精氨酸 精氨酸	U C A G
G	缬氨酸 缬氨酸 缬氨酸 缬氨酸	丙氨酸 丙氨酸 丙氨酸 丙氨酸	天冬氨酸 天冬氨酸 谷氨酸 谷氨酸	甘氨酸 甘氨酸 甘氨酸 甘氨酸	U C A G

整个生命界,从最简单的病毒到最高等的人类,其根本的生命活动即合成蛋白质的活动,都服从这张统一的密码表的规定。如果说上个世纪三十年代末德国细胞学家施来登、施旺确立的细胞学说是从细胞水平上论证了生命有机界的统一性,那么,二十世纪六十年代中期分子遗传学家所揭露的遗传密码表则是从分子水平上论证了生命有机界的统一性,因而具有重大的理论意义,有人把它比作是生物学上的“罗赛达石碑”^①。遗传密码的发现也具有重大的实践价值。试想,如果不了解核苷酸和氨基酸的对应关系,人们怎么可能人工合成基因和开展遗传工程的研究呢?

近两年来,分子遗传学在解析核酸的一级结构即核苷酸排列次序方面取得了重大进展。比利时肯特大学以菲尔斯为首的遗传学家们成功地解析了一种小型的叫 MS2 的 RNA 病毒的核苷酸排列次序。MS2 是大肠杆菌的噬菌体,它的遗传物质是由 3569 个核苷酸组成的单链 RNA 分子,分子量在 100 万左右。它总共只有三个基因,负责为三种蛋白质(A 蛋白质,外套蛋白和 RNA 复制酶的 β 亚基)编码。MS2 是核苷酸次序即遗传信息被人们完全搞清楚的第一个生物(图 8)。

一九七七年,分子生物学家又相继搞清了大肠杆菌的噬菌体 ϕ X174 和动物病毒 SV40 的全部碱基次序。两者都是

① 1799 年拿破仑远征埃及的时候,拿破仑的一个士兵在尼罗河口小城罗赛达附近发现的一块石碑,是公元前 195 年底比利斯祭司为歌颂埃及王托勒密五世即位所刻。石碑的上部为古埃及象形文字,中部为古埃及世俗体文字,下部为古希腊文,三种文字刻写着相同的赞美辞。这块石碑成为法国年青的有才华的学者 F. 商博梁在十九世纪初叶解开古埃及象形文字之谜的钥匙。