

现代农业科学讲座

XIAN DAI NONG YE KE XUE JIANG ZUO

基 因 工 程

谈家桢 著

农 业 出 版 社

中 国 农 学 会 编

现代农业科学讲座

基 因 工 程

谈家桢 著

中国农学会 编

现代农业科学讲座
基 因 工 程

谈 家 桢 著
中国农学会编

农业出版社出版 新华书店北京发行所发行
农业出版社印刷厂印刷

787×1092 毫米 32 开本 1 印张 14 千字
1979 年 4 月第 1 版 1979 年 4 月北京第 1 次印刷
印数 1—44,500 册

统一书号 16144·1871 定价 0.11 元

根据需要，有目的、有计划、有步骤地改造现有生物的遗传特性，进而创造新的生命类型，是人类的夙愿，也是多少科学家梦寐以求的理想。基因工程的诞生^{*}，标志着人类在实现这一宏伟愿望的征途上，迈出了决定性的一步。基因工程是建立在基因学说基础上的最新科学手段，揭开了定向改变生物类型的无限光明灿烂的前景。

一、基因工程的理论基础

分子遗传学是基因工程的理论基础。什么是分子遗传学呢？这里我们先简要地回顾一下遗传学的发展历史。

遗传学作为一门科学，严格地讲，是从1900年孟德尔规律戏剧性地再发现之后，才正式形成的。

孟德尔1866年提出的分离法则与自由组合律，以及以后由摩尔根等补充的连锁交换律，奠定了近代遗

^{*} “遗传工程”包括细胞工程、染色体工程和基因工程，比较狭义的、从分子水平来讲的遗传工程，就是“基因工程”。——编者注

传科学的基础。即便在分子遗传学取得惊人发展的今天，重温他们的著作，仍使人受益匪浅。

在孟德尔以前，人们虽然知道种瓜得瓜、种豆得豆，但在理论和实践上长期占支配地位的却是所谓“混合遗传”的概念，即认为父本和母本结合，后代的性状有的象父本，有的象母本，基本上是两亲的混合体。这一概念根深蒂固地把遗传现象视为血液的或类似于液态的某种混合。孟德尔在豌豆杂交试验中所暗示的颗粒遗传，即粒子遗传的概念，则是整个思想体系上的重大突破。他认为遗传不是两个亲本的混合，决定遗传作用的是存在于细胞中一种称为遗传因子的粒子，后来又发展称为基因（约翰逊，1909年）。

本世纪初，遗传学研究转入到细胞学水平，迎来了急风骤雨般的发展时期。基因理论在果蝇、玉米等动植物各类材料中都得到了广泛的验证。在大量试验的基础上建立了细胞遗传学。细胞遗传学的中心内容是，遗传粒子（即基因）主要存在于细胞核里面的染色体上，基因在染色体上呈直线排列，这就是基因染色体学说，也就是后来被批判的所谓“摩尔根遗传学”。

在这里我想顺便谈一点属于科学史范畴方面的问题。一门学科不同于定理或定律。定理、定律可以冠

以人名，如门捷列耶夫周期表，爱因斯坦相对论等等。但是用特定的人名来命名一门学科就不仅极不适当，而且是荒谬绝伦了。李森科把遗传学人为地分成所谓的“米丘林遗传学”和“摩尔根遗传学”，一捧一贬，一扶一打，即是出于政治上的别有用心，也是基于主观唯心主义的需要。是完全反历史唯物主义的科学史观的。邓付主席极其精辟地指出：“科学技术是人类共同创造的财富。”（1977年8月17日同丁肇中教授的谈话）任何一种科学都是社会历史的产物，是人们共同劳动的结晶。现代自然科学发展中的一个引人注目的特点是各门科学的相互渗透，彼此结合，因之更不可能是个别科学家单枪匹马的独创。遗传学作为遗传和变异规律的总结，当然也是许多人共同劳动的结果。任何科学部门的发展都不可能是一帆风顺的，在其发展的过程中，出现不同的学派，提出不同的理论，是正常的现象。但是，称得上规律的东西，则一定是源于实际，并经得起实际反复检验的客观真理。基因学说所以历尽风霜而不衰，就在于它具有客观真理的性质。

从混合遗传的概念发展为颗粒遗传、基因遗传的概念是一个重大的飞跃。它标志着遗传科学的黎明。当然，和世界上的任何事物一样，这一概念也有其发展

过程。比如人们曾把基因同蛋白质联系在一起,并且笼统地把基因当成是三位一体的、不可分割的单位(重组单位,功能单位,突变单位)。今天看来,这些见解都不对了。1944年,美国学者阿维利等通过肺炎球菌的转化实验,确定在遗传上起支配作用的化学物质成分并非是蛋白质而是核酸,特别是脱氧核糖核酸(简称DNA),从而澄清了过去长期认为基因多样性渊源于蛋白质多样性,并把基因当作是蛋白质的错误观念。到1953年,华生和克里克研究核酸的分子结构,确定DNA呈双螺旋结构,这是生物学上的重大发现,是继达尔文的进化论之后在生物学上第二个重大发现的里程碑。以此为开端,人们展开了分子生物学、分子遗传学一系列全新领域的研究,短短25年,成果累累,捷报频传,从研究内容到研究方法,使生物科学的整个面貌焕然一新。分子遗传学的研究,使得遗传基因彻底物质化了。基因就是特定DNA区段中核苷酸的排列顺序,遗传信息就贮存在特定的核苷酸顺序之中。

(一) DNA的双螺旋结构

在生物化学分子里,到现在为止只发现DNA分子结构是特殊的象麻花、油条一样扭在一起的双螺旋结构。

DNA的基本单位是核苷酸。核苷酸由磷酸、脱氧核糖和碱基构合而成。它们按一定的顺序首尾相连，联结成很长的DNA链，遗传特性的奥秘就在于碱基的排列顺序上。DNA的碱基有4种：即腺嘌呤（A）、鸟嘌呤（G）、胸腺嘧啶（T）和胞嘧啶（C）。四种碱基就决定了四种核苷酸的区别。这四种核苷酸在相互偶

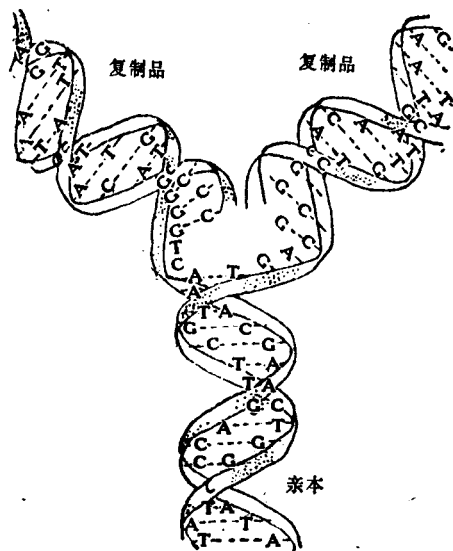


图1 DNA的复制

在复制过程中，亲代的双股螺旋结构的DNA(图的下半部)拆散为两个单链，然后，以这两个单链为模板，合成两个子代双链DNA分子(图的上半部)。

合时，A只能和T配对；C只能和G配对。

DNA的双螺旋结构能够复制，此种复制机理很重

要。它是确保细胞分裂使子细胞基本维持原有遗传结构不变的前提，也是在个体生长和繁衍后代的过程中遗传特性得以稳定传递的重要保证。DNA复制是在几种聚合酶的参与下，依据碱基互补法则完成的，是一类半保留复制。过程本身极为复杂，在这方面，日本学者冈崎曾作出重要贡献。他提出的“冈崎模型”，不论是在原核生物中还是在真核生物中，目前均已得到证实。

(二) 中心法则和遗传密码的统一性

1958年，克里克提出了遗传物质与遗传性状之间关系的中心法则，认为：生物体的遗传信息是经由DNA，传到RNA，再传到蛋白质而表现出来的，即DNA→RNA→蛋白质。

生物体的特性都是由蛋白质直接或间接地体现出来的。蛋白质和核酸都是生物高分子。蛋白质由20种氨基酸组成，两个氨基酸结在一起叫二肽，十个氨基酸结在一起叫十肽，更多的氨基酸连在一起就是蛋白质。不同的氨基酸以不同的形式组合起来，就成为多种多样的蛋白质（结构蛋白和酶）。不同的蛋白质具有不同的性质和作用。酶是一种催化剂。一个生物体内存在着许许多多的酶，新陈代谢的每一个化学反应

过程都要有特定的酶参与。例如，矮秆玉米、矮秆小麦的矮秆性状，就同体内的生长激素合成有关，而能否合成或合成多少某种特定的激素则又取决于特定的酶。有的人生下来眉毛、头发、皮肤都表现白化症。这是由于缺少一种酪氨酸酶，不能产生色素。它与正常人在遗传上的差别就在于能不能产生这种酶的差异。生物体内酶和蛋白质的差别大多是由遗传物质（基因）决定的。

现今关于DNA同蛋白质生物合成的关系，已经基本查明，都是以DNA为样板，根据“三体密码”法则，合成相对应的信使RNA，这个过程称为转录。下一步则是根据信使RNA在细胞质里精确无误地合成相对应的蛋白质，这个过程叫做翻译。适应细胞分化与发育过程的需要，DNA在不同的时间、空间条件下发出不同的信息指令，经过极为复杂的过程，最终合成各种不同的蛋白质，从而使生物体表现出各种生物学机能。

构成DNA的基本核苷酸单位有四种（A、T、C、G），现已发现，核苷酸总是三个连在一起成为一组，这就叫做“三体密码”。四种核苷酸应包含有64种不同的三体密码。当由DNA向信使RNA转录时按照A—U

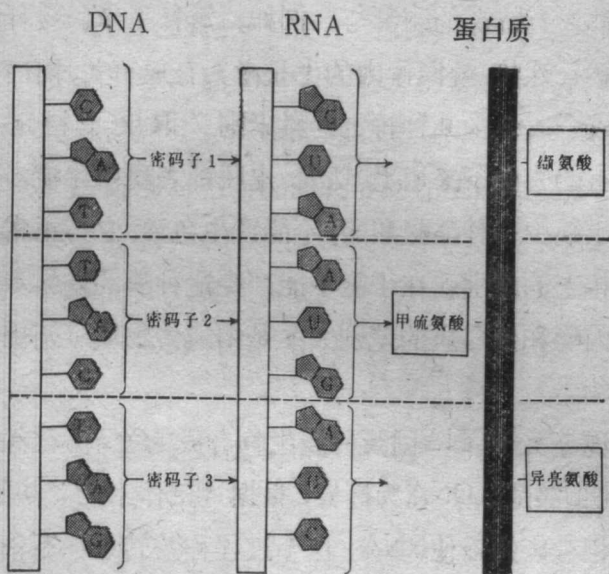


图2 蛋白质合成中DNA作用示意图

图解说明蛋白质合成中DNA的作用。核苷酸顺序，如CATTACTAG所含的遗传信息是“写”在三联体“密码子”CAT、TAC、TAG之中。每一个密码子通过一系列包括核糖核酸(RNA)在内的中间步骤，确定一个蛋白质分子中特定氨基酸的位置。于是，DNA分子中特定的核苷酸顺序决定了蛋白质中相应的氨基酸顺序，每一个三联体核苷酸密码子把20个可能的氨基酸中一个氨基酸置于蛋白质链中正确的位置上。因为氨基酸的顺序既确定了蛋白质的结构，也确定着蛋白质的功能，所以DNA的核苷酸顺序本质上决定着生物的每一个性质。

(RNA中以尿嘧啶U代替胸腺嘧啶T, 配对时A—T亦相应改为A—U), C—G配对法则, 在mRNA中也会相应包含64种三体密码。

这种以每三个核苷酸组成一单位的三体密码组合

表1 遗传密码表

	中 间 碱 基				
5'—末端碱基	U	C	A	G	3'—末端碱基
U	苯 丙 氨 酸	丝 氨 酸	酪 氨 酸	半 胱 氨 酸	U
	苯 丙 氨 酸	丝 氨 酸	酪 氨 酸	半 胱 氨 酸	C
	亮 氨 酸	丝 氨 酸	终 止	终 止	A
	亮 氨 酸	丝 氨 酸	终 止	色 氨 酸	G
C	亮 氨 酸	脯 氨 酸	组 氨 酸	精 氨 酸	U
	亮 氨 酸	脯 氨 酸	组 氨 酸	精 氨 酸	C
	亮 氨 酸	脯 氨 酸	谷 氨 酰 胺	精 氨 酸	A
	亮 氨 酸	脯 氨 酸	谷 氨 酰 胺	精 氨 酸	G
A	异 亮 氨 酸	苏 氨 酸	天 冬 酰 胺	丝 氨 酸	U
	异 亮 氨 酸	苏 氨 酸	天 冬 酰 胺	丝 氨 酸	C
	异 亮 氨 酸	苏 氨 酸	赖 氨 酸	精 氨 酸	A
	甲 硫 氨 酸 或 甲 硫 氨 酸	苏 氨 酸	赖 氨 酸	精 氨 酸	G
G	缬 氨 酸	丙 氨 酸	天 冬 氨 酸	甘 氨 酸	U
	缬 氨 酸	丙 氨 酸	天 冬 氨 酸	甘 氨 酸	C
	缬 氨 酸	丙 氨 酸	谷 氨 酸	甘 氨 酸	A
	缬 氨 酸	丙 氨 酸	谷 氨 酸	甘 氨 酸	G

形式在生物界中普遍存在，同出一辙。三体遗传密码的发现和遗传密码表的确定，是分子遗传学在六十年代中期获得的最重大的突破。现已证明，整个生命界从病毒到高等植物，从变形虫到人类，其最根本的生命活动都服从于这张密码表的规定。因此，如果说上个世纪三十年代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

末，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。此外，遗传密码表的发现，也具有重大的实践价值。只有在了解核苷酸和氨基酸的对应关系的基础上，人们今天才有可能去着手解决人工合成基因与开展基因工程的研究。

1970年，梯明、巴尔地摩等在有些生物中发现一种新的酶，叫逆转录酶。在逆转录酶的作用下，RNA可逆转复制成DNA ($\text{RNA} \rightarrow \text{DNA}$)。这一发现进一步丰富了中心法则的内容。到现在为止，还没有发现蛋白质也能逆转复制成RNA或DNA的事例。因此中心法则可以作些修订，用图3来说明之。图中实线箭头表示遗传信息的一般流向，虚线箭头表示特殊流向，其中RNA的复制和由RNA转录成为DNA已经发现，由DNA直接流向蛋白质虽然还没有发现，可是认为是可能的。

另外，有些病毒，如烟草花叶病毒，本身完全

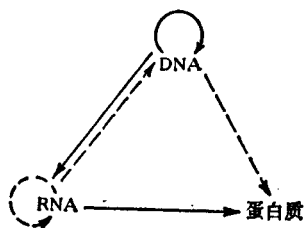


图3 中心法则示意图

DNA 转录为RNA，RNA 亦可反向转录为DNA，RNA 转译为蛋白质，DNA 不能直接转译为蛋白质。

没有DNA,它们的遗传信息是记载在核糖核酸的分子上(RNA)。在这里核糖核酸代替脱氧核糖核酸担负遗传的主角,但核酸决定氨基酸组成特定的蛋白质的这一基本点则完全不变。

(三) 基因表达和控制调节

分子遗传学的第三个重大成就就是阐明了原核生物性状表达的调控机制。任何生物的遗传信息都只有通过转录、转译才能表达出来。在生命体的生长发育过程中遗传信息根据时间、空间以及内外环境条件的不同,有选择地、有次序地进行表达。对于这样一个错综复杂而又高度自控的问题,1961年雅各布和莫诺提出了大肠杆菌操纵子模型。乳糖操纵子包括三个结构基因,分别携带合成 β -半乳糖苷酶、乳糖透性酶和乙酰化酶的遗传信息。三个结构基因前面有一个操纵基因,上面有同阻遏物结合的位置,在阻遏物附着时,结构基因失去活性,不能合成三种酶,操纵基因前面有一个启动基因,操纵基因上没有阻遏物结合时,RNA聚合酶就同启动基因接触,向操纵基因和结构基因移动,于是合成出三种酶。启动基因前面还有一个调节基因,它控制阻遏物的合成,来调节结构基因的活性。这一套相互制约的基因共组成一个功能单位,使生物在不

同环境下,表现出不同的特性。这一理论基本上就原核基因表达的调控机理作出了令人信服的阐明(图4)。从六十年代末以来,关于真核类生物基因表达的调控机理,虽有几个模型,但均未取得根本性进展。这个问题不突破,遗传工程的效用势必要受到很大限制。

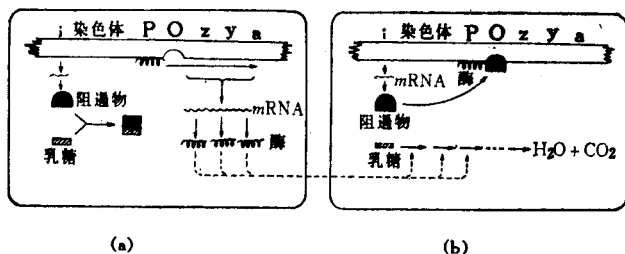


图4 大肠杆菌的乳糖分解自动控制“线路”图

i 调节基因; *P* 启动基因; *O* 操纵基因; *z*, *y*, *a* 三个结构基因分别决定 β -半乳糖酶、乳糖透性酶和乙酰化酶。

(a) 表示细胞中有乳糖存在时,促使和乳糖代谢直接有关的三种酶合成; (b) 表示当乳糖通过这些酶的作用而消失时,有关的酶的合成便停止。

二、基因工程的基本方法

由于分子遗传学确定了各种生物的遗传物质都是核酸。碱基配对的规则是一致的,遗传密码也是统一的。六十年代末,又相继发现了一些工具酶:限制酶、连接酶、末端转移酶等,它们能够切割和连接遗传物质。在七十年代初一个崭新的科学技术部门——基因

工程，终于在分子生物学发展的基础上脱颖而出。

基因工程的概念是：在体外或试管中借助酶促反应将目的基因或异源DNA片段与适当的载体进行重组，造成杂种DNA分子。然后，再将杂种DNA分子输入到受体细胞中去，使之增幅与表达。基因工程又可称重组DNA分子技术。

（一）供体基因的分离与制备

把所要的基因从供体细胞中提取出来或人工合成出来，是基因工程的第一步。从1969年底夏皮罗第一次从大肠杆菌中分离出来乳糖操纵子的部分DNA以来，通过生物学方法、生物化学方法、酶促合成等方法，已分离和人工合成了许多基因。

六十年代中期，别耐斯梯尔（1966）分离出爪蛙的rRNA的基因以后，许多人运用物理化学方法（分子杂交技术、聚赖氨酸法和密度梯度超速离心）进行了分离基因的工作。1972年布朗分离出非洲爪蛙的5sRNA的基因。

1972年三个实验室（巴梯摩尔、斯别戈尔曼、列捷尔）几乎同时利用反向转录酶合成了家兔和人的球蛋白的互补DNA（cDNA），这是第一批在试管中获得真核类基因的开端。以后，这种方法被广泛地应用于研

究基因的结构与功能以及基因工程方面。

1977年春，拉特和古德曼从大白鼠的细胞中分离出胰岛素的mRNA，然后合成其互补的DNA片断，再将此DNA片断同质粒pMBq (Cole与psclol的衍生质粒，3.5—3.9MD，松弛型复制，抗TC，EcoRI及Hind对它有一个切点)的DNA相结合形成杂合分子，当将这个杂合分子引进到大肠杆菌×1776株（不能在实验室外生存）中去时，却并未得到表达。乌德和李也使用在试管中得到的mRNA—cDNA杂合分子使之同质粒组合并引进大肠杆菌中增殖，结果，获得一个球蛋白的基因株系。

在用化学合成的方法人工地合成基因方面，1970年合成了77个碱基对的酵母丙氨酸的结构基因，因缺少其它部件，不能表达。1973年合成了126个碱基对的大肠杆菌酪氨酸tRNA的结构基因，同样没有生物活性。1976年8月科拉纳报告，他们又成功地合成了在细胞中有表达功能的人工基因—大肠杆菌酪氨酸基因。它具有199个碱基对，除酪氨酸tRNA的结构基因部分（126个碱基对）外，还包括启动子部分（52个碱基对）和终止信号部分（21个碱基对）。这项研究成果，不仅对研究生命分子结构与功能有重要意义，而且为基因