

遺傳之秘澤



目 录

遗传工程和质粒	1
真核细胞 mRNA 的提纯.....	6
分离真核细胞基因的新方法	15
化学合成生长激素释放抑制因子的基因 在大肠杆菌中的表达	20
蚕丝基因导入大肠杆菌	31
自然界中不常见的 DNA 重组	40
鸡 DNA 中卵白蛋白基因是分隔开的	41
人绒毛膜生长激素的重组 DNA 的建成与分析	49
有关遗传工程的国际会议一览表 (1972~1978. 3)	57

遗传工程和质粒

B. Mach

«*Experientia*» 33(1): 105~109, 1977.

从广意上来说, 遗传工程不是新词。许多年来, 遗传学家曾通过变异和选择, 修饰了病毒和细菌以及在某些例子中高等生物的基因组。最近, 高等生物细胞的溶合和核移植还使遗传上有差异的细胞株得到创造。但是, 本文所述的遗传工程是指一种新发展的技术。它使得基因, 基因组或一部分基因离体地从一个有机体向另一个有机体转移, 特别是穿过种间屏障的转移成为可能。

这种基因转移是通过 DNA 片段的切割, 和将它们“移植”到另一有机体的 DNA 中的生物化学技术来完成的。这种实验导致在离体情况下产生带有另一物种基因的新基因组(并使这些基因永远存在下去)。这些发展对生物学研究, 特别是对高等生物的基因结构、功能和调节的研究有重要关系。它们对类如大规模生产特异蛋白这样问题可能也具有实际意义。

根据 DNA 的来源和被插入的受体的性质, 应该把这些实验区分为三大类。第一类是把一种原核生物 DNA 转移到另一种原核生物的基因组, 例如, 把噬菌体 DNA 插入到细菌质粒的 DNA。第二类是把一种真核生物的基因转移到一种原核生物受体的 DNA。这种实验可能涉及到例如把哺乳动物基因引入细菌质粒和噬菌体。第三类将是通过病毒载体把基因插入到真核生物的 DNA。

遗传工程技术

当前存在只有发展了新技术才能解决的二类不同问题。一方面是 DNA 片段的处置、切割和将它们重新插入另一个 DNA (基因转移的生物化学); 另一方面是基因特异性包括选择、筛选等问题, 而在复杂基因组情况下, 这是一个特别难以解决的问题。

(一) 基因转移的生物化学

这些实验多数是以通过特异核苷酸碱基对互补实现的 DNA 链的分子杂交原理作为依据而进行的。可以为这些实验制定二项基本原则:

(1) 首先为将要被结合的不同 DNA 分子配备这样一种单链尾部(粘合末端), 即一个 DNA 分子的单链尾部已互补于另一个分子的单链尾部这样的结构。因此, DNA 片段的结合是通过二个 DNA 片段上的短的互补单链段落的配对而实现的。

(2) 参与这种 DNA 结合反应的一方面是能够感染细菌宿主, 并且独立地或者整合到细菌染色体中之后再复制的自身复制的自主 DNA。因而, 参加结合反应的二 DNA 中

的一个是受体并是载体，例如，一个细菌噬菌体或质粒。

因此，DNA 转移实验所指的，是把一个外源 DNA 片段转移到一个噬菌体或质粒载体的环状 DNA。不同 DNA 分子的末端与末端连结涉及两种生化机制：

(甲)均聚尾的连结。可以通过用小牛胸腺制备的一种酶(末端转移酶)聚合一条 DNA 链的 3' 末端的脱氧核苷酸。就双链 DNA 而言，如果每条 DNA 链的 3' 末端是可以容易地达到的话，这种酶同样能聚合核苷酸。在这种反应中，如果只提供四种三磷酸盐中的一种，那末，末端转移酶就能将均聚尾(单链)加在一个 DNA 分子的每一边。例如，一段 λ 噬菌体的 DNA 碎片能接上若干段聚 dA “尾”。独立地，一种无关的 DNA，如病毒 SV40 的 DNA，也能运用相同程序延长，但是采用聚 dT 均聚尾。当将这二种无关的 DNA 分子混合时，聚 dA 和聚 dT 段落配对而含有 SV40 和 λ 噬菌体的 DNA 的环状杂交分子得到形成。因而，这种原理涉及人造“粘性末端”的酶促合成。通常它对任何 DNA 片段的结合都适用。

(乙)通过限制性内切酶产生的“粘性末端”进行结合。限制性 DNA 内切酶是能在准确的位置上切断 DNA 的一组酶。产生特定限制核酶的微生物通过修饰限制酶的这些特异目标部位而保护自己的 DNA。特定限制性内切酶在完全相同的位置上把 DNA 的二根链切断，因而产生“相等的分裂”。但是，另一些限制性内切酶则在相互间隔几个(大约 3—5)核苷酸的不同位置上切断二根 DNA 链。因而这种酶产生了“参差的分裂”。这种类型的切开的主要特征是，由特定的这类限制性内切酶产生的 DNA 片段的末端将具有互补的单链凸出。结果，这种 DNA 片段的末端可通过碱基对互补而相互韧化，并可能形成含有二种不同来源的 DNA 的环状 DNA 分子。非常有趣，不同的限制性内切酶产生不同的单链凸出(不同的核苷顺序)，但是，不管 DNA 的来源如何，特定酶将产生同样的结构，因而，提供了无亲缘关系的种的 DNA 结合的通用途径。

作为受体分子以插入基因使用的载体 DNA，是根据一系列标准选择的。受体分子的 DNA 应能在其宿主细胞中独立复制。可能用裸露的 DNA “感染”宿主细胞(转化，转染)。环状载体 DNA 在一点上切裂后即应成一条直线。外源 DNA 片段的插入不应使载体分子丧失在其宿主细胞中复制的能力。最后，从实际的目的出发，载体 DNA 应带有一个基因，其产物将授于被感染细胞一个特别的表型，如对一种抗菌素的抗性。这种依赖载体的表型，将有助于对载体分子已在其中复制的宿主细胞菌落进行选择 and 筛选。因为很明显的原因，在过去几年中，基因转移的实验极大地依赖质粒 DNA 和噬菌体 DNA 作为载体分子。

质粒 这些是环状 DNA 分子，只要选择适当，能满足上述对基因转移载体的所有要求。到目前为止，微生物学家主要利用质粒作为抗菌素抗性的基因的载体，特别是它们把这些抗性传输到其他细菌中去的能力。质粒的这些微生物特性使得它们成为遗传工程领域中非常有用的工具。为了这一目的，选择了一些质粒，它们在不影响抗菌素抗性的基因情况下被特定限制性内切酶在一点上切裂。已经证明，甚至引入非常大的外源 DNA 片段，接着使质粒分子重新环化，并不影响杂交质粒在宿主细胞中的复制和生长。对采用质粒 DNA 作为转移实验的载体感兴趣的另一原因是在某种情况下，只要抑制宿主细胞 DNA 的复制，就可使质粒在宿主细胞环境中扩增。这样就能够进行每个细菌含有几千个质粒 DNA 分子的细菌培养。一些实验室，特别是 Cohen 实验室，正在积极进行适合基因转移实验的质粒的选择和构造。

噬菌体 已证明 λ 噬菌体是一项极好的载体分子，主要是因为它适用于转染实验。

由于众所周知的 λ 的遗传学,所以能设计一些可能的选择和筛选程序。与质粒 DNA 情况相反,噬菌体可容纳的外源 DNA 的大小有严格限制。 λ DNA 的大小大约是三千万道尔顿,大于正常值 10% 的 λ DNA 杂交体通常是不能生存的。但是,由于大部分 λ DNA 本身对生长是无关紧要的,所以,设计了一些方法采用缺失的 λ 分子,例如,仅含有 3/4 λ DNA 质量的分子进行试验以便插入更大的外源 DNA 片段。

DNA 转移实验具体步骤总结如下:

(1) DNA 片段(待插入的外源 DNA)和载体 DNA 的制备。这包含制备均聚物尾(末端转移酶)或者由限制性内切酶产生的参差末端形成的粘性末端。

(2) 由于粘性末端的互补性,载体 DNA 和外源 DNA 相互钝化。

(3) 用聚核苷酸连接酶(从 T₄ 噬菌体感染的大肠杆菌中制备的)进行连接并形成由载体分子和外源 DNA 片段组成的共价环。

(4) 将杂交体或者嵌合的载体引入其宿主细胞,从而复制,甚至可能扩增外源基因。

(二)选择性和特异性

一个重要的选择问题是在二种水平上发生的。首先,当用质粒 DNA 或噬菌体 DNA 作为载体进行 DNA 插入实验时,产生的分子包括重组 DNA(含有外源 DNA 片段)和已经成环而没有外源 DNA 的质粒或噬菌体。因此,发展了一些重组 DNA 分子的选择或筛选技术。当用适当的均聚物尾预先延伸的 DNA 进行连接反应时,运载体分子不能自行成环,而只有通过含有互补同聚尾外源 DNA 片段配对时,才能这样做。因此在这种有利情况下,产生的生存质粒(或噬菌体)都是带有一个外源 DNA 片段的重组体。另一方面,当使用限制性内切酶产生的粘性末端进行连接时,则必须对重组体进行选择或筛选。例如,如果开始的 λ DNA 分子太小以致于不能产生能生存的噬菌体时,那么就能获得 λ 噬菌体重组体的选择。DNA 片段的插入将使 DNA 恢复适当大小与 λ 噬菌体生存能力一致。如果外源 DNA 片段的插入改变了受体质粒或噬菌体中的特定表型,那么就可以筛选到重组体分子。现在已发展了几种这样的系统,特别就 λ 噬菌体来说,甚至在成千噬菌斑或菌落中也能容易地鉴定重组体。

更重要和更困难的是一个给定基因的一定重组体的特异性的总问题。这是遗传工程的一个关键方面,而在复杂基因组和特定 DNA 片段只占总 DNA 的百万分之一不到的高等生物中则特别难以解决。到目前为止,设想了四种不同的方法以解决 DNA 转移实验中基因特异性这一问题。

(1) 在一些罕见的实例中,一个基因或一组基因可能有产生特别的物理化学行为的总的化学成份。在不同密度的基础上,核蛋白体 RNA 的基因和组蛋白基因的异常鸟苷胞嘧啶含量使得人们可以根据密度的不同把他们从大量 DNA 中富集出来。但是,在密度不同的基础上,纯化基因不是一个一般适用的原理。

(2) 解决特异性问题的一个方法是大量的 DNA 片段混合物,如用完整基因组不分部的 DNA 的无规则地插入。然后,让所有重组体在细菌中生长和选择或筛选带有特异基因的重组体。这种所谓的‘散弹枪’方法是否可行,极大地依赖于是否具有选择和筛选技术。理论上说,如果一个新插入的基因在一个产物中表达,从而允许在重组体中进行选择,这

样选择就能够得到实现。筛选的方法主要是，通过通常是固定在硝化纤维素膜滤器上的 DNA 的分子杂交，鉴定带有特定基因的菌落。这些技术适合于成千个重组体的筛选。但是，就复杂的 DNA 有机体而言，由于需要在几百万重组体中筛选，用含有未分部的 DNA 进行基因插入看来是一种不切实际的方法。

(3) 更先进和有希望的方法是特定基因的纯化(或浓缩)，接着有意识地将特定基因插入到合适的受体质粒或噬菌体。特异基因的纯化方法大都依靠与特异探针(信使 RNA 或互补 DNA)杂交的技术。完全变性的 DNA 片段或部分变性的 DNA 都可被用以进行杂交纯化。这一重要研究领域的进展在极大程度上取决于是否具有大量纯核酸探针，很可能一些带特异基因和用另一种方法(见下面 4)制备的质粒将证明在这方面是非常有用的。

(4) 信使 RNA 指导的特异基因顺序的插入。如果能用编码特定蛋白质的信使 RNA 作模板，合成双链 DNA，那么就有可能选择地插入那个特定蛋白质的基因顺序。根据这一原理进行的实验将在下面一节中详细描述。

(三) 信使 RNA 指导的特异基因顺序的插入

已经知道，信使 RNA 能作为合成单链互补 DNA 的样板使用。最近观察到，互补 DNA 一旦分离，它就能被用作为以大肠杆菌的 DNA 聚合酶，或在特殊情况下，用鸟成髓细胞血症病毒的 DNA 聚合酶合成第二

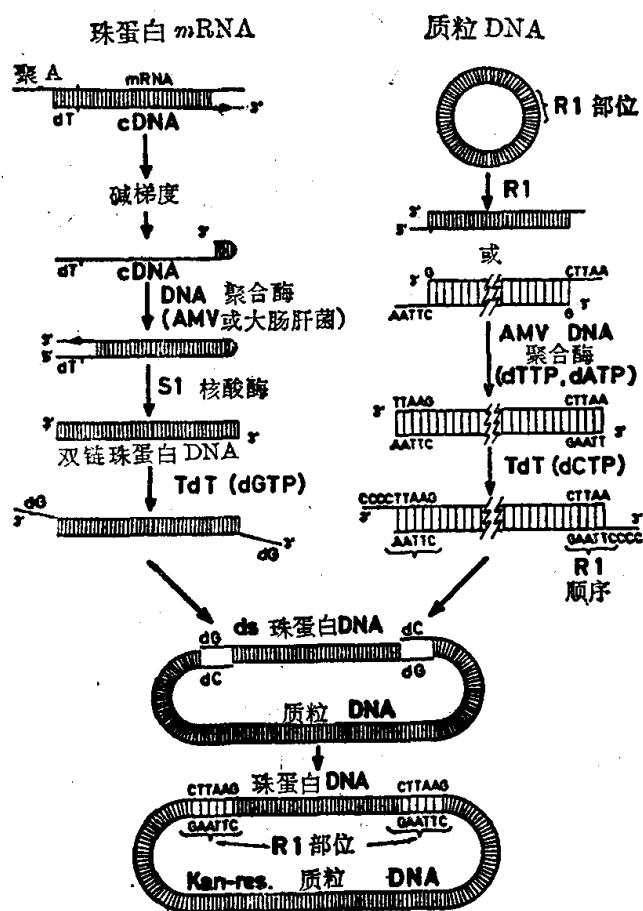


图 信使 RNA 指导的珠蛋白基因顺序插入大肠杆菌质粒的说明图。

AMV—髓细胞血症病毒
Kan-res—抗卡那霉素

根 DNA 链的样板。因此，从信使 RNA 开始获得同样顺序的双链 DNA 分子是有可能的。图以兔珠蛋白信使 RNA 为实例说明这种实验。珠蛋白 DNA 可用一条聚 dG 均聚尾延伸，并分别为载体 DNA 分子制备一条聚 dC 均聚尾。这种实验挑选的载体是大肠杆菌质粒 pCR1，它能在一点上由酶 EcoR1 打开(转换成线形)。这种质粒带有一个抵抗抗菌素卡那霉素的基因。因此，pCR1 成功地感染的细菌可以作为抗卡那霉素菌落而被选择。

如图所示，通过杂交产生了分别带有聚 dG 和聚 dC 聚合尾的珠蛋白特异 DNA 和线形质粒 DNA，并且把这些杂交分子直接用来使大肠杆菌细胞转化成抗卡那霉素。这些实验证明不需聚核苷酸连接酶就可获得共价环。DNA 是从抗卡那霉素菌落中制备，并且对珠蛋白特异顺序的存在进行了试验。这些试验是通过用由兔珠蛋白信使 RNA 合成的具有 α -或 β -珠蛋白特异性的放射性互补 RNA 作探针在微

孔滤膜或在液体中进行杂交而完成的。因而，获得大量含有 α -或 β -珠蛋白基因顺序的细菌无性繁殖系，并且用其中一部分制备了质粒 DNA。这些珠蛋白特异的嵌合体质粒的详细研究示出，由于进行了无性繁殖系实验，它们或者含有 α -或者含有 β -珠蛋白基因顺序，但从来不同时包含二者。珠蛋白特异插入顺序的长度大约是 400~600 碱基对。特别就 β -珠蛋白特异质粒来说，获得了带有珠蛋白的全部结构基因以及在结构顺序二边的大部分不翻译地区的无性繁殖系。最后，观察到当用限制性内切酶 EcoR1 消化特定珠蛋白特异质粒时，珠蛋白特异基因顺序能被切下来，并且易于用凝胶电泳纯化。这一最后实验说明，如图所预示的，在插入珠蛋白顺序二边确实重建了 EcoR1 特异部位。

上述就免珠蛋白基因而发展的方法使我们能够制备带有小鼠 α -或 β -珠蛋白基因和带有小鼠免疫球蛋白轻链基因顺序的重组质粒。原则上说，该程序可以应用于一切可以获得足够纯度的信使 RNA。将证明这些对珠蛋白特异的或对免疫球蛋白特异的质粒是非常有用的工具，特别是在各种生物学条件下对基因特异顺序(信使 RNA 或 DNA)进行试验，以及从细胞 DNA 中纯化(通过杂交)特异基因。

遗传工程的前途和可能的用途

由于采用本文在前面概括的生物化学和生物学技术以及其他技术，近几年已能够将一个噬菌体或质粒的 DNA 转移到另一个噬菌体或质粒。然后，高等生物的 DNA 如编码爪蟾核糖体 RNA 的基因被转移到在大肠杆菌中生长的质粒。随后又对编码爪蟾组蛋白的基因进行了同样实验。就哺乳动物基因来说，曾把兔和小鼠珠蛋白以及小鼠免疫珠蛋白的基因顺序插入质粒 DNA，但是，这些实验限于由信使 RNA 编码的部分。

按从这些不同实验获得的有趣观察之一，虽然自然界总是设计了非常严格的障碍防止不同的种类之间交换遗传信息，但人们发现，不管 DNA 的来源是什么，DNA 分子在离体的情况下对交换非常敏感。在离体情况下，DNA 对其来源的这种完全“中性”与进化过程中保持的严格种属特异性是显著不同的。

遗传工程实验在两种明显的水平可能证明是有用的：

(1) 细菌载体中特异基因的无性繁殖大概是研究基因结构、基因调节和基因表达的不同生物学领域的一项重要进展。因而，这种新技术代表了对现代生物学研究的一个决定性工具和重要贡献。

(2) 对重组 DNA 分子明显感兴趣在于表达外源基因的可能性。除了这一现象的生物学意义以外，它提出了利用重组质粒或噬菌体大规模生产特异蛋白质的可能性。目前，遗传工程的这个第二方面只是假设。但是，它已成为广泛研究的课题。现已可设想到培养能够合成大量这样的蛋白质，如胰岛素原或病毒抗体(作为疫苗用)以及一系列具有药物学意义的蛋白质的可能。甚至能想象细菌培养中，具有抗体活性的特异免疫珠蛋白分子的重链合成的可能。

(3) 遗传工程的第三个方面涉及更加假设性的可能，即把特异基因从病毒载体转移到高等生物，包括人的染色体。由于现已具有把 DNA 片段引入环状 DNA 分子(如 SV40 DNA)的技术，这一领域的工作正集中在发展能将它们的 DNA 整合入宿主细胞基因组的安全病毒的载体。如果这一类 DNA 的转移实验涉及体细胞，这样的基因 (下转第 39 页)

真核细胞 mRNA 的提纯

佐野 浩

《蛋白质 核酸 酵素》 22(12):1303~1311,1977.

在基因表达的研究中,手头如有 mRNA(信使 RNA)则在许多方面都是有利的。首先是能运用 DNA-RNA 的杂交法以鉴定专一的基因。其次是能够由 mRNA 合成互补的 DNA(cDNA),再经过无性繁殖而使其大量扩增。还有,以 cDNA 作为亲和层析的底物或许能提取全部基因。着眼于此类应用,本文以真核细胞 mRNA 的提纯法为中心,对迄今所知各种 mRNA 的饶有兴趣的方面予以阐述。

一、引言

人染色体(单倍体)中有 2.8 微微克的 DNA,相当于 2.8×10^9 个碱基对。假设一个基因由 1500 个碱基对组成,那么每个单倍体就有 2×10^6 个基因。而根据遗传学的分析认为,每个单倍体约有 50,000 个基因,这只不过是全部 DNA 的 2.5%。这 50,000 个基因中,在发育、分化的各阶段中被活化的充其量不过 10%,余下的 90% 则处于非活性的即所谓关闭的状态。在完成分化的细胞中更是只有极少的基因在活动,例如在浆细胞中除了为抗体编码的基因之外,几乎全都是非活性的。使大量基因沉默,或者说只让特定的基因表达的机制究竟若何?

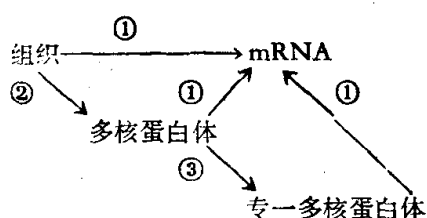
在基因表达即 mRNA 生成之中,从基因方面观察似有两种机制:被动的与主动的。前者由外部因子决定,例如由细胞质内的种种因子及糖类等环境因子予以调节,或由激素激活特殊基因等等。至于后者,除基因扩增之外,也有人认为基因的转录起始部位在特定的时间给出“开”或“关”的讯号,还就此提出了各种模型。真核细胞的 DNA 内有不少“空位”区域(不是基因的部份)分散在基因之间,但间距极小,每隔 10 个左右的碱基顺序便出现一个。有一种说法颇有吸引力,即认为在一次细胞分裂中一个“重复单位”的某个碱基被酶修饰(如胞嘧啶的甲基化),那么,只在细胞分裂数与重复单位数相符时,某些基因才会处于开(或关)的状态,可以说,碱基顺序的变化起着钟表的作用。不过,虽说有种种实验报告但却尚未发表明确的数据。但无论如何,毋庸置疑的是对基因表达的各方面进行的这种实验研究,其先决条件或许便是首先抽提特定的基因。

真核细胞的基因除极少数例外几乎都是单拷贝的(每个单倍体一个基因)。以人为例,所谓抽提某种特定的基因也就是说专一地抽提 $1/50,000$ 的基因即基因总数的 0.002%,DNA 总数的 0.00005%,这种可能性是微乎其微的。而最近开创的基因无性繁殖技术使这成为可能的了。将特定的 DNA 片段插入质粒并以此引入大肠杆菌之类的细菌中,使目的 DNA 大量扩增,其方法有两种。其一是用限制性内切酶将含有目的基因的 DNA 切断后插入质粒,经无性繁殖后再筛选出含有目的基因的菌落。常用的筛选法是使与基因的转录产物即

mRNA 杂交——菌落杂交。其二，从 mRNA 合成双链 DNA，使其插入质粒。但不论采用何法，mRNA 都是不可缺的。因此本文在以基因分离为主要目的的前提下，按照迄今所知事例对 mRNA 的提纯法加以阐述，而关于无性繁殖本身以及关于 mRNA 的一般问题，则请分别参阅本刊 5 月号专集《基因操作研究》、Brawerman 著的综述。

二、mRNA 的提纯法

mRNA 被定名为“信使 RNA”是在 1961 年。其后虽然使用细菌的系统而确认了各种 mRNA 的存在，但由于绝对量少、难以与其他 mRNA 区别、结构脆弱等不利因素凑在一起而致使提纯极为困难。可是由于应用了生成单一蛋白的组织系统，在真核细胞方面，兔网状红血球的珠蛋白 mRNA 的抽提于 1967 年宣告成功。此后随着技术水平的提高，至今已能分离出大约 20 多种 mRNA。但其大部均系结构蛋白与分泌蛋白的 mRNA，明显地表明所选用的材料是只能合成单一蛋白的系统(参阅表 2)。从这些系统中抽提 mRNA 较为容易，即可从组织直接分离，又可在得到多核蛋白体后立即按常规提纯。在分离出组织中含量极少的 mRNA 方面，新近开创的方法与此不同，采用了抗原抗体反应法。作成目的 mRNA 的产物即蛋白质的抗体，将它加入到含有从组织获得的全部多核蛋白体溶液中，利用抗原抗体反应而只专一地分离出含有目的 mRNA 的多核蛋白体。这些方法的关系如下：



其中①表示利用 RNA 的物理化学性质进行提纯的过程；②表示多核蛋白体的抽提过程；③表示从全部多核蛋白体中专一地分离目的多核蛋白体的过程。以下就此扼要叙述。

1. 多核蛋白体

在真核细胞的 mRNA 方面，首先得到的是为珠蛋白编码的 mRNA，这可说是由于材料的选择而带来的一大成功。网状红血球的 mRNA 具有许多优点：几乎全是为珠蛋白编码的单一品种，容易从血液中粗抽提，最重要的是其中全无核酸酶，等等。但这种理想的系统不可能很多，大部分组织不仅含有较多的核酸酶，还有多余的 DNA、RNA 之类，不得不为此采取相应的措施。特别成问题的是，从组织的破碎到多核蛋白体被顺利沉淀为止，如何避开核酸酶的袭击。肝素、聚硫酸乙烯酯是有效的核酸酶抑制剂。肝素尤佳，只须在抽提液中加入 100 微克/毫升左右即能使 RNA 酶活性下降 20~40 倍，所以现已广为应用。另外有报道说，焦碳酸二乙酯由于能使氨基酸的氨基与咪唑基烷化而使 RNA 酶失活，但 mRNA 也被烷化而失掉翻译活性。另一方面也有报道说这是能顺利提纯的，不过要依赖于反应的条件、浓度及 mRNA 的种类，所以在使用前须先确定各种条件才行。

总之，在动物细胞方面如从胚胎、肝脏、癌细胞等(参阅表 2)，在植物方面如从玉蜀黍、菜豆、豌豆、衣藻等等迄今都已得到各种多核蛋白体。其方法一般是先将组织急速冷冻成粉末，加抽提液使成匀浆。抽提液由 50mM 左右的中性缓冲液、5mM 氯化镁、0.1~0.5 毫克/毫升肝素、1%左右的 TritonX-100 之类可溶剂以及少许蔗糖组成。植物组织的核酸酶活性不很高，不用肝素也能作充分抽提，而且在很多场合用比 TritonX-100 还要温和的 Nonidet P40 等作为可溶剂。在适当地除去细胞碎片后，将匀浆放在 2M 左右的蔗糖垫上，经 10 万 g 离心 3 小时左右即得多核蛋白体沉淀物。也有报告说，成匀浆后加 0.1M

氯化镁使多核蛋白体沉淀则得率较高。为了观察多核蛋白体的沉降图谱,可使匀浆在10~40%蔗糖密度梯度中离心,测定各分层的吸光度。这个实验最好是先在各种条件下制成匀浆,观察多核蛋白体的图谱,确定最佳抽提条件。图1表示在两种盐浓度下抽提的多核蛋白体的图谱。

这样得到的多核蛋白体,按理可不分动物植物就进入提纯阶段,但如核酸酶活性过强而连多核蛋白体都无法顺利抽提,或只能得到hnRNA(mRNA的前体),则采取从组织直接抽提RNA的方法。前者的例子是从胰腺抽提胰岛素mRNA的试验。用含有SDS-苯酚的抽提液直接处理已成冷冻粉末的组织,用寡聚(dT)纤维素从所得的全部RNA中分离出含有多聚腺苷酸聚(A)的RNA(后述)。但只为胰岛素编码的mRNA的分离未获成功。后者可举从红血球样细胞抽提珠蛋白mRNA的前体为例。用SDS(十二烷基硫酸钠)、肝素等与蛋白酶K一起处理含细胞的溶液,用苯酚抽提RNA,获得在15S出现峰值的mRNA的前体。在SDS存在下用蛋白酶K的方法确实有效,今后可能大量使用。

2. mRNA

自从1970年发现牛痘病毒mRNA的3'末端上有聚(A)连接以来,今已阐明至少真核细胞的mRNA几乎都有聚(A)。连过去一直被认为无聚(A)的组蛋白mRNA上最近也发现有聚(A),因此可以认为聚(A)是真核细胞mRNA的固有属性。所以从多核蛋白体的全部RNA中分离聚(A)⁺RNA就等于只分离mRNA,用寡聚(dT)纤维素作亲和层析是最适合这个方法而被广泛采用。1972年用此法提纯的照例还是珠蛋白mRNA。

为了从多核蛋白体中分离出具有这种性质的mRNA,首先用酚只将RNA分离出来,其次用寡聚(dT)纤维素(或多尿苷酸-Sephrose)得到聚(A)⁺RNA,这已成一般规律。酚抽提时,为使蛋白质与RNA分离完全,须在各种条件下检验。如只用水合了的酚,聚(A)⁺RNA会被吸附在蛋白相上。据说在这种情况下或用pH9抽提,或按1:1将氯仿加入酚中,则效果良好。此外还有加少量异戊醇以提高分离效率的,有加甲酚以作抗冻剂的,有加羟基喹啉以防酚氧化的种种方法。残留于水层中的RNA可用乙醇沉淀,方法简单,效果亦佳。用寡聚(dT)纤维素作层析分离的操作方法极简单,使在高盐浓度中(0.2~0.5M氯化钾或氯化钠)的全部RNA溶液通过柱,以使聚(A)与寡聚(dT)结合,冲洗掉所有杂质后,用低盐浓度(10mM Tris-HCl)使其解离而溶出。但用这样简单的方法不能分离对mRNA亲和力强的rRNA,而且碱基数在20以下的聚(A)又不能结合。作为必要措施,提出了如在层析分离前在60°C加热数分钟^[27],或在缓冲液中加SDS^[28],或在较

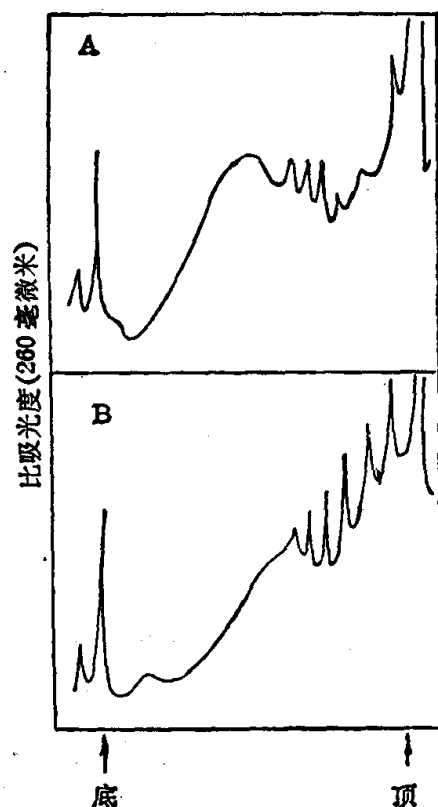


图1 从衣藻抽提的多核蛋白体的蔗糖密度梯度离心分析绿藻的一种——衣藻经液体培养,所得之植物体用Nonidet-P40平稳溶化,除去细胞碎片后置于10~40%的蔗糖密度梯度中,以40,000转/分钟离心75分钟(50mM Tris-HCl, pH7.8, 5mM氯化镁、2.5mM β -巯基乙醇,氯化钾浓度: A为100mM, B为25mM)。高盐浓度显示多核蛋白体解离较少。(W. G. Burton, H. Sano 原图)

低的盐浓度下使在 0°C 结合再于室温中溶出^[29]等等改进方法。聚(U)-Sepharose 原理相同,但用 70%以上的甲酰胺溶出^[30]。采用何法为佳,这是各人爱好的问题,虽然用寡聚(dT)纤维素的人较多,但也有报导用聚(U)-Sepharose 得率较高^[30]。

如此所得的聚(A)⁺RNA 再用蔗糖密度梯度离心及聚丙烯酰胺凝胶电泳等方法提纯至分子量单一的 RNA 组分。这种方法已大体确立,可参阅合适的实验书。

3. 双重抗原抗体法

自 1960 年以来,已经知道某种蛋白的抗体对多核蛋白体上未完成的多肽链也有良好的作用。但当时的兴趣集中在核蛋白体上的蛋白合成,曾用这一反应将专一的蛋白质从全蛋白合成系统中分离出来^[31-33]。

1961 年, Cowie 等人就已说明了用双重抗原抗体反应(后述)能有效地分离合成的 β -半乳糖苷酶的多核蛋白体,并指出将来可能成为分离“专一的核蛋白体”的有效方法^[31]。不过, mRNA 这个概念的形成恰恰也在 1961 年,当时并未能同 mRNA 的提纯联系起来,从进入 70 年代以来直到 Schimke 等人的研究组积极应用这一反应以分离卵清蛋白 mRNA 之前的约十年之间,谁也没有注意到它的实用价值^[34,35]。

从原理上说,这种反应就是把目的蛋白质的抗体加进从组织抽提的多核蛋白体溶液中使其专一地结合在多核蛋白体的未完成的多肽链上,用适当的方法予以分离,但由于这种复合体不沉淀、非专一的多核蛋白体结合较多等问题的产生,初期的实验不能顺利进行^[36]。其后研究出新方法,即对已结合在多核蛋白体上的抗体(第一次抗体)再作抗体反应(第二次抗体),可从反应溶液中简便地分离出沉淀物。从这种沉淀物中提纯的 RNA 当然是为目的蛋白质编码的 mRNA 与 rRNA 的混合物。事实上若选取适当的条件,此法一步就能提纯到 90% 以上的纯 mRNA,现在称作双重抗体抗原反应法(double immuno precipitation),已普遍使用。此外还开创了将第二次抗体结合在纤维素等基质上、根据亲和层析原理有效地分离目的多核蛋白体的方法^[38]。原理上虽如上述,而在实验中却要花不少时间、精力作抗原抗体的准备工作等等。以下就实际问题扼要介绍,双重抗原抗体法的概况如图 2 所示。

A. 抗原及抗体的准备

不言而喻,抗原就是目的 mRNA 所编码的蛋白,越纯越好。如果混有杂蛋白则以后会使与之相当的抗体脱落。抗原的量虽因使用的动物而异,但若用的是兔则最低量需 2 毫克。分 2~3 次给与,并用沉淀试验测试抗体的产生状况。快则四个星期,慢则十个星期左右便形成抗体。详情可参阅合适的实验书。如沉淀试验为阳性、血清也能抽提,就必须提纯的抗体即免疫球蛋白。粗血清虽也产生反应,但由于会得到非专一的多核蛋白体、或核酸酶活力强、或与第二次抗体反应的效率差而不理想。粗血清用 40% 饱和硫酸铵反复处理,大致可得球蛋白组分,再用 CM-DEAE 纤维素的复合层析法将 γ -球蛋白从其他蛋白及核酸酶中分离出来。所得到的组分再通过与原先的抗原蛋白结合的 Sepharose 的亲和层析法分离出抗体。至于基质,用 Cr-Br 处理过的 Sepharose 4B 操作简易、结合率高。抗体一般是在降低 pH 时溶出而得,但也可用高浓度的氯化镁。用氯化镁得率较高,但如果抗原方面的蛋白是由几个亚基构成的,那么不仅抗体解离,亚基也被解离,故须注意。一般用盐酸及醋酸,但用 pH2.8 的甘氨酸缓冲液也足以达到目的^[19,35]。第二次抗体也用同法制得。但用沉淀法所需量为第一次抗体用量的 40 倍以上,最好从大动物取得。不过近

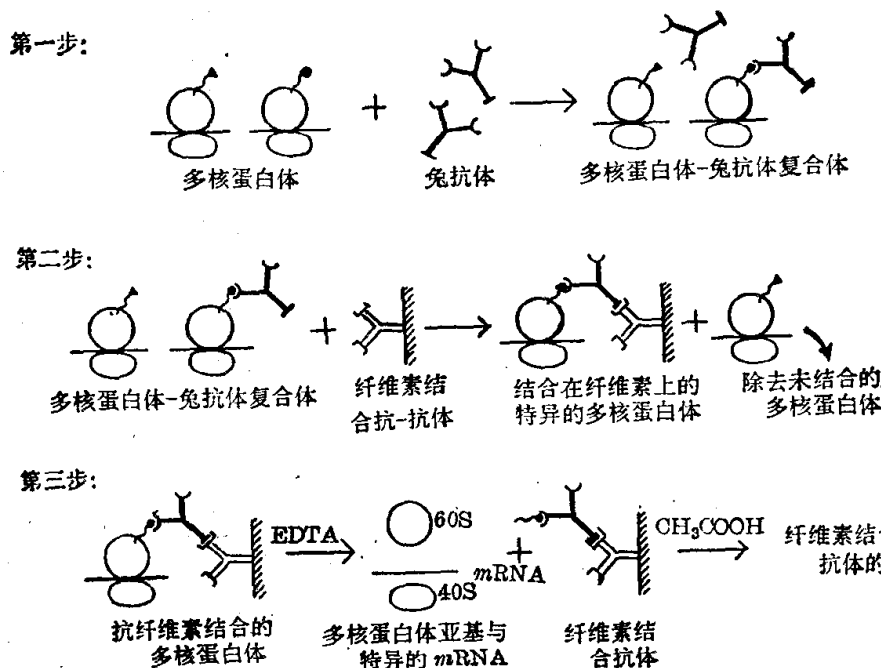


图 2 用双重抗原抗体反应分离专一的多核蛋白体在第二步中一般也可用如下方法: 不使第二次抗体结合在纤维素上而直接加入到多核蛋白体-抗体复合体, 从反应溶液中取出沉淀物。在这种场合, 由于沉淀物不能被 EDTA 解离, 所以必须用 SDS-酚等直接抽提 RNA。

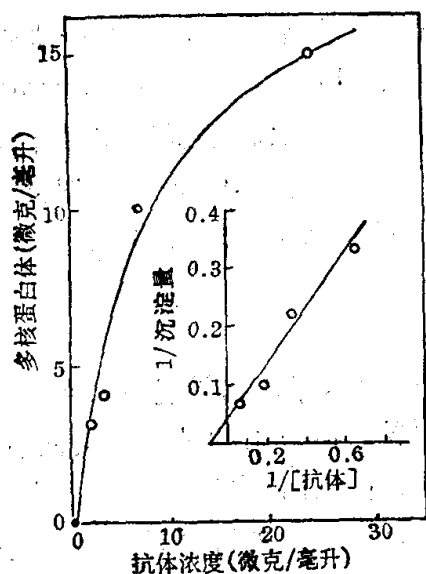


图 3 依赖于抗体浓度的多核蛋白体沉淀试验
在体内用氘标记的衣藻多核蛋白体溶液中加入各种浓度的抗-核酮糖-2-磷酸羧化酶 (LS) 后, 再加 1.8 毫克/毫升的第二次抗体, 检查沉淀的多核蛋白体的量。从倒数作图计算出这个反应系中羧化酶使多核蛋白体一半沉淀所需抗体量 (K_m) 和理论上可能的最大沉淀量 (V_{max}) 各为 10 微克/毫升及 22 微克/毫升。

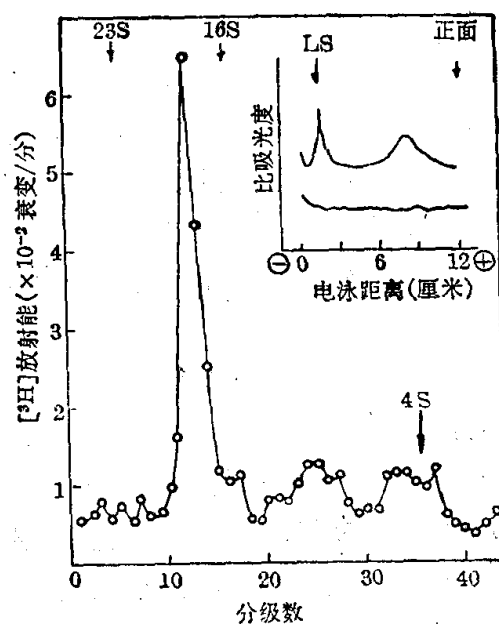


图 4 用双重抗体抗原法精制的核酮糖-2-磷酸羧化酶 LS-mRNA 的蔗糖密度梯度离心的分析

将表 1 所示之提纯的 mRNA 置于含 SDS 的 10~20% 蔗糖中, 以 49,000rpm 离心 4 小时。mRNA 于 16S 处表现为单一的峰值 (箭头表示作为标记物的 tRNA 与 rRNA)。插入图表示将如此提纯的 mRNA 给与体外蛋白合成系中, 用抗原抗体反应抽提产物, 用 10% 聚丙烯酰胺凝胶-SDS 电泳进行分析的结果。a 系未给与 mRNA 的对照组; b 系给与 3mRNA 的。LS 之外产生了相当多的未完成的多肽链。

来已成商品，容易买到。例如第一次抗体若是用兔制得，则用山羊·抗兔·免疫球蛋白为好。市场上出售的均系用硫酸铵及 CM-DEAE 纤维素处理过的。

B. 多核蛋白体的分离

多核蛋白体与第一次抗体的反应极为简单，只须将两者混合，在 0°C 放置 30 分钟左右即可。浓度比例虽按目的多核蛋白体的多少而定，但一般是对全部多核蛋白体的一个光密度单位只需 10 微克左右的抗体即够。不过在实验之前最好改变两者的浓度而找出最佳条件，举例如图 3。其次加以 40 倍于第一次抗体以上的第二次抗体，同样在 0°C 放置 30 分钟。根据是否立即产生肉眼可见的混浊以判明第二次抗体加入之后反应的好坏。沉淀物通过 1M 蔗糖垫后离心收集之，使溶于含 SDS 的缓冲液中。如在 60°C 加热数分钟则易于溶解，但也会引起 RNA 的分解，必须注意。将第二次抗体结合于 PAB·纤维素作为亲和层析的基质以分离多核蛋白体和第一次抗体复合物的方法也有效。详见 Schutz 等人的报导^[38]。无论用何法所得之专一的多核蛋白体均须用酚处理，按前面二、2 所述方法提纯 mRNA。图 4 所示即为用此法提纯 mRNA 之例。

三、体外蛋白合成

如此得到的 mRNA 是否真正是目的 mRNA 还须确定。最普通的方法是对用体外蛋白合成系所得产物进行分析。使用的合成系有 Krebs I 腹水瘤^[40]、网状红血球^[41]、麦胚等^[42]。但据说前二者有翻译效率差、有非专一的掺入等缺点，目前一般都用麦胚系统^[43]。产物也可就原样进行电泳分析，但如利用其抗体，在其从全部反应溶液中分离后立即进行分析则灵敏度较高。反应虽与分离多核蛋白体的方法基本相同，但大多加少量抗原蛋白作载体并加以与标记氨基酸相当的非标记氨基酸。如此所得产物先用 SDS-聚丙烯酰胺凝胶电泳从分子量方面检验其是否为目的蛋白质。但有时目的蛋白并不一定明确显示峰值，有时由于反应条件所决定也会只得到未完成的肽链(参阅图 4 的插入图)。在这种场合，可用胰蛋白酶将产物与原有蛋白消化，如离子交换层析的溶出图谱显示二者一致，则首先是个有力的

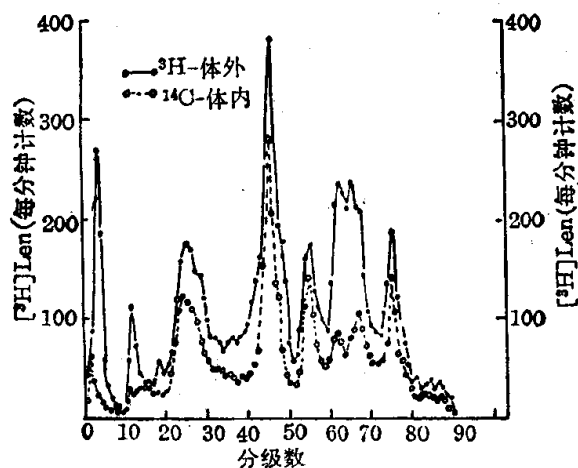


图 5 体外合成的蛋白的肽链分析
以免疫球蛋白(L链)mRNA 为模板所得的蛋白质(·)与体内标记的 IgG (L链)(○)同时用胰蛋白酶消化后，用 Technicon-P 柱分离为各个肽链进行比较。从大体一致的情况可以证明所得的 mRNA 即为 IgG(L链)编码的 mRNA。

的证据(图 5)。此外，产物若是酶，则可检测其活性；若是诱发的 mRNA，则在非诱发条件下产生的 mRNA 表示其蛋白合成活性低，这也是很好的证明。如手头没有电泳装置与观察胰蛋白酶消化的装置，则可在蛋白合成系中加入原有蛋白，单单观察它与抗体的竞争性沉淀作用亦能得到大致的证明。在目的蛋白质的氨基酸组成已知的场合，既可比较体外各种氨基酸的掺入量，也可直接间接地测定 mRNA 的顺序，看有否相当于氨基酸顺序的碱基顺序存在。后者不只是个证明，它本身就是一项独立的研究，胰岛素基因与 mRNA 就是用此法鉴定的^[46]。总之，用这种种方法都能确定是否为 mRNA，不过这不限于最终提纯了的样品，

最好是对各个提纯阶段的粗抽提液都同样观察活性。表 1 的数值虽不能说是最佳者，但它给出了平均的效率*。

表 1 核酮糖磷酸羧化酶 LS-mRNA 的提纯

	RNA 量 (微克)	%	活 性 (单位) %		每 微 克 RNA 的活性	
全部多核蛋白体 RNA	1800	100	720	100	0.4	
由抗体作用所得全部沉淀 RNA	145	8	430	60	3.0	
寡聚 (dT) 纤维素	非吸附 RNA	135	7.5	30	4	0.2
	吸附 RNA	7	0.3	73	10	10.4

LS: large subunit, 大的亚基

1 个单位是指每小时 [³H] 标记亮氨酸的掺入量是 10⁻¹⁴ 克分子

四、迄今已分离的 mRNA 的种类及其应用

自 1967 年获得珠蛋白 mRNA 以来至今 10 年之间，又分离出 20 余种 mRNA (表 2)。同这 10 年之久的时间以及世界上星罗棋布的研究室数目相比，其为数之少、多样性之差，实在令人惊叹。这就不能不充分认识到是由于专一分离 mRNA 的把握不大、抽提技术不高明而造成的。还必须补充一点的是，在各种 mRNA 的提纯方面，所付出的劳力是非同寻常的。以下试按表 2 的顺序就各种 mRNA 的特性、应用等饶有兴趣的方面予以说明。

清蛋白被选作模型以便确立双重抗原抗体反应法，基本技术方法完备^[30]。最近得到了 cDNA^[64]。酪蛋白是牛奶的组成蛋白质，有几个种类，它既由各种催乳激素之类的激素诱发，又显示出黄体酮抗催乳激素的作用。现已知道这些激素是在转录水平上起作用的，为研究其作用机制，已抽提了 mRNA^[47]。δ- 晶体蛋白是占发育初期的晶状体 80% 的蛋白质，而在发育后期减少到 30%，因此是研究分化时基因表达形式的有效材料^[49]。丝心蛋白是丝绸的结构蛋白，它的 mRNA 早就分离出来了，每个基因组只有一个基因 (single copy)^[65]。珠蛋白是在网状红血球中合成的唯一的蛋白质，它构成血红蛋白，有 α 链与 β 链，其 mRNA 是最先从真核细胞中分离出的 mRNA，据说全世界有 20 个以上的研究室竞相使用之作为基因无性繁殖的材料，最近对从无性繁殖出来的 cDNA 确定了全部碱基组成^[66]。组蛋白 mRNA 也早就被获取作为 9~10S 的 RNA，但分离出 5 种 mRNA 则是最近的事^[52]，知其每个基因组有 10 个 H5 的基因^[53]。

至于免疫球蛋白，有关它的 V-C 区域的基因表达方面的研究甚广，可参阅合适的参考书^[2]。角蛋白是鸟类羽毛的结构蛋白，在发育的第 12 天出现，第 14 天它的合成达到高峰。利用 cDNA 探明了鸡(毛)的每个基因组有 100~240 个基因。α 乳清蛋白与酪蛋白一样，是牛奶的组成蛋白，占全蛋白的 3% 左右，是乳糖合成酶 (E.C.2.4.1.C) 的亚基^[68]。催乳激素与绒毛膜促性腺激素一起组成胎盘激素，前者多在妊娠前期合成，而后者在妊娠后期合成，是研究不同生理条件下激素产生的合适材料。有报导说^[58]，cDNA 已被合成，

* 注：此处原文是 balance sheet, 今改译成提纯效率。——译者。

每个单倍体有两个基因。肌球蛋白是肌肉的结构蛋白，由两个重链(H链)与三个轻链(L链)组成。原以为H链 mRNA 欠缺聚(A)，最近表明它也是有聚(A)的。卵清蛋白占卵管蛋白的一半以上，它的合成常常需要雌激素。所以它是研究激素作用机制与细胞分化的好材料，有几个研究室早就在尝试 mRNA 的分离了。最近 cDNA 也已合成，并有两个实验室报导了它的无性繁殖^[69,70]。胶原的基本结构体称为原胶原，由三个亚基(分子量均为95,000)构成。据说在第16天的鸡胚中胶原占全蛋白的60% (其前体为酸溶胶原，分子量150,000)，它们的亚基合成速度因分化状态而异。鱼精蛋白是结合在高等动物精子 DNA 上的碱性蛋白，它的 mRNA 有两类：一类是聚(A)⁺，一类是聚(A)⁻^[71]。核酮糖 2-磷酸羧化酶占叶绿体可溶蛋白的60%，是由8个大亚基(LS)与8个小亚基(SS)组成的复合酶，催化光合作用的初期反应——固定 CO₂。LS 由叶绿体的 DNA 编码，SS 由核的 DNA

表 2 真核细胞 mRNA—览表

编 码 的 蛋 白 质	分子量 × 10 ⁻⁴	材 料	mRNA大小 (S)	分子量 × 10 ⁻⁵	碱 基 数	纯度(%)
清蛋白	6.37	鼠 肝	18	7~9		>90
酪蛋白 α ₂	2	母羊乳腺		3.7		75~80
酪蛋白	2.5	母鼠乳腺	12	3.2		>90
酪蛋白	3.0, 4.6	母鼠乳腺	15	4.5		>90
β-晶体蛋白	5.0	鸡 胚	17	6.8	2200	
A ₂ -α-晶体蛋白	2.0	鸡 胚	14	3.6	1200	
丝心蛋白	17	家 蚕	47		16000	>95
珠蛋白	1.6	兔	9	2	630	>90
组蛋白 H-1	2	HeLa细胞		2.4	720	
组蛋白 H-2a, H-2b, H-3	1.5	”		1.6~1.8	480~550	
组蛋白 H-4	1.1	”		1.4	420	
组蛋白 H-5	2.1	鸡 胚	10			纯
免疫球蛋白 L 链 (λ)		骨 髓 瘤	13~14	4	1150	纯
免疫球蛋白 L 链	2.37	”	14	3.8		
免疫球蛋白 H 链	5.3	”	16~17			
角蛋白	1.0	鸡 胚	12	2.5	800	
α-乳清蛋白	2.25	母鼠乳腺	10.5			85
催乳激素	2.5	人 胎 盘	12~13	3	1000	
肌球蛋白 H 链	20	鸡 胚	26			
卵清蛋白	4.3	鸡 卵 管	16	6.5	1890±180	>99
酸溶胶原	15	鸡 胚	27~30	17~18		30
鱼精蛋白	0.5	虹鳟精腺	6		165~170	纯
核酮糖 2-磷酸羧化酶 LS	5.5	衣 藻	18	8		纯
卵黄蛋白的前体	27	鸡 肝	28	25	7220	纯
玉米醇溶蛋白	总4.1	玉 米	16	5.4		30

编码, 在它的生物合成与组合时, 还牵涉到植物色素, 对于研究基因的表达很是有趣。卵黄蛋白的前体是由雌激素诱发的。雄爪蟾肝脏也会因注射这种激素而在数小时内开始合成卵黄蛋白的前体, 所以可作为研究激素作用机制的有用材料^[62, 28]。玉米醇溶蛋白是占玉米粒 50% 的结构蛋白, 由几种成分组成。有一种含赖氨酸较多的非玉米醇溶蛋白, 能大量产生上述蛋白并被叫作 Opaque 2 的突变株在农业上有用, 因而正以分析它们的基因为目标而开始分离 mRNA, 无性繁殖的试验也在小量进行着^[72]。

五、结 语

分散在巨大的 DNA 分子上的许多基因虽说已用遗传学方法对其相互之间的位置关系进行了相当多的研究(绘制基因图), 而物质的, 即为特殊蛋白编码的 DNA 片段, 至少在真核细胞方面直至最近也还未获得。是因为没有方法。但是, 正象科学本身就是千方百计地探索解决看来根本不可能解决的问题的方法那样, 人们还是通过从 DNA-RNA 杂交的动力学到用电子显微镜观察 R 环形成等等方法, 使得基因的定性成为可能。到目前为止, 与 rRNA, tRNA 的基因已被鉴定的同时, 好些结构基因也用此法予以鉴定了(参阅第四节)。由限制性内切酶切成的 DNA 片段给与转录-翻译公用系统, 再分析形成的蛋白, 也鉴定了一些基因, 并发现线粒体 DNA 含有细胞色素 C 氧化酶之类的基因。但是, 即使鉴定了基因, 而就此便抽提单一的 DNA 也还是困难的, 为此不得不有待于基因无性繁殖方法的开创。众所周知, 无性繁殖有两种方法(参阅第一节), 在大量获取纯基因 DNA 方面, 以 mRNA 作材料的嵌入法(tailing)较为优越。但是这样得到的基因仅是为蛋白质编码的部分, 而缺少对基因表达极有意义的 RNA 聚合酶的结合部位, 乃是一大缺点。不过, 以这种基因片段作为亲和层析的基质、一次抽提全部基因的方法的出现, 多少弥补了这个缺陷。这个方法再稍加改进, 或许将能够在各种生理条件下抽提特定基因, 在基因表达的研究方面成为非常有用的方法, 尽管这是一项需要许多时间和精力以及坚强毅力的工作。

[补记: 最近, 免疫球蛋白 V λ 基因的无性繁殖已告成功, 且已阐明其碱基顺序的一部分。参见 Proc. Natl. Acad. Sci. USA. 74, 3518(1977), 以及 Gordon Conference, 1977.]
(参考文献略)

(鱼金涛译 盛沛根校)

分离真核细胞基因的新方法

A. П. РЫСКОВ

«Природа» 6: 85~91, 1977.

由于分子生物学和分子遗传学的重大进展,出现了现代生物学的新方向——基因工程。基因工程的最初成果已表明,与利用杂交分子有关的新方法是有前途的。将来这个领域的发展,使人们可能以制备量获得某些非常缺乏的物质。特别指出的是与工业方面提取一系列以人体作来源的具有生物活性的蛋白质和激素(如:免疫球蛋白、生长激素及人的胰岛素)有关。所有这一切都取决于在科学研究中迅速推广基因工程方法,以及广泛的舆论界对它感兴趣。

本文主要阐述的中心问题之一,也是现在很少研究的方法学的问题——为杂交分子、或所谓重组分子提取真核细胞基因。

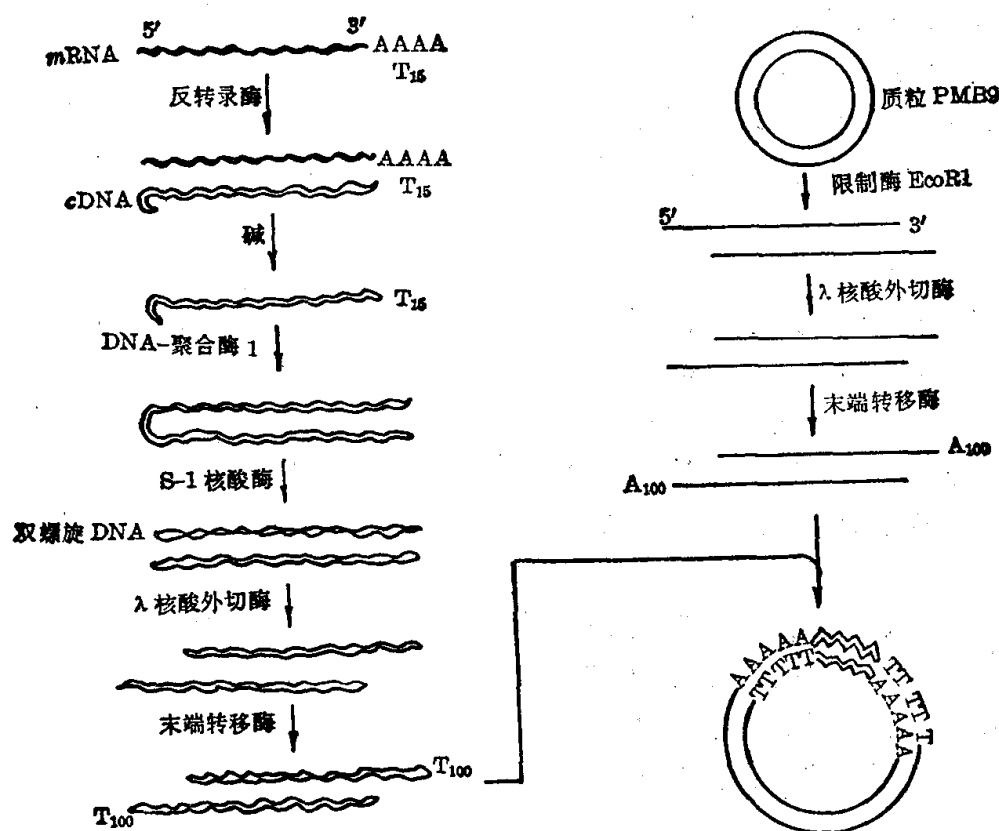


图1 珠蛋白基因的合成及其嵌入质粒DNA图(据T.Маннатиc等人, 1976)

图左——借助反转录酶在珠蛋白 mRNA 样板上合成在末端上具有发夹状结构的单链互补 DNA (cDNA)。然后,用碱脱去 mRNA,而 cDNA 在存在 DNA 聚合酶 1 时被制成双螺旋珠蛋白 DNA。借助末端转移酶得到在末端上具有胸腺核苷酸的珠蛋白 DNA 片段。接着,把用这种方法分离出的珠蛋白基因与预先用类似方法处理过的 PMB9 质粒 DNA 杂交(图右),及获取杂交分子 A_{100} , T_{15} 和 T_{100} ——腺嘌呤核苷酸和胸腺核苷酸的数量。