

讲 座

分 子 克 隆 技 术

敖 世 洲

(中国科学院上海生物化学研究所)

分子克隆技术是七十年代初发展起来的新技术,在十年时间里已经得到了广泛地应用。它不仅对分子生物学、遗传学、细胞学等学科研究起了很大的推动作用,而且在工业、农业和医学等实际部门也将作出应有的贡献。

分子克隆技术涉及面很广,包括限制性内切酶的应用,基因分离,载体组建,DNA 片段连接,细胞转化,重组体的筛选和鉴定,基因表达和物质的生产等方面。限于篇幅本文仅就大肠杆菌中的分子克隆技术基本原则,作简要的介绍。

一、限制性内切酶的应用

作用双链 DNA 分子的限制性内切酶到 1980 年已经发现 258 种,至少有 69 种不同的专一性。它们的识别序列为 4—6 碱基对。识别 4 碱基对的限制性内切酶作用大分子 DNA,一般说产生几百碱基对长的片段,而识别 6 碱基对的限制性内切酶则产生几千碱基对长的片段。这些片段有的是平头末端,有的带粘性末端。限制性内切酶有所谓“生物手术刀”之称,是分子克隆不可缺少的工具。无论是 DNA 重组,还是基因结构分析,都要借助于限制性内切酶的剪切作用。国外已有四十多种限制性内切酶作为商品出售,国内也有二十多种生产或试制。表 1 列举了一些常用的限制性内切酶的来源,识别序列和作用位置。

二、基 因 分 离

获得特异的基因有以下几种方法:

从染色体 DNA 中分离特异的基因,但是

表 1 常用限制性内切酶的识别序列

名 称	识别序列	来 源
Alu 1	↓ AGCT	Arthrobacter luteus
Ava 1	↓ CPyCGPuG	Anabaena variabilis
Bam HI	↓ GGATCC	Bacillus amyloliquefaciens
Bgl II	↓ AGATCT	Bacillus globigii
Eco RI	↓ GAATTC	Escherichia coli/R1
Hae III	↓ GGCC	Haemophilus aegyptius
Hind III	↓ AAGCTT	Haemophilus influenzae
Hpa 1	↓ GTTAAC	Haemophilus parainfluenzae
Kpn 1	↓ GGTACC	Klebsiella pneumoniae
Pst 1	↓ CTGCAG	Providencia stuartii
Sal 1	↓ GTCGAC	Streptomyces albus
Sma 1	↓ CCCGGG	Serratia marcescens
Xba 1	↓ TCTAGA	Xanthomonas badrii
Xho 1	↓ CTCGAG	Xanthomonas holicola

↓ 表示切点位置; 序列方向 5'—3'。

真核细胞的单拷贝基因占染色体 DNA 的比例很小,大约为 10^{-5} — 10^{-7} ,即使多拷贝基因也只占 10^{-3} 。因此直接从染色体 DNA 中克隆目的基因必须经过某种程度的提纯。最简便的方法是染色体 DNA 先经限制内切酶消化,根据分子量大小用琼脂糖凝胶电泳分离目的基因所在的片段。此种方法回收量较低,而且有些人的经验认为用一般琼脂糖分离的片段不适于重组连接,要用 Seaplague 琼脂糖 (Marine Colloids,

Rockland, ME) 才有较好的连接效果。最近采用 RPC-5 树脂反相层析方法能有效地大量分离限制性内切酶的作用片段。另外用梯度超离心方法也可以获得不同大小的 DNA 片段。

从 mRNA 反转录合成 cDNA 是获得基因的另一途径。mRNA 是基因转录, 经过加工的产物。分离 mRNA 一般是先用双抗免疫方法分离含特异 mRNA 的多聚核糖体。真核细胞的 mRNA 一般 3'端都具有多聚腺嘌呤核苷酸 (polyA) 的尾巴, 可以用寡聚脱氧胸腺嘧啶核苷酸 (oligo-dT) 纤维素柱从多聚核糖体总 RNA 中分离 mRNA, 或者再经过梯度超离心纯化。mRNA 用 oligo-dT 作引物, 在反转录酶作用下可以合成 cDNA。图 1 列出了 cDNA 合成的流程。双链 cDNA 或 RNA-DNA 杂链 cDNA 都可以进行克隆。

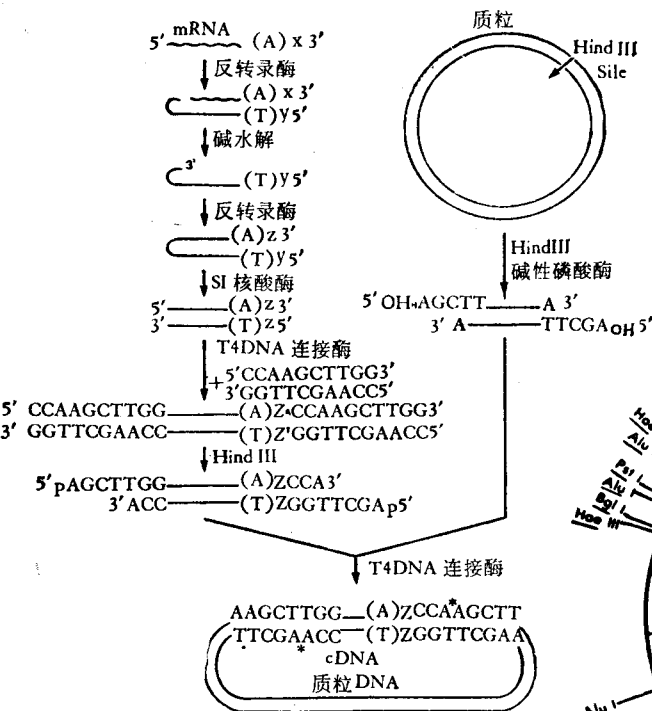


图 1 cDNA 的合成和 *Hind* III 识别序列人工接头在重组中的应用

化学合成特异的基因只限于核酸或肽链结构已经阐明, 分子量较小的基因, 大的 DNA 片段合成难度较大, 耗费也昂贵。

三、载体的选择

特异的基因片段转入大肠杆菌扩增和表达必须有能自主复制的载体携带。用作分子克隆载体的有细菌质粒和病毒 DNA。质粒载体如 pSC101, pMB9, pCR1, RSF2124, pBR322 等, 都是经过人工改建的, 带有一定的抗药标记, 具有一种或几种单切口的限制性内切酶的切点。国内外最普遍采用的载体质粒是 pBR322, 图 2 为其限制性图谱。它的分子量 2.6×10^6 道尔顿, 具有抗氨基青霉素 (Ap^r) 和抗四环素 (Tc^r) 的标记, 是含有 *ColE1* 复制起始区, 能为氯霉素扩增的多拷贝质粒, 具有 *Pst* I, *Eco* RI, *Hind* III, *Bam* HI 和 *Sal* I 多种限制性内切酶的单切点。《*Pst* I 切点插入外源基因, 使 Ap^r 基因破坏, 从 *Hind* III *Bam* HI 或 *Sal* I 切点接入外源基因, 使 Tc^r 基因破坏。这些抗药标记的改变有利于重组体的筛选。

用于大肠杆菌的病毒 DNA 载体是 λ 噬菌体的衍生株, 如 λ gtWES- λ B, charon 4A 等。这些噬菌体 DNA 中间有一段是其生活非必须片段, 用 *Eco* RI 切去, 再插入外基因片段。插入

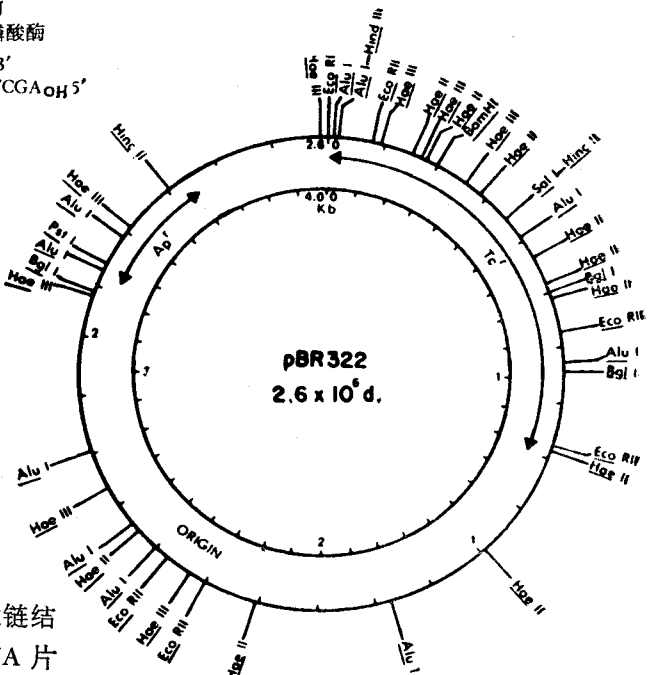


图 2 质粒 pBR322 的限制性图谱

片段要求在 15—20Kb 范围内。 λ DNA 载体广泛被用来组建各种真核细胞的基因库。

另一类载体 Cosmide 是由质粒和 λ DNA 粘性末端 (cos 区) 组建而成, 既具有质粒的性质, 又能为 λ 噬菌体外壳蛋白包装而具有感染作用, 能将重组体注入寄主细胞。这种载体可组装很大的外源 DNA 片段, 大约有 30—40Kb 长度。

四、DNA 片段连接

含有特异基因的 DNA 片段和载体在试管内连接有多种途径。最为常用的方法是用相同的限制性内切酶作用产生的粘性末端, 在 DNA 连接酶作用下形成共价键结合。图 3 所示 lac 重组体就是用 *Eco*RI 分别作用载体 PBR322 和 λ plac5 DNA, 用 T_4 DNA 连接酶作用组建而成的。这种方法的优点是重组体可以用同样的限制性内切酶把插入片段切下。它的缺点是可能发生同一 DNA 片段的环化 (DNA 浓度较大可以减少这种情况的发生) 或自身形成二聚体, 因

而重组的机率变低。

第二种是平头连接。平头切割的限制性内切酶作用片段, 或者载体和外源基因具有不同的限制性粘性末端先经 S_1 酶作用形成平头末端, 都可以在 T_4 DNA 连接酶作用下进行连接。但是此种方法也有很多非特异的重组, 形成重组体的机率比粘性末端连接还低。一般说提高 DNA 底物的浓度, 增加连接酶的用量, 可以达到较好的重组效果。另外化学合成的各种限制性内切酶识别序列的人工接头, 用平头连接法接在插入基因片段上, 再用限制酶切成粘性末端与载体重组 (见图 1) 也已被广泛采用。

第三种方法是借助于末端转移酶的作用。这种酶能在 DNA 引物 3'-OH 端加长单链脱氧核苷酸链。如果在载体的两个 3'-OH 端加上 dA 链, 在插入片段两端加上 dT 链, 两者混合退火就可以得到环状重组体。分别加上 dG 链和 dC 链也有同样作用。一般说 dA-dT 需要上百个核苷酸链, dG-dC 只要十多个核苷酸就能起同样的作用。早期报道这种酶只作用 3'-OH 突出末端, 以后的研究表明, 在 Co^{++} 存在下, 5' 端突出, 内陷的 3'-OH 端也能作用。使用这种方法, 要注意防止末端转移酶内污染有核酸酶, 否则就得不到理想的结果。

五、细胞转化

体外构建的重组体只有输入细胞内进行扩增和表达, 才有研究和实用意义。使用的载体不同, 引入细胞的方法也不一样。以质粒为载体, 常用细胞转化的方法。大肠杆菌受体菌都是 K-12 衍生株, 如 C600, HB101, X1776 等。这些菌的细胞经过 Ca^{++} 的处理, DNA 即可进入。用培养基稀释 Ca^{++} 的浓度, 基因便进行表达。细菌质粒的转化效率大约为 10^5 — 10^7 转化体/ μ g DNA。 Ca^{++} 处理时间较长, 可以提高转化效率, 但活细胞总数减少。一般说质粒的分子愈小转化效率愈高; 环形分子比线形分子转化效率要高 1000 倍。用 λ DNA 为载体, 可以用重组 DNA 直接转染 Ca^{++} 处理的受体细胞, 转染效率可达 10^5 — 10^6 噬斑/ μ g DNA。重组体

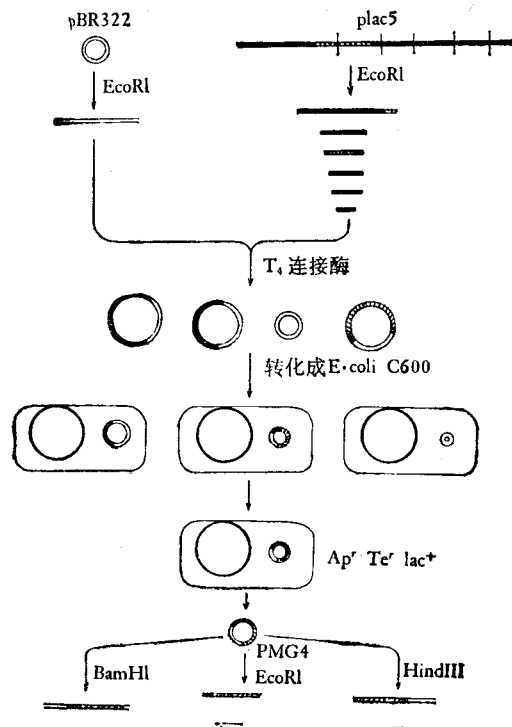


图 3 lac 重组体 pMG4 的构建、筛选和鉴定程序

在体外用 λ 噬菌体外壳蛋白包装，感染受体细胞可以大大提高效率。

六、重组体的筛选和鉴定

转化或转染以后，获得了大量的菌落或噬斑，可以用以下方法筛选克隆株：最为方便的是平板筛选，抗药基因，营养缺陷互补，某些酶的显色反应（如 β -半乳糖苷酶可用 λ gal 或 ONPG 显色）等都可此种办法。其次是菌落的原位杂交或噬斑杂交，采用 Grunstein 和 Hogness 程序将菌落转移和固定在硝酸纤维膜上，与 ^{125}I 或 ^{32}P 标记的 RNA 或 cDNA 杂交，从杂交的阳性显影斑上选出克隆株。

获得克隆株后，作重组体分离，进一步作限制性内切酶分析(图 4 为 lac 重组体限制性内切酶分析的凝胶电泳图)或作 Southern 墨迹转移，与探针杂交 (图 5 为中华大蟾蜍 rDNA 限制性酶切片段 Southern 转移放射自显影图)以确定插入片段的性质。如果插入片段的核苷酸序列或编码的氨基酸序列已经知道，可用 Maxam 和 Gilbert 方法作核苷酸序列测定，如果克隆基因能够表达，可测定基因产物，如某些酶的活力或

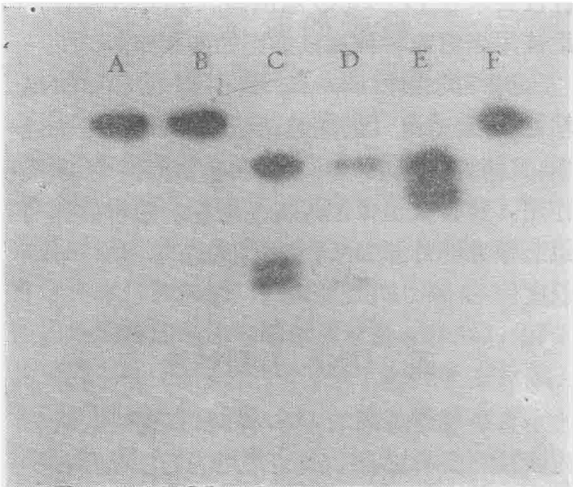


图 5 中华大蟾蜍 rDNA 限制性片段 Southern 转移放射自显影图

中华大蟾蜍 rDNA 重组体 (A, C, E) 和染色体 DNA (B, D, F) 的 *Bam*H1 (A, B), *Eco*R1 (C, D), *pst*I (E, F) 片段作 Southern 墨迹转移，与 ^{125}I -rRNA 杂交。

蛋白免疫反应等，都可以对重组体作出明确的鉴定。

七、真核基因在大肠杆菌中的表达

真核基因在大肠杆菌中表达，将分子克隆技术用于物质的生产已经成为现实。目前美国已经有近五十种蛋白、多肽、酶等医用药物计划投入生产。有些细菌产物如人胰岛素，人干扰素，人生长激素等已经提供临床试验。

要使真核基因在细菌中表达，实现物质生产的目的，需要解决以下几个技术环节：首先，引入细菌的真核基因不能有插入序列，常用 cDNA，因为细菌细胞内没有真核细胞的转录后加工系统。其次，真核基因的表达要用原核的启动基因控制，广泛使用的是 lac 启动基因。有人发现色氨酸启动基因具有更高的效率，可能在实际应用上有较大的价值。另外 β -内酰胺酶的启动基因也有人应用，但效率较低。在真核基因与原核启动子连接时，要注意核糖体的结合位点和翻译起密码之间的核苷酸长度，即所谓 SD 序列，一般应保持 7—12 个核苷酸，这是翻译过程中的影响因素。第三，某些真核基

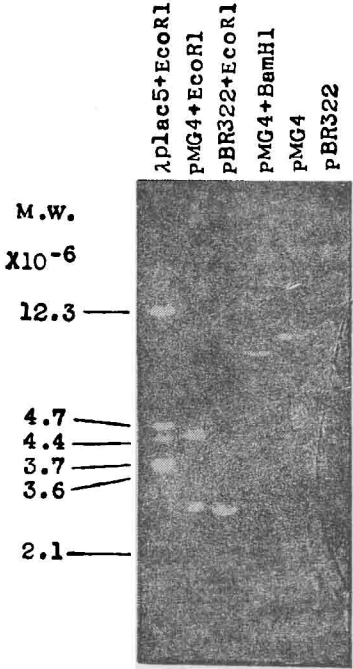


图 4 lac 重组体 pMG4 琼脂糖凝胶电泳鉴定

因的产物在细菌细胞内很不稳定,易被蛋白水解酶分解,需要用细菌的肽段加以保护,因此细菌的初级产品是杂合蛋白,如 β -半乳糖苷酶-生长激素释放抑制因子, β -内酰胺酶-前胰岛素原等。要使两个肽段特异地断开,可以用甲硫氨酸作接点,在溴化氰作用下两条肽链断开。但是如果基因产物中也有甲硫氨酸成分就不能应用。有人研究利用某些蛋白水解酶特异专一的肽链作为接点,这可能有广泛应用的价值。以上这些方面有的还可能进一步改进,潜力很大,值得深入研究。

最后关于分子克隆技术应用的安全问题,也不应忽视。在这一技术刚刚出现的时候,曾经引起社会各方面的关注,出现不少忧虑。近年来国外对重组分子的潜在危害估计已经降低,安全限制也有放宽,但是短时间内还不可能对其危害程度作出定论。我们在应用中可以结

合国内条件,参照国外条例,做到有备无患。

参 考 文 献

- [1] Roberts, R. J.: *Nucleic Acids Research*, 9, r75, 1981.
- [2] Wu, R. (ed.): *Methods in Enzymology*, Vol. 68, Academic Press, New York, 1979.
- [3] Ullrich, A. et al.: *Science*, 196, 1313, 1977.
- [4] Cohen, S. et al.: *Proc. Natl. Acad. Sci. U.S.A.*, 70, 3240, 1973.
- [5] Grunstein, M. and Hogness, D.: *Proc. Natl. Acad. Sci. U.S.A.*, 72, 3961, 1975.
- [6] Southern, E. M.: *J. Mol. Biol.*, 98, 503, 1975.
- [7] Maxam, A. M. and Gilbert, W.: *Proc. Natl. Acad. Sci. U.S.A.*, 74, 560, 1977.
- [8] Guarente, L. et al.: *Science*, 209, 1428, 1980.
- [9] Goeddel, D. V. et al.: *Proc. Natl. Acad. Sci. U.S.A.*, 76, 106, 1979.
- [10] Goeddel, D. V. et al.: *Nature*, 281, 544, 1979.
- [11] Goeddel, D. V. et al.: *Nucleic Acids Research*, 8, 4057, 1980.

[本文于1981年6月18日收到]

书刊消息

“分子生物学杂志(J. Mole. Biol.)”内容将有重要变化

“分子生物学杂志(J. Mole. Biol.)”已创刊21年。自创刊以来,这本杂志的报道,一直位于本研究领域的前沿,并对建立和明确分子生物学领域的界限作出了贡献。今年就这本杂志的内容和编排进行了一些修改,以保证及时而全面地报道继续以惊人速度发展着的分子生物学领域中所出现的新成就。

报道范围:以前的传统是偏重于分子遗传学、复制和转译机理,以及生物结构方面。这种传统将继续保留。另外鉴于从分子的角度研究细胞生物学方面的问题,已取得许多成果,增加了细胞生物学方面的内容。

编排结构:为使读者便于查找有关文章,同时又避免把杂志分成几个部分,每期的文章将按照主题领域依次编排,遗传学在最前,接着是细胞生物学和生物化学,最后是物理化学和生物大分子的结构。

编辑部成员及分工:

主编: J. C. Kendrew

副主编: S. Brenner

基因——基因结构、基因修饰、基因表达、基因调控: S. Brenner, P. Chambon, M. Gottesman.

细胞——细胞发育、细胞功能: W. Franke, K. Simons., 细胞器结构、大分子聚集体: H. Huxley, A. Klug.

分子——大分子结构: R. Huber, J. Kendrew., 物理化学: M. Gellert

致编者的信——一般内容: S. Brenner X-射线晶体研究初步资料: R. Huber, J. Kendrew

助理编辑: M. Bretscher, C. Cantor, G. Gilbert, V. Luzzati, M. Moody, S. Weissman

每位编辑分工单独负责本杂志报道范畴内一个特定领域,并在每篇文末注明责任编辑。作者可直接把文稿寄给有关编辑。

速度——将尽可能缩短收文和发稿时间,但不能降低审稿、编辑、印刷和图片复制的标准。

每年发行36期,新JMB有样本可供索取。样本内印有对作者的要求和注意事项及编辑部关于更改内容和编排的正式通知。

消息来源: *Cell Biol. Int'l Rep.* Vol.

5 (8) 1981 (情)