

核酸研究技术

下 册

蔡良琬 主编

科学出版社



中科院植物所图书馆



S0014753

2006.10.11

58.17425
721

核酸研究技术

下册

蔡良琬 主编

朱文彦 原籍刻



科学出版社

1990

内 容 简 介

本书包括分子克隆和基因表达两大部分。分子克隆部分介绍有关的基本知识,如常用的受体菌、载体及有关的核酸酶;分子克隆的具体操作,即DNA的剪切、重组、转化及重组DNA克隆的筛选和鉴定;基因文库和cDNA库的建立。基因表达部分重点介绍外源基因在大肠杆菌、酵母及哺乳动物细胞三种体系中的表达。可供生物学、医学、农学的有关研究人员和大专院校师生参考。

核 酸 研 究 技 术

下 册

蔡良婉 主编

责任编辑 吴铁双

科学出版社出版

北京东黄城根北街16号

邮政编码: 100707

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1990年7月第一版 开本: 787×1092 1/16

1990年7月第一次印刷 印张: 19 1/2

印数: 0001—1 400 字数: 443 000

ISBN 7-03-001650-5/Q·242

定价: 16.50 元

前 言

基因工程是近十年发展起来的新技术,由于它与工业生产密切相关,可促进生物产品的工业改造,故一开始就为发达国家所重视,有些企业家为此创办了相应的基因工程公司,从而有力地推动了这一技术的发展。基因工程现已成为生物工程中的核心部分,在某些国家还被列入重点研究与开发项目。

基因工程的核心技术是分子克隆,其次是外源基因的高效表达,而基因表达也离不开分子克隆技术。从1973年出现第一个分子克隆以来,应用的有关载体与核酸酶类在品种上已大量增加,应用新的分子克隆与基因表达原理构建了许多适用的载体,又建立了三大外源基因表达体系,使基因工程技术日趋完善。从事这一新领域工作的科研人员不断增加,实验室也不断扩展,这种技术已应用到整个生物学、医学与农学等各学科的研究中。

基因工程技术受到广泛重视,除了它与工农业生产密切结合,可提供某些生物产品外,另一个突出的优点是应用分子克隆技术可纯化任一生物基因,并可通过大量培养受体菌(通常是大肠杆菌)得到纯化的基因,使研究者可用此技术得到其所需的基因;分子克隆技术也因此广泛应用,并逐步达到简便易行,成为分子生物学中即将普遍应用的技术。现在所用的载体、限制性内切酶、常用的核酸酶以及人工合成的寡核苷酸等都有商品供应,其中不少国内已生产供应,这也大大促进了分子克隆技术的推广。为了进行复杂的真核生物特别是高等动植物的基因研究,又建立了基因文库与cDNA库技术,操作虽然繁杂一些,但也不难掌握。本书首先介绍了一般分子克隆基本技术,也对以上两种文库作了详细的介绍。

近年来通过基因工程研究,已有一些生物产品,如蛋白质制品、活性多肽、酶等开始出售,并产生了一定的经济效益和社会效益。此外,基因工程的产品还显示出其特殊的优点,即提高了产品纯度,降低了销售价格。同时,还获得了一些很难工业生产的药物与疫苗等,使外源基因高效表达研究理论仍为当前研究的一个重点,本书也介绍了三大基因表达体系。

本书作者都是从事有关研究的同志,所写内容是几年来的经验总结,方法又都是国内比较切实可行的。可供从事核酸或基因研究、教学及生产人员参考。

本书的编写由蔡良琬同志主持,参加编写的有中国医学科学院基础医学研究所杨善荣、强伯勤、余曙华、王申五、蔡良琬、李进、罗刚、袁建刚、许彩民、吴石君同志,南开大学生物系蒋如璋同志,中国预防医学科学院病毒学研究所侯云德、任贵方同志,中国科学院生物物理研究所申同健同志。

蔡良琬

一九四九年十月二十二日
（一）

（二）

（三）

（四）

（五）

目 录

前言

第一部分 分子克隆

第一章 大肠杆菌的基本知识和技术	1
第一节 大肠杆菌的特征及其细胞结构	1
一、大肠杆菌的分类特征	1
二、大肠杆菌的细胞结构	2
第二节 大肠杆菌基因的命名规则	3
一、与代谢有关的结构基因	3
二、与代谢无直接关系的结构基因	4
三、与药物或噬菌体抗性有关的基因	3
四、抑制基因	4
五、染色体结构突变	5
六、其他	6
第三节 大肠杆菌的遗传性和变异性	6
一、大肠杆菌的变异性	6
二、大肠杆菌的菌株及其特性的鉴定	8
三、大肠杆菌菌株的保存方法	10
第四节 大肠杆菌的营养和生长	11
一、大肠杆菌的营养	11
二、大肠杆菌在固体培养基上的生长	12
三、大肠杆菌在液体培养基中的生长	13
第二章 基因工程的载体体系	16
第一节 质粒载体	16
一、细菌质粒的基本概念	16
二、基因工程中质粒载体的选择	20
三、介绍几类质粒载体	23
第二节 λ 噬菌体载体	28
一、 λ 噬菌体的生物学特性	28
二、 λ 噬菌体载体的构建	33
三、常用的 λ 载体种类	37
四、 λ 载体的选择和应用	43
第三章 基因工程中常用的核酸酶类	47
第一节 DNA 和 RNA 连接酶	47
一、T4 DNA 连接酶	48

二、大肠杆菌 DNA 连接酶	50
三、T4 RNA 连接酶	51
第二节 DNA 聚合酶	51
一、大肠杆菌聚合酶 I	51
二、DNA 聚合酶的 Klenow 片段	53
三、T4 DNA 聚合酶	54
四、逆转录酶	56
五、末端转移酶	58
第三节 磷酸激酶和磷酸酶	60
一、T4 多核苷酸激酶	60
二、碱性磷酸酶	63
第四节 核酸水解酶	65
一、外切酶 III	66
二、S1 核酸酶	67
三、绿豆核酸内切酶	68
四、外切酶 VII	69
五、 λ 外切酶	69
六、Bal 31 核酸酶	70
七、核糖核酸酶 H	72
第五节 甲基化酶	73
一、原核细胞中的甲基化酶	73
二、存在于大肠杆菌 K12 株中的甲基化酶	74
三、甲基化酶的用法	75
四、甲基化酶的实际应用	76
第四章 分子克隆技术(一)	
——基因与载体的剪切及重组	81
第一节 重组技术概述	81
一、剪切获取特定的基因片段	81
二、载体的选择	82
三、载体的剪切	82
四、基因与载体的重组	82
第二节 基因剪切常用方法	82
一、基因剪切的特点	82
二、应用限制性内切酶水解基因与载体形成粘末端	83
三、形成带有平末端的基因	88
四、从平末端改造成粘末端	97
第三节 剪切后基因与载体的连接	104
一、带有粘末端的 DNA 片段之间的连接	105
二、平末端及 3' 突出末端的连接特点	106
三、基因片段大小与连接的关系	107
四、防止载体自身连接的脱末端磷酸基团	108
五、连接后限制性内切酶识别位点的恢复	110

六、同聚物加尾后的连接	110
第五章 分子克隆技术(二)	
——重组 DNA 的转化、筛选和鉴定	111
第一节 重组 DNA 对大肠杆菌的转化	112
一、氯化钙法	112
二、氯化钙-氯化钾法	113
三、大肠杆菌 X1776 的转化	114
四、重组 DNA 克隆的初步筛选	116
第二节 重组 DNA 克隆的核酸杂交筛选方法	117
一、菌落杂交	117
二、重组噬菌体 DNA 的杂交筛选	119
第三节 免疫化学筛选方法	120
一、抗体的制备	121
二、抗体的纯化	121
三、抗体蛋白质的 ¹²⁵ I 标记	122
四、溴化氰活化滤纸法	122
五、固相筛选法	124
第四节 重组质粒和 λ 噬菌体 DNA 的小量快速制备方法	125
一、碱裂解法制备质粒 DNA	125
二、煮沸法制备质粒 DNA	126
三、从单菌落中制备质粒 DNA	127
四、平板裂解法制备 λ 噬菌体 DNA	127
五、液体培养基裂解法制备 λ 噬菌体 DNA	128
第五节 重组 DNA 的酶切分析和鉴定	129
一、单酶解分析	129
二、双酶解分析	130
三、多次酶解和电泳分析	130
四、部分降解酶切分析	131
五、人工合成探针在鉴定中的应用	132
第六节 SDS 聚丙烯酰胺凝胶电泳和 Western 技术在克隆鉴定中的应用	133
一、SDS 聚丙烯酰胺凝胶电泳	134
二、电泳转移	136
三、抗体的放射性检测法	137
第六章 基因文库	139
第一节 引言	139
一、基因的分离	139
二、什么是基因文库	140
三、基因文库构建的基本步骤	141
第二节 供体 DNA	143
一、染色体 DNA 的抽提	143
二、染色体 DNA 的切割	145
第三节 载体 DNA	148

一、 λ 替代型载体的结构特点	149
二、载体 DNA 的制备	151
三、制备载体的两臂和去除中央片段	152
第四节 重组连接、体外包装和感染方法	154
一、重组连接	154
二、重组噬菌体 DNA 的体外包装	158
三、重组噬菌体的感染	161
四、基因文库的构建规模	165
五、基因文库的扩增和贮存	166
第五节 考斯质粒为载体的基因文库	168
一、考斯质粒作载体的优缺点	168
二、基因文库的构建过程	168
结语	170
第七章 cDNA 克隆	172
第一节 cDNA 克隆的一般方法和步骤	173
一、cDNA 克隆方法	173
二、cDNA 克隆的具体步骤	174
第二节 特异 cDNA 的克隆	179
一、插入到表达载体中的 cDNA 克隆的筛选	179
二、选择杂交法筛选特异 cDNA 克隆	179
三、cDNA 克隆操作实例	180
第三节 cDNA 克隆应注意的问题	182
一、cDNA 克隆所使用的酶制剂	182
二、作为 cDNA 合成模板的 mRNA	183
三、载体 DNA	185
四、器械的消毒处理	185

第二部分 基因表达

第八章 基因表达与几种表达体系	187
第一节 基因表达的基本原理	188
一、乳糖和色氨酸操纵子的特点及其转录翻译元件	188
二、lac 与 trp 操纵子的基因表达调节	191
三、真核细胞中的基因表达与多级调控	193
第二节 常用的基因表达质粒	194
一、lac 启动子的应用	194
二、trp 启动子与人工合成 tac 启动子的应用	196
三、 λ 噬菌体中 P_L 与 P_R 两个启动子的应用	199
第三节 翻译框架与融合蛋白	199
一、基因表达与翻译框架	200
二、融合蛋白	200
第四节 几种基因表达体系的比较	201

第九章 大肠杆菌体系中的基因表达	202
第一节 真核基因与原核基因在结构和表达调控上的差异	202
第二节 影响真核基因在大肠杆菌体系中表达的因素	203
一、启动子	203
二、基因剂量	205
三、核糖体结合位点	206
四、基因产物的稳定性	207
五、编码多肽密码子的组成情况	207
六、表达产物的构象	208
七、表达产物的分子量大小	208
八、原核增强序列	208
第三节 真核基因在大肠杆菌中表达的方式	209
一、融合基因-融合蛋白	209
二、非融合基因-非融合蛋白	210
三、融合基因-非融合蛋白	211
第四节 真核基因在大肠杆菌体系中的表达设计	211
一、lac 启动子.....	211
二、trp 启动子.....	216
三、tac 启动子.....	219
四、P _L 启动子	220
五、β 内酰胺酶启动子	225
六、lpp 启动子.....	225
第十章 酵母体系中的基因表达	232
第一节 酵母表达系统概述	232
一、载体	232
二、启动子及其他控制序列	234
三、寄主酵母菌株	236
第二节 酵母表达系统的构建	236
一、酵母启动子的提取	236
二、表达载体的构建	239
三、外源基因的引入和表达质粒的构建	243
第三节 酵母体系中的基因表达	248
一、酵母菌的转化	248
二、酵母质粒的检测	250
第十一章 哺乳动物细胞受体系统中的基因表达	255
第一节 哺乳动物细胞表达外源基因概述	256
第二节 哺乳动物细胞表达外源基因的必需条件	256
一、外源 DNA 或供体 DNA	258
二、载体	259
三、重组质粒的改建	260
四、受体细胞	267

五、选择标记和选择培养基	269
六、重组 DNA 转染细胞	271
七、图示几种转化表达 HBsAg 的系统	275
附录一、大肠杆菌的常用培养基	279
附录二、大肠杆菌的常用菌株	281
附录三、大肠杆菌遗传学中常见的基因型	282
附图	283

第一部分 分子克隆

第一章 大肠杆菌的基本知识和技术

蒋如璋

大肠杆菌 (*Escherichia coli*) 是现代生物学尤其是分子遗传学和生物化学研究中简单而又极富成果的模式实验系统, 是生命科学前沿领域研究的主要实验生物之一。大肠杆菌分布广, 容易得到, 营养要求简单, 生长迅速, 便于实验操作, 被人们广泛而深入地研究过。它在近代生物学研究的各个阶段, 以及在解决重大基础理论上起了关键性的作用。

Delbrück 和 Luria (1943) 以大肠杆菌及其噬菌体为材料完成的彷徨变异测验 (fluctuation test), 证明原核生物的性状也是由基因决定的, 从而使微生物遗传学最终确立为一门独立的学科。J. Lederberg (1946) 证明大肠杆菌也有有性生殖, 可以用与高等生物相同的遗传学原理, 研究原核生物的遗传和变异问题, 从此遗传学得到了飞速的发展。又由于后来遗传学方法和生物化学方法的卓有成效的结合, 从而解决了许多单用遗传学或单用生物化学方法不能解决的许多重大问题, 比如代谢途径的研究、DNA 复制、蛋白质合成、乳糖操纵子模型, 等等。

70 年代出现的新技术——重组体 DNA 技术的基础研究, 例如寄主限制性变异和 DNA 限制性内切酶的发现等, 最初也都是在大肠杆菌上进行的。大肠杆菌是目前基因工程研究中的最为完善的载体受体系统, 因此, 凡应用大肠杆菌为实验材料的科学工作者对它的遗传背景作一定的了解是必要的。

第一节 大肠杆菌的特征及其细胞结构

一、大肠杆菌的分类特征

大肠杆菌属肠细菌科 (Enterobacteriaceae), 革兰氏染色阴性, 菌体杆状 ($2-3\mu\text{m} \times 0.4-0.6\mu\text{m}$), 是需氧或兼性厌氧细菌, 不形成芽孢和荚膜。大肠杆菌在分类上的其他特征是周身被鞭毛, 能运动, 在琼脂上形成圆形、光滑、白色菌落, 有臭味。典型的大肠杆菌菌株能利用乳糖和发酵糖产气, 能由色氨酸产生吲哚。在伊红美蓝培养基上菌落圆形, 中央黑色; 在远藤氏培养基上菌落圆形红色, 有金属光泽。

大肠杆菌除了少数菌株外, 通常都不致病。它是肠细菌区系中的优势兼性细菌, 与其他多种肠细菌如沙门氏菌属 (*Salmonella*)、克氏杆菌属 (*Klebsiella*)、志贺氏菌属 (*Shigella*) 和弧菌属 (*Vibrio*) 等细菌共存于肠道环境中形成肠细菌区系。它们在肠道中就

像在实验室的连续培养装置中一样生长着，同时又以高密度存在。从遗传学和进化论的观点来看，肠道细菌区系中的不同类细菌的基因可以通过噬菌体转导、溶源化和接合作用等过程互相转移，从而促进了基因在亲缘关系较远的不同物种之间的水平分布，增强了肠细菌的变异性及其对环境的适应能力。携带多种抗药基因的R因子在临床上有特殊的重要性。该因子最初是在志贺氏菌属中发现的，现在已知它们广泛分布于大肠杆菌和沙门氏菌属等肠细菌中。已经证明，它们无论是在体外或肠道内都很容易互相转移，所以，肠细菌与其他菌种相比表现出更大的代谢及抗原型上的变异性，以及遗传上的不稳定性。正因为这样，即使大肠杆菌本身并非致病菌，但它仍属于一种潜在的病原菌。

二、大肠杆菌的细胞结构

和其他革兰氏阴性细菌一样，大肠杆菌具有比革兰氏阳性细菌更复杂的细胞包被(图 1.1)。

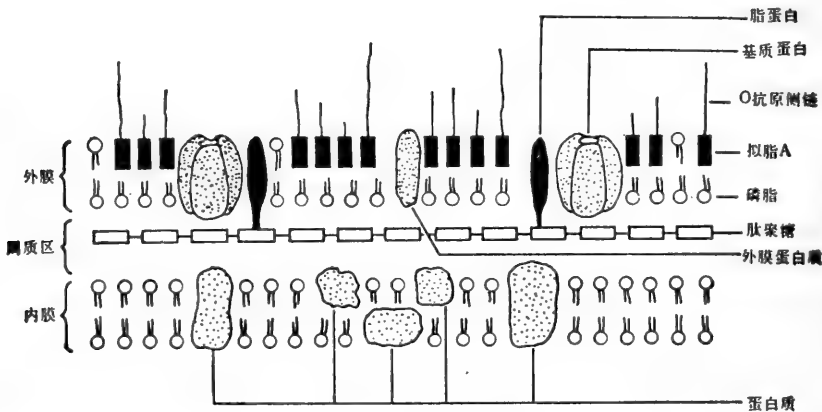


图 1.1 大肠杆菌包被模式图。

外膜中的基蛋白与脂蛋白和拟脂 A 连接，而脂蛋白与肽聚糖共价结合使外膜与细胞壁连成一体。外膜除了能让小分子扩散透过外，疏水的和亲水的脂分子都很难通过，因此，通常革兰氏阴性菌对抗生素较不敏感，可分泌蛋白质只能聚积在周质区。

大肠杆菌在细胞膜外面有一层与革兰氏阳性细菌相似的以肽聚糖为主要成分的细胞壁。与革兰氏阳性细菌不同的是，在细胞壁外面还有一层脂多糖外膜。这种结构成分与细菌的抗原型有关。脂多糖是热源物质的主要来源。外膜也是阻挡可分泌蛋白分泌到细胞外的屏障，使分泌蛋白质只能积累在周质部位。

细菌染色体总是附着在间体 (mesosome) 上。被细胞膜包裹着的细胞浆是细胞代谢活动的场所，约有 30,000 个核糖体散布在其中。

细胞壁在维持细胞形状和抵抗外界不良环境方面起着重要作用。在高渗溶液 (如 20% 蔗糖) 中，用溶菌酶处理大肠杆菌细胞，可将大部分细胞壁成分去除，这时在相差显微镜下可以看到原来为杆状体的菌体变成球形，即球质体 (spheroplast)。如果将细胞壁成分全部去除，则为原生质体。除去细胞壁的大肠杆菌细胞失去了对渗透压的抵抗力，一旦遇水立即破裂。

噬菌体可以感染完整的大肠杆菌细胞，使之裂解，但不能感染原生质体。说明噬菌体

感染寄主的早期阶段需要细胞壁的某种成分作为它的受体。原生质体虽对噬菌体缺乏感受性,但可被噬菌体 DNA 转染或被外源 DNA 转化。原生质体转化已成为克服完整细胞不出现感受态的一种好方法,因而被广泛应用于多种细菌的遗传学研究。

第二节 大肠杆菌基因的命名规则

据估计大肠杆菌的整个基因组可以编码 2,000—3,000 个基因,已经做过基因定位的约占 1/3,即将近 1,000 个基因。为了叙述方便必须有一个大家共同遵循的基因命名规则。为此, M. Demerec (1966) 提出了一套细菌遗传座位的三字命名规则,其中包括应标示出的与代谢有关的基因,虽无直接关系但又是代谢中较重要的基因,药物或噬菌体抗性基因,某些与该细菌生长有关的生长抑制基因,以及突变所发生的遗传结构改变(如基因缺失,转座基因的出现等)。有关的表示法将以下面叙述的办法来命名。

一、与代谢有关的结构基因

与合成和分解代谢有关的结构基因座位,以与基因编码的酶所参与的代谢途径的三个小写英文字母表示。例如用 lac 表示与乳糖利用有关的基因, trp 表示与色氨酸合成有关的基因等。三字符号后面加上一个大写的字母用来区分与同一代谢途径有关的不同的遗传座位,例如 lac Z 和 lac Y 分别表示 β 半乳糖苷酶结构基因和乳糖透性酶的结构基因; trp A 和 trp B 分别表示色氨酸合成酶的两个亚基的结构基因等。对于同一基因内的不同点突变则以阿拉伯数字表示,比如 trpA1、trpA26 等。如果只知道突变基因所影响的表现型,而不知道确切的遗传座位,则在基因符号后以一短线代替大写的英文字母,以数码表示分离子 (isolate) 的编号,例如 thr-15, met-20 等。

在大肠杆菌的菌株列表中可以看到用诸如 *E. coli* K12 LE 392: sup E44, sup F58, hsd R514, gal K2, trp R55, met B1, lac Y1, 这样来描述一个菌株。它几乎记述了有关菌株的全部信息: ①该菌株属大肠杆菌 K12 的一个衍生株, ②LE392 是衍生菌株的代号, ③七个基因符号表示该菌株的特性或基因型。通常为了方便只需用大肠杆菌 LE 392 就可以了。

除了大肠杆菌 K12 系统的菌株外,最常用的还有大肠杆菌 B 和 C 菌株。它们之间的区别除了在生理代谢上略有不同外,对不同噬菌体感受性限制修饰酶系统等也有区别。

一定的表现型是一定的基因型在特定的条件下发育的结果,因此表现型与基因型之间存在着相应关系,一般以同样的三字符号来表示表现型,只是第一个字母为大写,并在其右上角加“+”或“-”号以表明有关基因的表现型状态。如 Lac⁺和 Lac⁻分别表示能或不能以乳糖为唯一碳源生长; Trp⁻表示该菌株的表现型只有在含有色氨酸的基本培养基上才能生长。

二、与代谢无直接关系的结构基因

有许多结构基因,其基因产物并不是代谢链的某种酶,而是 RNA 或结构蛋白,或者

是与 RNA 转录、DNA 复制、蛋白质合成等生物学过程有关的酶或多肽等。在这种情况下仍以三字符号表示，不同的是往往用几个英文单词的字头缩写。例如表示与核糖体大亚基中的某种多肽结构有关的基因时用 *rpl* (ribosomal protein large)。表示与辐射损伤修复有关的酶的基因用 *uvr* (ultra violet repair)。dna 表示与 DNA 合成有关的基因。rec (recombination) 表示普遍性(同源性)重组有关基因的符号。met T, ala T 分别表示蛋氨酸和丙氨酸转移 RNA 基因。rrn 是核糖体 RNA 操纵子的基因符号,等等。

三、与药物或噬菌体抗性有关的基因

这类基因通常直接以所抗的药物或噬菌体名称的缩写表示。例如 *str* 是与链霉素抗性有关的基因; *rif* 是与利福平抗性有关的基因; *ton* 和 *tsx* 是与抗噬菌体 T1 (和T5) 和 T6 有关的基因。

抗性表现型与前面介绍的表示法相似,只是在符号的右上角加上“r”或“s”,以表示抗性和敏感表现型的区别。例如 *Str^r/Str^s* 分别为链霉素抗性和敏感表现型。

随着抗性机制研究的进展,对有些抗性突变的实质已经了解。例如 *str* 是因为核糖体小亚基的第十二号蛋白质的变化,所以也可以表示为 *rpsL*(ribosomal protein small)。

四、抑制基因

也有些基因不符合上述基因型-表现型命名规则,例如抑制基因就是这样。野生型菌株不携带抑制基因,在遗传学上应为 *sup⁺*,而其表现型效应为 *Sup⁻*; 携带抑制基因的突变型菌株,其基因型应为 *sup⁻*,表现型效应为 *Sup⁺*。为了避免此基因型与表现型的混淆,现在常用数字表示特定抑制基因的表现型,以字母表示相应的基因型。例如 *Su1* 表示特定抑制基因的表现型,而以 *sup D* 表示其基因型(表 1.1)。如 *Su⁻* 及其基因型则不以任何符号表示。抑制基因的实质是 tRNA 基因的突变。

表 1.1 大肠杆菌无义密码系的抑制基因

抑制基因	表型符号	被抑制的无义密码子	插入的氨基酸
<i>sup D</i>	<i>Su1</i>	UAG	丝氨酸
<i>sup E</i>	<i>Su2</i>	UAG	谷酰胺
<i>sup F</i>	<i>Su3</i>	UAG	酪氨酸
<i>sup C</i>	<i>Su4</i>	UAG,UAA	酪氨酸
<i>sup G</i>	<i>Su5</i>	UAG,UAA	赖氨酸
—	<i>Su6</i>	UAG	亮氨酸
<i>sup U</i>	<i>Su7</i>	UAG	谷酰胺
<i>sup V</i>	<i>Su8</i>	UAG,UAA	—
—	<i>Su9</i>	UGA	色氨酸
<i>sup B</i>	<i>SuB</i>	UAG,UAA	谷酰胺

摘自 Gorini, L., *Ann. Rev. Genet.*, 4:107, 1970.

五、染色体结构突变

除了点突变外,还有一类与染色体结构改变有关的突变,常遇到的有缺失和易位(translocation)两种类型。表示缺失突变的符号是“ Δ ”。 Δ 符号后面标记的基因符号为所缺失的基因,例如 Δuvr 表示缺失了与辐射损伤修复有关的酶的基因。如果缺失范围包括几个连锁基因,则将所缺失的基因放在括号内,如 $\Delta(\text{lac pro})$ 。如果同一区域分离到多个独立的缺失突变型时,便在缺失基因后面加上分离子的编号,如 $\Delta(\text{gal uvr B})47$ 。染色体中的易位,从遗传因子看,实际是基因的转座(transposition)。

可转移遗传因子是 50 年代初美国细胞遗传学家 McClintock 在玉米遗传学研究中发现的。这种因子可以跳跃转移(转座),影响插入部位邻近基因的表达和引起染色体重排。但对转座作用的分子基础的研究则是 60 年代末 70 年代初在大肠杆菌和低等真核生物(酵母菌)上完成的。

在大肠杆菌中发现有两类可转移遗传因子:插入序列(IS)和转座子(Tn)。二者的不同点仅在于后者除了具有与转座功能有关的基因外,还携带其他功能基因,比如抗性基因等。不同的插入序列的大小不同,DNA 序列不同,转座时要求的靶序列的碱基数也不同。在大肠杆菌中已发现六种不同的插入序列,它们分别表示为 IS1, IS2…… IS5 和 $\gamma\delta$ 。

不同转座子所携带的功能基因不同,按其结构及转座特性又可分为三种类型:IS 型复合转座子, Tn3 型转座子和噬菌体 μ 。

IS、Tn 和噬菌体 μ 的共同特性是,它们都能以一定的机率或多或少随机地插入细菌染色体或其他复制子中,中断插入部位基因的功能,引起基因突变;或者影响邻近基因的表达。表示转座作用的符号为“::”例如 $\text{gal}::\text{Tn10}$ 表示 Tn10 插入大肠杆菌基因组的 gal 操纵子内并引起突变。

由于重组体 DNA 技术的广泛应用,人为的转座作用几乎彼彼皆是。实际上重组体 DNA 分子也可以用“::”表示。例如将 λ 噬菌体基因组的一个已知限制性内切酶 EcoRI 酶切片段 C 重组到质粒 pBR322 上,所得到的杂种质粒也可以记为 pBR322:: λ EcoRI-C。

无论是缺失或转座都有可能得到基因融合突变型。如果相邻近的两个结构基因之间缺失一个片段,这时两个基因的部分结构序列将拼接在一起,出现基因融合。当该 DNA 片段转录为 mRNA 时,融合基因可以产生一个杂种的 mRNA 分子,从而有可能翻译成具有其中一种功能或同时具有两种功能的融合蛋白。表示融合的符号为“ ϕ ”,例如 $\phi(\text{Trp-lacZ})$ 表示由色氨酸合成有关的某一酶与 β 半乳糖苷酶基因形成的一个融合基因,转录方向为 $\text{trp}_p \rightarrow \text{lacZ}$ 。

如果一个基因的遗传调控部位(促进子和操纵子)与任何一个结构基因融合称为操纵子融合,其表示符号基本相同。例如 $\phi(\text{trp-lacZ}^+)$,意即有活性的 β 半乳糖苷酶是在色氨酸操纵子的控制下合成的。

操纵子融合可以用来研究一些不便检测的操纵子表达的调控机制。上例中可以通过 lac Z 基因的表达来研究 trp 操纵子的表达控制规律。

六、其 他

此部分包括原来不是细菌染色体上的基因,如性因子,其他质粒或噬菌体等,它们通过整合,转移到细菌染色体中。

1. 表示细菌性别状态采用的符号

F^- 表示无 F 因子(性因子)。

F^+ 表示携带质粒状态的 F 因子。

F' 表示携带有宿主染色体片段(基因)的 F 因子。

Hfr 表示所携带的 F 因子以整合在细菌染色体上的状态存在。

2. 表示携带原噬菌体或质粒

E. coli K12 (λ) 表示该菌株携带 λ 原噬菌体; *E. coli* K12 (pBR 322) 表示携带质粒 pBR 322。

第三节 大肠杆菌的遗传性和变异性

遗传性和变异性是一切生物的基本属性。大肠杆菌作为一个物种而区别于其他物种,在形态、生理、生化、遗传特性上是很稳定的。然而当我们仔细研究它的表现型特征时,发现这种稳定性又是相对的。

一、大肠杆菌的变异性

各种生物的特性都蕴藏在其遗传物质——DNA 分子的一级结构中,由于 DNA 的半保守复制机制才使得各种生物的特性得以世代相传,表现为遗传性的稳定性。即使在自然状态下 DNA 的复制也会因为个别核酸碱基的互变异构体的出现,或正在复制中的 DNA 双链出现局部瞬时环突,而引起的碱基的置换或个别碱基的缺失和插入,从而导致基因的自发突变(图 1.2)。诱变剂的作用只是对这类变化起了促进作用而已。

当我们检测不同基因的稳定性时,会发现在一定条件下不同基因表现出不同的突变率(表 1.2)。可见基因的自发突变率(即在自然情况下基因在细胞经历一次细胞分裂时发生突变的机会)通常是最低的。

如果以表现型特征的改变,而不是以特定基因的突变为依据来考查突变率时,突变率虽然是相当恒定的,但这一表现型相似的特征的变化可能是所研究的性状有关的某一个基因突变的结果。如以大肠杆菌由不依赖组氨酸 (His^+) 突变为依赖组氨酸 (His^-) 的表现型改变为例,已知大肠杆菌组氨酸的生物合成过程是由前体 5-磷酸核糖焦磷酸开始经相继的 9 个反应步骤完成的。 $His^+ \rightarrow His^-$ 的突变率 (1×10^{-6}) 实际上是合成途径有关的 9 个基因的所有可能发生突变的位点的突变效应的总和。由 $His^- \rightarrow His^+$ 的突变率 (3×10^{-8}) 则不同,它通常只反映组氨酸合成途径的与某一特定反应步骤有关的基因的回复突变的结果。这也可以解释正向突变 ($His^+ \rightarrow His^-$) 比回复突变 ($His^- \rightarrow His^+$) 的突变率高 300 倍的原因(表 1.2)。

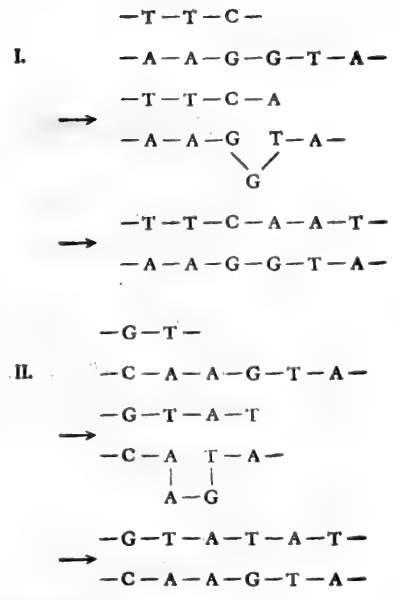
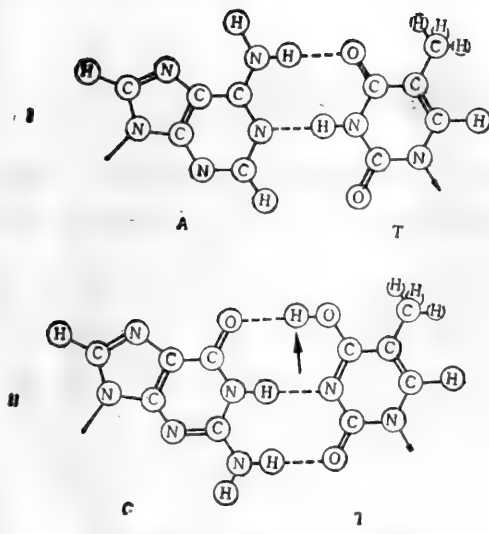


图 1.2 自发突变的机制。

左图 I 示出正常情况下酮式构型的胸腺嘧啶与腺嘌呤配对；II 为烯醇式构型的胸腺嘧啶与鸟嘌呤配对，从而引起转换突变。右图为正在复制中的 DNA 由于瞬息的环突变化引起碱基的缺失和插入。I 示单个碱基环突；II 示两个碱基环突。

表 1.2 大肠杆菌某些基因的自发突变率

发生突变的基因	突 变 率	资 料 来 源
抗链霉素 (1,000μg/ml)	1×10^{-10}	<i>J. Bact.</i> , 57, 565, 1949
抗辐射	1×10^{-7}	<i>Genet.</i> , 32, 639, 1951
抗噬菌体 T3	1×10^{-7}	<i>Genet.</i> , 30, 119, 1945
抗噬菌体 T1	1×10^{-4}	<i>Genet.</i> , 28, 491, 1943
组氨酸缺陷突变 ($his^+ \rightarrow his^-$)	1×10^{-4}	<i>Genet.</i> , 36, 460, 1951
组氨酸回复突变 ($his^- \rightarrow his^+$)	3×10^{-4}	<i>Genet.</i> , 34, 72, 1949
半乳糖不发酵 ($gal^+ \rightarrow gal^-$)	1×10^{-10}	<i>J. Bact.</i> , 85:244, 1963

正因为突变不时发生，所以在实际工作中即使我们的操作十分仔细，也有可能出现菌株丢失或不纯的现象。

对于生物的变异现象要有所分析，首先要将遗传性变异(基因型变异)和环境变异引起的(对环境的适应)这两类不同的变异现象区分开来。例如我们经常做大肠杆菌的转化实验，为了得到高感受态细胞，除了要掌握好菌的生长期外，还要控制好培养条件。通常我们都采用 LB 培养基，并在 37℃ 通气培养，便可达到目的。如果一切条件都不变，而只是在 30℃ 下通气培养；或者虽在 37℃ 通气培养，但不注意菌的生长期，都不能得到良好的感受态细胞。这是一种变异现象，但它并不是遗传性变异，而是环境变异引起的。因为当用上面所得到的感受态不好的细菌悬液重新接种并在 37℃ 下培养，控制生长期，仍可得

到良好的感受态细胞。这说明在实验中操作规范化的重要性，只有操作规范化才能得到可重复的实验结果。

二、大肠杆菌的菌株及其特性的鉴定

通常在进行某一细菌实验时并不需要考虑其在分类上的地位，只要选用一定实验目的特定菌株即可。当手头或搜集保藏中没有自己所需要的菌株时，可以自己动手筛选或者构建某一菌株。

上述所指的菌株就是携带一定遗传标记的克隆。两个克隆之间只要有一个遗传标记不同，它们便属于不同菌株。例如正常的大肠杆菌能发酵乳糖 (Lac^+)，但可以从这个细胞群体中分离得到不发酵乳糖的突变型克隆 (Lac^-)，这时将它们分别表示为大肠杆菌 lac^+ 和大肠杆菌 lac^- 两个菌株，所以菌株是以特定的特性或性状来区分的。

为了达到预期的实验结果，不仅在实验开始前要选用合适的菌株，而且在实验过程中要防止菌株的混杂或丢失。

如何选择合适的菌株呢？例如研究 $recA$ 基因对紫外线致死效应的影响，就应当选用 $recA^-$ 和 $recA^+$ 两个菌株做实验。为了排除其他基因对实验结果的干扰，二者还应是等基因菌株，即所选的菌株除了要研究的个别或几个基因彼此不同外，它们的遗传背景应尽量相同。否则所得结果将缺乏可比性。例如图 1.3 中的大肠杆菌 HB101 和 1046，虽然都是 $recA^-$ ，但对紫外线辐射的反应能力不尽相同，后者比前者敏感得多，这主要是它们的遗传背景不同的缘故。

构建上面所讲的等基因菌株是比较麻烦的。严格地讲，经诱变筛选得到的 $recA^-$ 菌株，未必是 $recA^+$ 野生型菌株的等基因菌株。所以，为得到相应的 $recA^-$ 等基因菌株，就需以经诱变筛选得到的 $recA^-$ 菌株为供体，通过转化或转导使 $recA^+$ 野生型菌株变成为 $recA^-$ 菌株。这样得到的 $recA^-$ 菌株，一般说来除了 $recA$ 基因不同外，其遗传背景才是相同的。

由于大肠杆菌已被人们广泛而深入地研究过，有十分丰富的菌种保藏。但是，由于基因突变等原因，即使从工作细致的实验室索取的菌株，也难免因回复突变或混杂等原因而出现差错，所以在正式用作实验菌株之前，仍应对其有关特性一一进行鉴定。各类突变基因的鉴别方法如下。

对药物抗性标记不仅需要检测其抗性标记是否存在，还应确定其抗药水平，即测定药物对待测菌株的最低抑菌浓度。方法是先配制一含较高浓度抗生素的液体培养基，然后再作 $(\frac{1}{2})^n$ 的系列稀释，分装成含不同浓度抗生素的培养液的系列小管。将待测菌以终浓度约 10^3 (菌落形成单位)/ml 接入各管中， 37°C 温箱培养 18—24 小时后，观察生长情况 (图 1.3)。最低抑菌浓度应为出现生长的那一试管中抗生素浓度的二倍。检测时要以敏感菌为对照。

对营养缺陷型标记的检测可按 J. Miller 主编的《分子遗传学实验》有关部分进行。

限制修饰系统缺陷突变型的检测需借助于已知的限制作用和修饰作用的突变型和噬菌体进行。例如已知大肠杆菌 C 菌株不带有大肠杆菌 K12 菌株的限制修饰系统，所以是

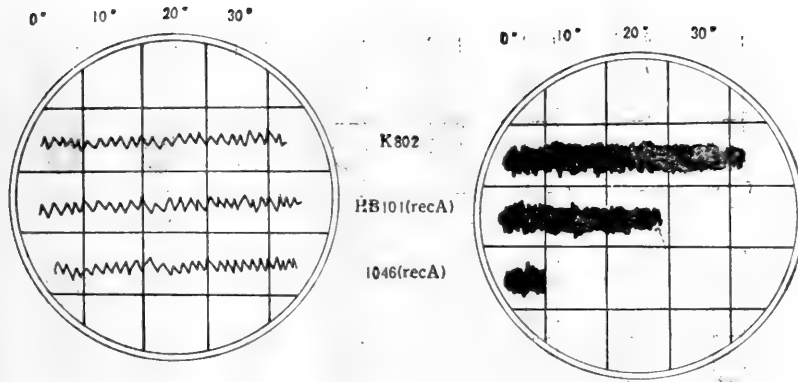


图 1.3 利用 *recA* 菌对紫外线的敏感特性来鉴别大肠杆菌重组缺陷突变型。图中大肠杆菌 K802 为野生型对照菌株。*recA* 菌株因对紫外线诱发的损伤无修复能力,而变得对紫外线很敏感。只要紫外线剂量合适,便可得到右侧的生长图相。HB101 和 1046 对紫外线敏感性的差异,反映出二菌株在遗传上的差异。

表 1.3 大肠杆菌不同菌株的限制修饰作用

λ 噬 菌 体	寄 生 菌 株		
	A. 大肠杆菌 K12 ($hsdR_K^+$ $hsdM_K^+$)	B. 大肠杆菌 K12 ($hsdR_K^-$ $hsdM_K^+$)	C. 大肠杆菌 ($hsdR_K^-$ $hsdM_K^-$)
λ_A	1	1	1
λ_B	1	1	1
λ_C	10^{-4}	1	1

表内数字表示平板效率。

λ_A 、 λ_B 和 λ_C 分别代表由寄生 A、B 和 C 产生的 λ 噬菌体。

大肠杆菌 K12 限制性内切酶和修饰酶基因的双重缺陷型 ($hsdR_K^- hsdM_K^-$)。以大肠杆菌 C 菌株为寄生繁殖产生的 λ 噬菌体,其 DNA 自然未被 $hsdM_K^+$ 基因产物修饰。所以,当以这种噬菌体感染大肠杆菌 K12 $hsdR_K^+ hsdM_K^+$ 菌株时,将不出现或极少出现 (10^{-4}) 噬菌斑。这是因为大肠杆菌 K12 菌株的 $hsdR_K^+$ 基因产物将进入细胞的未被修饰过的 λ 噬菌体 DNA 降解了。如果同样的噬菌体感染大肠杆菌 K12 $hsdR_K^- hsdM_K^+$ 菌株,则出现大体与以大肠杆菌 C 为指示菌时的噬菌斑数相等。这样便可推测出待测菌株的限制修饰系统的基因型。

重组缺陷突变型菌株 (*recA*⁻) 的检测可通过两种方法进行:第一种方法是利用重组缺陷菌株对紫外线辐射的高度敏感性。具体做法是将缺陷型与野生型菌株划平行线接种在平板上,用紫外线进行剂量梯度照射后放在温箱内培养过夜,观察生长情况。如果紫外线剂量合适出现如图 1.4 的图像。此法的优点是简便易行。

第二种方法是通过测定其同源重组能力的有无。比如大肠杆菌 LE 392 和大肠杆菌 1046 分别为 *recA*⁺ 和 *recA*⁻ 菌株。若以大肠杆菌 RR1 为宿主制备的 P1 噬菌体裂解物转导这两个菌株,并筛选链霉素抗性转导体 (transductant)。结果可以得到大肠杆菌 LE392

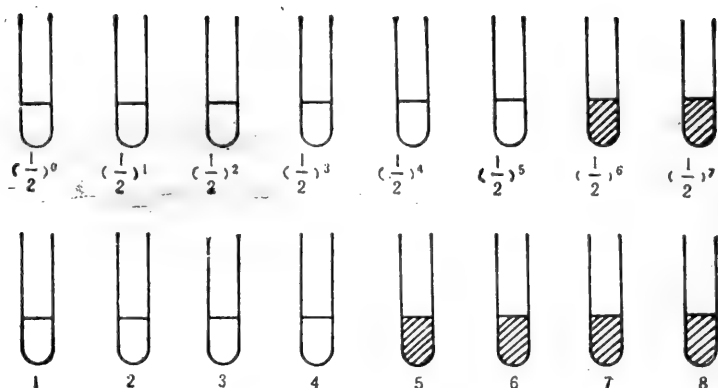


图 1.4 抗生素最低抑菌浓度 (MIC) 的测定。

$(\frac{1}{2})^0 \rightarrow (\frac{1}{2})^7$ 表示在培养基中抗生素的浓度。第一排为对照(敏感菌),第二排为待测菌。在培养 18—24 小时后,对照菌第 8、7 管生长;实验菌在第 8、7、6、5 管生长,所以 MIC 分别为 $(\frac{1}{2})^7$ 和 $(\frac{1}{2})^5$ 。

转导体,但不能得到大肠杆菌 AB 1157 的链霉素抗性转导体。说明后者是重组缺陷型菌株。有关普遍性转导的具体方法可参阅盛祖嘉主编的《微生物遗传学实验》一书。

一个菌株如果被其他杂菌污染,一般可以通过划线或稀释分离单菌落等微生物学技术来解决。在一个被杂菌严重污染以致大肠杆菌已成为劣势菌种的情况下,可采用肠细菌的鉴别培养基,如麦康凯 (McConkey) 培养基。在这种培养基上大多数革兰氏阳性细菌都不能生长,因此有利于肠细菌的分离。

三、大肠杆菌菌株的保存方法

菌种的变异主要发生在传代过程中,静息细胞发生突变的机率是非常低的。因此为了防止获得的菌株丢失,需及时将它们保存起来。保存方法可分为短期、中长期和长期三种。

1. 短期保存

所谓短期保存是指 1—3 个月时间内的保存。该法适用于正在使用中的工作菌株。通常采用斜面或划线分离的单菌落平板置 4℃ 冰箱保存。

短期保存的期限除了受培养基及保存条件影响外,也与菌株的特性有关。大肠杆菌 K12 菌株作为实验菌种已在人工条件下生活了 50 多年,生活力已经减弱,保存期限一般不超过三个月,即每隔三个月应转换一次,否则保存的菌株将因死亡而丢失。许多突变型菌株生活力更低,例如大肠杆菌 X1776 的斜面保存期不能超过一个月。

用于实验工作的短期保存最好采用划线分离的单菌落平板,这样一方面可达到短期保存的目的,同时每次实验从单一菌落开始,可以减少菌株污染或异质性而干扰实验结果的可能性。

2. 中长期保存

为弥补短期保存的缺陷,而又不具备长期保存的条件时,可采用中长期保存。保存方法有二。

(1) 穿刺琼脂保存。将无糖营养琼脂培养基(见附录一)分装于3或5ml的小型螺帽管,培养基高度约为瓶高的 $\frac{3}{5}$ 。灭菌后,用接种针将待保存菌种针刺接入管内,37℃培养过夜,然后以蜡膜或石蜡密封管口,室温保存。此法可保存1—3年。

(2) 低温甘油保存。将待保存菌株转接LB培养基,待生长至对数期,取等量悬液加入装有甘油的无菌小螺帽管内(50%甘油),混匀,密封后置-20℃低温冰箱冷藏。此法简单易行,可保存1—2年。

(3) 长期保存。方法同(2)。取菌悬液加入装有甘油的小螺帽管内(10%甘油),混匀后立即置干冰-丙酮浴冷冻,置-70℃超低温冰箱保藏。此法可保藏5—10年或更长。如有超低温冰箱,这是十分可取的方法。

如有冷冻干燥设备,采用冻干法亦可达到长期保存效果。

在实际工作中往往是短期保存与中长期或长期保存结合使用。

第四节 大肠杆菌的营养和生长

对大肠杆菌的营养要求和生长规律的了解是选择适当的人工培养基和控制实验菌株的生理状态的必要前提。

一、大肠杆菌的营养

生物都能从外界环境中摄取营养,并按其固有的方式,同化所吸收的营养物质构成它们的身体,并完成生长和繁殖过程。不同微生物因长期适应不同的生活环境而有不同的营养要求。大肠杆菌属腐生性细菌,进化过程中因长期适应于多变的生活条件,营养要求比较简单。野生型的大肠杆菌可以在无机盐及仅含一种有机碳源的合成培养基上生长。

合成培养基中的无机盐,一方面提供组成细菌菌体的元素,同时也起着维持渗透压平衡和pH的作用。

虽然大肠杆菌能在营养成分简单的合成培养基上生长,但为了一定的实验目的,人们常常利用天然成分组成的有机培养基来培养它,最常用的有机培养基便是LB培养基。

野生型大肠杆菌可以在合成培养基上正常生长繁殖,而营养缺陷型只能在合成培养基中加入所需物质(氨基酸、维生素或核酸碱基等)才能生长。LB培养基的营养成分比较完全,无论是野生型还是缺陷型菌株一般都可以生长。因为LB培养基中含有各种营养物质,因此当大肠杆菌在这种培养基上生长时,其大分子(蛋白质、核酸等)的合成无需从氨基酸、碱基等小分子成分的合成开始,所以生长速度要比在合成培养基中生长时快得多。尽管如此,在筛选缺陷型,或用同位素标记法研究蛋白质、核酸的合成等,仍须使用合成培养基。

二、大肠杆菌在固体培养基上的生长

在固体培养基表面,一个细菌细胞经过分裂最终将形成一个菌落。由于一个菌落中的 10^6-10^7 个细胞都是同一祖先细胞的后代,所以它实际上就是一个克隆。

在实验中为了进行活菌计数、培养物的分离纯化或进行菌落形态结构研究,都需采用固体平板培养法。

1. 活菌计数

实验中人们经常采用光密度法或比浊法估算菌悬液的菌体浓度。之所以能这样做是因为事先做过光密度或浊度与活菌数的相关曲线。可见固体平板法乃是光密度法或比浊法的基础。

在许多实验中,我们关心的不是总的细胞浓度而是活细胞浓度,这只能用固体平板法才能测定出来。固体平板法的具体做法是,将菌悬液作 1:10 的系列稀释,取样涂平板,然后经一定时间保温培养,取菌落数在 100—400 之间的平板进行计数,再按稀释倍数换算成原菌悬液的活菌浓度。例如取 0.1ml 10^{-3} 稀释度的菌悬液平板计数平均为 160 个菌落,那么原菌悬液的活菌浓度推算为 $1.6 \times 10^8/\text{ml}$ 。

在采用固体平板法测定活细胞数时,计数的准确性是有限的,这是因为取样有误差的缘故。在上述例子中标准差等于菌落计数的平方根,即等于 $\sqrt{160} = 12.6$ (或 7.9%)。即实际菌浓度在 $1.6 \times 10^8 \pm 1.3 \times 10^7/\text{ml}$ 范围内。为了使计数准确,除了需做若干重复平板外,还应注意所用的器皿要清洁,操作要严格。

2. 培养物的分离纯化

这里所说的纯化有两重含意:一种含意是指培养物是否被杂菌污染,即从物种概念上讲是否还是一个纯种。这可以从分类特征来鉴别。另一种含意是指所保存的菌株在遗传上是否还是一个克隆,即所有细胞是否仍具有相同的特性(即相同的基因型)。为了解决这两个问题都需采用固体平板法进行单菌落分离鉴定。具体做法可以用稀释分离,也可以进行平板划线分离(图 1.5)。为了达到良好的分离效果,通常采用三区划线法:第一

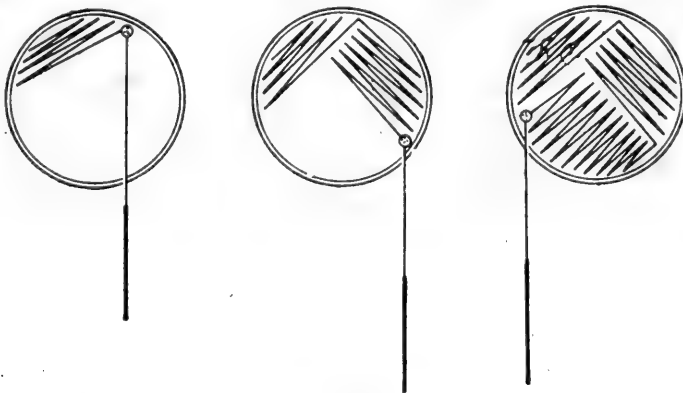


图 1.5 平板划线纯化法。

开始划第二区和第三区前应对接种环进行火焰灭菌,这样在第二区尤其是第三区将出现许多单菌落,达到分离纯化的目的。

区细菌密度很高,而且划完第一区后,接种环上仍携带大量待分离的细菌,所以应用火焰灭菌后,再由第一区划线的终点开始作第二、三区划线。这样操作的结果往往是第一区表现成片生长,而第二区、第三区出现许多单菌落。

如果保存的菌株被污染,那么至少要进行连续两次纯化操作,并同时对该菌株的遗传标记一一进行检验,将符合原菌株特性的克隆重新保存起来。如果纯化后原菌株的遗传标记已部分或全部丢失,那么此菌株也就不存在了。

三、大肠杆菌在液体培养基中的生长

对生物化学研究来说,细菌生长常被定义为细胞重量(生物量)的增长,而对遗传学研究来说,细胞数的增加是更为合适的指标。在恒态生长时,二者是一致的,而当环境条件变化时,两者就有所不同。下面着重讨论以细胞数量变化来研究生长的方法。

(一) 生长周期

大肠杆菌细胞呈杆状,在生长过程中杆状体的直径保持不变,通过杆的延长,然后由杆的中央一分为二,成为两个等价的子细胞。新产生的子细胞继续延长,当达到原长度的两倍时,它们再次一分为二,成为四个子细胞(图 1.6)。

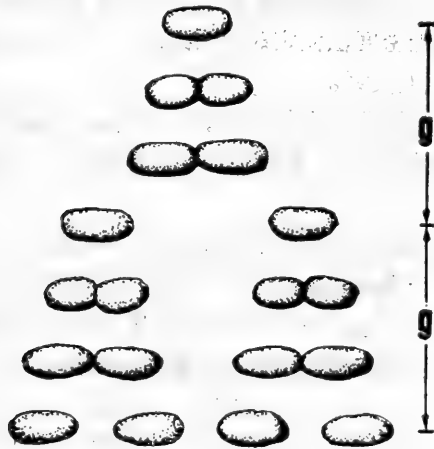


图 1.6 细菌的生长和繁殖方式。

世代时间 g 是指一个细菌自其形成到它分裂成二个细胞之间的时间间隔。

在最适条件下,细菌的生长呈指数特征,被称为对数生长期。然后终因营养条件的限制或因自身代谢产物的抑制作用,而更经常的是二者的共同影响使生长速率减缓,以至最终停止生长,达到了平衡期。

细菌细胞在由对数生长期转到平衡期的过程中,细胞的代谢活性、结构成分等发生许多深刻的变化,细胞体积也变得比较小。大肠杆菌在生理生化方面的许多变化是相似的。

当以平衡期细胞转接新鲜培养基时,细胞并不立即进入活跃的生长期,而要经历一个细胞重量增加而不表现细胞数量变化的时期——延滞期。这是一个适应时期,细胞内进行着活跃的代谢,并在环境中积累对生物合成所必需的二氧化碳。

在一定的培养基及培养条件下,细菌的世代时间是相对恒定的,可由图 1.7 直接求出。

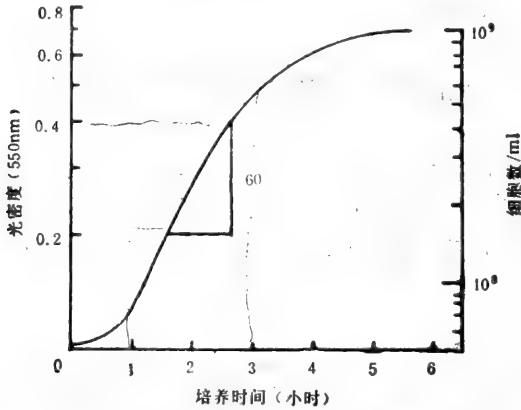


图 1.7 细菌在液体培养基中的生长曲线。

停滞期、对数生长期和平衡期分别为 0—1 小时、1—3 小时和 3 小时以后。图中世代时间为 60 分钟。

在实际工作中常用 490—550nm 波长范围内的光密度值表示菌的生长情况。采用这个范围的波长是因为波长越短光散射作用越大,而波长低于 490nm 时由消毒产生的黄色物质的吸收作用增强。采用光密度法测定时,菌浓度在 0.01mg/ml(约 10^7 细胞/ml)至 0.5 mg/ml 范围内与光吸收成线性关系。

由生长曲线很容易判断出细菌的生长阶段,以指导我们的实验工作。同时也便于直观判断细菌在特定条件下的世代时间(图 1.7),从而预期经历一定时间间隔的光密度或细胞数。

细胞处在不同生长期具有不同的生理状态,因此在实验中应十分注意掌握菌的生长期。为了得到大肠杆菌的感受态细胞,以取中对数生长期细胞为宜。平衡期细菌几乎不能得到感受态细胞。为了提取染色体 DNA 则以取晚对数期细胞为宜,因为这时不仅可以得到足够量的细胞,而且细胞对溶菌酶也较敏感,容易得到质量好的染色体 DNA。提取质粒则常采用平衡期细胞。一般采用过夜培养物,但培养时间不宜过长,例如不要采用超过 24 小时的老培养物,因为培养时间过长的老培养物中含有许多死亡细胞,并且由于自然选择的缘故,在一些存活细胞中质粒可能自发消除或发生其他结构上的变化。

(二) 对数生长期的动力学

细菌在经历延滞期后,随即便转入对数生长期,这时细菌细胞数的增加速率可用下式表示:

$$dN/dt = KN \quad (1)$$

或
$$N_t = N_0 e^{kt} \quad (2)$$

式中 t 代表时间; N_t 为时间 t 时的菌数; N_0 为实验开始时 (t_0) 的菌数。可见在对数

生长期,细菌数的增加是时间的指数函数。因此,若以菌数的对数对时间作图便可得到一条生长曲线(图 1.7)。

如果菌数由 N_0 增加至 $2N_0$ 所需时间为 g , 则

$$2N_0 = N_0 e^{Kt}$$

$$Kg = \ln 2, \quad K = \ln 2/g \quad (3)$$

如果以世代时间 g 为时间单位, 则(3)可变为

$$K = \ln 2$$

代入式(2), 于是

$$N_t = N_0 e^{t \ln 2}$$

$$\ln N_t = \ln N_0 + t \ln 2$$

$$N_t = N_0 2^{t/g} \quad (4)$$

如果某一大肠杆菌培养物在最适条件下的世代时间为 20 分钟 ($g = 20'$), 培养 3.33 小时后, 细菌将经历 10 个世代, 这时一个细胞将产生 1,024 个细胞, 这个数字可由式(4)直接算出:

$$N_t = N_0 2^{t/g} = 1 \times 2^{3.33 \times 60 / 20} = 1 \times 10^{3.33} = 1,024$$

同样, 我们还可以作进一步运算, 例如将大肠杆菌培养在 M9 培养液中, 其世代时间约为 60 分钟, 即 3 倍于 LB 培养液的世代时间。因而由一个单细胞形成一个可见的菌落(约 10^6 — 10^7 细胞)约需培养 20 小时。如果一个大肠杆菌细胞的平均体积为 $1 \mu^3$ 或 10^{-12}cm^3 , 而地球的体积等于 4×10^{39} 个大肠杆菌细胞。一个指数生长的大肠杆菌要达到这样大的体积也只需要 45 个小时! 然而这种事情是永远不会发生的。

参 考 文 献

- [1] 盛祖嘉, 微生物遗传学, 科学出版社, 1986。
- [2] 盛祖嘉, 微生物遗传学实验, 人民教育出版社, 1982。
- [3] Jeffrey, H., *Miller Experiments in Molecular Genetics*, published by Cold Spring Harbor Laboratory, 1972.
- [4] Maniatis, T. *et al.*, *Molecular Cloning*, published by Cold Spring Harbor Laboratory, 1982.
- [5] Davis K. W. *et al.*, *Advanced Bacterial Genetics*, published by Cold Spring Harbor Laboratory, 1980.
- [6] Davis, Bernard, D. *et al.*, *Microbiology (third edition)*, 60—70, 646—672, published by Harper and Row, 1980.
- [7] Stephen, J., O'Brien (Editor), *Genetic Maps, Vol. 3*, published by Cold Spring Harbor Laboratory, 1984.

第二章 基因工程的载体体系

杨善蓉 强伯勤

第一节 质粒载体

质粒是基因工程重要的载体之一。由于它具有独特的性质,随着基因操作技术的迅速发展,已成为分子生物学领域内倍受关注和研究异常活跃的对象。

质粒(plasmid)一词最早是 Lederberg 于 1952 年提出的,原意为任何细胞质中能自主进行复制的遗传因子,包括真核细胞中的线粒体和质体(plastid)等。现在的所谓质粒一般只指细菌胞质内的独立复制单位。研究得最早的质粒是决定细菌性别的 F 因子,对其结构、行为、性质的研究大大推动了微生物遗传学的发展。Lederberg 也因这项研究而获得诺贝尔奖。

50 年代末期,在细菌抗药性的研究中,发现了对抗菌素耐药的 R 质粒。这一发现,使人们对质粒的认识在质和量上都发生了根本变化。质粒所带有的抗药性基因使细菌能抵御抗菌素的杀菌作用而存活下来,从而揭开了耐药菌株之谜,于是质粒成为与人类日常生活密切有关的因子,对它产生了极大的兴趣。从分子水平上对质粒的基本结构、存在方式、复制方式等基本性质进行了大量研究,为 70 年代利用质粒作为重组 DNA 技术的重要载体打下了坚实的基础。

一、细菌质粒的基本概念

(一) 定义

现在的所谓质粒是指细菌细胞质中独立于细菌“染色体”之外能自主地进行复制的遗传单位,随着细菌分裂,它能稳定地把自己的遗传特性传递给后代细胞。

质粒的本质是环形双链 DNA。在空间位置上,它不在细菌“染色体”上(这里的“染色体”即指宿主 DNA 本身,因为原核生物没有核的结构,习惯上把功能上表现为连锁的细菌基因结构称作细菌“染色体”)。质粒在细胞质内,是细胞质复制子。但是,染色体外的核酸分子不一定是质粒。

由天然材料获得的质粒一般分子量较小,分子量在 $2-5 \times 10^6$ 左右,而拷贝数较多;也有分子量较大($5-10 \times 10^7$)而拷贝数甚少的大型质粒。当质粒大小接近宿主染色体大小时,如达 2×10^8 以上,这种质粒就不稳定了。

质粒对宿主细胞没有干扰,这与病毒、噬菌体不同。当细胞分裂时,质粒就将它的拷贝数准确地分配给每个子细胞。尽管质粒的复制是与染色体分开进行的,但一般情况下,它们是同步的,并且具备一套保持拷贝数恒定的机制。由此,我们对细菌质粒有了一个基本的认识。

(二) 质粒分子的形状、大小及拷贝数

迄今所知,质粒不论大小,毫无例外都以环形双链 DNA 的形式存在。分子链末端各自以共价键的形式和自己的另一粘末端相结合。这种环形 DNA 分子具有超螺旋的性质,发生扭曲,使分子处于紧张状态。一般有两种状态:未经切割能够产生(或已经产生)紧张性的分子,简称为闭合环(cc)或共价闭合环(ccc);一条链上发生断裂(不限于一处)而形成的松弛形分子叫做开环(oc)。由生物体内提取得到的质粒分子,往往两种形式同时存在。

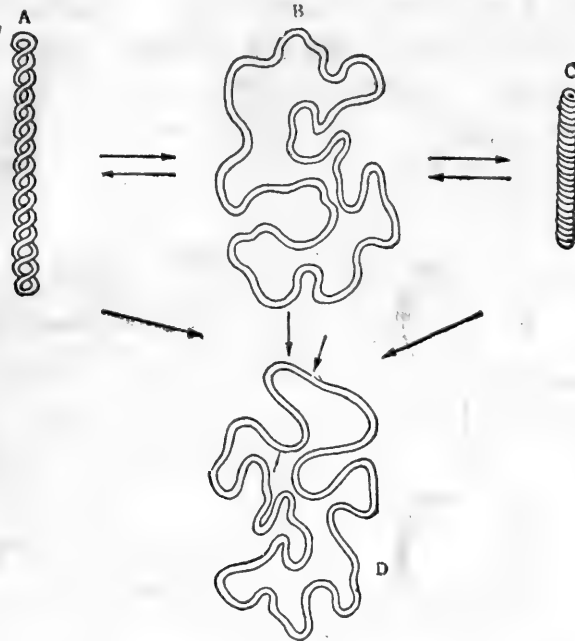


图 2.1 质粒 DNA 分子形状的模式图。

A、C. 分别表示处于方向相反的紧张状态的分子。B. 表示两条链闭合但不存在紧张性的分子。D. 一条链被切断而松弛的开环(oc)分子。

由于 ccDNA 和 ocDNA 结构上的差异,导致了性质的差别。例如吖啶橙、溴乙锭等染料分子能嵌入双链碱基对之间时,具有 cc 结构的分子,由于分子紧压,流体力学阻力比 oc 分子要小,因此插入染料量少。另外 ccDNA 在琼脂糖凝胶电泳时泳动度相对要快,离心时沉降也快,难以随 pH 或温度上升而发生解链,需要在更高的 pH 或更高的温度下才开始变性。在变性过程中如回复到原有条件,即迅速再复性。

质粒 DNA 的大小可以用电镜、超离心或琼脂糖凝胶电泳等方法进行测定。但不管用哪种方法,都必须加上已知分子量的 DNA,以校正测定值。在精确测定分子量的基础上,同时又知道它的含量,就可以算出拷贝数。最一般的方法是抽提整个细胞所含的 DNA,用氯化铯-溴乙锭密度梯度平衡离心法确定 ccDNA 占全菌 DNA 的百分率,求得质粒 DNA 与宿主染色体 DNA 的比率。假定大肠杆菌染色体的分子量为 3×10^9 ,质粒 DNA 的分子量为 4.7×10^6 ,质粒 DNA 占全菌 DNA 的 15%,即相当于染色体 DNA 的 $15/(100-15) = 0.176$,则拷贝数 $= 0.176 \times (3 \times 10^9) / 4.7 \times 10^6 = 112$ 个。

(三) 质粒给与宿主细胞的一些标记

质粒带有一些特殊的基因,表现出特有的性状,给宿主细胞带上了某种标记。如对抗菌素产生抗性,能产生肠毒素或溶血素等,能分解芳香族化合物,产生硫化氢,对重金属如 Hg 产生抗性、等等。最常用于基因工程的标记是质粒的抗药性基因。对某抗生素是耐药(存活)还是敏感(致死)可作为筛选是否插入外源基因的标志。那些还不知道给与宿主细胞什么表型的质粒,叫做隐蔽质粒。

例如,对氨苄青霉素 (Ap) 有抗性的质粒,一般写作 Ap^r ;对该药敏感的质粒一般写作 Ap^s 。细菌对氨苄青霉素产生耐药的原因是由于抗氨苄青霉素即 Ap^r 基因的产物是 β 内酰胺酶。此酶存在于细菌的细胞膜与细胞壁之间,壁破之后酶即释放,把 Ap 降解,降解的氨苄青霉素不能再杀死细菌。抗性基因必须完整存在,才能转录出有活性的酶去破坏抗菌素而表现出抗药的功能,如果把抗性基因切断、分开或改变,如接上一段外源基因,则抗性基因就不能表达出有功能的酶而使抗性基因失活,这种现象就叫做插入钝化或插入失活。

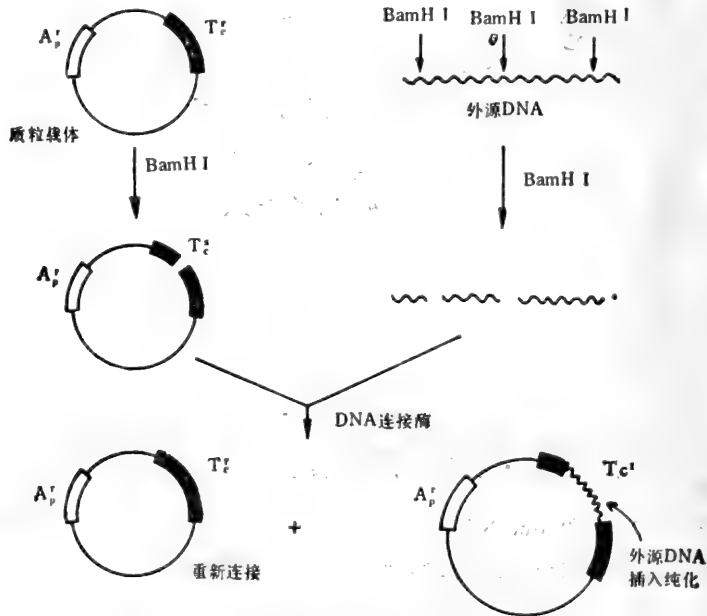


图 2.2 抗性基因插入钝化示意图。

Ap^r : 氨苄青霉素抗性。 Tc^r : 四环素抗性。 Tc^s : 四环素敏感。 $BamH I$ 为限制性内切酶。

重组 DNA 技术可利用质粒抗性基因的插入钝化使细菌由耐药变为敏感的特点来筛选该质粒是否插入了外源基因。

(四) 大肠杆菌中质粒的复制

质粒在大肠杆菌中的复制可以分为两大类。一类叫松弛型控制复制,另一类叫严紧型控制复制。这是由质粒本身复制子的性质决定的。所谓严紧型控制的复制是指拷贝数被限制在少量(几个)的复制方式,而松弛型控制复制的拷贝数可以很多(100—200 拷贝/

菌)。造成这种差别的主要原因在于质粒复制时是否需要大肠杆菌蛋白的不断合成,这些蛋白是与大肠杆菌自我复制有关的酶,如 RecA 等。如果质粒复制需要 RecA 等酶的不断合成,则为严紧控制复制,其复制拷贝数必然少,反之,则为松弛型控制复制,拷贝数多。

氯霉素是蛋白质合成的抑制剂。当菌生长到一定程度,一般是指对数生长期后期,如 1×10^9 细胞 /ml, 加入氯霉素,由于严紧型控制复制需要有关复制蛋白的不断合成,而蛋白质合成受抑,染色体与严紧型质粒的复制随着也停止。但松弛型质粒的复制不受大肠杆菌这些复制蛋白的影响,以至经过 10—12 小时,每个细胞最多能生成 1,000—3,000 个质粒,这样每个细胞中有 50% 以上的 DNA 是质粒 DNA。质粒拷贝数大大超过大肠杆菌的个数。质粒大量增加的过程就叫做质粒 DNA 的扩增。现在用作载体的质粒绝大多数是带有 ColE1 复制子的衍生物,是松弛型控制复制。例如 ColE1 质粒是一个自然发生的质粒,能产生大肠杆菌素 E1, 也携带对大肠杆菌素 E1 有免疫性的基因,其复制子是松弛型的。严紧型质粒拷贝数不能随意增加,严格受细菌生长的限制,例如质粒 pSC101。这类质粒不常用,只是在需要克隆一个对细菌本身有毒性的外源基因时,可试用 pSC101,因为它的拷贝数少,对细菌毒性相对就小,使细菌不致被有毒性的外源基因杀死。

如果质粒上存在氯霉素的抗性基因,那么扩增时不能采用氯霉素作蛋白抑制剂,可采用壮观霉素,同样有抑制细菌生长质粒扩增的作用。

松弛型控制复制说明质粒的复制与宿主的分裂、增殖相协调或同步的机制完全被破坏,然而为质粒复制所必需的酶系统还是稳定的。有人认为在自然状态下,宿主 dnaA 蛋白控制着 ColE1 的复制,当宿主蛋白质合成被抑制时,这种控制则消除, ColE1 能大量扩增。

(五) 质粒的转移和不相容性

将大肠杆菌在低温下用氯化钙处理,使细胞膜的通透性改变,形成能让质粒 DNA 进入的表面结构,即能发生质粒的转移,这个过程叫转化。这是重组 DNA 时非常有用的技术,一般转化率为 10^{-6} 。如果两种不同质粒在同一细胞内能够稳定地共存,两者的复制控制互不干扰,叫做质粒相容性。例如相容性 P 群质粒是唯一广宿主质粒,它们能在很多细菌属间转来转去。RP₁ (亦称 RP1、R1822、R18) 分子量为 36×10^6 , 作为克隆载体虽然用途不广,但其重组质粒保留了亲代质粒的广宿主性。

如果质粒之间具有相同的复制控制机制,互相产生干扰而不能共存,叫做质粒不相容性。一般在没有选择压力的情况下,两个不同的质粒不能共存于同一宿主细胞内,两种质粒中就有一种完全丧失掉。究竟脱落哪一个,从原则上讲机会是均等的,由其“强弱”决定。这种不相容性已作为传递性质粒分类的依据之一。分子克隆技术正是基于质粒的不相容性质,在只带有一种质粒的细菌中,提取到单一的质粒 DNA,作为载体进行 DNA 重组。

(六) 质粒的消除

有的质粒在自然条件下可以从宿主菌中消失;还有的可在培养细菌时施加适当处理,以促进其消失而获得失去质粒的细菌,这叫做质粒的消除。

如果细菌的某种性状由于质粒的消除而丧失,那么一般可以认为这种性状由质粒控制。通过质粒消除而丧失的性状,决不能发生回复突变,不能通过再感染细菌来恢复原来的性状。由此可进一步证明该性状由质粒决定。

对质粒消除有影响的物质如染料吡啶橙、结晶紫、溴乙锭等,还有利福平、表面活性物质 SDS、紫外线照射等。所以在扩增质粒时要避免这些因素的刺激,防止有价值的质粒自行消除。

二、基因工程中质粒载体的选择

(一) 质粒载体的选择

作为理想的克隆载体应具备以下性质:分子量小,拷贝数多,易转化。有明显的标志,即赋予宿主细胞一些易于检测的表型,供筛选用。对许多限制性内切酶有单一切点,且切点位于易被检测的标志基因上。质粒本身的性质是安全的。

分子量小的优点是:容易从宿主细胞中分离;通常拷贝数多,导致克隆的外源基因量的增加;DNA 碱基对也少,对任何一个限制性内切酶,在质粒上的作用位点相对减少。

质粒 DNA 作为克隆外源基因的载体在形成重组质粒后,必须转化到适宜的受体菌中。由于转化率很低,重组体上必须有一个容易被检测的表型,与其他没有重组的质粒相区别。而这种表型往往由质粒上某一标志基因决定。没有这种标志就无法筛选出哪一个是重组有外源基因的质粒。当然也可以由插入的外源 DNA 上的某基因表现出标志。最有利的情况是外源 DNA 的插入使质粒上易于检测的表型基因失活,也叫做插入钝化,从而有可能把重组子与虽经内切酶切割又自身连接上的质粒区别开。

未经体外构建的“天然”的克隆载体质粒有上面提及的 ColE1、pSC101,此外还有 RSF2124,它们的性质见表 2.1。

表 2.1 用于克隆 DNA 的一些“天然”质粒的性质

质粒	分子量($\times 10^6$)	拷贝数	单一切点的内切酶	选择转化子的标记	复制类型
pSC101	5.8	1—3	EcoRI	抗四环素	严紧
ColE1	4.2	20	EcoRI	产生大肠杆菌素 E1	松弛
RSF2124	7.4	10	EcoRI, BamHI	抗氨基青霉素	松弛

(二) 宿主菌的选择

一般都以大肠杆菌为宿主菌,对宿主菌的要求是:易于培养;质粒容易转化;菌本身安全。

也有以枯草杆菌作为宿主菌的,其优点是:枯草杆菌有芽孢,可用作原核生物分化的模型;为专性需氧的革兰氏阳性菌,与大肠杆菌兼性厌氧的革兰氏阴性菌内环境完全不同;可广泛用于发酵工业;极安全,与人、动物都没有已知的相互作用。

其缺点是:可检测的表型少,只有大量隐蔽质粒;多拷贝,但不为氯霉素扩增;很多抗

性决定子不携带内切酶位点,不能用插入钝化来筛选。

大肠杆菌被作为最常用的宿主菌。常用的大肠杆菌一般属于大肠杆菌 K12。标准菌株为 C600 和 HB101。其中 HB101 是 K12 与 B 杂交的菌株,含链霉素抗性基因,能大量生长。此外,符合生物防范的菌株还有 X1776,如果它们逸散到自然环境中,这个菌株连同它携带的任何质粒能自身破坏。曾作为 pSC101、pCRI、pMB9 及 pBR322 等质粒的宿主菌。

(三) 选择性标记

(1) 抗性基因。质粒上带有对某抗菌素耐药的基因,作为筛选的标志。除了常用的氨苄青霉素抗性 (Ap^r)、四环素抗性 (Tc^r) 或 Tet^r 外,还有氯霉素抗性 (Cm^r)、卡那霉素抗性 (Km^r)、新霉素抗性 (Neo^r) 等。以能否在含有一定浓度某抗菌素的培养基中长成菌落作为观察的标准。这些抗性基因具有某些酶编码的基因,这些酶可对氯霉素、链霉素等药物进行乙酰化、腺苷化、磷酸化或水解等化学修饰,从而破坏这些抗生药的作用,产生抗性。

如果质粒上只含有一个抗性基因,那么当外源基因插入使之钝化后,该质粒就不能抗该抗生素,只能在不含该抗生素的培养基上生长。如果质粒上带有两个抗性基因,外源基因插入钝化了一个抗性基因,保留另一个抗性基因,那么分别在两种抗生素培养基上都能生长的,就不是重组质粒的菌株,而只在未钝化的抗生素培养基上生长的才是真正带有重组质粒的菌株。若带有三种抗性基因的质粒,插入钝化的自由度就更大了。

(2) 菌落的颜色。带有乳糖操纵子的调控区域和基因的质粒在含有 IPTG (Iso-Propyl- β -D-Thiogalactopyranoside, 异丙烯基 β -D-半乳糖苷) 作为诱导物,以 X-Gal (5-Bromo-4-Chloro-3-Indolyl- β -D-Galactoside, 5-溴-4-氯-3-吲哚基- β -D-半乳糖苷) 作为显色剂的培养基中,由于乳糖操纵子中 Z 基因 ($lac Z$) 的产物半乳糖苷酶的释放,切割显色剂使吲哚基显蓝色。结果 lac^+ 的菌株产生蓝色菌落,被破坏的 lac^+ 及 lac^- 菌株产生白色菌落。把外源基因插入 lac^+ 的质粒上,即可通过菌落的颜色来区别重组的质粒和自身连接的质粒。另外还有一种叫麦康凯的培养基(作用原理与上同),其显色剂为中性红。由于 lac 基因的产物引起 pH 的改变显不同颜色。 lac^+ 为白色, lac^- 为红色。

(四) 载体的人工组建

前面提到的 ColE1、pSC101、RSF2124 等均为“天然”质粒,虽然可以作为载体,但是它们都有若干不足,为了达到理想克隆载体所具备的性质,人们花了很大的努力于体外改造已知的质粒,构建了优良的克隆载体,其中应用最广的是 pBR322。

质粒 pBR322 的构建见图 2.3。

pBR322 含有 RSF2124 的抗氨苄青霉素 (Ap^r) 基因和 pSC101 的抗四环素 (Tc^r) 基因,还结合了 pMB9 (一类 ColE1 质粒) 的复制子部分。其前身是 pBR313。pBR322 与 pBR313 的区别在于前者除位于 Ap^r 基因上有一个 PstI 位点,其余的 PstI 位点全部被去掉。因此在 PstI 位点上插入外源基因时,用抗四环素而不抗氨苄青霉素的特性来选择转化菌。

pBR322 是目前研究得最清楚的质粒之一。其分子量较小,为 2.6×10^6 ,含 4,632 bp,

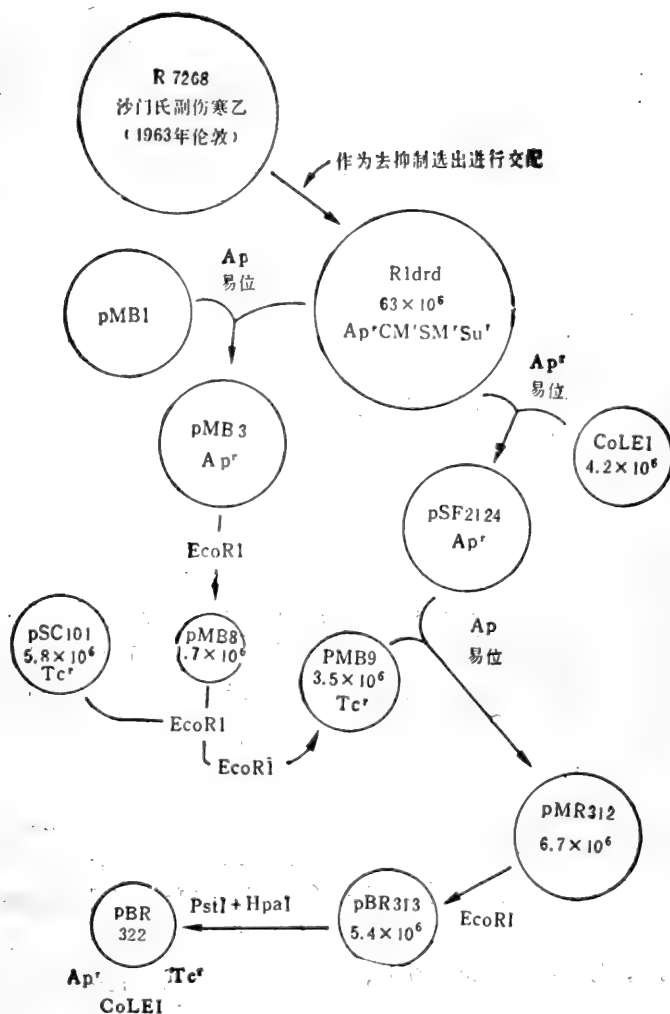


图 2.3 pBR322 的构建。
 Ap^r-氨苄青霉素抗性。Tc^r-四环素抗性。

全部核苷酸序列都已测出(见附录一、二、三)。在 pBR322 上有单一酶切位点的限制性内切酶, 为 *Ava*I, *Pst*I, *Bam*HI, *Pvu*II, *Cla*I, *Sal*I, *Eco*RI, *Hind*III。而 *Ava*X, *Bcl*I, *Bgl*II, *Eca*I, *Hpa*I, *Kpn*I, *Sma*I, *Sst*I, *Xba*I 和 *Xbo*I 在其上无切点。一般以 *Eco*RI 切口处作为“0”点位点。*Pst*I 可以在 Ap^r 上造成插入钝化, *Bam*HI, *Sal*I 可以在 Tc^r 上造成插入钝化, *Hind* III 位点虽然也在 Tc^r 上, 但不一定造成插入钝化, 因为它切在抗性基因非关键的位置上。

构建人工质粒的目的有以下几点: 改变质粒的大小; 引入适当的选择标志; 改变质粒的限制性内切酶的切割位置, 增加适宜的单一酶切点; 如用于测序或基因表达时, 增加有利于进行测序、基因表达的装置。总之, 使质粒改造得容易插入外源 DNA, 容易转化, 容

易筛选重组子以及测序或表达外源基因等。

载体的人工构建使人们有可能随心所欲地得到符合各种需要的特殊基因工程载体，来达到克隆、筛选、测序、表达等目的，为基因工程技术开辟了广阔的前景。

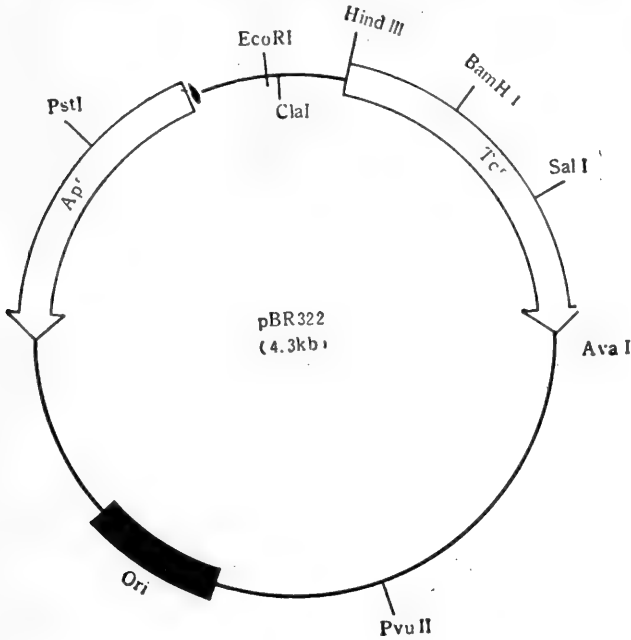


图 2.4 pBR322 结构图。

箭头表示 Ap^r 基因和 Tc^r 基因的转录方向。Ori 表示复制起始点，图中标出了单一切点内切酶的位置。

三、介绍几类质粒载体

(一) pBR322 系列载体

pBR322 系列是以 pBR322 质粒为基础人工构建的一系列载体。pBR322 质粒的构建在上面已经提及。由于其分子量小，拷贝数较多，同时具有 Ap^r 及 Tc^r 两个抗性基因，并有合适的单一酶切位点，因此，它是较理想的载体，被广泛应用。在此基础上目前已改造出十余种载体(见表 2.2)。

由表 2.2 可见，pBR321 与 pBR322 在表型和单一酶切位点上是完全相同的，但分子量大小有差异。pBR324 和 pBR325 大部分表型和内切酶位点也相似，但 pBR325 又引进了一个氯霉素的抗性基因，且分子量变大。pBR341 和 pBR345 又引进了乳糖操纵基因。

还有两个表上没有的质粒，一个是 pBR328，它是在 pBR325 的基础上构建的，分子量为 3.24×10^6 ，略小于 pBR325。利用 EcoRI 单一切点或 PvuII 和 Ball 位点插入外源基因片段，使氯霉素抗性被破坏。另一个是 pBR327。在 pBR327 不同位点上插入 Tn9 氯霉素乙酰转移酶基因分离物(Tn9: chloramphenicol acetyl-transferase gene)，可以得到三个起动机分析质粒，即 pCM-1、pCM-4 和 pCM-7。CAT 基因分离物含有完整的

表 2.2 pBR322 系列人工载体的性质

质粒	分子量($\times 10^6$)	表型	重要的单一酶切位点
pBR312	6.7	Ap ^r , Tc ^r , Ice	EcoRI, HindIII, HpaI, Sall, SmaI
pBR313	5.8	Tc ^r , Ice	BamHI, EcoRI, HindIII, HpaI, Sall, SmaI
pBR317	5.4	Ap ^r , Tc ^r , Ice	BamHI, EcoRI, HindIII, HpaI, Sall, SmaI
pBR318	3.8	Tc ^r , Ice	BamHI, EcoRI, HindIII, HpaI, PstI, Sall, SmaI
pBR320	1.9	Ap ^r	BamHI, BglI, EcoRI, HindIII, PstI, Sall
pBR321	3.2	Ap ^r , Tc ^r	BamHI, EcoRI, HindIII, PstI, Sall
pBR322	2.6	Ap ^r , Tc ^r	BamHI, EcoRI, HindIII, PstI, Sall
pBR324	5.5	Ap ^r , Tc ^r , Cel	BamHI, EcoRI, HindIII, Sall, SmaI
pBR325	3.6	Ap ^r , Tc ^r , Cm ^r	BamHI, EcoRI, HindIII, PstI, Sall
pBR333	1.2	Ap ^r	BglI, EcoRI, EcoRII, HaeII, HindIII PstI
pBR341	5.1	Km ^r (op)	HindIII, Sall
pBR345	0.7	(op)	EcoRI, HaeII

Tc^r: 四环素抗性 Ap^r: 氨苄青霉素抗性, Km^r: 卡那霉素抗性, Cm^r: 氯霉素抗性, Ice: 大肠杆菌 E₁ 免疫性, Cel: 产生大肠杆菌素, (op): 乳糖操纵基因。

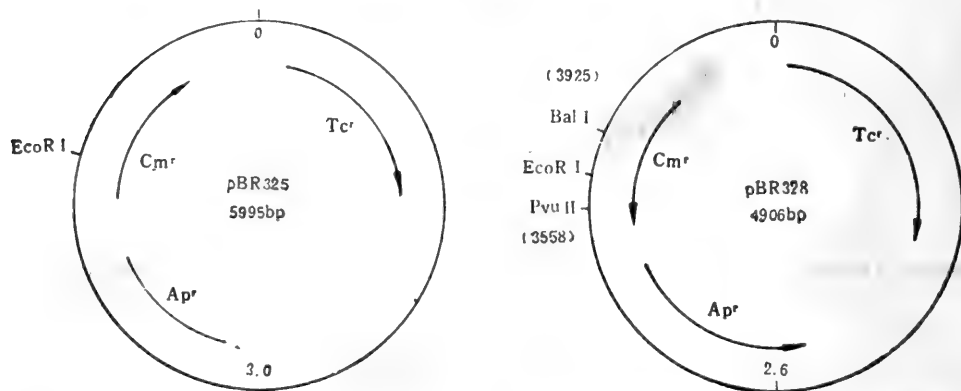


图 2.5 pBR325 和 pBR328 结构图及部分酶切位点。

Cm^r: 氯霉素抗性。Tc^r: 四环素抗性。Ap^r: 氨苄青霉素抗性。EcoRI, Ball 和 PvuII 为限制性内切酶。

多肽编码序列和一个大肠杆菌的核糖体结合位点,但是缺失了最后一个插入序列和 CAT 启动子。这些质粒可以作为细胞内启动子的分析。它们的结构见图 2.6。

pBR322 系列除可用于外源基因的克隆外,还作为基因探针的载体,便于保存,随时可以扩增。

(二) pUC 系列的载体

pUC 系列是 J. Messing 和他的同事于近几年内发展起来的一组对于大肠杆菌体系进行克隆、测序和表达外源基因非常有用的质粒载体系列。它们含有 pBR322 的 Ap^r 基因和复制起始点,并包含 lac 操纵子及 lac Z 基因的部分。在这些质粒的 lac Z 的近 5' 末端区域导入了各种单一内切酶的识别位点,从而得到各种 pUC 质粒。由于 lac Z 基因的存在,在插入外源基因后都能改变菌落的颜色,可用于筛选(见质粒的选择性标志所

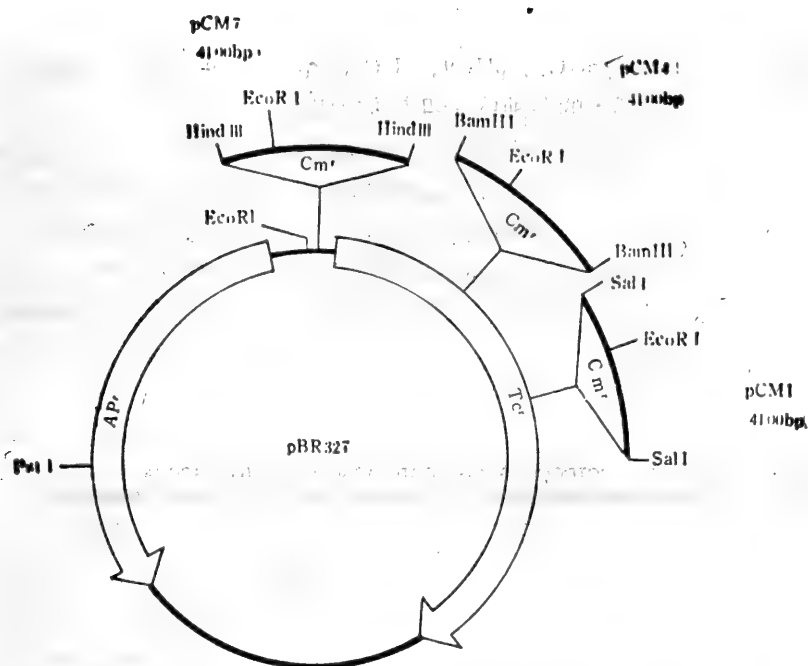


图 2.6 由 pBR327 衍生的 pCM1、pCM4 和 pCM7 质粒的结构。
 Ap^r·氨基青霉素抗性。Tc^r·四环素抗性。Cm^r·氯霉素抗性。

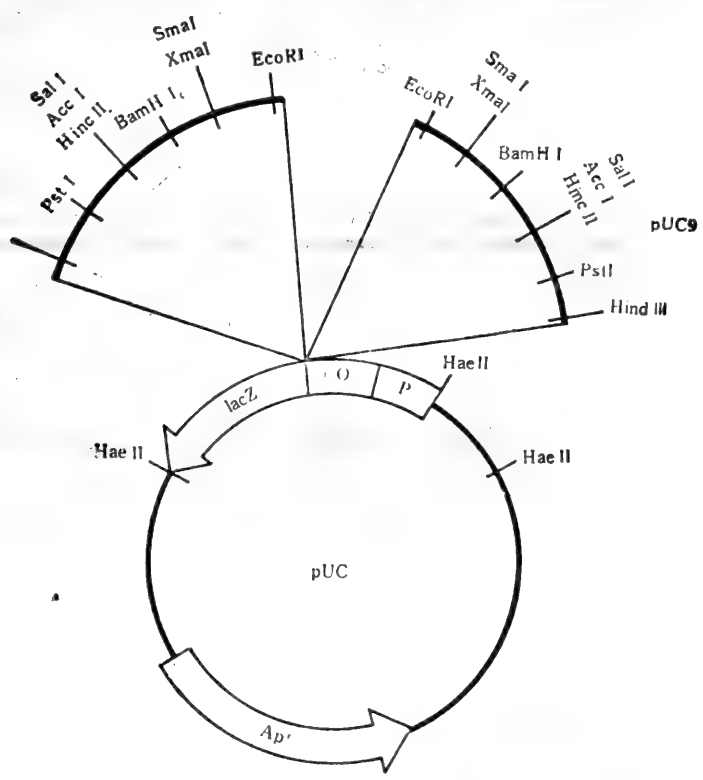


图 2.7A pUC8 和 pUC9 lac⁺区域单一切点酶切图。
 Ap^r·氨基青霉素抗性。lacZ·β半乳糖苷酶基因。O·乳糖操纵子的操纵基因。
 P·乳糖操纵子的启动子。

述)。

常用的 pUC 序列有 pUC8、pUC9、pUC18 和 pUC19。图 2.7 A 表示在 pUC8 及 pUC9 质粒的 lac⁺区域插入的各种单一切点的内切酶位点。

pUC 系列与单链噬菌体 M13 有密切关系,实际上 pUC 的 lac Z 区域中的各种酶切位点就是从相应的 M13 噬菌体的合适部位切下来再重组上去的,如图 2.7 B 表示 pUC-12、13、18、19 分别与 M13mp 10、11、18、19 之间的相应关系。

pUC 系列质粒一般除用于克隆外源基因外,还用于双脱氧末端终止法测定核苷酸顺序工作,作为相应的 M13 的引物。在筛选时还可利用 lac 基因的启动子表达插入的外源基因,作产物的免疫筛选。

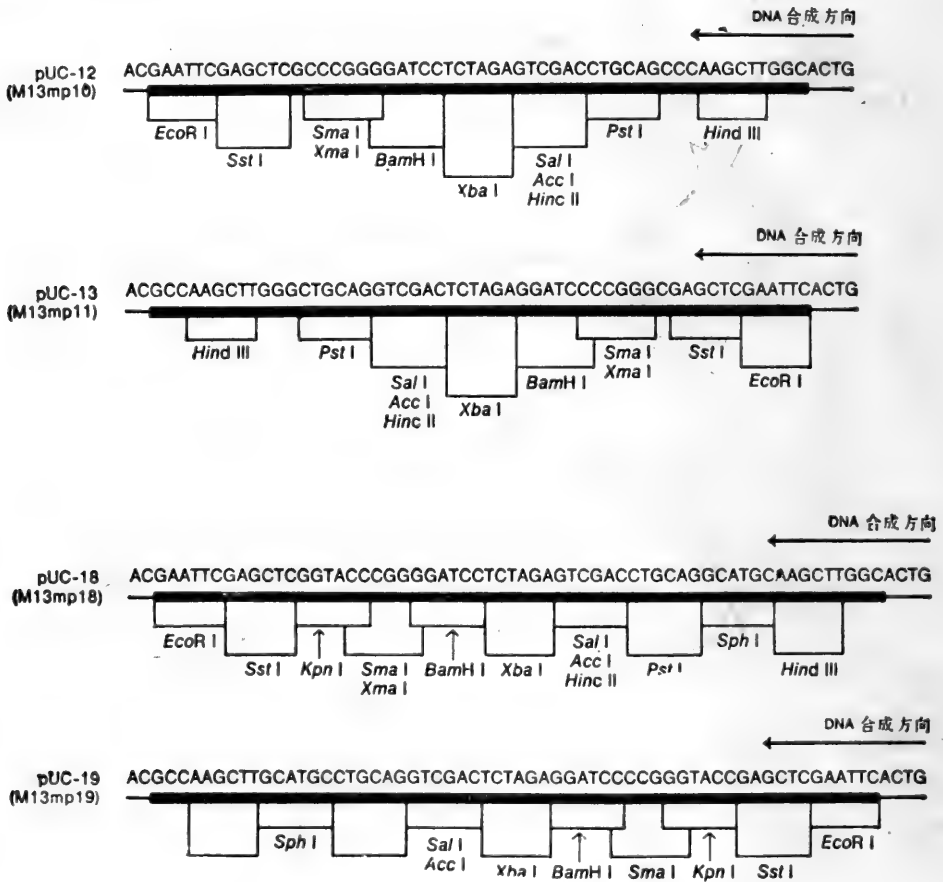


图 2.7B pUC 系列与 M13mp 系列部分结构示意图。

(三) 表达载体

所谓表达载体是一类带有某种启动子的质粒,如上述 pUC 质粒带有 lac 基因的启动子。当把外源基因插入到启动子下游的合适位置上,则在一定条件下可引起外源基因的表达。启动子越强,表达效率越高,得到的产物也应越多。

这类载体具有以下特点：①带有强有力的启动子；②在启动子下游一定位置有合适的酶切位点能连接外源基因片段；③具有筛选标记；④能转化到原核细胞(或真核细胞)中去以得到高效表达的产物。

例如质粒 pKK223-3,由 4585bp 组成,由 pBR322 改建.含 trp- lac (简称 tac) 启动子,及 lacUV-5 启动子的核糖体结合位点,其后是 pUC8 的多聚连接物,上面有多种内切酶的位点(图 2.8),可作为外源基因插入处。再后面连接着强的 rrnB 核糖体 RNA (rRNA) 转录终止子的 DNA 片段,它可以稳定宿主-载体体系,使外源基因在强的 tac 启动下得到高效表达。宿主为大肠杆菌 JM105。

表达载体种类很多,详细讨论请参阅本书第二部分基因表达。

(四) 特殊型质粒载体

为了使外源基因在体外可控制条件下得到高效表达,获得基因疫苗、激素及蛋白类药物等具有极大经济价值的产物,或为了更好地研究真核基因表达及调控,必须构建高效的表达载体。但是真核基因在原核细胞中表达有许多限制或不利的因素,要构建一些杂合质粒,同时能在原核细胞(大肠杆菌)和低等真核生物(酵母)或高等真核生物(哺乳动物细胞)中生存和扩增。这类杂合质粒叫做穿梭质粒(shuttle plasmid)。现举大肠杆菌-酵母穿梭质粒为例,看其结构特点：①含有酵母基因复制起始区；②含有某种营养缺陷突变的基因片段,用作标记能在酵母菌中筛选；③含有某些酵母基因；④有大肠杆菌质粒的复制起始区,使此质粒既能在大肠杆菌中扩增,又能转化到酵母细胞中,并在其中表达出基因产物。详见本书第二部分。

穿梭质粒的含义也指不同菌属之间质粒的重新组合,如大肠杆菌与枯草杆菌、放线菌等不同菌属质粒的重组所得到的杂合质粒也是一种穿梭质粒。

例如产生创新霉素的放线菌质粒 AJP1,由于该放线菌的培养、质粒的提取和转化均比大肠杆菌困难得多,所以试图构建一个既可存在于大肠杆菌,也可存在于放线菌中的载体。应用 AJP1 与 pBR322 重组,得到了重组的杂合质粒 pAX,既可转化到大肠杆菌 C600 中,又可在原放线菌中生长。

以上所述的特殊表达载体——穿梭质粒的构建,为基因工程的应用开辟了新的前景。

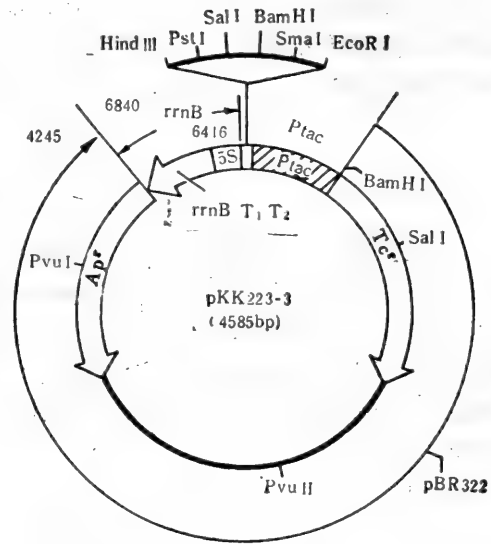


图 2.8 质粒 pKK223-3 结构图。

Ptac: (tac) 启动子。rrnB: 核糖体 RNA 转录终止子的 DNA 片段。

第二节 λ 噬菌体载体

在采用大肠杆菌体系克隆外源基因中,常见的另一类载体是大肠杆菌 λ 噬菌体。与质粒 DNA 相比, λ 载体的优点是:①可携带较长的外源 DNA 片段,其最大的容纳量可达 18—22kb (千碱基对);②重组后的 DNA 分子可在体外被包装成噬菌体颗粒,与 DNA 直接转染细菌的效率相比,包装成的噬菌体具有更高的感染宿主细胞的能力;③重组的噬菌体容易筛选和贮存。 λ 噬菌体的遗传学与分子生物学已有详细和深入的研究,并获得了一系列的突变体。这些工作为 λ DNA 成为一类实用的载体奠定了良好的基础。为了便于读者更好地了解和使用 λ 载体,本节先将大肠杆菌 λ 噬菌体的生物学知识作一简要叙述,然后介绍载体的构建、种类及其应用。关于 λ DNA 制备已在本书上册第二章介绍了,载体的用法将在本书第六章基因文库的有关部分叙述。

一、 λ 噬菌体的生物学特性

λ 噬菌体颗粒中的蛋白质和脱氧核糖核酸 (DNA) 大约各占一半。在形态上, λ 颗粒由头与尾两部分组成。呈六角形的头部装载了噬菌体的整个基因组 (即 λ DNA)。 λ DNA 是双链线性分子,其两末端的 5' 位各有一长 12 个核苷酸的单链,它们呈互补的核苷酸顺序,不论在体内或体外,两 5' 端的单链可借氢键粘合成环形分子。野生型 λ 基因组的全部核苷酸序列已被测出,若按末端相连成环形分子计算,其总长度为 48,502bp,分子量为 32×10^6 。

λ 噬菌体的基因组图谱示于图 2.9。它像其他噬菌体一样,生物功能上相关的基因连在一起。例如,密码噬菌体头部或尾部蛋白的各种基因定位在一起,集中于基因组的左臂部位;在其右臂区域集聚了与 DNA 复制、早期和晚期调节以及细胞裂解等作用有关的基因。

(一) λ 噬菌体的繁殖

λ 噬菌体感染大肠杆菌时,首先由噬菌体尾部吸附于宿主细胞表面。这种吸附作用是由尾部的 J 基因所编码的蛋白和宿主细胞膜的外部来自 LamB 基因的产物相互发生特异反应而引起的。宿主细菌 LamB 基因产物是一种麦芽糖转运蛋白,它的表达可由麦芽糖诱导。所以说,供 λ 感染用的宿主细胞应在含有麦芽糖的培养基中繁殖。

当噬菌体吸附于宿主细胞膜之后, λ DNA 被注入到细胞内。DNA 穿透细胞膜的作用是由噬菌体尾部和细菌内膜某成分发生专一性反应而触发的。细菌内膜上的这种成分很像是 ptSM 基因的产物,它是属于负责几种糖类转运的磷酸转移酶系统中的一种成分。

噬菌体的双链线性 DNA 分子进入细胞内不久,即自行环化。它是由 DNA 左、右两端的 5' 端单链以氢键互补粘连所致。可以配对相连的左右末端位点 (即 12bp 长的单链部分),加上 DNA 包装时所需的邻近核苷酸顺序,统称为粘性末端位点 (cos 位点)。

在宿主细胞内,噬菌体 DNA 的复制是由其自身的 O 和 P 基因负责,同时,还需要有宿主细胞的某些功能相配合。DNA 复制作用可分两期 (见图 2.10),开始时称为 θ 型的环

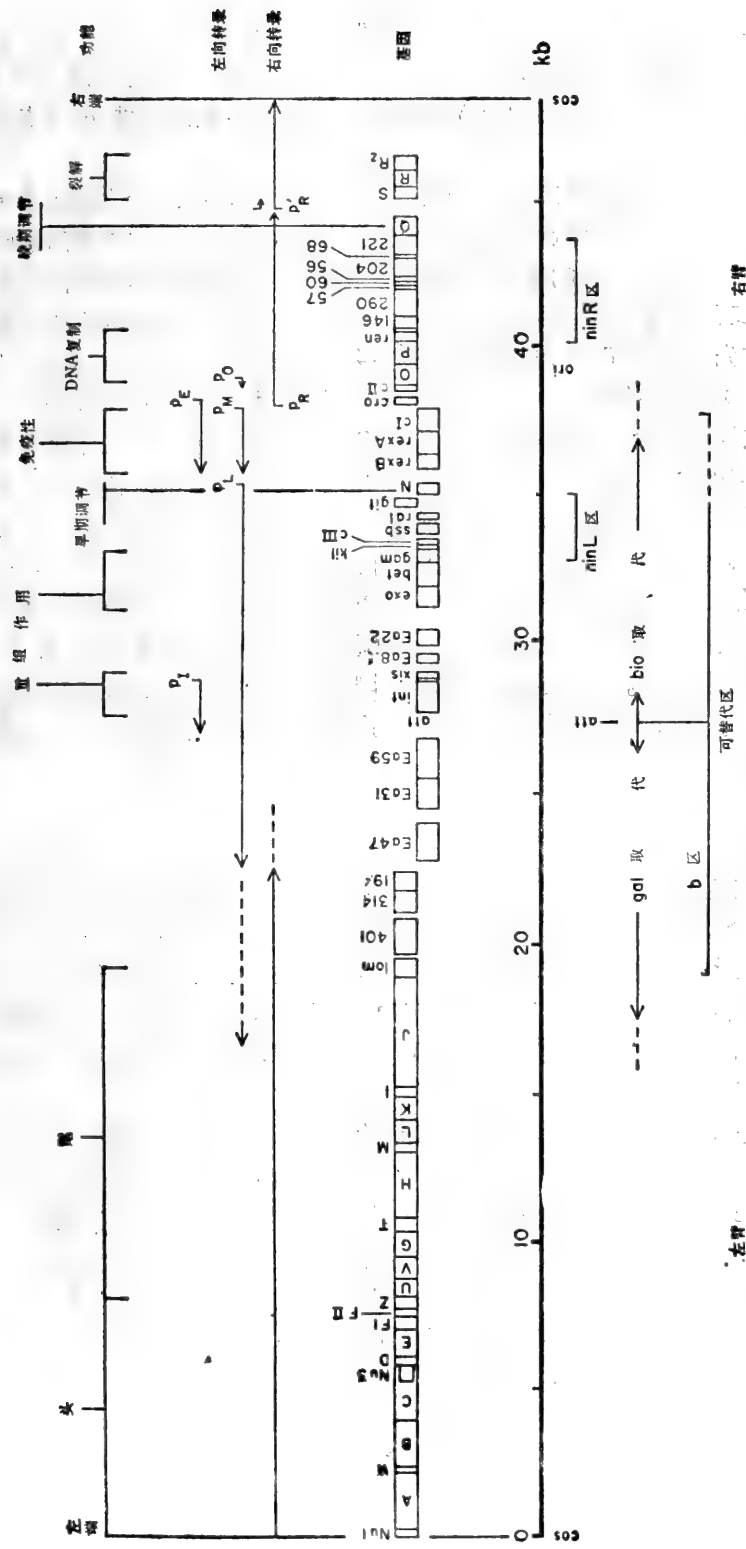


图 2.9 大肠杆菌 λ 噬菌体的基因谱。
 λ 基因组的全长为 48,502bp, 本图中以 kb 为单位标示其长度。基因谱开始于左端的 cos 位点, 终于右端 cos 位点, 各基因用英文字母表示。功能上相关的基因簇用方括弧示于图的顶端, 并用垂直线示出具有早期转录调节作用的 N 基因以及起晚期转录调节作用的 Q 基因的位置。图中标出了各个启动子转录作用的起始位置, 并且用箭头标出其转录方向, 箭头所示的终点表示转录作用延伸的范围。其中虚线表示转录还可能通过的区域。P_L:int 蛋白的启动子; P_E:cI 蛋白的确立启动子 (establishment promoter for cI); P_M:cI 的维持启动子 (maintenance promoter for cI); P_L:主要的左向转录启动子; P_R:主要的右向转录启动子; P_O:oop 转录产物启动子, 起始于复制起始点 (ori) 附近的一个小的左向转录启动子 (oop 产物在 λ 的 DNA 复制及转录调节中作用至今尚不清楚); P_R:晚期启动子。 λ DNA 可被替代或缺失的区域标于图的下端。b 区为基因组内对 λ 存活和感染作用非必需的编码区, 包括了 DNase, 膜蛋白等基因和 Int 基因表达的调节位点。在可被替代区的两侧为噬菌体基因的必需区, 以左臂和右臂示之。

形复制。复制作用起始于 *cII* 和 *O* 基因区域,复制弧沿 DNA 模板向两个方向延伸,直至新合成链的两端碰在一起。产物为两分子环形的子代双链 DNA。在复制的第二期, DNA 以滚筒方式增殖,产生呈串联状的 DNA 线形分子,多个 *N* 基因组借 *cos* 位点互相结合,排列成串。

在 DNA 复制的同时,基因组在受控制的条件下开始转录和表达,产生构成噬菌体颗粒所需的结构蛋白。经过一定的潜伏期,时间大约为自细菌受感染到被裂解所需时间的一半,宿主细胞内即可累积足够量的噬菌体结构蛋白,提供包装成噬菌体颗粒的全部成分。噬菌体颗粒的组装包括两个各自独立的途径:头部结构的组装和尾部的形成。在噬菌体头部的组装中,首先由基因 *E*(产物是噬菌体的主要外壳蛋白)、*Nu3*(核心蛋白的基因)、*C* 和 *B* 的各种产物在宿主细胞 *gro* 基因产物的作用下组装成头部前身物(称 *pitit* λ);第二步是经滚筒式复制产生的呈串联状线形多聚 DNA 分子被包装到头部前身物之中。其包装过程有噬菌体 *Nu1*、*A*、*FI* 和 *D* 等基因产物参加,同时,宿主细胞内的某些因子也参与此过程。在包装时,一个噬菌体头部只能容纳一分子全基因组的 DNA,并包括其完整的 *cos* 位点结构。包装有 DNA 的头部结构,再在基因 *W* 和 *FII* 产物的参与下,和已组装好的尾部相结合,构成成熟的子代噬菌体。当宿主细胞内积聚了相当数量的子代噬菌体之后,细胞会破裂,噬菌体颗粒被释放出来。细胞的裂解是由 λ 基因组的 *S*、*R* 和 *R₂* 基因的特定生物功能引起的。被释放出来的噬菌体颗粒可以再去感染其他宿主细胞,并使新的宿主细胞裂解。如此周而复始,噬菌体得到了繁殖。这样的繁殖过程称为噬菌体的裂解性生长。

(二) 溶源化作用

λ 是一种温和型噬菌体。它进入宿主细胞之后,除了能进行裂解性繁殖导致细胞破裂之外,在一定的条件下,它的基因组还可以被整合到宿主的染色体 DNA 上(图 2.10)。整合后的噬菌体基因组将随宿主的 DNA 一起复制,并将噬菌体基因组带入子代细胞,这种整合作用叫做噬菌体的溶源化作用。染色体 DNA 上整合有噬菌体基因组的宿主细胞,并不会由于噬菌体的感染而发生裂解,这样的现象称为溶源现象。整合有噬菌体的宿主细胞叫做溶源菌,被整合的噬菌体称为前噬菌体。

溶源现象的产生是因为噬菌体基因组晚期转录的有关基因不能被表达的缘故。噬菌体基因组中基因 *cII* 和 *cIII* 的产物能激活 P_E 和 P_I 两启动子的左向转录作用,使分别受此两启动子控制的 *cl* 和 *int* 基因随之表达。*cl* 基因产物是噬菌体基因早期转录作用的阻遏物,它会引起早期转录的停止,从而中断由早期转录进入晚期转录的过程,使晚期转录的各种基因不能表达。*int* 基因的蛋白产物能识别宿主细胞染色体 DNA 和噬菌体基因组上相应的 *att* 位点,并催化两者的断裂和再连接。*att* 位点称为附着位点,在宿主细胞上的称 *attB*,位于噬菌体上的称 *attP*。*attP* 和 *attB* 位点上互有一段 15 bp 长度的同源性核苷酸顺序,借同源性重组作用使噬菌体基因组插入到宿主的染色体 DNA 上。此外,有人还认为,邻近同源序列的非均一性核苷酸序列对于这种特异性位点的整合也起重要作用。

处于整合状态下的噬菌体基因组,除了 *cl* 基因被表达之外,几乎所有基因的转录作用被阻遏。*cl* 产物还调节自身的合成。当产物浓度低时,它催化维持启动子 (P_M) 的活

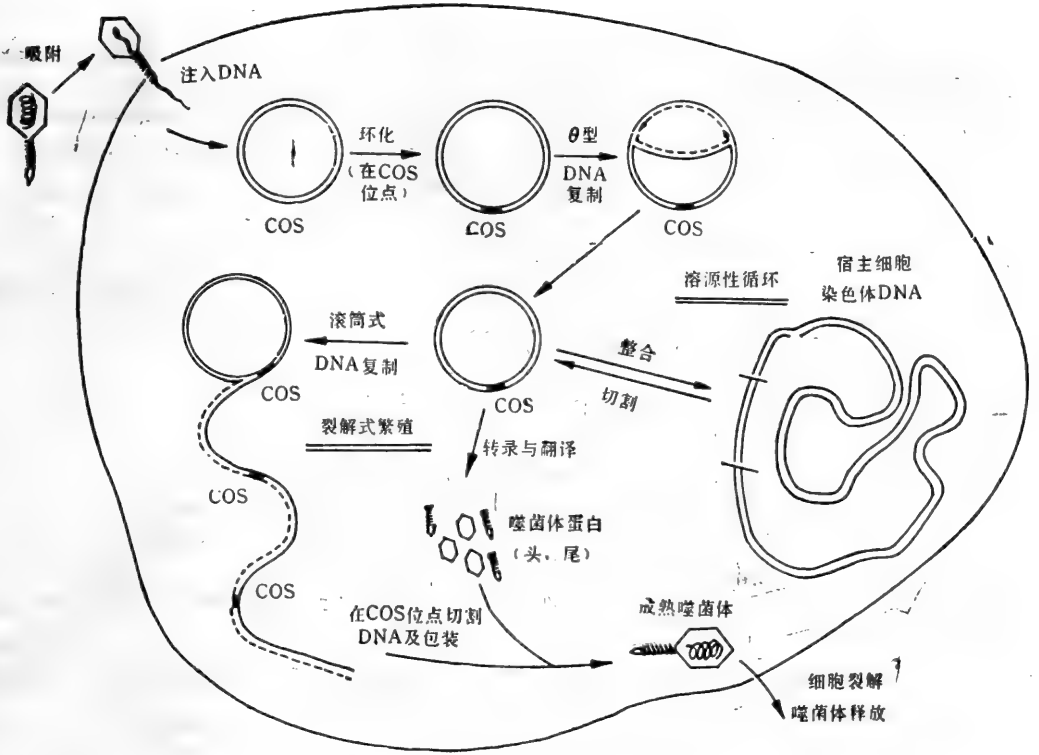


图 2.10 λ 噬菌体的裂解式生长及其溶源化作用。

性,以继续产生 cI 产物;而当产物浓度增高后,过高的浓度又会抑制 P_M 的转录作用。

λ 基因组被整合到受感染细菌的染色体 DNA 上的作用是可逆的。在一定条件下,被整合的噬菌体 DNA 又可离开细菌的染色体,而游离出来,恢复它的裂解性生长。这样的作用需要噬菌体基因组中 int 和 xis 两基因产物的参与,同时也受宿主细胞内的一些因子的控制。噬菌体 DNA 从宿主染色体 DNA 上释放出来的过程称为诱导。比较方便的诱导噬菌体的方法是用紫外线辐射或者丝裂霉素 C 处理,但常见的是选用 cI 基因的温度敏感突变株。这样的突变株在低温($30-32^{\circ}\text{C}$)下, cI 基因能正常工作,表达出 cI 产物(称为阻遏物),当温度升高到 42°C ,突变株中 cI 基因不能再产生阻遏物,原噬菌体基因组就会从染色体上切除下来,释放到细胞内,重新开始裂殖生长。

(三) λ 噬菌体基因表达的调控

λ 噬菌体感染了宿主细胞之后,其基因的表达顺序可简单地用图 2.11 来说明。在裂解生长过程中,只有三个启动子对基因的表达是重要的。在感染后不久,宿主细胞 RNA 聚合酶结合到 λ DNA 链上, P_L 和 P_R 启动子开始了转录作用。由 P_L 启动的转录作用是左向的,转录停止于终止区 t_{L1} 或 t_{L2} 位置,因为开始转录时噬菌体 N 蛋白缺乏,80%的转录反应终止于 t_{L1} 位点。此时, N 基因已被转录成 mRNA,并合成出 N 蛋白。在 P_L 转录的同时, P_R 启动子向右转录,转录产物有一半停于 cII 基因左侧的终止区 (t_{R1}),另一半可延续至第二终止区 (t_{R2})。 t_{R2} 位于 nin 基因区内。停止于 t_{R1} 和 t_{R2} 的两

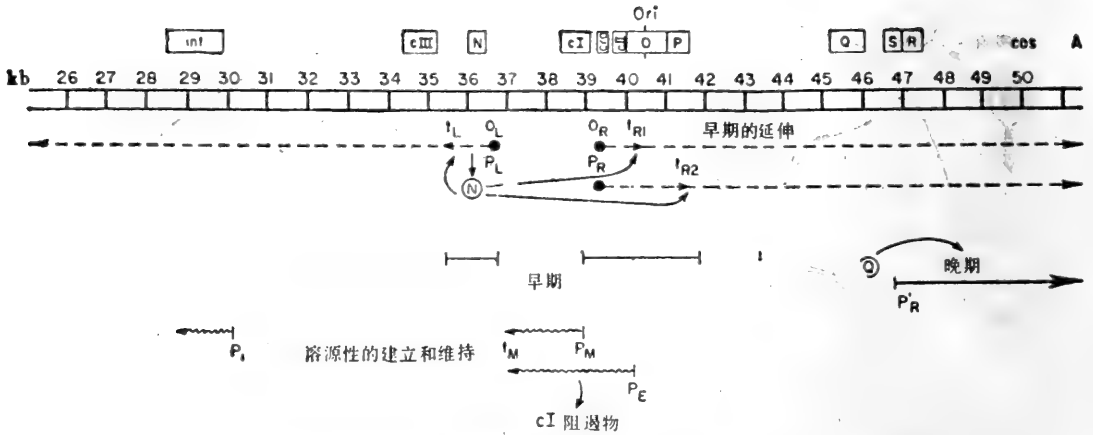


图 2.11 λ 噬菌体的基因表达及其调节。

与 λ 基因表达有关的基因在图的上部以方框标出。(●) 为启动子位点。--> 为转录产物, ⊙ 为 N 基因产物。⊙ 为 Q 基因产物。←←← 与建立和维持溶源性有关基因的转录产物。

个转录作用的相同产物是 *cro* 基因产物。反应延续至 t_{R2} 的转录作用还将 O 和 P 两基因转录并翻译。

N 基因的产物 N 蛋白称为抗终止转录物, 当积累到一定量时, 它能拮抗 t_{R1} 、 t_{R2} 和 t_{L1} 、 t_{L2} 的终止转录作用, 使 P_L 和 P_R 的转录作用越过各自终止区继续延伸下去。左向转录的延伸, 使与重组和整合作用的有关基因得到表达。与重组有关的基因为 *red*, *int* 基因与整合有关。右向转录的 P_R 作用的继续, 则使 O、P、Q 等基因获得表达, 导致 DNA 的复制及激活晚期的转录作用。Q 基因的产物能激活由 $P_{R'}$ 开始的晚期转录。Q 蛋白是正调节子, 它能克服 $t_{R'}$ 的终止转录作用, 把转录延伸到 $P_{R'}$ 的右端, 使和细胞裂解有关的基因 S 和 R 表达, 同样也使编码头和尾的 A 至 J 结构基因被表达。

cro 蛋白是操纵基因 O_L 和 O_R 的抑制物, 它的大量积累会阻遏 O_L 和 O_R 作用, 从而使由 P_L 和 P_R 开始的早期转录作用关闭。

以上简述了噬菌体裂解性生长时的基因转录及表达的调控。下面再介绍在溶源性繁殖中的转录和调控。N 蛋白的累积和它的去终止作用, 同样使 *cII* 和 *cIII* 基因表达。*cII* 和 *cIII* 产物共同激活由 P_1 (*int* 基因启动子) 和 P_E (阻遏确立启动子) 开始的转录作用。这两个启动子的作用与噬菌体的繁殖(即, 是继续保持裂解性生长还是宿主被溶源化) 有着密切关系。 P_E 的转录作用是向左进行的, 因为启动子定位于和 *cro* 基因相对的另一条 DNA 链上, 因此, P_E 对 *cro* 基因不起作用。 P_E 的转录会直至 *ci* 区域。

前面已提到, *ci* 产物是 λ 的阻遏物, 它像 *cro* 产物一样, 对操纵基因 O_L 和 O_R 有亲和性。但是在代谢上, *ci* 蛋白比 *cro* 蛋白更稳定, 它与 O_L 和 O_R 的结合位置不同于 *cro* 蛋白。当 *ci* 蛋白累积到足够的量之后, 通过阻止 O、P 和 Q 基因的表达中止噬菌体 DNA 的复制作用以及所有晚期转录基因的表达。由于 *int* 基因产物(称为整合酶) 的作用, λ 基因组被整合到宿主细胞的染色体 DNA 上, 变成了前噬菌体。整合到染色体上的噬菌体 DNA 将随宿主细胞一起复制。但是, 随着细胞的分裂传代, 最初由 P_E 启动产生的 *ci* 产物等于被稀释了, 要保持溶源化状态, 就要维持足够量的 *ci* 蛋白, P_M

就在此时被激活,提供 cI 蛋白。

(四) 几种突变株的表型

下面介绍几种与溶源现象有关的 λ 噬菌体突变株的表型。

(1) cI 基因的突变失活,噬菌体不能产生 cI 阻遏物,宿主细胞也不会溶源化,这种 λ 突变株形成清亮噬菌斑。但要提醒的是, cI 突变株同样保留着对野生型 λ 的 cI 阻遏物的敏感性,不能在溶源化的细菌株(例 K12)中生长,并对 λ 和 λ C 变异体有免疫性。

(2) λ vir 是指能在大肠杆菌 K12(λ) 菌株中生长的突变株。通常它们具有双突变(或者三突变): 在 O_L 上有一突变并在 O_R 上也有一突变。这些突变影响了操纵基因,它们不能与 cI 阻遏物结合。

(3) Nin 缺失作用指 λ 基因组中去除了定位于 P 和 Q 基因之间的 t_{R2} 终止区,使 Q 基因能在没有 N 蛋白的存在下进行转录。Q 蛋白的存在就允许晚期转录的基因表达。所以 λ nin 噬菌体的裂解性生长不取决于 N 蛋白的存在与否。 λ nin 不能正常生长,因为失去了与重组功能有关的 Red 和 Gam 的表达作用。Red 和 Gam 是与晚期 DNA 复制有关的基因。

(4) cII 或 cIII 的突变株很难表现其溶源性,除非能借助噬菌体的共感染来提供相应的基因产物。一旦溶源性已经确立, cII 和 cIII 的突变株能维持于前噬菌体状态。因为前噬菌体的维持是靠 cI 产物, cII 和 cIII 产物并不起作用。

(5) λi^{434} 是一杂合噬菌体。在 P_L 到 P_R 之间的基因来自噬菌体 434, 其余部分来自 λ 噬菌体。突变体 λi^{434} 对噬菌体 434 的阻遏物是敏感的,而对 λ 的 cI 阻遏物不敏感。 λi^{21} 也是杂合噬菌体。 λ 基因组中由 N 到 cII 之间的区域由噬菌体 21 的相应部位来替代。因此杂合体对 λ 的 cI 蛋白不敏感,而对来自噬菌体 21 的 cI 蛋白敏感。

二、 λ 噬菌体载体的构建

正如前面介绍的(图 2.9), λ 噬菌体的基因组可以分为三个部分: 左臂、右臂及中央区。左臂负责构建头和尾部结构,包装噬菌体 DNA 和产生具有感染性的噬菌体所必需的全部基因(从基因 A 到 J); 右臂是指由 N 基因开始向右直至右末端的区域,包括负责 DNA 复制(基因 O 和 P),使宿主细胞裂解(S 和 R 等基因),以及它们的调控顺序。中央区是从基因 J 之后到 N 之前的部位,这个部位的基因不是形成噬菌体所必需的,它们约占基因组总长度的 40%。若用外源 DNA 取代这个区域的基因,不会影响噬菌体的形成。这一特性是 λ 噬菌体 DNA 可作为载体携带外源基因的基础。

λ 噬菌体在宿主细胞中繁殖,复制的 DNA 要包装到事先已组装好的头部,并再接上尾巴后才成为成熟的噬菌体和具备感染的能力。噬菌体要具有感染性,被包装的 DNA 要有一定的长度。 λ 的野生型噬菌体 DNA(48.5kb) 包装的效率最高,并且所形成的噬菌体颗粒具有最强的感染力。实验证明,当被包装的 DNA 分子的长度在野生型 λ DNA 的 78—105% 的范围内,仍可被包装成具有感染能力的噬菌体颗粒,但其感染效率将低于野生型噬菌体。如果 DNA 的长度短于野生型 DNA 长度的 78% 或者超过了 105%, 便不能形成具有感染力的噬菌体颗粒。这就表明在选用 λ 作载体时,插入的外源 DNA 的大

小需有一定的限制。

把外源 DNA 插入到 λ 载体的可替代区域,常用的方法是借助特定的限制性内切酶位点的插入重组。可是,许多种限制性内切酶在野生型 λ DNA 上具有多个切点,而且有些位点正定位于与噬菌体生长繁殖密切有关的必需基因之中,这些位点的剪切重组将会严重影响噬菌体的繁殖。因此,野生型 λ DNA 不能直接作为载体,需要经过人工改建才能成为有价值的克隆载体。改建载体的工作包括①消除或改变 λ 基因组必需区内的不必要的限制性内切酶位点;②在可替代区构建所需的内切酶位点;③其他结构上的改建,以提高载体的效用。譬如,插入一个外源启动子,使载体具有表达外源基因的功能等等。

(一) 限制性内切酶位点的消除或构建

在 λ DNA 上进行限制性内切酶位点的消除或构建,目的在于使载体 DNA 的可替代区域具有单一或两三个某一种或几种特异的限制酶位点,并且消除这一种或几种酶在 λ DNA 左、右臂上的位点,以使外源 DNA 定位地插入到载体上。

1. EcoRI 位点的消除

根据微生物限制修饰特性,可以通过体内实验,达到消除 λ 基因组必需区(即左、右臂)内 EcoRI 位点。把 λ 噬菌体循环交替地在大肠杆菌 K 株和含有 EcoRI 酶基因质粒的 K 株中繁殖,最终可以分离到只有在基因组非必需区含有单个或两个 EcoRI 位点的 λ 突变株。Murray、Rambach 和 Thomas 等三个实验室于 1974 年都报道了此结果。不过,这种方法只适合于宿主细胞能产生的那种限制酶的位点消除,即消除 λ DNA 中的 EcoRI 位点。但要消除 λ DNA 中的其他酶切位点像 BamHI、HindIII 等,显然不能采用此方法。

2. 遗传突变

构建 λ 载体最常用的方法是利用遗传学的点突变、取代或缺失等,使野生型 λ 发生突变,改变其基因组中限制性内切酶位点的分布,或消除基因组必需区部位的酶切点。

3. λ 同类噬菌体之间的基因片段置换

$\phi 80$ 、 $\phi 21$ 和 $\phi 434$ 等都是 λ 的同类噬菌体,称为 Lambdoid 噬菌体,它们的基因片段可以和 λ DNA 中相应的基因片段进行替换。由于选用的片段含有所希望的限制酶的位点或者不含有要从野生型 λ DNA 中消除的酶位点,用它去取代 λ DNA 上的相应部位,即可构建成符合要求的载体。

体外包装噬菌体的成功,为用诱变构建 λ 载体提供了快速、有效的筛选方法。用某种能消除特定位点的限制酶来降解在构建中的载体 DNA,再进行体外包装和感染宿主细胞,只有未被酶切的载体 DNA 才能高效地包装成噬菌体颗粒和感染宿主细胞。如此循环筛选突变体数次,就可选择性地消除不需要的限制酶位点,提供有效的载体。

采用上述的数种方法,迄今已建成了一系列的 λ 载体(见本节第三部分)。一般说来,一种载体主要适用于某一种特定内切酶位点的插入或替代重组。但目前发展的是希望在几种限制酶的位点上接受外源 DNA。当载体 DNA 的非必需区内只有一个酶切位点,此部位可以插入外源基因片段;当载体有两个以上的限制酶的位点,用酶切之后,非必需区的 DNA 可用外源 DNA 取代,构成杂合的重组噬菌体 DNA。

(二) 其他目的的改建

除了消除 λ DNA 必需基因区内某些限制酶位点或构建所需位点之外, 还需进行其他目的的改建工作。为的是①运载不同大小的外源 DNA; ②有更多的插入位点; ③克隆之后便于从噬菌斑的表型来区分重组体和自身的载体; ④表达外源基因; ⑤使重组后的克隆有较高的产量, 容易回收插入的片段, 等等。下面列举载体的几种改建。

1. 乳糖操纵子的引入

乳糖操纵子存在于大肠杆菌中, *lacZ* 是乳糖操纵子中的一段 DNA 片段, 它包含有启动子的序列以及半乳糖苷酶结构基因。由于 *lacZ* 中的半乳糖苷酶基因能和完整乳糖操纵子一样获得表达, 因此也显示有半乳糖苷酶活性。当把 *lacZ* 片段重组到 λ 载体上的合适部位(例如中央区), 载体会具有该基因的表型。这就是说, 带 *lacZ* 片段的 λ 载体在乳糖操纵子阴性(*lac*⁻)的大肠杆菌中, 可显示半乳糖苷酶活性。被此种载体感染后的宿主细菌, 当在含有 5-溴-4-氯-3-吲哚基- β -D-半乳糖苷(5-bromo-4-chloro-3-indolyl- β -D-galactoside, 简称 X-gal) 的培养基中生长时, 形成的噬斑呈现蓝色。但是, 当载体 DNA 含此 *lacZ* 片段的区域被外源 DNA 取代, 重组体在同样条件下感染宿主菌, 形成的重组体噬菌斑则为无色透亮的斑点, 显示无半乳糖苷酶活性。通过 *lacZ* 的表型检测, 有利于重组体的辨认和选择。Williams 和 Blattner 等构建的 charon 系列载体中的 charon3A 和 4A, 或 Murray 等构建的 λ MN762, 781 或 816 等均有此特性。

同样, 虽然不将含 *lacZ* 基因的片段取代, 只要使 *lac* 基因的正常表达受阻, 不能产生有活性的半乳糖苷酶, 也可导致对 X-gal 的阴性反应。具体地说, 若把外源基因片段插入到 *lacZ* 的酶结构序列中, 使之得不到有功能的酶产物, 重组体就呈阴性反应。本节第四部分要介绍的新载体 λ gt11 就属这种类型的载体。

2. Spi⁻ 基因表型的构建

野生型 λ 噬菌体不能在携带有 P2 前噬菌体的大肠杆菌宿主细胞中生长, λ 的此种表型称为 Spi⁺(对 P2 噬菌体干扰的敏感性)。但是, 当 λ 基因组中与 DNA 体内重组作用有关的两个基因 *red* 和 *gam* 丢失之后, λ 的突变体就可以在 P2 溶源菌中生长, 此时噬菌体的表型称为 Spi⁻。*red* 和 *gam* 基因是定位于基因组的非必需基因区域内。当选用 147.1、11059 以及 EMBL3 或 4 等取代型载体时, *red* 和 *gam* 随中央片段的除去而丢失, 加入了外源 DNA 的重组体显示 Spi⁻表型, 能在大肠杆菌 *recA*⁺ 的 P2 溶源菌(例如大肠杆菌 Q359 等)中繁殖。因此, 当选用合适的宿主菌(即 *recA*⁺ 的含 P2 前噬菌体的大肠杆菌), 就可以来区分和选择重组克隆。

Spi⁻表型的重组噬菌体不能在 *recA*⁻ 的宿主细胞中繁殖。另外, *gam*⁻ 的噬菌体感染 *recA*⁺ 的细菌时, 由于 *gam* 基因产物缺乏, 宿主细胞的重组体系就异常活跃, 导致克隆 DNA 的重排。这是 Spi⁻噬菌体在 *recA*⁺ 宿主细胞中繁殖的一种弊端。为了防止这种弊端, 在采用 Spi⁻的表型选择重组体时, 有人提出使用 *recB, C*⁻ 的菌株作宿主细胞。Jeffreys 等在 1982 年报道了将携带有哺乳动物血红蛋白基因的 Spi⁻ 重组噬菌体引入到 *recB, C*⁻ 的大肠杆菌宿主细胞中繁殖, 并通过对插入基因的核苷酸序列分析, 表明 DNA 序列中未发现任何重排现象。

3. *chi* 突变

正如上述, Spi^- 表型的噬菌体是由于 red 和 gam 两基因的缺失 (red^- 和 gam^-), 它们在 P2 溶源菌或野生型的大肠杆菌中的生长均很贫乏, 形成的噬菌斑很小。如果噬菌体 DNA 中引入 chi 位点, Spi^- 的噬菌体在上述宿主细胞中的繁殖力会增强, 噬菌斑变大。 chi 位点是指一段长度为八核苷酸的 DNA 序列, 有人测出其核苷酸序列为 5'-G-C-T-G-G-T-G-G-3'。野生型的 λ 噬菌体 DNA 并不存在此序列, 但可在 λ 的某些突变体中发现它。大肠杆菌染色体 DNA 中存在 chi 序列, 野生型的大肠杆菌染色体中, 平均约每 5,000bp 出现一次 chi 序列。在真核细胞基因组中都含有此序列。 chi 的功能是激活由宿主细胞 $recBC$ 旁路引起的重组作用, 但是在 $recA$ 、 $recB$ 或 $recC$ 突变的宿主细胞中就不能测出 chi 的活性。另外, 对于 red^+ 或 gam^+ 的 λ 噬菌体, chi 的作用会明显减弱, 因为前者是由于高的 red 重组旁路的活性, 后者是 gam 基因产物抑制了 $recB$ 、 C 酶的活性。对于 red^-gam^- 的突变体, 包括重组体因插入外源 DNA 而呈现的 Spi^- 表型 (也是 red 和 gam 基因的丢失), chi 位点的存在会使噬菌体生长繁殖力变强。不论 chi 是原来就存在于载体中还是随外源 DNA 带人的, 均有相同的效果。这也表明了用不带 chi 序列的载体克隆外源基因时, 在外源 DNA 片段中带有 chi 序列的重组体的繁殖能力强于不含 chi 序列的重组体。因此在重组体扩增时, 后者很容易丢失。所以在改建 λ 载体时, 使其基因组的任一臂区 (左臂或右臂) 带有 chi 序列很有必要, 尤其是用作构建基因文库的载体。为确保得到的重组体均为 chi^+ , 防止重组体在扩增时丢失, 要尽可能保持基因文库的完整性。 $\lambda 1059$ 和 $\lambda L47.1$ 分别在其右臂或左臂区含有 chi 位点, 可使重组体噬菌斑的大小趋向相对一致。

4. imm^{434} 片段的引入

噬菌体 $\phi 434$ 是 Lambdoid 噬菌体中的一种, 它形成的噬菌斑混浊不清亮。这是在噬菌斑区域内的细菌对噬菌体感染的一种溶源性反应。噬菌斑变混浊, 是因为噬菌体阻遏物 (cl 蛋白) 的合成。 cl 蛋白对噬菌体的其他基因的表达具有负控制作用, 并且使溶源性细菌对持有相同免疫特异性的噬菌体的感染具有免疫力。当外源 DNA 插入 cl 基因时, cl 蛋白的合成受阻, 重组的噬菌体失去了上述混浊噬菌斑的性状, 呈现完全清亮的噬菌斑。因此, 噬菌斑的混浊度可以作为重组体和原载体的鉴别标志。

由于噬菌体 434 基因组上 cl 基因只有单一的 $EcoRI$ 和 $HindIII$ 位点, 若用它的 cl 基因去取代野生型 λ 基因组的 cl 区域, 得到的载体可在 cl 基因上的 $EcoRI$ 或 $HindIII$ 单一一位点插入外源 DNA, 然后, 借噬菌斑的混浊度来区分重组体和自身载体。带有 $\lambda imm^{434}cl$ 基因的载体被称为免疫性插入型载体。外源 DNA 片段插入 λimm^{434} 的 cl 基因中, 重组噬菌体呈现清亮型噬菌斑。Murray 等构建的 $\lambda NM641$ 和 728 载体以及 Blattner 等报告的 $charon7$ 等均可接受外源 DNA 的 $EcoRI$ 或 $HindIII$ 降解片段, 载体的最大插入容量为 10kb。插入后形成的重组噬菌体在宿主菌中的生长比原载体更健壮, 噬菌斑为清亮型。新近介绍的 $\lambda gt10$ 载体, 也是 imm^{434} 的免疫性插入型载体。

5. $cl857$ 突变体

cl 基因温度敏感性 (ts) 的突变体用 $cl857$ 表示。此种突变体在低温 (30—32°C) 下, 前噬菌体能稳定地嵌合于宿主的染色体 DNA 中, 呈溶源性繁殖; 然而, 当温度升高到 42—45°C, 前噬菌体基因组就会从宿主的染色体上释放出来, 进入裂解性生长, 最终导致宿主细胞的裂解。把此温度敏感性突变体 ($cl857$) 引入载体中, 载有外源 DNA 的杂

合噬菌体也将具有对温度敏感的溶源特性。表达型载体 λ gt11 就具有此特性。

6. 其他突变体

λ DNA 中的 S 和 R 两个基因是宿主细胞裂解所必需的。S100 是指 S 基因的琥珀型 (amber) 突变体, 它使宿主细胞不被裂解, 噬菌体颗粒可大量地累积于细胞内, 最后, 靠加入氯仿来导致细胞的裂解和噬菌体的释放。所以, S100 突变体可以提高噬菌体产量。

nin5 是指 λ 基因组中去除了转录终止位点 t_{R2} 的缺失型突变作用; 突变体的裂解性生长将不受 N 基因产物的控制。

三、常用的 λ 载体种类

随着 DNA 分子克隆技术的诞生, Murray 和 Murray, Rambach 和 Tiollais 以及 Thomas 等几个实验室在 1974 年就提出了 λ 噬菌体可以改建成克隆外源基因的载体。此后, 不断有各种 λ 载体被构建成功。迄今为止, 已报道的载体已超过 100 种, 表 2.3 摘录出部分常用的载体。一般说来, λ 载体可分为替代型和插入型两类。插入型载体在其基因组的非必需区内, 有一种限制性内切酶的单一位点供外源基因的插入, 重组时并不丢失载体片段, 得到的重组 DNA 分子都比原载体要长。插入型载体可容纳的外源 DNA 片段都比较小。重组分子的长度不超过野生型 λ 分子长度的 105%, 过长的分子 (>105%) 就不能被包装成噬菌体颗粒。插入型载体也可构建成含几种酶的单一切点, 以便接受几种不同限制酶的降解片段。

取代型载体往往在 λ 基因组可替代的中央区具有同一种限制酶的二个以上的切点, 载体酶切之后, 需把可替代的中央片段分出, 用载体的左右两臂和外源 DNA 重组。重组之后, 载体失去了原有的中央片段, 而接纳了外源基因片段。一般说, 最大的可插入片段为 22kb, 最短的片段不能短于 9—10kb。有些载体可兼有插入和替代两种重组形式, 这取决于载体所具有的可供插入外源 DNA 的限制酶的种类和切点数目。

(一) 由 λ gt 衍生来的载体

λ gt 是野生型 λ DNA 经过消除基因组中央非必需区上二个 EcoRI 位点并引入缺失性突变衍生来的。产生的突变体称为 λ gt· λ B 和 λ gt· λ C, 它们有两个 EcoRI 位点。当用 EcoRI 降解后, 产生的三个片段中, 除了左、右两臂之外, 还有中央片段。其中央片段为野生型 λ 的 EcoRI B 片段时称 λ gt· λ B, 为 λ EcoRI C 片段时称 λ gt· λ C。中央片段均可用外源 DNA 片段取代。此外, 考虑到载体在生物学上的安全性, λ gt 噬菌体只能在实验室用的宿主菌内存活, 因此引入了三种缺失型突变 (W, E, S)。W 基因是噬菌体组装所必需的, E 和噬菌体外壳蛋白合成有重要关系, S 是引起宿主细胞裂解所不可少的。当 λ gt 被导致这三个基因的缺失后, 缺失型噬菌体称为 λ gtWES。向 λ gt· λ B 和 λ gt· λ C 引入上述三种缺失突变后, 载体就分别称为 λ gtWES· λ B 和 λ gtWES· λ C。这种缺失型的 λ 噬菌体衍生物只能在可校正上述三基因的大肠杆菌突变体 SuII 和 SuIII 的宿主细胞中生长。 λ gtWES· λ B 载体的总长度为 40.4kb, 较多地应用于中等大小的 EcoRI 降解片段的克隆。

表 2.3 几种λ载体的特性

载体	大小 (kb)	适用的插入位点	载体类型 ¹⁾	可插入外源片段的容量(kb)	重组子的选择标记 ²⁾	参 考 文 献
λgtWES·λB	40.0	EcoRI	R	2.1—15.1		Leder, P. <i>et al.</i> , <i>Science</i> , 196 , 175 (1977).
λNM540	39.2	HindIII	I	0—11.8	清亮噬斑	Murray, K. <i>et al.</i> , <i>J. M. B.</i> 98 , 551 (1975).
λNM641	39.3	EcoRI	I	0—11.6	清亮噬斑	Murray, N. E. <i>et al.</i> , <i>Mol. Gen. Genet.</i> , 150 , 53 (1977).
λNM728	38.7	HindIII	I	0—12.2	清亮噬斑	同上
λNM762	43.3	HindIII	R	2.6—15.6	lac ⁻	同上
λNM816	44.3	EcoRI	R	1—13	lac ⁻	Wilson, G. G. <i>et al.</i> , <i>J. M. B.</i> 132 , 471 (1979).
λNM1033	42.7	EcoRI	R	6—19	lac ⁻ , Bio ⁻	Murray, N. E., in "LAMBDA II" Ed. by Hendrix, R. W. <i>et al.</i> , 1983, pp395—432.
λNM1149	39.7	EcoRI HindIII	I I	0—11 0—11	清亮噬斑	同上
λNM1151	39.2	EcoRI HindIII BamHI	I I I	0—12 0—12 0—12		同上
Charon4A	45.4	EcoRI	R	7.1—20	lac ⁻ , Bio ⁻	Blattner, F. <i>et al.</i> , <i>Science</i> , 196 , 161 (1977).
Charon21A	41.8	HindIII	I	0—9.1	用 πVX 筛选	同上
Charon22	46.1	XmaI	R	3.8—16.8		Williams, B. G. <i>et al.</i> , in "Genetic Engineering" Ed. by Setlow, J. <i>et al.</i> , 1980, Vol. II, p201
Charon28	40.2*	BamHI	R	6—19		Rimm, D. L. <i>et al.</i> , <i>Gene</i> , 12 , 301 (1980).
Charon30	46.7	EcoRI HindIII BamHI	R I R	4.5—17.5 0—11.7 6.1—19.1	Spi ⁻	同上
Charon34	46.7	poly-L ³⁾	R	8—21		Loenen, W. A. M. <i>et al.</i> , <i>Gene</i> , 26 , 171 (1983).
Charon35	45.8	poly-L ⁴⁾	R	9—22		同上
λL47	40.4	EcoRI BamHI HindIII	R R R	8.4—21.4 4.2—17.2 6.8—19.8	Spi ⁻ Spi ⁻ Spi ⁻	Loenen, W. A. M. <i>et al.</i> , <i>Gene</i> , 10 , 249 (1980)
λ1059	42.5	BamHI	R	8—21	Spi ⁻	Karn, J. S. <i>et al.</i> , <i>PNAS</i> , 77 , 5172 (1980).
λBF101	46	BamHI	R	8—21	Spi ⁻	Maniatis, T. <i>et al.</i> , "Molecular Cloning", 1982, p44.
λEMBL3/4	43.2	EcoRI BamHI Sall	R R R	9—22 9—22 9—22		Frischauf, A-M. <i>et al.</i> , <i>J. M. B.</i> , 170 , 827 (1983).
λgt10	43.3	EcoRI	I	0—7.6	清亮噬斑	Young, R. A. <i>et al.</i> , <i>PNAS</i> , 80 , 1194 (1983).
λgt11	43.7	EcoRI	I	0—7.2	lac ⁻	同上

1) 载体类型分取代型(R)和插入型(I)。

2) 重组子选择标记: lac⁻, 因为重组子丢失了 lac 操纵子, 在含 X-gal 平板上显示无色噬斑; 在 λgt11 中, 因 lacZ 结构基因中插入外源 DNA 而导致半乳糖苷酶不能正常表达, 使呈现 lac⁻; Bio⁻, 生物素测试阴性; Spi⁻ 见正文。

3) 载体来自 stack2221, λDNA 长度为 39.39kb, 来自 2270, DNA 长度是 40.2kb。

4) poly-L 为 -60 核苷酸左右的寡核苷酸多用连接器, 含有 EcoRI, BamHI, HindIII, Sall, SstI, XbaI 六种可供插入用的限制性内切酶位点。

(二) Charon 系列载体

这是由 Blattner 和他的同事为主构建的系列克隆载体, 后来也有其他实验室报道了这类载体。到目前为止衍生物有 30 多种, 包括有插入和取代两种类型, 其中常用的几种列于表 2.3 中。为了使载体在使用上更简便, 生物学上更安全, 以适合于携带真核基因,

还作了一系列的改建。Charon 系列载体常用 Lambdoid 噬菌体 $\phi 80$ 中的同源基因,象 att^{80} 、 imm^{80} 、 QSR^{80} 等片段替代野生型 λ DNA 中相应的基因,有时还将引入的 $\phi 80$ 基因进行诱变。例如遗传标记 KH53 就是指 $\phi 80$ 上与 ci 基因类似区域的缺失型突变,以有效地防止溶源化作用。有的衍生物中引入了大肠杆菌 $lac5$ 片段,可以借噬菌斑在含 X-gal 染料的培养平板上所显的颜色来鉴别重组体和未重组的原载体。噬菌体外壳蛋白基因 A 和 B 在引入 amber 突变(突变标记为 Aam32, Bam1)后,可以增加载体的安全性。凡引入上述突变的 Charon 载体,均在其载体号码后加有大写字母 A。例如 Charon4A, 就是指 Charon4 载体上引入了 Aam32、Bam1 突变。其余的基因表型相同于 Charon4。常被选用的载体有 Charon4A、Charon21A、Charon34 和 35。

Charon4A。载体总长度为 45.4kb, 含有大肠杆菌 $lac5$ 基因, 需要用 $SuIII$ 的宿主细胞, 选用 $EcoRI$ 位点, 可容纳 15—20kb 长度的基因片段。

Charon21A。总长 41.7kb, 需要含校正基因 $SupE$ 或 $SupF$ 的宿主菌。是 $HindIII$ 位点的插入型载体, 容纳的片段可大至 9.1kb。

Charon28。可有两种长度略有差别的载体, 来自 stock2221 的 Charon28, 长度为 39.39kb, 而 stock2270 来源的为 40.29kb, 在其左臂多了一段长 900bp 的未知来源的片段。这是 1980 年发表的可借 $BamHI$ 位点携带外源片段的取代型载体, 容量 6—18kb, 适用于 $MboI$ 或 $Sau3AI$ 部分降解作用产生的外源 DNA 片段的克隆。也有人选用 $HindIII$ 位点插入片段长度较小的片段(小于 11kb)。在 Charon28 载体中不含 chi 位点, 因此, 重组的噬菌斑大小的差别比较大, 在扩增时容易丢失某些重组体, 使基因文库不完整。近年来常用 L47.1、 $\lambda 1059$ 和 $\lambda BF101$, 尤其是 EMBL3 和 4 载体来取代 Charon 28, 用于 $BamHI$ 位点的取代性插入重组。和 Charon28 一起构建成的 Charon30 也是 $BamHI$ 位点的取代型载体。

Charon34 和 35。是由 Loenen 和 Blattner 于 1983 年构建成的新载体(图 2.12)。二个载体的中央片段的长度稍有不同, 在中央片段和左、右两臂相连处各有一 60bp 的多聚核苷酸接头, 接头提供了五种不同限制酶的插入位点 ($EcoRI$ 、 $BamHI$ 、 $HindIII$ 、 $SstI$ 和 $XbaI$)。 $Sall$ 的位点也可用, 但其中一个 $Sall$ 位点得选用靠近多聚核苷酸接头但定位于右臂上的位点。接头中的 $PstI$ 和 $SmaI$ 位点不能用于克隆外源基因, 但它们可用于切割插入的外源 DNA。

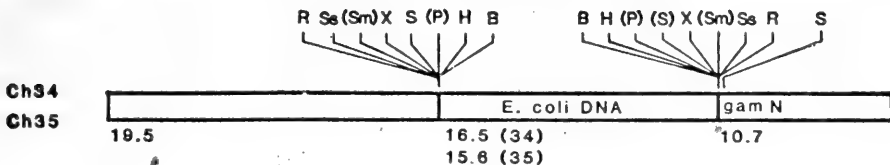


图 2.12 λ 载体 Charon34 和 35。

这是 Charon 34 和 35 载体的简单示意。在中央的非必需基因区的左右两侧各有一个长度为 60bp 的多聚连接子, 其核苷酸序列相同而连接的方向正相反。多聚连接子共含八种限制性内切酶位点, 依次为 $BamHI$ (B)、 $HindIII$ (H)、 $PstI$ (P)、 $Sall$ (S)、 $XbaI$ (X)、 $SmaI$ (Sm)、 $SstI$ (Ss) 和 $EcoRI$ (R)。它们是取代型载体, 可提供六种酶(见图示, 除 $SmaI$ 和 $PstI$ 位点之外的)的插入位点。

(三) λ NM 系列载体

这是由 Murray 和他的同事为主构建的一组系列载体。早年(1975年)获得的二个载体命名为 λ NM426 和 540, 分别是 EcoRI 和 HindIII 位点的插入型载体。以后, 他们把来自 Lambdoid 噬菌体 434 的 *ci* 基因 (*imm434*) 引入到载体中, 构建成免疫性插入型载体: λ NM641 和 728, 前者(641)为 EcoRI 位点, λ NM728 为 HindIII 位点的插入载体。其载体因 *imm434* 基因而呈现混浊噬菌斑, 当插入外源基因片段时, 两种载体的重组噬菌体均转为清亮噬菌斑。 λ NM816、762 和 781 是 *lac*⁺ 表型的载体, 具有大肠杆菌半乳糖苷酶基因活性。但当插入外源片段时, 重组体呈现 *lac*⁻ 表型, 可凭借 X-gal 的琼脂平板上的噬菌斑颜色来鉴别重组体和载体。

(四) λ 1059 及其衍生物

λ 1059 是 BamHI 位点的取代型载体。BamHI 降解 λ 1059 产生的片段为 20kb 的左臂、9kb 的右臂及 13kb 左右的中央片段。在中央片段中含有 *red* 和 *gam* 基因, 载体为 Spi⁺ 表型, 能在 *recA* 宿主菌 (例大肠杆菌 Q358 等) 中繁殖。当外源 DNA 片段取代了中央片段之后, 重组体为 Spi⁻ 表型, 可在 P2 溶源性宿主菌 (大肠杆菌 Q359) 中生长。 λ 1059 的基因组右臂含有 pBR322 DNA 复制起始区序列, 使载体在 λ 溶源菌中可以质粒的复制方式繁殖。但是 pBR322 片段的存在对用克隆 DNA 作探针的筛选工作会带来麻烦, 杂交时易产生假象。即使用限制酶切除 pBR322 部分, 把作探针的基因进行纯化也不理想, 因为难于除尽残存的 pBR322。

为了克服 λ 1059 的弊端, 后来又改建成了 λ BF101, 从 λ 1059 中除去了一个 HindIII 片段, 即把含有 pBR322 的序列去除。同时在中央的非必需基因区再增加 λ 1059 的 8.3 kb 片段, 并且该片段与载体左右臂的连接方向正好与 λ 1059 中的片段方向相反, 使片段中的 *int* 基因灭活。以防止 Spi⁺ 的载体自发性转变为 Spi⁻ 的表型。在 λ 1059 中, 这种自发性转变率为 0.1%。载体这种自发性的 Spi⁺ → Spi⁻ 转变会干扰重组体 (Spi⁻) 的鉴别。

(五) λ EMBL3 和 4

λ EMBL3 和 4 是从 λ EMBL1 衍生而来。当 λ 1059 的含 pBR322 序列的一个 HindIII 片段用来自大肠杆菌含 *trp* 基因的片段替代时, 构建成了 λ EMBL1 载体。以后, 在此基础上又改建成 EMBL3 和 4 二种载体。改建过程包括: 采用突变方法去除靠近右臂的 EcoRI 位点, 然而在中央片段与两臂相连的部位各加入一段长 21bp 的寡聚核苷酸接头。在接头上有三个酶切位点: EcoRI、BamHI 和 Sall。同一载体两接头上的三种酶的位点呈对称排列(图 2.13)。 λ EMBL3 和 4 是一对结构、大小和功能都相同的载体, 但两者的寡核苷酸接头的连接方向相反。载体的全长为 43,225bp, 主要为 BamHI 位点的取代型载体, 也可选用 EcoRI 或 Sall 位点。这两个载体的优点是可采用双酶解制备载体两臂, 不用去除中央片段, 即可直接用于和外源 DNA 片段的重组。例如 λ EMBL3 载体, 可采用 BamHI 和 EcoRI 的联合降解, 产物为左、右两臂带 BamHI 末端, 中央片段带 EcoRI 末端, 除三个大片段外, 还有 BamHI 和 EcoRI 切点间的两个寡核苷酸(九核

苷酸长度)。如果把此寡核苷酸产物分走,载体 BamHI 和 EcoRI 双酶解产生的左、右两臂就不能与中央片段自身连接,因为片段的末端结构不相同。因此在分子克隆时不必将中央片段和载体两臂分离,简化了载体两臂的纯化步骤。选用 λ EMBL3 的另一优点是插入的外源基因片段可用邻近两臂的 SalI 位点回收,得到的外源片段将不含载体序列,避免了在外源片段进一步分析和应用时载体序列的干扰。至于 λ EMBL4,可以采用

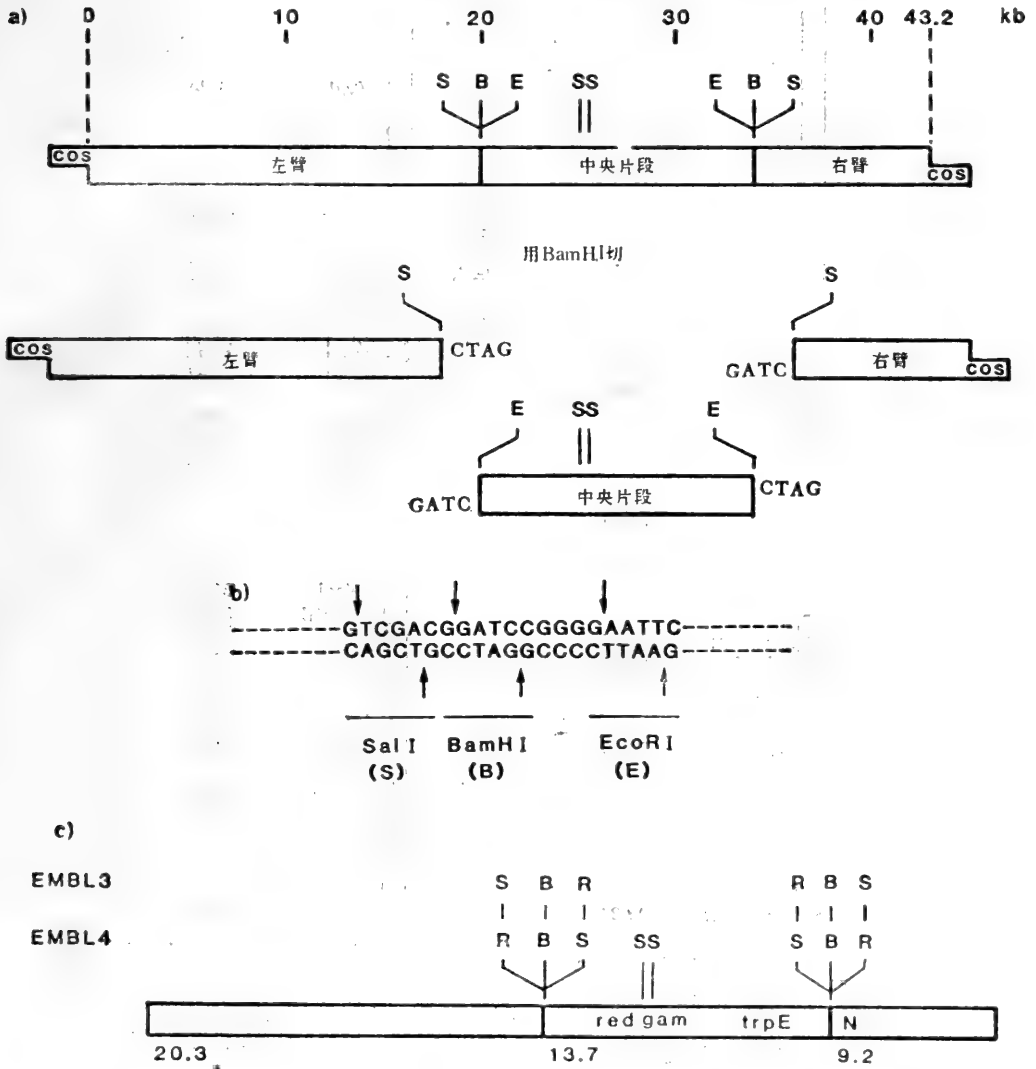


图 2.13 λ 载体 EMBL3 和 4。

λ EMBL3 和 4 均为取代型载体 a)。在中央的非必需基因区的两侧各有一长 21bp 的多聚连接子,其中包含三种限制性内切酶位点 b)。该连接子与噬菌体 DNA 的左右二臂相连时,方向相反,因此三种酶的两个位点呈对称分布;除多聚连接子上的三种酶的位点之外,在中央片段中还额外有二个 SalI 切点。因此用 SalI 降解载体 DNA 时,中央片段会被切成三个大小不等的片段。在 c) 中标出了 EMBL3 和 4 所含的三种内切酶位点的相对位置的不同。除此差别外,两者的结构是相同的。

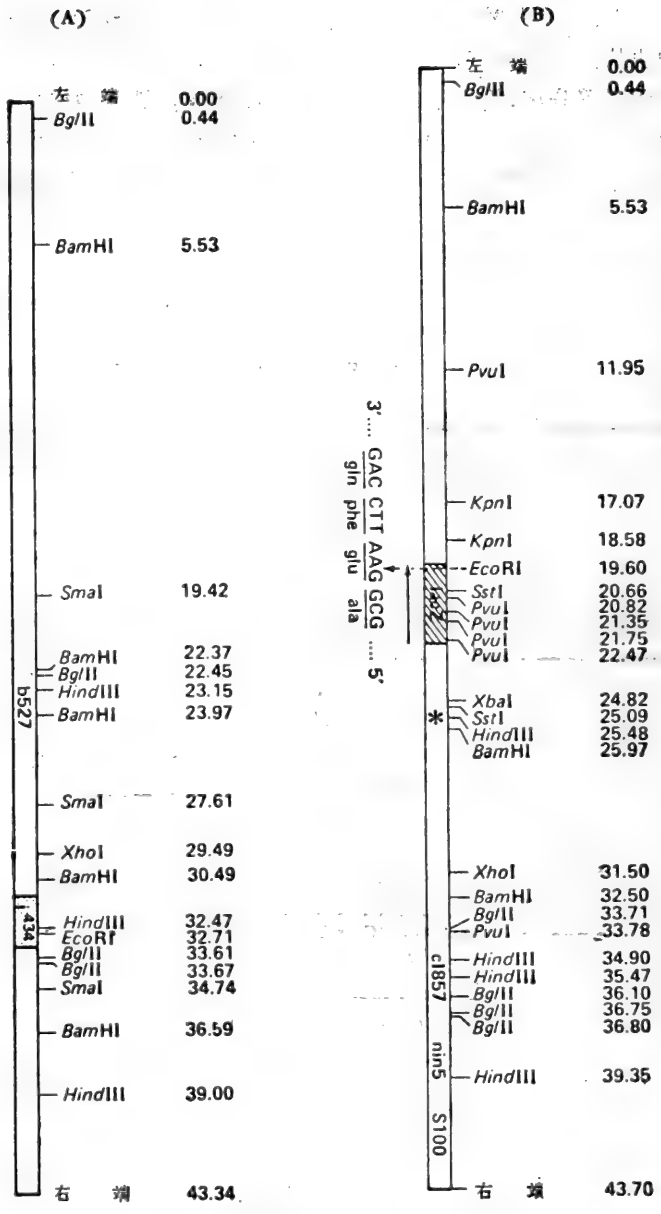


图 2.14 λ 载体 λ gt10 和 λ gt11 的图谱。

图上方的数字表示每个限制酶位点的定位,以 kb 为单位。(A) λ gt10。其中 b527 的缺失突变是去除了野生型 λ DNA 上从 49.1—57.4%的核苷酸序列, imm⁴³⁴片段的取代作用发生于 72.9—79.3%的部位。(B) λ gt11。*标记了 λ 附着位点,水平箭头指示了 lacZ 的转录方向。

BamHI 和 Sall 的联合降解来制备克隆用的载体两臂,克隆后外源片段的切出可选 EcoRI 位点。这一对载体的第三个优点是重组体为 Spi⁻表型,而且由于载体构建中 int 基因已被灭活,载体 (Spi⁺表型)自发的 Spi⁻突变率大为降低。

λ EMBL3 和 4 是目前应用较多的 BamHI 位点取代型载体,其中 L3 优于 L4。当用 BamHI 和 Sall 双酶降解 L4 载体时,不易得到彻底的降解而保证中央片段的末端

结构 (Sall 粘末端)与两臂不一致(两臂为 BamHI 粘末端),因此自身连接的比例高于 L3 载体。而 L3 的 BamHI 和 EcoRI 的联合降解结果满意,因此自身连接的本底很低。

(六) λ gt10 和 λ gt11

这两种载体均为插入型。 λ gt10(43.34kb)有单一的 EcoRI 位点,它定位于 imm^{434} 的 cl 基因上。当外源 DNA 片段插入 EcoRI 位点时, cl 基因失效,重组体呈现 cl^{-} 表型,产生清亮型噬菌斑,而 λ gt10 本身的噬菌斑为混浊状,这有利于重组体的筛选。 λ gt10 可插入的最大片段为 7.6kb 左右。在使用 λ gt10 时,为了尽可能降低载体本身的连接,在 EcoRI 切开 λ gt10 之后,要用碱性磷酸酶(例小牛肠碱性磷酸酶, CIP)除去 EcoRI 粘末端 5'位上磷酸单酯。还有,在重组体繁殖时要选用高频率的溶源性突变株的大肠杆菌(例 hflA150)作宿主菌。载体 λ gt10($cl^{+}imm^{434}$)不能在此宿主菌中形成噬菌斑,而重组体($cl^{-}imm^{433}$)能以正常效率感染宿主细胞,从而有效地消除了载体自身形成的噬菌斑。

λ gt11 是表达型的插入载体。载体 DNA 中引入了大肠杆菌 $lacZ$ 基因片段,在半乳糖苷酶结构基因终止密码上游的 53bp 处有 EcoRI 的单一位点,供外源基因片段的插入。当插入的外源基因的密码阅读框架正确,外源基因可以随半乳糖苷酶结构基因的表达而表达,产生一种含半乳糖苷酶肽链及外源基因肽链产物的融合蛋白。然后用抗体来筛选重组克隆。

为了增加融合蛋白在宿主菌中的稳定性,宿主细胞要选择缺失蛋白水解酶活性的突变株(lon^{-}),以利于用抗体探针筛选。常用的宿主菌是大肠杆菌 Y1088、Y1089 和 Y1090。其中 Y1089 为繁殖溶源菌的宿主细胞。抗体筛选采用噬菌斑的筛选方法。噬菌斑转到含 IPTG (半乳糖苷酶表达的诱导剂)的硝酸纤维素膜上繁殖一定时间,然后从硝酸纤维素膜进行抗体筛选实验。操作较方便,筛选量也大。

一般, λ gt10 和 λ gt11 主要适用于构建 cDNA 库。当有核酸探针时,大多数情况下选用 λ gt10 载体。 λ gt10 可提供大小均一的大的重组噬菌斑,而且通过选择合适的宿主菌,可以只允许重组噬菌体繁殖,减少载体干扰。 λ gt11 的特点是能使外源基因有较高的表达水平。凡能获得抗体作为筛选基因的探针时,应采用 λ gt11 载体。

四、 λ 载体的选择和应用

上面介绍了几种具有代表性的 λ 载体,它们通常适用于克隆较大的外源 DNA 片段,作为构建基因文库的载体。真核细胞的基因结构比原核细胞复杂,有外显子和内含子等结构,使基因的实际长度远比编码结构基因所需的核苷酸序列长得多。选用质粒作克隆载体,不能接纳过长的基因片段,要选择 λ 载体,因为 λ 载体可容纳相当长的外源 DNA 片段。在真核基因的结构研究和染色体 DNA 的步移(chromosome walking)中, λ 的突变体衍生物是有用的载体。另外,用 λ 噬菌体作载体得到的重组克隆是噬菌体,在筛选重组体时,噬菌斑的原位杂交筛选比细菌菌落的原位杂交方法更灵敏,因此噬菌体载体的应用会更广泛。

在实际应用中如何选择合适的 λ 载体? 这取决于实验的目的, 一般说来可遵循以下几点:

(1) 从染色体 DNA 来构建基因文库, 以选择取代型的噬菌体载体为好, 因为它可容纳大的 DNA 片段, 最大容纳为 22kb 左右。由于染色体 DNA 采用 *Sau3AI* 或 *MboI* 的部分降解作用提供随机切割的大片段, 由片段的粘末端结构决定了要选用 *BamHI* 插入位点的载体, 因此 *charon28*、 *λ 1059*, 尤其是 *λ EMBL3* 和 4, 成为较多地用于构建基因文库的替代型载体。

(2) 要从染色体 DNA 中分离某种基因, 而又不需要构建完整基因文库时, 选取载体就比较灵活, 主要决定于用什么限制酶来切割 DNA, 即根据限制酶来选择相应的载体。如果某基因含于 *EcoRI* 酶解物中, 染色体 DNA 或基因大片段可用 *EcoRI* 降解, 载体要选择含 *EcoRI* 插入位点的载体, 选择插入型或替代型载体要视含所需基因的待克隆的 DNA 片段的大小。

(3) cDNA 片段的克隆, 过去多数选用质粒载体, 现在较多采用 *λ gt10* 或 *λ gt11* 插入型载体, 尤其是构建 cDNA 库。当拟用抗体作探针时, *λ gt11* 是有效的表达型载体。

(4) 不少载体具有选择性标记, 例如重组前载体有 *lac⁺* 标记, 重组后变为 *lac⁻*; 或者, *Spi⁺* 表型的筛选, 由重组前的 *Spi⁺* 变成重组后的 *Spi⁻*; 还有噬菌斑混浊度的检测, 由重组前的 *cl⁺imm⁴³⁴* 成为重组后的 *cl⁻imm⁴³⁴* (呈清亮噬菌斑) 等等, 都有助于重组体的鉴别。选用载体时要优先考虑带标记载体。

选用噬菌体载体克隆外源基因, 重组的 DNA 分子要先在体外包装成噬菌体颗粒, 然后去感染相应的宿主细胞, 可以得到高的重组效率。和质粒载体不同的是, 使用 λ 噬菌体载体要提供高效率包装蛋白, 因此克隆过程要复杂些。关于载体的包装及包装蛋白的制备在本书第六章介绍。

λ 载体均是野生型 λ 噬菌体 DNA 的突变体, 不同突变体的存活、繁殖, 对宿主细胞有一定的要求。 λ 噬菌体中基因的诱变大多是琥珀型突变, 但这些琥珀型突变可以被校正, 即对校正基因是敏感的, 称 *sus* (suppressor-sensitive)。 λ 琥珀型突变株需要在“容许宿主细胞” (permissive host) 中生长, 因为只有在这样的细胞中, 琥珀型突变可被校正。容许宿主细胞的校正基因突变能出现于许多基因, 像 *supI*、*supII* 或 *supIII*, 这些校正基因还分别称为 *supD*、*supE*、和 *supF* (*tyrT*)。这些基因是编码 tRNA 的突变体, 它们能翻译链终止密码体, 转运成特定的氨基酸, 使肽链合成继续下去, 而不停止于终止密码上。*supE* 是最有效的校正基因, 而通常的 λ am 突变 (指 λ 基因的 amber 突变) 可以被 *supE*, 或 *supD* 或 *supF* 所校正。偶尔的情况下, 某种琥珀型突变只能被以上三者校正基因中的某一种校正。例如 λ 的 *sam7* 突变体, 只能由携带了 *supF* 的宿主细胞校正。大多数 λ am 突变体能在大肠杆菌 C600 (*supE*) 宿主细胞中生长。除此之外, 要注意的是, 像 *gam⁻* 和 *red⁻* 的 λ 重组体只能在 *recA⁺* 或 *recBC⁻* 宿主中繁殖, 不能选用 *recA⁻* 细胞。若要检测 λ 载体或重组体的 *lac* 基因活性, 一定要选择 *lac⁻* 的宿主菌。总之, 在使用 λ 载体时必须选用合适的宿主细胞。表 2.4 列出了部分供选用的宿主细胞。在第六章基因文库中也介绍了有关宿主细胞的重要性。

最后, 要提及外源基因在 λ 载体中的表达作用。当外源基因和某功能性启动子一起克隆于 λ 载体中, 该基因有可能被表达。随重组噬菌体 DNA 在宿主中复制, 外源基因的

表 2.4 λ 载体的常用宿主菌株(大肠杆菌)

菌株	基因型	主要特点和用途	参考文献
C600	F ⁻ , thi-1, thr-1, leuB6, lacY1, tonA21, supE44, λ^- .	容许细胞, λ 噬菌体指示菌, 供裂解性生长用	Appleyard, R. K., <i>Genetics</i> , 39 , 440(1954).
CSH18	Δ (lac pro), supE, thi, (F' ⁺ lacZ, proA'B')	校正基因型宿主菌, 用于筛选带 lac 基因的 λ 载体, 重组后, lac 片段被取代, 呈 lac ⁻	Williams, B. G. <i>et al.</i> , <i>J. Virol.</i> , 29 , 555(1979)
Kro	recA	为 λ BF101 的宿主菌。	Miniatis, T. <i>et al.</i> , "Molecular Cloning" 1982, p505, Cold Spring Harbor Lab.
K802	hsdR ⁻ , hsdM ⁺ , gal ⁻ , met ⁻ , supE	λ 载体的常用宿主菌	Wood, W. B., <i>J. M. B. 118</i> (1966).
NK5486	thyA ⁻ , rha ⁻ , Sm ^r , lacZam	为 π VX 筛选方法中用菌株, 可测试含 amber 校正基因的噬菌体	Miniatis, T. <i>et al.</i> , "Molecular cloning" 1982, p505, Cold Spring Harbor Lab.
Q358	hsdR _g ⁻ , hsdM _g ⁺ , supE, ϕ 80'	适合 λ 1059 载体用 Su ⁺ 宿主。	Karn, J. <i>et al.</i> , <i>PNAS</i> , 77 , 5172(1980).
Q359	hsdR _g ⁻ , hsdM _g ⁺ , supE, ϕ 80, P2	为 Spi ⁻ 重组体的 Su ⁺ 宿主。	同上
Y-mel	supF	λ SamT 指示菌	Hendrix, R. W. <i>et al.</i> "LAMBDA II", 1983, p434, Cold Spring Harbor Lab. Goldberg, A. R. <i>et al.</i> , <i>Virology</i> , 38 , 200(1969)
WA803	supE, hsd ^s	r ⁻ , m ⁻ , λ 指示菌和繁殖用宿主菌	Wood, W. B., <i>J. M. B. 118</i> , 118(1966).
ED8654	supE, supF, hsdR, trpR, lacY.	r ⁻ , m ⁺ , λ 指示菌和繁殖用宿主菌	Brock, K. <i>et al.</i> , <i>Mol. Gen. Genet.</i> , 146 , 199, (1976).
ED8767	supE, supF, hsdR, lacY, racA56	r ⁻ , m ⁻ , Rec ⁻ 指示菌 (适用于 (cosmid ⁺))	Murray, N. E. <i>et al.</i> , <i>Mol. Gen. Genet.</i> , 150 , 53, (1977).
NM514	hsdR, lyc7 (hfl 等位基因)	r ⁻ , m ⁺ , λ imm ⁴³⁴ cl ⁻ 的选择菌株	Lecocq, J. P. <i>et al.</i> , <i>Arch. Inst. Biochem. Physiol.</i> 81 , 803,(1973).
Y1088	Δ lacU169, supE, supF, hsdR ⁻ , hsdM ⁺ , metB, trpR, ton21, proC::Tn5(pMC9)	λ gt11 宿主菌	Young, R. A. <i>et al.</i> , <i>Science</i> , 222 , 778(1983).
Y1090	Δ lacU169, proA ⁺ , Δ lon, araD139, strA, supF[TrpC22::Tn10] (pMC9)	μ gt11 宿主菌	同上
Y1089	Δ lacU169, proA ⁺ , Δ lon, araD139, strA, hflA150[Chr::Tn10](pMC9)	λ gt11 溶源菌	同上
BNN93	hsdR ⁻ , hsdM ⁺ , supE, thr, leu, thi, lacY1, ton21	λ gt10 宿主菌	同上
BNN102	=BNN93, hfl150[Chr::Tn10]	λ gt10 宿主菌	同上
BHB2688	N205, recA ⁻ , [λ imm ⁴³⁴ , clts, b2, red ⁻ , Eam, sam/ λ]	制备噬菌体包装抽提物 (冻融抽提物用)	Hohn, B. <i>et al.</i> , <i>PNAS</i> , 74 , 3259(1977).
BHB2690	N205, recA ⁻ , [λ imm ⁴³⁴ , clts, b2, red ⁻ , Dam, sam/ λ]	制备噬菌体包装抽提物(超声抽提物用)	同上

拷贝数可大为增加,这意味着基因的表达也将扩增。再者,噬菌体的 S^- 突变作用,使宿主细胞不易裂解,噬菌体颗粒大量积累,也有助于累积基因的表达产物。 Q^- 的噬菌体突变比 S^- 的突变作用更有效。因为 λ 的 Q^- 突变使晚期基因的表达受阻,宿主细胞不会被裂解,同时因噬菌体 DNA 的包装被封闭,即可延长 DNA 的转录作用。一般认为, Q^- 突变对由外源基因自身启动子或借助噬菌体 P_L 启动子的基因表达很奏效。而缺失 S 和 E 基因的重组体有利于由 $P_{R'}$ 启动的基因表达。带 cl 基因热诱导突变体也有益于基因表达,外源基因在原噬菌体阶段为每一溶源菌染色体含一个拷贝。当 $42-45^\circ\text{C}$ 热诱导,噬菌体 DNA 被释放,并开始复制和基因的转录,外源基因也随之表达。虽然 λ 载体可使某些基因表达,但是它更多地作为克隆载体,而不是表达载体。因为来自真核细胞的基因结构较复杂,由于内含子的存在,染色体 DNA 很难提供不间断的结构基因片段;另外,噬菌体的 RNA 聚合酶,不能识别真核细胞 DNA 链中的酶结合位点。启动子结构不同;再者,真核和原核细胞之间启用的密码子频率上有差别,所以 λ 载体主要是作为克隆载体。

参 考 文 献

- [1] [日]松原謙一,质粒,(程光胜、王清海)译,科学出版社,1979。
- [2] Mamiatis, T. *et al.*, Molecular Cloning, Cold Spring Harbor Laboratory, 1982.
- [3] 陈关君译,作为克隆载体的质粒,中国科学院图书馆编 遗传工程,3,48,1981。
- [4] 马贤凯等译,遗传工程,人民卫生出版社,1981。
- [5] Hendrix, R. W. *et al.*, Lambda II, Cold Spring Harbor Laboratory, 1983.
- [6] Dahl, H. H. *et al.*, The Use of Genomic Libraries for the Isolation and Study of Eukaryotic Genes, *Genetic Engineering 2*, Ed. by Robert Williamson, 49-127, Academic Press, 1981.
- [7] Mamiatis, T. *et al.*, Molecular Cloning, A Laboratory Manual Cold Spring Harbor Laboratory, 1982.
- [8] Kaiser, K. and Murray, N. E., The Use of Phage Lambda Replacement Vectors in the Construction of Representative Genomic DNA Libraries, *DNA Cloning*, I, Ed. by Glover, D. M., 1-47, IRL Press, 1985.
- [9] Brammar, W. J., Vectors Based on Bacteriophage Lambda, *Genetic Engineering*, 3. Ed. by Williamson, R., 53, 1982. Academic Press.

第三章 基因工程中常用的核酸酶类

余曙华 王申五 强伯勤

核酸酶类在基因工程中占有非常重要的地位,是基因工程中必不可少的工具酶,分子克隆的许多步骤都需用工具酶。例如,提纯 DNA 时往往需要用 RNase 去除样品中的 RNA;进行基因剪切需要用限制性内切酶;DNA 片段重组需要用连接酶;调整 DNA 片段长度需要用核酸水解酶类或 DNA 聚合酶;进行 DNA 末端标记可用多核苷酸激酶或末端转移酶;cDNA 合成中要用逆转录酶,等等。清楚地了解这些工具酶的作用机制,就能够灵活地应用它们;实验中出现问题也可以从酶的方面进行分析。

有些核酸酶随着被广泛应用和不断研究,可能还会发现它们新的功能。例如绿豆核酸酶,最初认为这个酶只能水解单链 DNA,没有发现其降解位点有什么特点。随着研究的深入,发现绿豆核酸酶在甲酰胺存在下可以很好地水解双链 DNA。有人在做疟原虫 DNA 降解时,又发现这种酶的降解位点多在结构基因的两端,而不在结构基因内部。这一特点如果有普遍性,是很有意义的。

有些核酸酶制剂具有几种酶的活性,因此需要特别注意这些酶制剂在各种不同条件下,主要表现哪种活性。例如大肠杆菌 DNA 聚合酶 I,它主要具有 5'→3'聚合酶活性,但又有 3'→5'外切酶和 5'→3'外切酶的活性,还有焦磷酸解和焦磷酸交换反应的活性。因此,了解大肠杆菌聚合酶制剂的特点,就能灵活地应用它。

在核酸酶中,有一类在基因工程中非常重要的工具酶,即限制性内切酶。关于限制性内切酶,在本书上册中已有详尽介绍。另外,蛋白酶、脱氧核糖核酸酶(DNase)、牛胰核糖核酸酶 A(RNaseA)核糖核酸酶 T1(RNaseT1)在上册的附录中也已经介绍了。蛇毒磷酸二酯酶和牛脾磷酸二酯酶过去多用于 RNA 序列测定,现在应用较少。在这一章里主要介绍 DNA 和 RNA 连接酶、DNA 聚合酶、磷酸激酶和磷酸酶、核酸水解酶类和甲基化酶。

第一节 DNA 和 RNA 连接酶

连接酶是核酸分子重组中所必需的一类酶,当进行 DNA 片段重组时,如果两个 DNA 片段具有相同的粘末端[如图 3.1(1)所示],在 A 和 T 之间或在 C 和 G 之间都可以形成氢键,使两分子相连,但缺少两个共价键[见图 3.1(2)],这样的分子连接不够稳定,必须再用连接酶使 DNA 的 3'端的羟基和 5'端的磷酸基之间形成 3'-5'磷酸二酯键[如图 3.1(3)],这种共价连接的片段才是稳定的。

分子克隆中最常用的是 T4DNA 连接酶,本节将重点介绍。这种连接酶不仅适用于大片段 DNA 的粘末端连接,亦可用于平末端的连接。T4RNA 连接酶虽然不能使 DNA 末端连接,但它可以帮助提高 T4DNA 连接酶连接 DNA 平端的效率。大肠杆菌 DNA 连接酶目前主要用于某些类型的 cDNA 分子克隆,本节也将介绍。

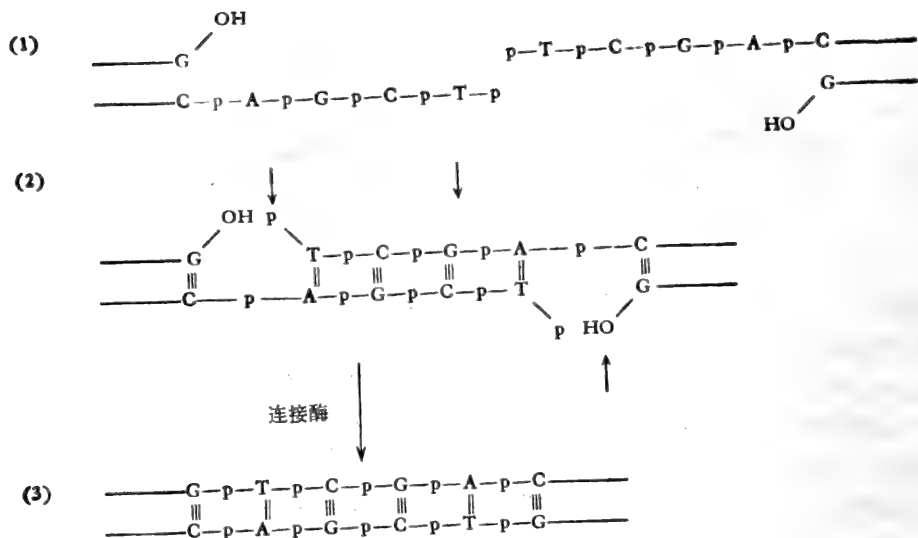
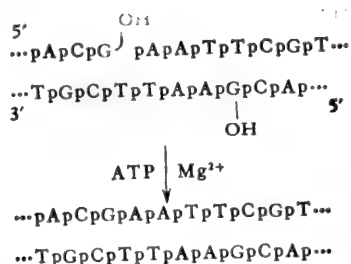


图 3.1 连接酶的作用。

一、T4 DNA 连接酶

T4DNA 连接酶是分子克隆中最常用的一种连接酶，它可以催化如下几种反应。

活性 反应
粘末端连接 举例:



平末端连接 举例:

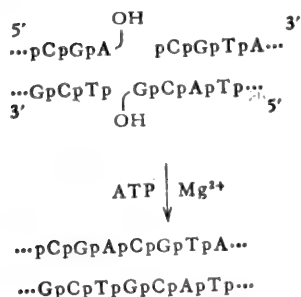


图 3.2 T4DNA 连接酶所催化的反应。

(1) T4DNA 连接酶要求底物为双链核酸,没有核苷酸的缺失,只是磷酸键断裂。在断点的3'端带有羟基,断点的5'端带有磷酸基。T4DNA 连接酶催化它们之间形成3',5'-磷酸二酯键。

(2) 当核酸分子断点的两端为3'-磷酸和5'羟基、3'羟基和5'羟基、3'二脱氧核苷酸和5'磷酸或3'羟基和5'三磷酸基末端时,都不能作为T4DNA 连接酶的底物。

(3) T4DNA 连接酶是已知的唯一可以连接平末端DNA 的连接酶。平末端连接的效率较低,需要大量的酶和高浓度的底物。加入适量的T4RNA 连接酶可以使T4DNA 连接酶连接平末端的效率提高大约20倍。但T4RNA 连接酶本身不能催化平末端连接。有人报道聚乙二醇(PEG)也可以提高平末端的连接作用。在4—8%PEG 存在下T4DNA 连接酶的活性没有什么改变,当PEG 浓度增至12—16%后连接作用可加强。

(4) T4DNA 连接酶的核酸底物必须是双链DNA,或是DNA 和RNA 的杂合链。当以双链RNA 或单链的寡聚或多聚的嘌呤或嘧啶核苷酸链作底物时,连接效率很低。

(5) T4DNA 连接酶反应的最适pH为7.2—7.8,常用的反应液为pH7.6的Tris-HCl 缓冲液。当pH 改变至6.9和8.0时,酶活性分别为pH7.6时的46%和65%。

(6) T4DNA 连接酶需要ATP 参加反应。它所需要的最适 Mg^{2+} 浓度为 1×10^{-2} mol/L。在最适 Mn^{2+} 浓度下,其效率为 Mg^{2+} 的25%,当以核糖同聚物与脱氧核糖同聚物作为底物时,最适 Mn^{2+} 浓度下的效率为 Mg^{2+} 的两倍。

(7) 巯基化合物可以促进T4DNA 连接酶的连接作用,其中二巯苏糖醇(DTT)的效果比二巯基乙醇的效果更好。

(8) 高浓度的 Na^+ 、 K^+ 、 Cs^+ 、 Li^+ 和 NH_4^+ 几乎完全抑制T4DNA 连接酶的活性。多胺、精胺和亚精胺也可以抑制反应。精胺浓度达1mmol/L 时大约抑制酶反应活性的90%。当增加DNA 浓度时可以解除以上抑制作用。

(9) 从琼脂糖凝胶上回收的DNA 片段中容易混杂硫酸粘多糖类物质。这些物质能抑制T4DNA 连接酶的活性。亚精胺可以与硫酸粘多糖结合,降低它对T4DNA 连接酶活性的抑制作用。因此,当用琼脂糖胶上回收的DNA 作连接反应的底物时,往往加入0.25—1mmol/L 的亚精胺以提高连接效率。

(10) 常用的连接反应条件如下:0.5—10 μ g DNA 反应体积为5—50 μ l,加入0.1体积的10倍的连接缓冲液,1u/ μ g 的DNA 的连接酶,在4—16 $^{\circ}$ C 反应12—16小时(EcoRI 粘末端在12 $^{\circ}$ C以下)。

10倍的连接缓冲液制备方法:

0.66mol/L	Tris-HCl(pH7.5)
50mmol/L	$MgCl_2$
50mmol/L	二巯苏糖醇
10mmol/L	ATP

制备好的连接缓冲液分成小管保存在-20—-70%的低温冰箱内。

其中ATP 常常先配成0.1mol/L 的储存液。方法是称60mg ATP 溶于0.8ml 水中,用0.1 mol/L NaOH 调整pH 到7.0,再加水到1ml,分装成小管储存在-20 $^{\circ}$ C 冰箱内。如果ATP 保存不好,其高能磷酸键很容易断裂变成ADP 或AMP,这样会严重影响连接反应的效力。

各种情况下的的连接反应将在第四章基因的剪切与重组中详细介绍。

二、大肠杆菌 DNA 连接酶

来自大肠杆菌的 DNA 连接酶是分子量为 74,000 的单一多肽链,它催化 DNA 片段连接反应的机制与 T4DNA 连接酶的作用相同:①底物必须是双链 DNA;②被连接的片段具有互补的粘末端;③酶催化同一条链上 5'磷酸单酯和相邻的 3'羟基形成 3'、5'磷酸二酯键(见图 3.1)。两者的不同在于大肠杆菌 DNA 连接酶的催化反应需要烟酰胺腺嘌呤二核苷酸(辅酶 I,即 NAD)的参加,而 T4DNA 连接酶需要的是 ATP。反应的第一步是酶的腺苷酰化大肠杆菌 DNA 连接酶释放出烟酰胺单核苷酸(NMN),而 T4DNA 连接酶是释放出焦磷酸(ppi)。第二步是把酶复合物上的 AMP 转移到 DNA 底物的 5'磷酸基上,产生含高能磷酸键的焦磷酸酯键。第三步是与相邻片段的 3'-OH 端相连形成磷酸二酯键并释放出 AMP。

目前纯化大肠杆菌 DNA 连接酶的菌株,是选用 Panasenke 等构建的带有前噬菌体 λ gt4-lop11 ligts7 的大肠杆菌 594(sù) 株。大肠杆菌连接酶的基因被重组到紧挨噬菌体上突变 lop11 启动子上。诱导后,随着噬菌体的复制,产生许多大肠杆菌连接酶的基因拷贝,酶蛋白的量可以达到整个细菌蛋白量的 2—3% 左右。

像 T4 诱导的 DNA 连接酶一样,大肠杆菌 DNA 连接酶能催化双链 DNA 分子间或分子内部的连接反应。它可以把用多聚(dA)作模板的寡聚(dT)连接在一起,也能使以多聚(dT)为模板的寡聚(dA)连接。不过后者的连接效率比前者低很多。大肠杆菌 DNA 连接酶还可以催化 DNA 链 5'磷酸基团和 RNA 链 3'-OH 末端间的连接,对 3'末端为尿嘧啶核糖核酸(U)的多聚[d(A-T)], U 的底物,大肠杆菌 DNA 连接酶或 T4DNA 连接酶可将其环化。

但是大肠杆菌 DNA 连接酶在催化各种底物的连接反应时有一定限制。一些 T4 DNA 连接酶可以催化连接反应的底物,大肠杆菌的 DNA 连接酶对它们不能起作用。例如 DNA 片段间的平末端连接, RNA 链 5'磷酸基和 DNA 链 3'-OH 间的连接,以及 RNA 链之间的连接,等等。所以在分子克隆中,大肠杆菌 DNA 连接酶的应用不如 T4DNA 连接酶广泛。目前,有人应用于 cDNA 的第 2 条 DNA 链的合成。当 RNA 借逆转录酶合成出单链的 cDNA 之后,他们用 RNaseH 来消除 DNA-RNA 杂合双链中的 mRNA,并用 DNA 聚合酶 I 以单链 cDNA 为模板,合成出一系列互补短链的 DNA,这些短链再借大肠杆菌 DNA 连接酶互相连接成第二条 cDNA 链,而得到双链 cDNA。

用大肠杆菌 DNA 连接酶催化的 DNA 连接反应条件如下:

30mmol/L	Tris-HCl(pH8.0)
4mmol/L	MgCl ₂
1.2mmol/L	EDTA
1.0mmol/L	DTT
26 μ mol/L	NAD
50 μ g/ml	牛血清白蛋白
300—600 μ g/ml	DNA 底物

在 16°C 下保温 30—60 分钟。

三、T4 RNA 连接酶

T4RNA 连接酶催化单链 DNA 或 RNA 的 5'磷酸基与单链 DNA 或 RNA 的 3'羟基共价键连接。

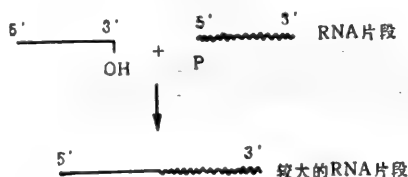


图 3.3 T4RNA 连接酶的功能。

T4RNA 连接酶用于：提高 T4DNA 连接酶在连接双链平末端 DNA 时的效率；寡聚核苷酸的合成；RNA 分子的修饰。在分子克隆中主要是第一方面。

第二节 DNA 聚合酶

分子克隆中常用的 DNA 聚合酶有大肠杆菌聚合酶 I 和它的 Klenow 片段,后者实际来自前者,由于它失去了前者具有的 5'→3'外切酶活性,因而在应用上很有些不同。此外还有 T4DNA 聚合酶。这些聚合酶各有不同特点,在使用时要不同对待。它们都有外切酶活性,但是只有大肠杆菌聚合酶 I 具有 5'→3'外切酶活性。5'→3'外切酶活性是缺口翻译标记 DNA 时所需的,所以在缺口翻译标记 DNA 时常用大肠杆菌聚合酶 I。以上三种聚合酶都有 3'→5'外切酶活性, T4DNA 聚合酶的 3'→5'外切酶活性又远大于其他两种。选用时需要注意这些特点。此外,我们将 AMV 逆转录和末端转移酶也放在这一节里叙述,因为它们也起合成 DNA 的作用。

一、大肠杆菌聚合酶 I

此酶来源于噬菌体 NM964 整合的溶源性大肠杆菌,是一种多功能酶。

(一) 大肠杆菌 DNA 聚合酶 I 的功能

1. 5'→3'聚合酶的功能

在 DNA 模板 3'羟基末端、引物及足量的四种脱氧核苷三磷酸存在时,在合适的离子强度和 pH 条件下,此聚合酶催化 5'→3'端的聚合反应,按模板的碱基配对,从引物的 3'OH 将不同的核苷酸逐个从 5'→3'的方向聚合。常见的聚合反应分以下几种情况:

(1) 间隙填补。将核苷酸加到 DNA 链的间隙处的 3'OH 基端,使双链 DNA 的单链间隙填补起来,形成完整的双链。

(2) 缺口平移。它是通过 5'→3'聚合作用和 5'→3'外切作用相互协调完成的。5'→3'

外切酶从双链 DNA 的缺口处 5'端将核苷酸逐个切下，而5'→3'聚合酶则将核苷酸从 3'羟基端逐个补上，形成沿 DNA 双链 5'→3'方向的缺口移动。这种核苷酸的切补和缺口移动过程被用于将同位素标记的核苷酸标记在 DNA 链上。双链 DNA 的缺口是由 DNaseI 完成的，这个反应常用于双链 DNA 分子的同位素标记。

(3) 链替换。在链替换反应前，如无 5'→3'外切酶将 DNA 链从5'端水解，则在聚合酶的作用下将核苷酸逐个加在 3'羟基端，从而将前方的单链游离，形成换链后的双链 DNA(见图 3.4)。

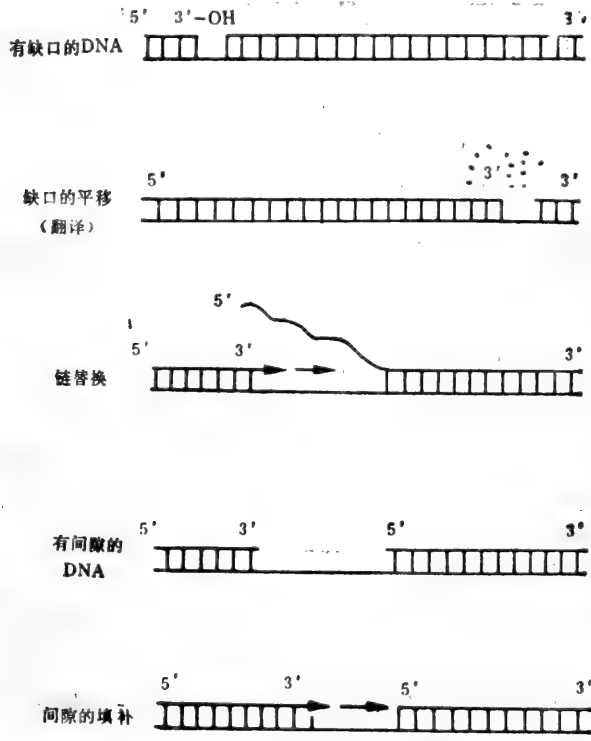


图 3.4 大肠杆菌聚合酶 I 的聚合作用。

2. 3'→5'外切酶的功能

大肠杆菌 DNA 聚合酶 I 的 3'→5'外切酶专一作用于单链 DNA，降解 DNA 成 5'核苷一磷酸(见图 3.5)。在体内，当聚合反应进行时，3'→5'外切酶活性受到抑制，一旦出现错配碱基对，聚合反应停止。3'→5'外切酶能将错误的碱基除去，这在机体内修复中起重要作用。

3. 5'→3'外切酶的功能

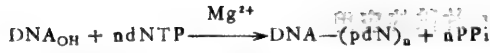
此酶以 5'→3'方向从双链 DNA 的缺口游离 5'端处开始降解双链中一条链的 DNA，但不降解单链 DNA。产生 5'磷酸为末端的寡聚核苷酸或单核苷酸，它能降解 DNA 与 RNA 杂合体中的 RNA 成分。如前所述，5'→3'外切酶在缺口翻译中起重要作用(见图 3.5)。

此外，大肠杆菌 DNA 聚合酶还可以催化焦磷酸酸解反应，即聚合反应的逆反应；以及焦磷酸交换反应，即由焦磷酸解和聚合反应连续重复多次。

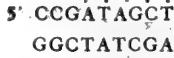
活性

反应

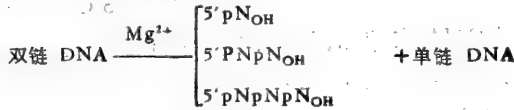
5'→3' 聚合酶



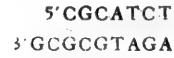
例如:



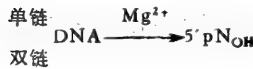
5'→3' 外切酶



例如:



3'→5' 外切酶:



例如:

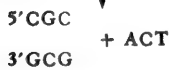
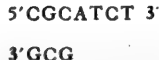


图 3.5 大肠杆菌 DNA 聚合酶 I 的几种作用。

(二) 大肠杆菌 DNA 聚合酶 I 的应用

大肠杆菌 DNA 聚合酶 I 目前常应用于: ①缺口平移法标记 DNA 探针; ②进行 DNA 复制反应; ③将有部分单链缺失的 DNA 修补完整。

缺口平移法标记 DNA 的详细实验方法详见本书上册的核酸的分子杂交部分。

二、DNA 聚合酶的 Klenow 片段

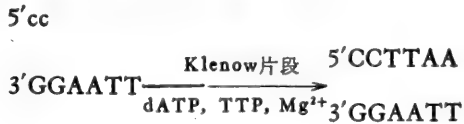
(一) Klenow 片段 (DNA 聚合酶 I 大片段) 的功能

枯草杆菌蛋白酶或胰蛋白酶可将大肠杆菌 DNA 聚合酶 I 降解成分子量为 76,000 和 36,000 的两个活性片段,其中带有 C 末端的大片段叫做 Klenow 片段或 DNA 聚合酶 I 大片段。它具有 5'→3'聚合酶活性和 3'→5'外切酶活性,而带有 N 末端的小片段肽只具有 5'→3'外切酶活性。

(二) Klenow 片段的应用

Klenow 片段由于失去了 5'→3' 外切酶活性, 故在应用上不同于全酶, 一般应用于下列反应中:

(1) 将 5' 突出末端补齐使成平末端。例如:



(2) 进行 DNA 末端标记。基本反应同上, 但需选用 1—3 种同位素标记的 dNTP。由于失去了 5'→3' 外切酶活性, 不会再被切下来, 只会把放射性 dNTP 标记在 DNA 末端上, 可做一般分析用。

(3) 在 cDNA 克隆后进行第 2 条 DNA 链的合成中, 也常应用 Klenow 片段。

(4) 用于 Sanger 氏的双脱氧核苷酸链末端终止法测定 DNA 序列。

(5) 偶尔也用 Klenow 片段的 3'→5' 外切酶活性降解 DNA 的 3' 突出端使成平末端。但是显然不如用 T4DNA 聚合酶好, 因为 T4DNA 聚合酶的 3'→5' 外切酶活性远远大于前者。

用 Klenow 片段进行 5' 突出末端补齐和末端标记采用的是同一反应条件, 具体实验方法参见 DNA 核酸序列分析中末端终止法。

三、T4 DNA 聚合酶

(一) T4 DNA 聚合酶的功能

T4DNA 聚合酶具有 5'→3' 聚合酶活性和 3'→5' 外切酶活性。

(1) 聚合酶活性, 在有单链 DNA 模板、带 3' 羟基的引物和四种三磷酸核苷酸存在时, T4DNA 聚合酶可将核苷酸加在 3' 羟基末端使 DNA 按 5'→3' 方向合成。与大肠杆菌 DNA 聚合酶 I 相比, T4DNA 聚合酶不能利用双链 DNA 的磷酸酯键断裂点(缺口)进行 DNA 链的合成, 因为它不含 5'→3' 外切酶活性。如用平末端的双链 DNA, 常常需用外切酶 III 先将 3' 端切去一段, 才能成为适合于 T4DNA 聚合酶作用的模板和引物(图 3.6)。

(2) 3'→5' 外切酶活性。T4DNA 聚合酶的 3'→5' 外切酶作用于单链 DNA 的活性远远大于作用于双链 DNA 的活性, 因为后者受到 5'→3' 聚合酶活性的影响。T4DNA 聚合酶的 3'→5' 外切酶活性很高, 大约比大肠杆菌聚合酶 I 中的 3'→5' 外切酶的活性高 200—1,000 倍。T4DNA 聚合酶的 3'→5' 外切酶活性大约比自身的 5'→3' 聚合酶高 3 倍。3'→5' 外切酶从 3' 端开始降解 DNA, 降解速度取决于链的长度, 当链的长度增加 10 倍时, 降解速度降低大约 100 倍。

(3) T4DNA 聚合酶的聚合作用与 3'→5' 外切酶活性之间的关系。脱氧核苷三磷酸的浓度以及模板和引物的状态, 都是影响聚合酶和外切酶活性的因素。当脱氧核苷酸缺乏时, T4DNA 聚合酶只表现出 3'→5' 外切酶活性, 而当有充足的脱氧核苷酸存在时, 5'→3' 聚合酶活性远远大于 3'→5' 外切酶活性。在有足量的脱氧核苷酸存在时, 是进行聚合反应

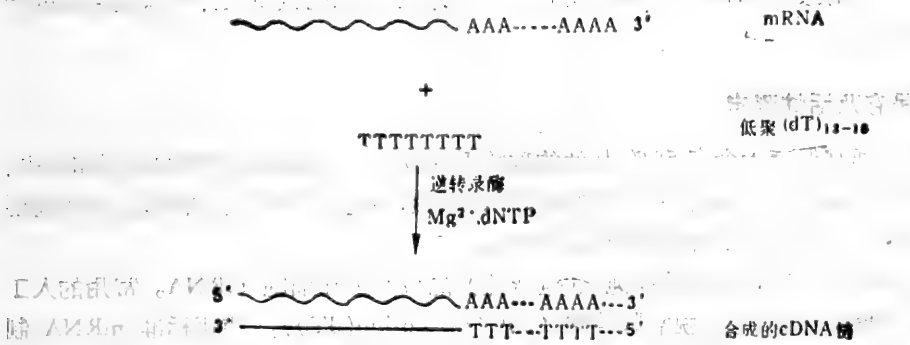


图 3.9 以逆转录酶合成 cDNA。

dCTP	0.1—0.5mmol/L
Tris-HCl pH8.3	100mmol/L
MgCl ₂	10mmol/L
KCl	50mmol/L
DTT	10mmol/L
放线菌 D	50μg/ml
oligo(dt) ₁₂₋₁₈	20μg/ml
RNasin	100u/ml
³ H-dCTP(或 ³² P-dCTP)	50μCi
AMV 逆转录酶	50u

反应在 37℃ 进行，保温 60—90 分钟。加入 0.2mol/L EDTA-Na₂ 至终浓度 0.01 mol/L 终止反应。以 tRNA 为载体，测定三氯乙酸不溶物同位素计数，以计算合成产量及反应效率。

(二) 用于 cDNA 第 2 条链合成

逆转录酶同样可以用单链 DNA 作模板。据报道互补合成的准确性比使用 DNA 多聚酶 I 更高，因此也常用于双链 cDNA 合成。单链 cDNA 合成完成后，得到 RNA-DNA 杂交分子，再以碱解法或加入 RNaseH 去除杂交分子中的 RNA 链，得到单链 cDNA。

双链 cDNA 合成反应溶液组成：

单链 cDNA	0.5μg 左右
dATP, dCTP, dGTP, dTTP	各 1mmol/L
Tris-HCl pH8.3	100mmol/L
MgCl ₂	10mmol/L
KCl	50mmol/L
DTT	10mmol/L
AMV 逆转录酶	50u

反应在 37—42℃ 进行，保温 60—120 分钟。由于单链 cDNA 已带有同位素标记，因

此无须加入标记化合物,只须在反应完成后,用 S1 酶消化单链部分。加入 0.2mol/L EDTA-Na₂ 至终浓度 0.01mol/L 终止反应。以同位素计数的减少来计算双链合成的产量及反应效率。

(三) 酶的保存及活性测定

逆转录酶通常保存于含有 50% 甘油的磷酸缓冲液中。为避免酶活性的降低和丧失,在保存和运输过程中必须保持在 -20℃ 环境下,时间不得超过三个月。否则,在使用前要以标准 mRNA 制剂或人工合成模板鉴定活性。

常用的标准 mRNA 制剂是经纯化(电泳纯)的兔网织红细胞 mRNA。常用的人工合成模板是多聚腺苷酸-寡聚脱氧胸苷酸 (poly(r)A-oligo(dT))。使用标准 mRNA 制剂为模板时,反应溶液与单链 cDNA 合成反应溶液相同,但不加未标记的 dCTP。以多聚腺苷酸-寡聚胸苷酸为模板时,反应溶液为:

Poly(r)A-oligo(dT)	0.1—0.5 μg
³ H-dTTP	50 μCi
Tris-HCl pH8.3	100mmol/L
Mg ²⁺	10mmol/L
KCl	50mmol/L
DTT	10mmol/L
AMV 逆转录酶	50u

37℃ 保温 30 分钟,测定同位素掺入并计算合成效率。

无论使用标准 mRNA 还是人工合成模板,合成效率都应高于 30% (指新合成 cDNA 量与 mRNA 模板量的比值)。

另外,由于逆转录酶和 mRNA 模板都很不稳定,因此所有使用的器皿都必须经过严格的高压消毒或以重碳酸二乙酯(0.1% 浓度)处理。玻璃器皿还应涂布硅油以防止吸附。在单链 cDNA 合成中往往还需要加入核糖核酸酶的抑制剂 (100u/ml)。

五、末端转移酶

末端转移酶的全称是末端转脱氧核苷酰酶 (terminal deoxynucleotidyl transferase, 简称 TdT)。目前商品提供的末端转移酶是从小牛胸腺中分离纯化的。

(一) 末端转移酶的功能

此酶能催化 DNA 链的 3'OH 端加脱氧核糖核苷酸的聚合反应。在 Mg²⁺ 的参加下,末端转移酶把来自脱氧核糖核苷三磷酸 (dNTP) 的核苷酸基加合到单链 DNA 3'OH 端,或者加到双链 DNA 3' 突出单链的末端 OH 上,同时释放出焦磷酸 (ppi)。发现此酶的 Lobban 和 Kaiser 曾认为 3' 和 5' 末端为平末端的完整双链 DNA 分子不能是末端转移酶的有效底物。它们需要用 λ 外切酶预处理以暴露出 3' 单链末端之后,才能接受核苷酸的加合。后来 Roychoudhury 等和 Brutlag 等两个实验小组分别报道,当改变反应条件使双链 DNA 处于不稳定,松弛两端的氢键,末端转移酶也可以在双链 DNA 的两个 3'

末端同时加合核苷酸。反应条件的改变包括降低反应体系中的离子强度和用低浓度 Co^{2+} 来代替 Mg^{2+} 。酶的催化反应如图 3.10 所示。

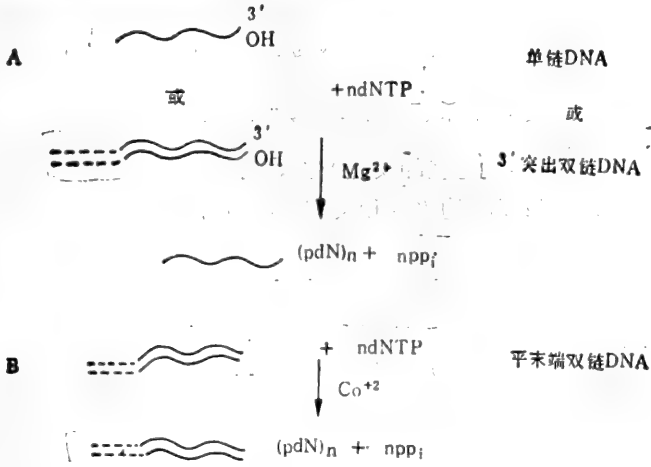


图 3.10 末端转移酶的功能。

(二) 末端转移酶的用途

(1) 在分子克隆实验中，末端转移酶的主要用途是使载体和待克隆的外源基因片段接上互补的同聚体尾部，以便两者能在体外重组。例如采用 PBR322 的重组实验：把载体 DNA 用 PstI 切成线形，在 PstI 的 3'单链突出末端基上加合寡聚 G 的尾部；同时把待克隆的基因片段接寡聚 C，载体和外源 DNA 片段借 G-C 的配对而粘连一起，并用 T4DNA 连接酶在体外共价连接，借尾部寡聚 G 与 C 间氢键粘连的重组分子也可以直接转化进入宿主细胞，达到克隆的目的。不过后者的转化效率低于前者。

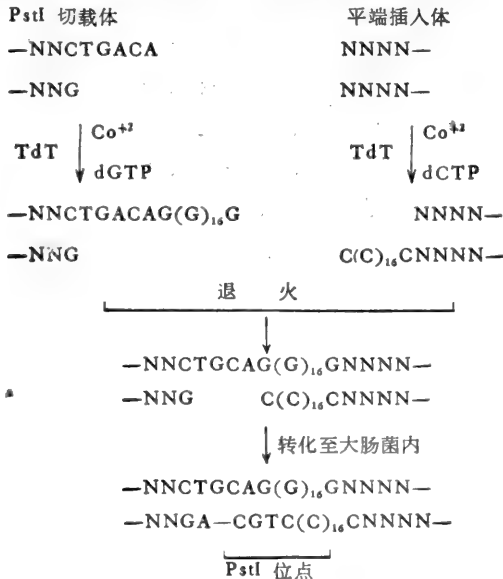


图 3.11 末端转移酶用于 cDNA 克隆。

一般说来, cDNA 的克隆经常选用接互补同聚体尾的方法。当然, 随机切割的 DNA 片段的克隆也可采取接尾的办法。它可以克服随机切割的片段末端不具有特定限制性内切酶位点而不便于克隆的弊端; 同时, 转化到宿主细胞后所得到的转化系将全部为重组的分子, 不可能存在载体自身连接的产物。因为它的尾部带上了相同的核苷酸单链, 不能借氢键互补相连。

四种脱氧核苷三磷酸加合到 DNA 链 3'OH 末端的反应不尽相同。为了获得一定平均长度的同聚体接尾, 需要摸索适宜的反应条件, 包括改变反应的时间、速率和限制脱氧核苷三磷酸的浓度。通过降低酶的浓度和反应温度, 可以有效控制核苷酸的加合速率。具体条件可以参照有关文献。

(2) 末端转移酶的另一个用途是在 DNA 的 3'OH 端标记带同位素的核糖核苷酸, 从而进行 DNA 序列测定。Maxam 和 Gilbert 介绍了把 α - ^{32}P -ATP 加到 DNA 链 3'OH 端的方法, 得到的 DNA-(prA) 产物经碱处理使末端只带一个腺嘌呤核糖核苷酸, 而与此腺苷酸相连的 3' 和 5' 位上磷酸酯均为同位素标记 (DNA- $^{32}\text{PrAP}^{32}$)。

末端转移酶用于 DNA 加尾的反应溶液:

二甲砷酸钾	100mmol/L
Tris-HCl pH7.6	25mmol/L
CoCl ₂	1mmol/L
DTT	0.2mmol/L
[α - ^{32}P]dNTP (加尾核苷酸)	0.1mmol/L
末端转移酶	75u

37°C 反应 10 分钟, 酚、氯仿抽提后回收放射性标记 DNA 部分。

第三节 磷酸激酶和磷酸酶

磷酸激酶催化底物的加磷酸基反应, 而磷酸酶是催化底物的脱磷酸基反应。这两个酶在基因工程中是常用的工具酶。磷酸激酶能在核酸的 5' 末端羟基上加上磷酸基, 而磷酸酶则能从核酸的 5' 末端去掉磷酸基。它们已广泛应用于核酸的结构和功能的研究。例如在 DNA 序列分析和核酸探针的制备中, 需要对双链或单链核酸的 5' 末端标记上同位素 ^{32}P , 就要由磷酸激酶催化, 常用的是 T4 多核苷酸磷酸激酶; 再如双链 DNA 片段进行重组反应中, 载体 DNA 的末端会发生自身连接, 使重组效率降低, 本底增高, 大大降低转化效率。若将载体 DNA 的 5' 末端磷酸基先切掉, 消除载体自身连接的影响, 可使转化效率提高。去掉核酸的 5' 磷酸基就需要由磷酸酶催化, 常用的是牛肠碱性磷酸酶 (CIP)。也可以用大肠杆菌碱性磷酸酶 (BAP)。

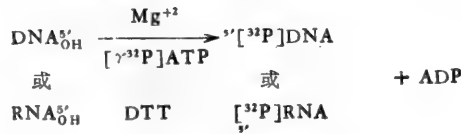
一、T4 多核苷酸激酶

(一) T4 多核苷酸激酶的功能

T4 多核苷酸激酶有多种功能: 催化核酸分子的 5' 羟基磷酸化; 催化核酸分子 5' 磷酸基去磷酸化; 催化 DNA 磷酸基的交换反应; 催化 DNA 3' 磷酸基去磷酸化。

1. 核酸分子5'末端羟基磷酸化的活性

这是 T4 多核苷酸激酶最主要的功能。它催化三磷酸核苷酸的 γ 磷酸基取代核酸 5'末端的羟基。当 γ 磷为 ^{32}P 时,就可以通过这一反应把核酸分子的 5'末端标记上同位素。



举例

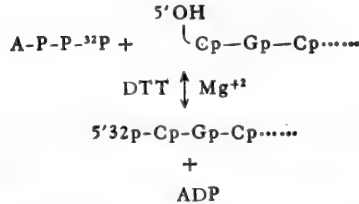


图 3.12 T4 多核苷酸激酶催化的磷酸化反应。

此反应有以下特点:

- (1) 底物为 5'末端带有羟基的单链或双链的 DNA 或 RNA 和寡核苷酸。
- (2) 在标记反应中单链 DNA 的 5'羟基末端比相应的双链 DNA 的两链间隙处的 5'羟基末端标记率大约高 10 倍。
- (3) 高浓度的 ATP 有助于标记上间隙处的 5'羟基末端,但不能将单链间缺口处的 5'羟基末端标记上 5'磷酸基。
- (4) 带 5'突出端的双链 DNA 比带 3'突出端和平端的双链 DNA 更容易标记上。如果在高浓度的 ATP 条件下,后两者的 5'羟基末端也可以被磷酸化。
- (5) 过量的 KCl 可以抑制磷酸化反应。T4 多核苷酸激酶也可以使碱基带保护基团的核酸 5'末端磷酸化,这可以用于多核苷酸链的化学合成中。
- (6) 反应中所需的 5'三磷酸核苷酸常规使用 ATP,实际上 CTP、UTP、GTP、dATP 和 dTTP 也有同样效果。
- (7) T4 多核苷酸激酶的 pH 适应范围为 7.4—8.0,最适 pH 为 7.6。
- (8) 在缺乏 Mg^{2+} 时,该酶失去活力。最适 Mg^{2+} 浓度为 $1 \times 10^2 \text{ mol/L}$ 。
- (9) 反应中加入 5mmol/L DTT 可获得高酶活性。10mmol/L 的 2-巯基乙醇和 10 mmol/L 谷胱甘肽使酶活力只能分别达到 5mmol/L DTT 的 80% 和 70%。在缺乏巯基化合物时,活性只为 5mmol/L DTT 时的 2%。
- (10) 当单链 DNA 作为底物时,NaCl 可以促进反应。而当双链 DNA 作为底物时,NaCl 则抑制反应。
- (11) 精胺一类的多胺化合物可以促进磷酸化反应,1.7mmol/L 精胺可以提高反应速度大约 30 倍。

2. 催化 DNA5' 磷酸基去磷酸化的活性

在有过的 ADP 时, T4 多核苷酸激酶可以催化多核苷酸的 5'末端磷酸基,脱磷反应产生 5'羟基末端和 ATP。这一反应用咪唑-盐酸缓冲液, pH 在 6.2 比较好。去磷酸化的效率比磷酸化的效率低得多。

3. 交换反应

由于去磷酸基反应的可逆性，反应中有过量 ADP 时，多核苷酸激酶可以催化 ATP 的 γ - ^{32}P 和多核苷酸的 5'-P 之间的交换。这种反应可用于带有 5'-P 的多核苷酸的标记。其反应见图 3.13。

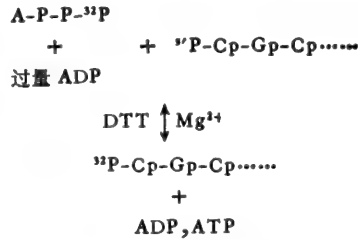


图 3.13 T4 多核苷酸激酶催化的交换反应。

交换反应需要 Mg^{2+} 和 2-巯基乙醇。这一反应可以被磷离子抑制，50mmol/L 磷酸钾可以抑制 60% 的酶活性。最适缓冲条件为咪唑-盐酸缓冲液，pH6.2。带 EcoRI 末端的 DNA 片段，在 pH6.6 时，交换反应中的 ATP 浓度为 $10\mu\text{mol/L}$ ，而 ADP 为 $300\mu\text{mol/L}$ 。

尽管标记反应是不完全的，各种核苷酸的 5' 磷酸末端都可以按交换反应标记上。交换反应进行的完全程度与底物量和酶量有关，而与底物分子大小 ($4 \times 10^3 - 14 \times 10^6$) 无关。单链寡核苷酸更易标记；3' 突出末端的 DNA 标记比较困难，在足够量的酶作用下，标记率可达 70%；双链 DNA 缺口处的 5' 末端磷通过交换反应的标记率只有 5' 突出末端的 $\frac{1}{30}$ 。

4. 催化 3' 末端磷酸基去磷酸化的活性

T4 多核苷酸激酶中除了含有 5' 磷酸化酶的活性外，还具有 3' 末端去磷酸化的酶活性，催化 3' 末端磷酸基水解成为羟基。其底物可以是 3' 磷酸脱氧核苷，3', 5' 二磷酸脱氧核苷和 3' 磷酸多核苷酸 (见图 3.14)。

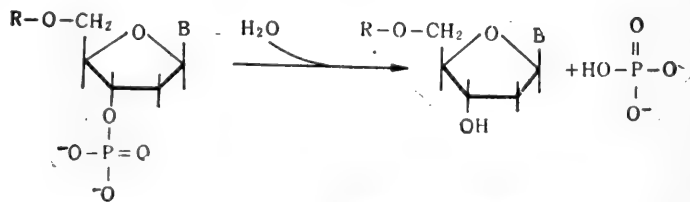


图 3.14 T4 多核苷酸激酶的 3' 磷酸酶活性。

B 为碱基。R 为 H、 PO_5^{2-} 、单核苷酸或多核苷酸。

3' 磷酸酶的反应最好用 Tris-顺丁烯二酸，pH5.8—6.2 缓冲体系。在缺乏 MgCl_2 时，仅可得到 2% 的酶活性，最适 Mg^{2+} 浓度为 8mmol/L。反应需要巯基化合物 (见图 3.14)。

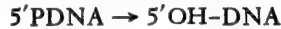
(二) T4 多核苷酸激酶的使用

T4 多核苷酸激酶常应用于：在化学法测定 DNA 的核苷酸序列时进行 γ - ^{32}P 的末端标记；在不含 5' 磷酸基团的片段连接之前用多核苷酸激酶加上 5' 磷酸基；磷酸化合成限

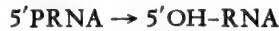
制性内切酶的接头；DNA一端³²P 标记，再加部分降解法绘制 DNA 物理图谱； γ -³²P 标记寡核苷酸探针。在实际应用中多是利用 T4 多核苷酸激酶使核酸 5' 末端羟基磷酸化的活性。其磷酸化反应的具体条件，请见本书上册的化学法测定核苷酸序列部分。

二、碱性磷酸酶

碱性磷酸酶 (BAP、CIP) 催化核酸的 5' 末端脱去磷酸基团的反应。其底物可以是单链或双链的 DNA 和 RNA。三磷酸核糖核苷或三磷酸脱氧核糖核苷。



或



常用的小牛肠碱性磷酸酶 (CIP) 的分子量为 140,000，是一种糖蛋白，由两个亚基组成。此酶是一种含锌的酶，每分子含 4 个锌原子。在 pH7.5—9.5 等碱性条件下很稳定，而在酸性条件下则很快失活。该酶的硫酸铵悬液在 4℃ 存放是稳定的。

实验中，在碱性磷酸酶处理后进行下一步反应前，往往需要将碱性磷酸酶完全失活。CIP 在 SDS 中加热至 68℃ 可以完全失活，而 BAP 则不能。因此往往在 68℃ 下进行反应以抑制 BAP 制剂中残留的核酸酶活性。为了去除操作中微量的 BAP 活性，需要进行多次酚-氯仿抽提及凝胶电泳纯化 DNA 片段。因此一般多用 CIP。

碱性磷酸酶常用于以下两种实验：①在 DNA 或 RNA 5' 末端标记 γ -³²P 时，先用 CIP 去除 DNA 或 RNA 5' 末端的磷酸基，这样带同位素标记的磷酸基可以与 DNA 或 RNA 相连接，使它们末端标记上 ³²P 同位素；②在载体和目的基因重组的反应中，由于载体两端有时是同样的粘末端，进行连接反应时载体可能自身连接，这样不利于外源基因插入；在筛选重组 DNA 时，要将其与自身连接的载体相区别，也很麻烦。如果用 CIP 去除载体末端的磷酸基，就可以防止载体自身连接。由于目的基因末端带有磷酸基，它可以与载体相连。这样有利于外源基因插入并容易筛选到重组基因。

商品碱性磷酸酶一般是悬浮在硫酸铵溶液中，在使用碱性磷酸酶之前必须除硫酸铵，否则将影响碱性磷酸酶反应之后的进一步实验。前面已经讲过，碱性磷酸酶去除 5' 末端磷之后往往用于进行两种实验，一是用多核苷酸激酶将带放射性标记的 γ -³²P 加到 DNA 的末端使其带上末端标记，二是用连接酶将目的基因插入载体。硫酸铵既影响多核苷酸激酶的活性，也影响连接酶的活性，因此在使用碱性磷酸酶之前需要用 Sephadex G-75 层析柱将硫酸铵除去。

除去硫酸铵的方法：例如，3mol/L (NH₄)₂SO₄ 悬浮的商品酶 BAP 或 CIP (活性为 400u/mg, 共 160u)。将样品用 Eppendorf 离心机离心 5 分钟后去上清液。沉淀溶解在 100 μ l 20mmol/L Tris-HCl (pH8.4) 和 100ml KCl 中，加在预先用同样缓冲液平衡好的葡聚糖凝胶 G-75 柱 (0.7 \times 30cm) 上，每 150 μ l 收集一管，用分光光度计 280nm 波长监测流出液的 OD 值，可画出一条曲线，并测定曲线高峰处各管的碱性磷酸酶活性。

碱性磷酸酶 (CIP 或 BAP) 活性的测定方法如下。

反应管中加入

0.5ml 0.05mol/L Tris-HCl (pH7.4)

0.5ml 0.01mol/L	对磷酸硝基酚
0.2ml 0.01mol/L	MgCl ₂
1.05ml	H ₂ O

2-10 CIP 样品流出液(空柱以柱平衡代替)

于 37°C 保温 1 小时后加入 1.75ml 0.2mol/L NaOH 终止反应。在 430nm 测定 OD 值。在此条件下,以每分钟水解 1 μ mol 的对磷酸硝基酚为一个单位。

为了使酶活性保持稳定,需要将有酶活性的各管加入等体积的甘油。

还有一种分子生物学专用的碱性磷酸酶制剂 (CIP),于三乙醇胺缓冲液, pH7.6 (三乙醇胺 30mmol/L, NaCl 3mol/L, MgCl₂ 1mol/L, ZnCl₂ 0.1mmol/L)。可不经处理直接应用,该酶制剂可保持在 4°C, 6 个月内活性不会降低。

碱性磷酸酶脱磷的条件可参照下面的方法。

(1) 限制性内切酶完全降解 DNA 以后, 用与样品等体积的酚-氯仿-异戊醇 (25:24:1) 溶液抽提一次, 加入 2.5 倍的乙醇和 1/10 体积的 3mol/L NaAc, pH5.0。在 -20°C 过夜或在液氮中冻 10 分钟, 然后在 Eppendorf 离心机中离心, 去掉上清液, 沉淀用 70% 乙醇洗一次(一般用 1ml 70% 的乙醇沿管壁轻轻加入, 不要把 DNA 沉淀冲起来), 然后再离心, 弃去乙醇。洗的目的是为了去除 3mol/L NaAc。

(2) 将 DNA 溶解在 20 μ l 10mmol/L Tris-HCl (pH8.0) 的溶液中。加 10 倍的 CIP 缓冲液 5 μ l, 加灭菌蒸馏水到 48 μ l, 加入 CIP 0.01 单位/pmol 与 5' 末端 DNA。

10 倍的 CIP 缓冲液:

0.5mol/L	Tris-HCl (pH9.0)
10mmol/L	MgCl ₂
1mmol/L	ZnCl ₂
10mmol/L	亚精胺

(3) 保温: 5' 突出端的 DNA 在 37°C 保温 30 分钟后, 加入同样一份 CIP, 继续保温 30 分钟。平末端或 3' 突出端的 DNA, 在 37°C 保温 15 分钟后 56°C 保温 15 分钟, 然后再加入同样一份 CIP, 并同样保温一次。

(4) 终止反应: 加 40 μ l 灭菌蒸馏水, 10 μ l 10 倍的 STE 缓冲液和 5 μ l 10% SDS, 加热 68°C 15 分钟。

10 倍的 STE 缓冲液 (TNE)

100mmol/L	Tris-HCl (pH8.0)
1mol/L	NaCl
10mmol/L	EDTA

(5) 样品用等体积的酚-氯仿-异戊醇 (25:24:1) 抽提两次, 再用等体积的氯仿-异戊醇 (24:1) 溶液抽提两次。

(6) 这样处理后的样品可能仍含有某些物质会影响下一步的连接反应, 需要将 DNA 样品走琼脂糖凝胶电泳, 再从胶中回收 DNA 片段。经过这样处理后的样品才合适进行连接反应。也有人将样品过一个葡聚糖凝胶 G50 的离心柱, 离心后的样品按上面 (1) 的方法用乙醇沉淀。

经过上面处理的 DNA 可以进行连接反应。通过连接反应和转化的结果可以分析

GIP 处理是否恰当。

进行连接反应和转化时要有适当的对照。例如用噬菌体作载体时可以做如下几种实验：① 不用内切酶降解的噬菌体的感染率。这一实验可以看出完整噬菌体的质量如何；② 用内切酶降解并用 CIP 处理后的噬菌体的感染率。这一步应该没有细菌被感染；如有少量细菌被感染，可能是某些 DNA 没有被内切酶降解干净的结果；③ 经内切酶降解，CIP 处理的 DNA 又进行自身连接反应后的转化率。这一实验可看出 CIP 处理的结果如何，如 CIP 处理得不够，那么连接反应后感染率可能很高。但是也不能用 CIP 处理到完全没有自身连接反应，因为 CIP 处理过份之后，外源 DNA 不易插入；④ 经内切酶降解，CIP 处理后又加入外源 DNA 进行重组连接反应的 DNA 的感染率。如果 CIP 处理过头，这一实验的感染率将很低，甚至没有感染的细菌。实际工作中 5' 末端 DNA 的 CIP 的用量并不是在任何条件下任何 DNA 都按上面所说 0.01 u/pmol，往往需要经过一定的摸索，才能掌握好 CIP 处理一步中 CIP 的用量。

第四节 核酸水解酶

核酸水解酶是一类可以降解核酸的酶。基因工程中的所谓基因剪切，就是用核酸水解酶降解核酸成较小的片段。从整个染色体基因组中分离纯化出所需要的特定基因，就需要用核酸水解酶把核酸降解成较小的片段，将这些片段克隆化，再从众多的克隆中筛选出我们所需要的基因。另外，在做基因表达时，也常需要核酸水解酶，用它来水解 DNA，以调整结构基因和启动基因之间的距离，调整核苷酸框架处于合适的三联密码位置等，争取获得高效表达。

核酸水解酶的种类很多，根据其作用机制、对底物的要求以及作用方式的差别，可以有不同的分类标准。

(1) 按照对底物作用的专一性可将核酸水解酶分成三类：只作用于 RNA 的叫做核糖核酸酶(简称 RNase)，只作用于 DNA 的叫做脱氧核糖核酸酶(简称 DNase)，既作用于 RNA 又作用于 DNA 的就叫做核酸酶。

(2) 按照对底物二级结构的专一性也有三类：作用于单链的核酸酶，作用于双链的核酸酶，以及既可作用于单链又可作用于双链的核酸酶。

(3) 按照对底物的作用方式又可将核酸酶分成三类：核酸内切酶、核酸外切酶以及具有内切外切双重功能的核酸酶。核酶内切酶的降解位点在多核苷酸链的内部；核酸外切酶则是从多核苷酸链的末端起，将核苷酸逐个水解下来。

(4) 按照磷酸二酯键的断裂方式可将核酸酶分成两大类：3'核酸酶和 5'核酸酶。3'核酸酶的作用点在 3'羟基和磷酸残基之间连接键上，其产物是以 3'羟基和 5'磷酸为末端的单核苷酸或寡核苷酸，如图 3.15(1)所示。5'核酸酶的作用点在 5'羟基和磷酸残基之间的连接键上，其产物是以 5'羟基和 3'磷酸为末端的单核苷酸或寡核苷酸，如图 3.15(2)所示。

此外，酶对核苷酸的碱基的识别也可能有选择性，有些酶专门识别嘧啶，另一些专门识别嘌呤。

常用的核酸水解酶类有各种限制性内切酶、SI 核酶酶、外切酶 III、Bal31 核酸酶、绿豆核酸内切酶。另外还有 λ 外切酶、外切酶 VII 和 RNaseH。其中限制性内切酶在本书

上册已有专题介绍,其他各酶本节将一一介绍。

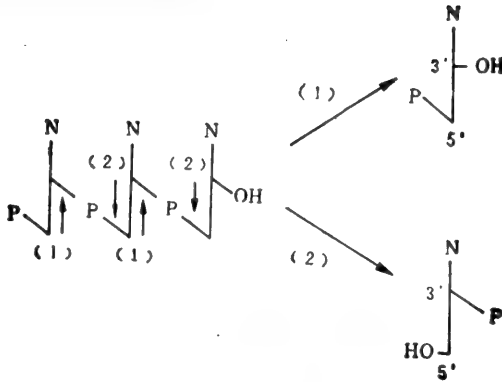


图 3.15 3'核酸酶和 5'核酸酶的作用点。
(1) 3'核酸酶, (2) 5'核酸酶。

一、外切酶 III

外切酶 III 的作用底物为带 3'羟基端的双链 DNA, 它按照 3'→5' 方向逐个降解核苷酸, 其产物是 5'单核苷酸。这个酶还有其他三种酶活性, 即对无嘌呤 DNA 的特异性内切酶活性、RNaseH 的活性和 3'磷酸酶的活性。

当用高活性的外切酶 III 降解 DNA 时, 在 90mmol/L KCl 和室温条件下, 大约每分钟从 3'末端降解 10 个核苷酸。

另外, 外切酶 III 还用于制备单链的 DNA, 作为 DNA 聚合酶的模板。

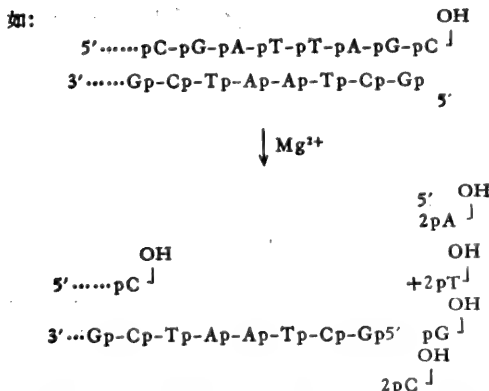
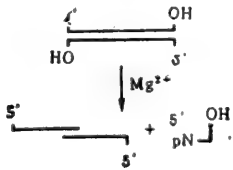


图 3.16 外切酶 III 的外切反应。

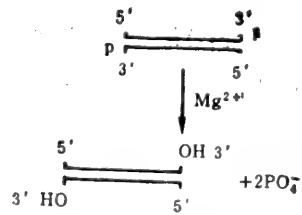


图 3.17 外切酶 III 中的 3'磷酸酶反应。

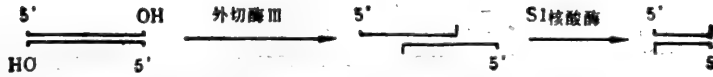


图 3.18 外切酶 III 和 S1 核酸酶联合将双链 DNA 片段缩短。

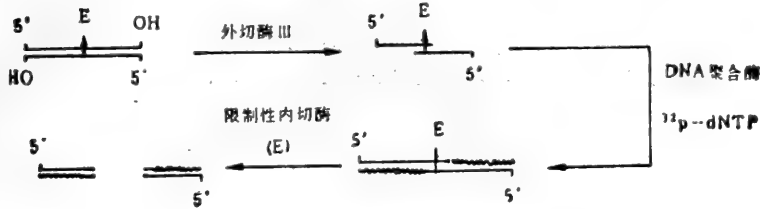


图 3.19 双链 DNA 的同位素标记。

E 代表某一限制酶的位点。

二、S1 核酸酶

S1 核酸酶降解 RNA 和 DNA,降解 DNA 的效力大约是降解 RNA 效力的 5 倍。纯化的酶对单链 DNA 有高度的特异性。只有存在大量 S1 核酸酶时,双链 DNA 才可能被完全降解。其降解位点在双链 DNA 缺口处或者小片段单链部分。S1 核酸酶的产物是 5' 磷酸单核苷酸或寡核苷酸。

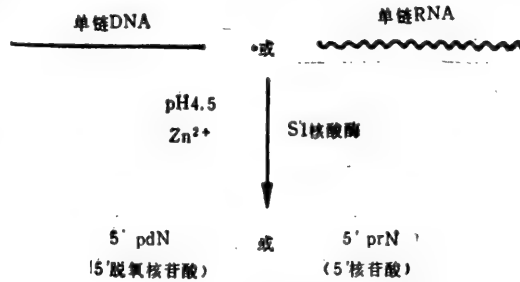


图 3.20 S1 核酸酶降解单链核酸。

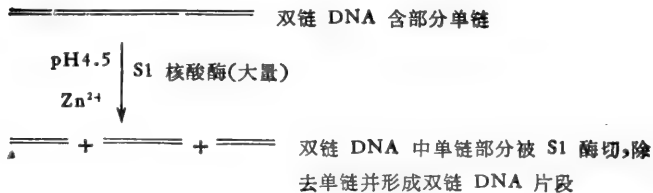


图 3.21 S1 核酸酶可切割存在于双链 DNA 中的单链部分,最终把 DNA 切成小片段。

S1 核酸酶和外切酶 III 联合使用可以缩短 DNA 的长度,在经典的 cDNA 克隆方法中也使用 S1 核酸酶去除双链 cDNA 中的单链部分(见图 3.22)。

此外, S1 核酸酶还用于将 DNA 的粘末端切成平末端。

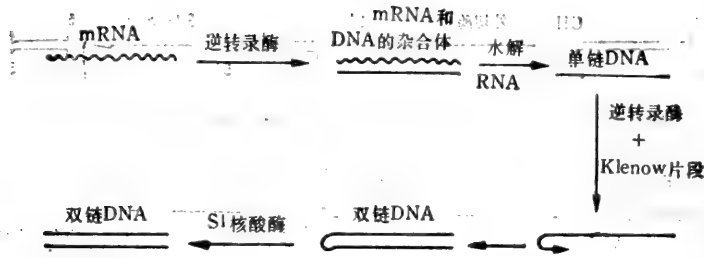


图 3.22 在 cDNA 合成过程中 S1 核酸酶的用途。

三、绿豆核酸内切酶

绿豆核酸内切酶的作用底物为单链的多聚核苷酸和多聚脱氧核苷酸。它以大约同等的速度降解这两种底物，产生 5' 磷酸为末端的单核苷酸和寡聚核苷酸。它还具有 3' 核苷酸酶的作用，这种核苷酸酶也作用于单核苷酸和寡聚核苷酸的 3' 末端磷酸基。

绿豆核酸内切酶并不识别专一的核苷酸序列，降解位点一般为 AT 丰富区。T. F. Mccutchan 等人在酶切疟原虫 DNA 时发现，绿豆核酸内切酶降解位点在结构基因前面或者后面，它不在结构基因内部降解 DNA。也就是说能使我们得到完整的结构基因。

Mccutchan 等人用绿豆核酸内切酶降解疟原虫基因组 DNA 及克隆的疟原虫 DNA，然后用合成的寡核苷酸等作探针进行杂交，可以得到近似大小的片段。通过核酸序列测定，得知这个酶的降解位点在结构基因的前后。他们推测这个酶的降解位点可能与 DNA 的二级结构有关，与一级结构关系较小。

绿豆核酸内切酶的最适 pH 为 5。但是酶活性在 pH5 时会迅速地丧失。加入 0.1 mmol/L Zn^{2+} 和 1 mmol/L 巯基化合物(半胱氨酸、二巯苏糖醇、谷胱甘肽)，可以稳定酶活性。

当用绿豆核酸内切酶降解双链 DNA 时，要在反应液中加入相当量的甲酰胺。甲酰胺可以使双链 DNA 解链，这时绿豆核酸内切酶就可以发挥其作用。在没有甲酰胺的条件下，绿豆核酸内切酶水解双链 DNA 的活性仅仅是单链 DNA 的 1/1000。

所用的甲酰胺量随不同的 DNA 和所需产物片段大小而变化，甲酰胺的含量一般在 30—50%，甲酰胺浓度越大，DNA 越易降解，生成的片段越小。

特定的 DNA 在特定的条件下，用绿豆核酸内切酶降解，可以产生特异性的区带。这一点与 S1 核酸酶不同，S1 酶作用后的电泳区带往往显示一系列非特异大小的片段，看不出集中的区带。

绿豆核酸内切酶的降解条件如下：

NaCl	0.2mol/L
ZnSO ₄	1mmol/L
NaAc	30mmol/L
酶	1u/ μ g DNA
甲酰胺	30—50%

反应体积 100 μ l, pH4.6, 在 50 $^{\circ}$ C 保温 30 分钟。反应后用 0.01mol/L EDTA 将反应液稀释 4 倍。经酚提乙醇沉淀回收 DNA, 然后再进行下一步反应。

所用的甲酰胺需要用 AG501-88 树脂处理。取 100ml 甲酰胺加入 6g 树脂, 搅拌 3 小时后静置使树脂沉淀, 倾出上层的甲酰胺即可使用。

所用的 ZnSO₄ 不易溶解, 一般配成 50mmol/L ZnSO₄ 储存液。方法是: 称好 ZnSO₄ 后, 加入适量水, 然后加入少量 HAc 并搅拌, 直到 ZnSO₄ 全部溶解为止, 最后加水达到所需体积。

四、外切酶 VII

外切酶 VII 可以从单链 DNA 的 3' 和 5' 两端降解 DNA, 释放出小寡核苷酸(详见图 3.23)。

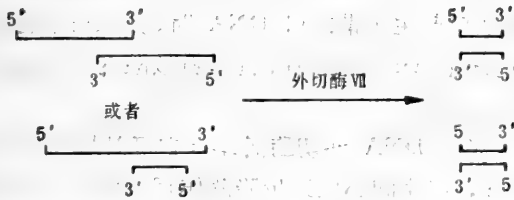


图 3.23 外切酶 VII 降解 DNA 的单链(部分)。

外切酶 VII 可以用于将带粘末端的双链 DNA 末端消平, 它还可以与外切酶 III 或 λ 外切酶联合使用将双链 DNA 缩短。

五、 λ 外切酶

λ 外切酶以具有 5' 磷酸末端或者具有 5' 突出端的双链 DNA 为底物, 从 5' 磷酸末端逐个降解核苷酸, 产生 5' 单核苷酸(见图 3.24)。

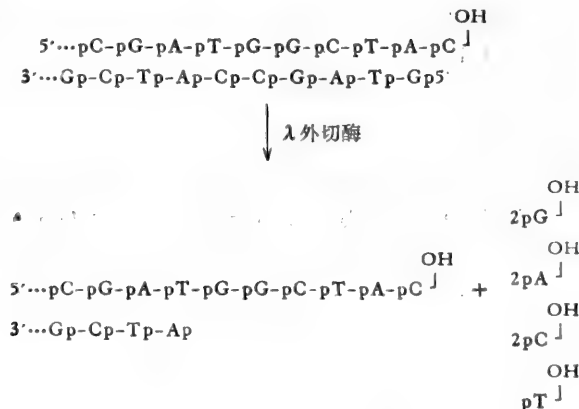


图 3.24 λ 外切酶的作用机制。

外切酶在双脱氧方法测定核苷酸序列中用来处理 DNA。它可以为外切酶 VII 联合使用缩短双链 DNA 的长度(图 3.25)。

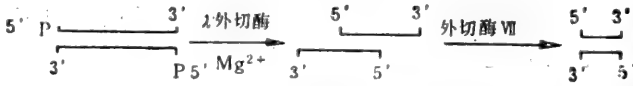


图 3.25 用 λ 外切酶切割 DNA 片段。

λ 外切酶还用来降解突出的 5' 末端。

六、Bal 31 核酸酶

Bal 31 核酸酶是由 *Alteromonas espejiana* Bal. 31 菌的培养液中分离得到的, 在细菌繁殖过程中被分泌到细胞外。它具有单链脱氧核糖核酸内切酶活性, 特异地切割单链 DNA, 产生寡核苷酸和单核苷酸, 这与 S1 核酸酶的作用相类似。同样, 当共价闭环双链 DNA 中出现单链缺口产生所谓缺口 DNA 时, 此酶可以将缺口部位相对的另一条单链部位切开, 使之成为线形 DNA。还有, 此单链核酸内切酶能特异地作用于双链共价闭环 DNA 分子上弱的氢键区域。

Bal 31 核酸酶还具有双链 DNA 外切酶活性。它可以从线性双链 DNA 分子的 3' 和 5' 两端同时切下单核苷酸, 产生向心性逐步降解作用。酶对双链 DNA 的末端结构形式无选择性, 就是说无论平末端还是单链突出末端均可同样地被外切酶所作用。鉴于 Bal 31 核酸酶的上述这些作用, 它可以把双链线性 DNA 分子从 3'、5' 两端逐个地降解, 削短 DNA 的长度(见图 3.27)。不过本酶并不具有专一作用于双链的核酸内切活性。

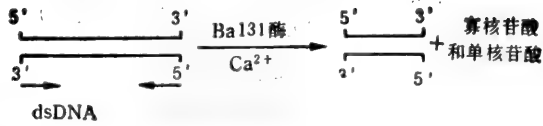
Bal 31 核酸酶的活性需要二价阳离子 Ca^{2+} 和 Mg^{2+} 。缺乏 Mg^{2+} 会降低酶的活性, 缺乏 Ca^{2+} 就会不可逆地丧失酶的全部活性。同时, 反应还需要一定的中性盐 (NaCl), 最适 NaCl 浓度为 0.6mol/L。盐浓度提高到 4mol/L, 酶仍可显示活性。SDS 和尿素不抑制酶的活性。EGTA 有强烈地抑制酶活性的作用, 因为它选择性地整合了反应体系中的 Ca^{2+} , 从而使酶失活。EGTA 不会与 Mg^{2+} 起作用。酶反应常用的缓冲体系为:

20mmol/L	Tris-HCl pH7.2 (有人用 8.0)
600 mmol/L	NaCl
12 mmol/L	CaCl_2
12mmol/L	MgCa_2
1mmol/L	EDTA

反应温度一般选用 30℃。有报道, 当温度下降到 20℃ 时, 酶的活性降低到 30℃ 下活性的三分之一。酶对 DNA 链中 A-T 富集区域的降解作用快于 G-C 密集区, 甚至酶的活性会暂停止在 G-C 密集区。这种表现在降低温度后更为突出。作用机制见图 3.26。

Bal 31 核酸酶是分子克隆实验中一个有用的工具。在进行基因片段的次级克隆和表达时, 需要把特定长度的有功能的基因片段接到启动子后面并具有正确的密码阅读框架, 使之克隆和表达。采用限制性内切酶切割出来的片段, 有时因位点定位不合适, 不能得到上面所述的特定长度和序列的基因片段。可是再用 Bal 31 核酸酶来处理片段, 通过控制

(1) 双链DNA (末端可为齐头或单链突出)



(2) 单链DNA

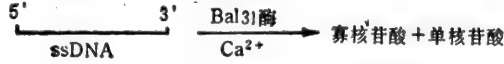


图 3.26 Bal 31 核酸酶的作用示意图。

酶的浓度、反应时间、温度以及中性盐浓度等，可以得到一定长度或一定 DNA 顺序的基因片段，以供克隆所用。图 3.27 示出了用 Bal 31 核酸酶切削 DNA 的过程。

图 3.27 还示出了应用 Bal 31 核酸酶可去除一些不希望存在的限制性内切酶位点，并可以用于内切酶位点的定位。

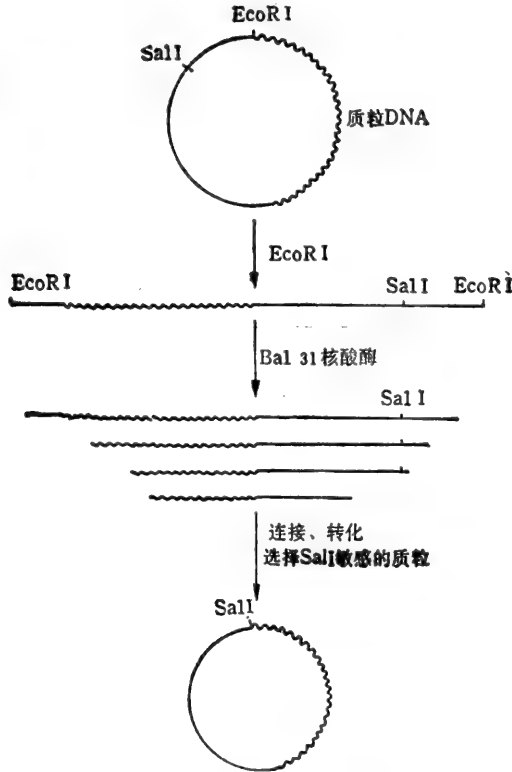


图 3.27 用 Bal 31 核酸酶切削 DNA。

在应用 Bal 31 核酸酶时有几点要注意：①在每一特定实验中，事先一定要细心地用 DNA 制剂检测酶的活性，并用琼脂糖凝胶电泳观察产物被降解的程度(图 3.26)，从而确定实验所需的条件；②在 T4DNA 连接酶连接之前，Bal 31 酶的酶解产物如果只用酚处理使酶失活，常常得不到满意的连接效果。若片段用琼脂糖处理，连接效率比单用酚

处理要高;③当用 Bal 31 酶来定位 DNA 链上的限制性内切酶的位点时,经 Bal 31 酶降解后的 DNA 片段产物只要用 EGTA 灭活 Bal 31 酶作用即可。加 EGTA 之后, Mg^{2+} 不影响浓度,因此又可直接用来进行内切酶的降解反应,不必进行其他处理。

七、核糖核酸酶 H

核糖核酸酶 H (RNase H) 最先是在小牛胸腺组织中发现的。它具有内切酶活性,特异地降解 DNA-RNA 杂合双链上的 RNA 链,产生带 5'磷酸单脂的寡核苷酸和单核苷酸。RNase H 不能降解单链的 RNA 或 DNA、双链 RNA 或 DNA,也不降解 DNA-RNA 杂交双链中的 DNA 部分。已从大肠杆菌、Ustilago、小牛胸腺、大鼠肝脏、鸡胚细胞和 KB 细胞等原核和真核细胞中纯化了 RNase H。酶反应需要 K^+ 和 Mg^{2+} 的激活。目前商品供应的 RNase H 是来自大肠杆菌。在分子克隆实验中,大肠杆菌 RNase H 的主要用途是与大肠杆菌聚合酶 I 和大肠杆菌 DNA 连接酶一起参加 cDNA 中第二条 DNA 互补链的合成。它是 Okayama-Berg cDNA 克隆方法中的关键工具酶之一。当用 AMV 逆转录酶以 mRNA 为模板合成第一条 cDNA 链之后,不用碱处理去除 mRNA,而是用 RNase H 大肠杆菌 DNA 聚合酶 I 和大肠杆菌 DNA 连接酶的联合作用去除 mRNA,以及合成第二条 cDNA 链。RNase H 的部分降解作用使 mRNA 链中产生缺口, RNA 片段就象冈崎片段一样,作为大肠杆菌聚合酶 I 的引物,以 dNTP 为原料合成一系列 DNA 片段,直至 mRNA 全部被 DNA 替代。然后这一系列的 DNA 片段在大肠杆菌连接酶的作用下,把缺口封闭而形成第二条 cDNA 链(见图 3.28)。

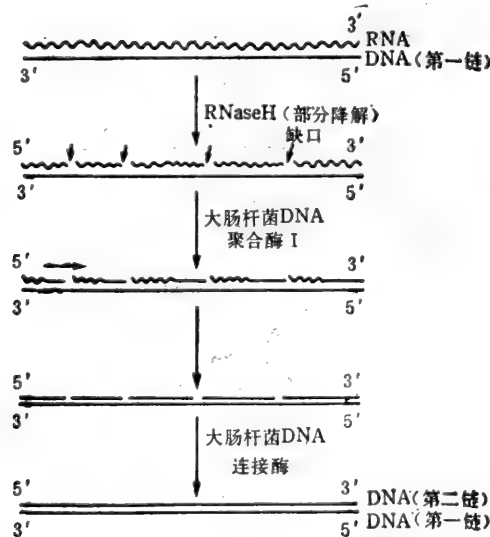


图 3.28 RNase H 参与 cDNA 第二条链的合成。

RNase H 还可以用于除去 mRNA 分子中 3' 端多聚 A 的尾巴。方法是用寡聚 (dT) 或多聚 (dT) 和 mRNA 杂交,然后加 RNase H,它专一地降解与 dT 聚合物杂交部分的多聚 A 链,而把 mRNA 尾部的多聚 A 去除。

第五节 甲基化酶

许多生物体(包括原核生物和真核生物)细胞的 DNA 在体内会被甲基化。这类甲基化作用一般出现在胞嘧啶碱基的第 5 位或者腺嘌呤的 N⁶ 位置。真核细胞中 DNA 的甲基化作用极为复杂,其中有一种现象需要指出的是:来自植物和脊椎动物的 DNA 对 Hpa II (CCGG) 和 Sma I (CCCGGG) 这一类在识别序列中含 CG 结构的限制性内切酶有高度的抗阻作用。

一、原核细胞中的甲基化酶

微生物中存在着限制修饰体系,识别特异序列的甲基化酶起着修饰自身 DNA 作用,使之不被自身的限制性内切酶所降解。一般说来,对于一种特异的限制性内切酶,会有其相同特异性的甲基化酶存在。正如在限制性内切酶一章中提到的, I 型限制性内切酶的酶蛋白同时具有甲基化酶活性;而 II 型限制性内切酶不具有甲基化酶功能,其甲基化作用由专门的酶蛋白——甲基化酶来完成,它们可从微生物中单独被纯化。在体外,当有甲基供体和 S 腺苷酰甲硫氨酸(S-adenosylmethionine, 简称为 SAM)存在时,可将特异识别序列中一定位置上的胞嘧啶甲基化,生成 5-甲基胞嘧啶 (^mC),或者使腺嘌呤甲基化生成 N⁶-甲基腺嘌呤 (^mA)。例如 EcoRI 甲基化酶(以 M·EcoRI 表示)就是与限制性内切酶 EcoRI 具有相同识别序列特异性的甲基化酶,它可以使双链 DNA 中的 GAATTC 序列中的 5' 末端起的第 2 个 A(腺嘌呤)上的 N⁶ 甲基化(图 3.29)。甲基化后的序列 (GA^mATTC) 就能抗阻 EcoRI 的降解作用。象 M·HaeIII 就是识别特异性相同于 Hae III 的甲基化酶,它可使双链 DNA 的 GGCC 序列中的 G-C 的 C 甲基化 (GG^mCC)。这样一类具有特异性的甲基化酶已被用作工具酶应用于基因工程技术中。目前已有商品供应的几种甲基化酶以及它们甲基化位点列于表 3.2 中。

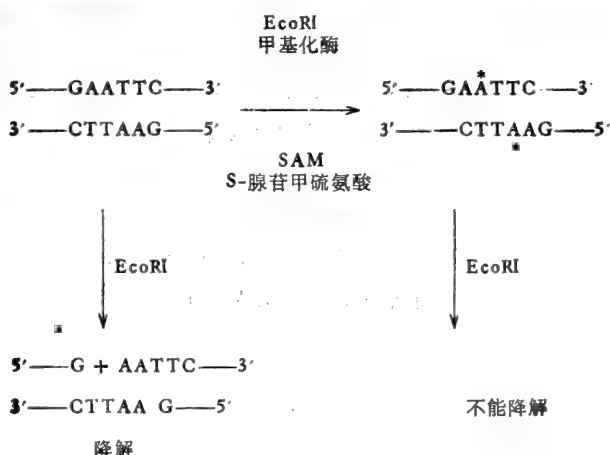


图 3.29 EcoRI 甲基化酶的作用。
* A 表示 N⁶-甲基化腺嘌呤核苷酸。

表 3.2 几种常用的甲基化酶

名 称	识别序列与甲基化位置	名 称	识别序列与甲基化位置
M · AluI	AGCT*	M · HpaII	CCGG*
M · BamHI	GGATCC*	M · HphI	TCACC*
M · ClaI	ATCGAT*	M · MboII	GAAGA*
M · EcoRI	GAATTC*	M · MspI	CCGG*
M · HaeIII	GGCC*	M · PstI	CTGCAG*
M · HhaI	GCGC*	M · TaqI	TCGA*

二、存在于大肠杆菌 K12 株中的甲基化酶

大肠杆菌 K12 是原核细胞分子克隆体系通常采用的菌株,有必要了解它体内的甲基化作用。大肠杆菌 K12 细胞中至少存在三种不同的甲基化酶。它们是 dam 酶、dcm 酶和 hsd 酶。

1. dam 酶

大多数大肠杆菌 K12 株都具有 dam 甲基化酶活性。它使 DNA 链中 5'-GATC-3' 序列中的腺嘌呤甲基化,称为 DNA 腺嘌呤甲基化酶 (DNA adenine methylase, 简称 dam)。dam 和其他细菌中甲基化酶不同,它在大肠杆菌 K12 株中并无相应识别特异性的限制性内切酶存在。有人认为 dam 酶的甲基化作用,可能在碱基错误配对的修复 (mismatch repair) 时,与链的选择有关。由于 dam 的甲基化作用,从 K12 (dam⁺) 菌株中抽提出来的 DNA 将不会被 BclI 酶切。这是因为 BclI 识别序列 (TGATCA) 中 GATC 的 A 被 dam 酶甲基化,产生了 TGATCA* 序列, BclI 的作用受阻遏。当双链 DNA 中有 TCGATC 序列时, TaqI 限制性内切酶的识别序列 TCGA 和 dam 酶的特异序列 GATC 相重叠, dam 的甲基化产生了 TCGATC*, 从而封闭了 TCGATC 序列中 TaqI 的酶切序列。同样,在 ATCGATC 序列中由于 dam 的作用,形成了 ATCGATC*, 使识别 ATCGAT 的 ClaI 酶不能切割第 5 位上甲基化 A 的序列。当某些限制性内切酶与 dam 的识别序列重叠时,由于 dam 的作用而导致该位点上的识别序列抵抗限制性内切酶切割的特性,可在基因工程中用来人为地封闭某种限制酶的切点。dam 酶已经被纯化并作为商品提供。

2. dcm 酶

dcm 是指 DNA 胞嘧啶甲基化酶 (DNA cytosine methylase, 简称 dcm), 它和限制性内切酶 EcoRII 具有相同识别特异性,使 CC(A/T)GG 序列上的第 2 位 C 甲基化,从而阻止了 EcoRII 的降解作用。其他一些内切酶像 BamHI (GGATCC)、BclI (TGGCA)、KpnI (GGTACC)、NarI (GGCGCC) 和 StuI (AGGCCT), 当其后的序列出现与 dcm 特异性的重叠序列时,均可因 dcm 的甲基化作用,使相应的限制酶识别序列表现为抵抗限制酶的切割。例如序列 GGATCCAGG 或 GGATCCTGG, 均含有 BamHI 和 dcm 的识别序列,两者中的 BamHI 位点可因 dcm 的甲基化作用,而抵抗 BamHI 的切割。

3. hsd 酶

hsd 酶是 K12 株中 I 型限制修饰体系中的甲基化酶,它是 hsdM 基因的产物,其识别序列同于限制性内切酶 EcoK (I 型酶):



当 HpaI 的识别序列 (GTTAAC) 和它重叠时,即在 DNA 链中出现 GTTAACNNNNNNGTGC 序列,由于 hsd 酶的作用而封闭了 HpaI 对 GTTAAC 的降解。

在分子克隆中所用的一些宿主细胞是大肠杆菌 K12 的突变株,它们带有 dam⁺ 或 dcm⁺ 基因型,因此,当重组的质粒 DNA 或噬菌体 DNA 转入这类基因型的菌株中进行复制后,从中抽提的 DNA,在其相应的序列上会被一定程度的甲基化,继而影响某些限制性内切酶的活性。现在已有 dam⁻ 或 dcm⁻ 的 K12 突变株,使用这样的突变株作宿主菌,可防止上述甲基化现象的发生。hsd 的影响极少见到,因为用于分子克隆中的 K12 宿主菌大多是 hsd⁻ 的菌株。

三、甲基化酶的用法

表 3.2 列出了已有商品出售的十几种甲基化酶制剂,它们大都是从产生相应限制性内切酶的菌株中分离得到的,只有少数酶制剂来自基因工程的高表达菌株。在试管内进行 DNA 甲基化的反应体系通常为:

50 mmol/L Tris-HCl (pH7.5)

10mmol/L EDTA · Na₂

5mmol/L 2-巯基乙醇

80μmol/L SAM (S 腺苷酰甲硫氨酸)

底物 DNA。

反应过程不需要镁离子。有些甲基化酶需要有中性盐 (NaCl) 的存在,例如 M · MspI 的反应需要 100mmol/L 的 NaCl。

最常用的 EcoRI 甲基化酶 (M · EcoRI) 的反应条件为 100mmol/L Tris-HCl (pH 8.0)、10mmol/L EDTA · Na₂、1.1μmol/L SAM (贮存于 0.01NH₂SO₄ 中)、400μg/ml 的 BSA。底物 DNA 浓度为 100μg/ml, 37℃ 保温 20—30 分钟即可。反应用酚抽提终止,用乙醇沉淀回收被甲基化的 DNA。

甲基化酶的酶单位定义,不同厂家出品的制剂有所不同。例如美国 NEB (New England Biolabs) 公司产品的酶单位定义为: 在体积为 10μl 的特定反应条件下, 37℃ 保温 1 小时,使 1μg λDNA 或 φX174 RFI DNA 得到保护不受相应的限制酶作用所需用的酶量为 1 个酶单位。

一般实验室所用的甲基化酶都来自商品,反应条件和酶单位定义可参照酶制剂的技术说明书提供的资料。

四、甲基化酶的实际应用

(一) 配合连接子的使用, 阻断 DNA 片段中限制酶位点

在分子克隆实验中, 为了重组的需要, 往往在 DNA 片段末端要加连接子, 以提供合适的粘末端。在加连接子和进行限制酶酶切反应前, 首先要把 DNA 片段内部可能存在的相应限制酶的识别序列甲基化, 以防止加上连接子后进行限制性内切酶切割时, 会把 DNA 片段打断。其简单过程示于图 3.29。所以在 DNA 片段加连接子建新末端结构的过程中, 往往都需要甲基化酶的参与。

(二) 构建新的酶切识别序列

DpnI 是唯一能识别 GATC 的限制性内切酶, 其 A 为 N⁶-甲基化腺嘌呤。当双链 DNA 分子中二条链上的 GATC 序列的 A 被甲基化后, DpnI 才可以降解甲基化的 GATC。M·TaqI, 使其识别序列 TCGA 中的 A 甲基化, 产生 TCGA^{*}。如果在 DNA 链存在序列 TCGATCGA, 即二个 TaqI 酶识别序列紧挨一起, 当用 M·TaqI 甲基化, 上述序列就成为 TCGA^{*}TCGA^{*}, 其中正好产生一个 DpnI 酶的识别序列 GATC^{*}。因此, 当采用先 M·TaqI 甲基化再 DpnI 酶切时, 可将 DNA 链中的 TCGATCGA 切开。这就是说, M·TaqI 和 DpnI 结合可构建出新的酶切特异性序列 (TCGATCGA)。McClelland 等人于 1984 年报道了上述试验的成功。他们还提出了 M·ClaI 和 DpnI 的结合可以切开 ATCGATCGAT。即首先在 M·ClaI 的作用下变成 ATCGA^{*}TCGA^{*} (提供了 DpnI 作用所需的序列 GATC^{*}), DpnI 再将上述十核苷酸序列切开。

除了上述介绍的某些甲基化酶与 DpnI 搭配的联合作用可构建八至十核苷酸的高特异性识别序列外, 甲基化酶与其他限制性内切酶配合也可以提供新的酶切识别序列。譬如 BanII 酶识别 G^{*}PuGCPy^{*}C, 这表示图 3.30 中(1)至(4)四种序列均可被切开。不过其中(1), (4)二种为互补序列, 应视为相同识别序列。当先用 M·HaeIII 甲基化(GGCC),

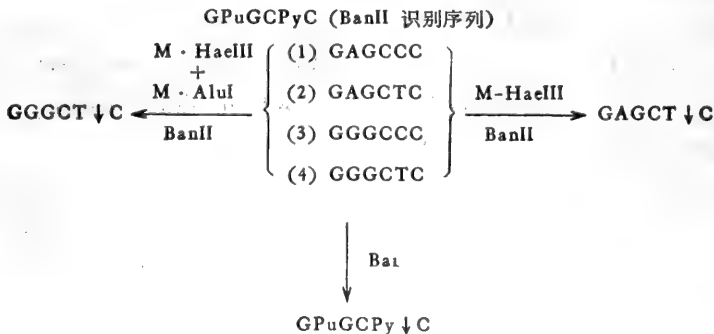
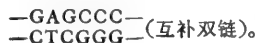


图 3.30 甲基化酶和 BanII 联合作用。

M·HaeIII 使 GGCC→GGC^{*}C。M·AluI 使 AGCT→AGC^{*}T。

BanII 识别序列可分为(1)至(4)四种, 其中(1)和(4)为互补序列, 应视为相同的识别序列:



使序列(3)抵抗 *Ban*II 的作用,再用 *Ban*II 切割,识别序列就为 *GPuGCTC*,若采用 *M*·*Hae*III、*M*·*Alu*I (*AGCT*) 和 *Ban* II 的联合作用,识别序列就只有 *GGGCTC*。这意味着构建了新的识别特异性。表 3.3 介绍了数种甲基化酶和限制性内切酶的联合作用所提供的新的识别位点。

表 3.3 甲基化酶和限制酶的联合作用构建新的识别特异性

限制酶	甲基化酶	识别特异性*
<i>Acc</i> I[GT(AC)(GT)AC]	<i>M</i> · <i>Taq</i> I[TCGmA]	GT(AC)TAC
<i>Aha</i> II[GPuCGPyC]	<i>M</i> · <i>Hha</i> I[GmCGC]	GPuCGTC
<i>Ava</i> I[CPyCGPuG]	<i>M</i> · <i>Taq</i> I[TCGmA]	CPyCGGG
<i>Ava</i> I[CPyCGPuG]	<i>M</i> · <i>Hpa</i> II[CmCGG]	CPyCGAG
<i>Ava</i> I[CPyCGPuG]	<i>M</i> · <i>Hpa</i> II[CmCGG] + <i>M</i> · <i>Taq</i> I[TCGmA]	CCCAG
<i>Ban</i> II[GPuGCPyC]	<i>M</i> · <i>Alu</i> I[AGmCT]	GPuGCCC
<i>Ban</i> II[GPuGCPyC]	<i>M</i> · <i>Hae</i> III[GGmCC]	GPuGCTC
<i>Ban</i> II[GPuGCPyC]	<i>M</i> · <i>Hae</i> III[GGmCC] + <i>M</i> · <i>Alu</i> I[AGmCT]	GGGCTC
<i>Bsp</i> 1286[G(AGT)GC(ACT)C]	<i>M</i> · <i>Alu</i> I[AGmCT]	G(GT)GC(ACT)C
<i>Bsp</i> 1286[G(AGT)GC(ACT)C]	<i>M</i> · <i>Hae</i> III[GGmCC]	G(AGT)GC(AT)C
<i>Bsp</i> 1286[G(AGT)GC(ACT)C]	<i>M</i> · <i>Hae</i> III[GGmCC] + <i>M</i> · <i>Alu</i> I[AGmCT]	G(GT)GC(AT)C
<i>Hgi</i> AI[G(AT)GC(AT)C]	<i>M</i> · <i>Alu</i> I[AGmCT]	G(AT)GCAC
<i>Hinc</i> II[GTPyPuAC]	<i>M</i> · <i>Taq</i> I[TCGmA]	GTPyAAC
<i>Sau</i> 96I[GGNCC]	<i>M</i> · <i>Hae</i> III[GGmCC]	GGACC
<i>Scr</i> FI[CCNGG]	<i>M</i> · <i>Hpa</i> II[CmCGG]	CCAGG

* 用 *M*·*Taq*I 修饰 DNA 之后,再用 *Acc*I 酶切,酶的识别序列由 GT(A/C) (G/T)AC 变为 GT(A/C)TAC。其他酶依次类推。

(三) 封闭 DNA 链中某些识别位点

当 DNA 链的序列中有两种酶的识别序列出现重叠时,用其中某一甲基化酶使相应的序列甲基化,可抵抗重叠序列中被另一限制酶的切割。图 3.31 中显示的一段 DNA 片段,它包含有二个 *Bam*HI 的识别序列,其中一个与 *Msp*I 序列 (CCGG) 相重叠。当片段直接用 *Bam*HI 降解时可产生三个片段,若片段先经 *Msp*I 甲基化酶(*M*·*Msp*I)作用,导致重叠的 *Bam*HI 序列中的一条链被甲基化,再用 *Bam*HI 切割时,甲基化的 *Bam*HI 识别序列就不再被酶切割。就是说由于 *Msp*I 甲基化酶的作用,使 CCGGATCC 序列中的 *Bam*HI 位点消失。*Christ* 和 *Nelson* (1984) 称上述作用为交叉保护作用 (cross-protection)。

表 3.4 中列举了二十多种序列,借甲基化酶作用可导致抵抗限制酶的切割。在封闭某些限制酶酶切位点的实验时,都要先进行甲基化酶反应,然后再用限制性内切酶降解。上面介绍的三方面应用都表明了甲基化酶在分子克隆技术中的实用价值。可以预见甲基化酶的使用在今后会上广泛起来。

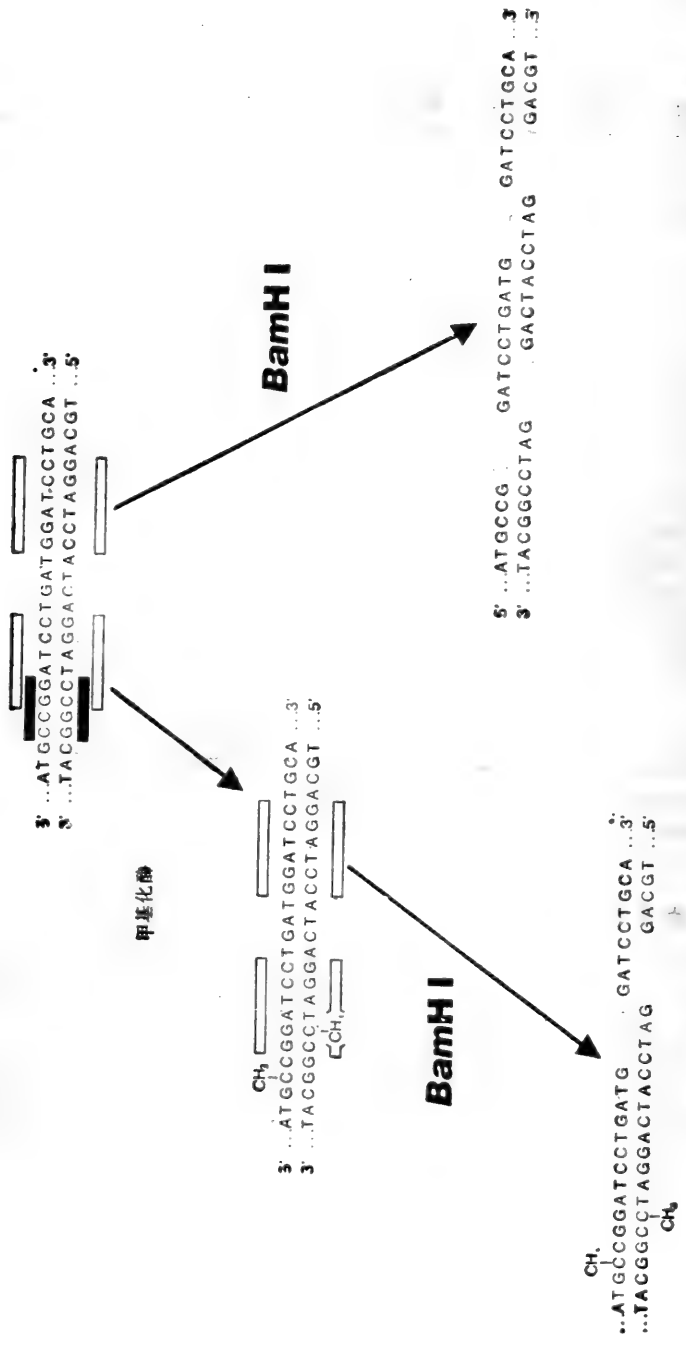


图 3.31 MspI 甲基化酶封闭 BamHI 切点。

表 3.4 甲基化酶修饰后被封闭的限制酶识别序列

限制酶	甲基化酶	被封闭的序列*
AhaII[GPuCGPyC]	M. HpaII[CmCGG]	CCGGCGPyC
AluI[AGCT]	M. PstI[CTGcmAG]	AGCTGCAG
AvaII[GG(AT)CC]	M. HpaII[CmCGG]	CCGG(AT)CC
BamHI[GGATCC]	M. MspI[mCCGG]	CCGGATCC
BglII[GCCN5GGC]	M. HaeIII[GGmCC]	GGCCN5GGC
BstXI[CCAN6TGG]	M. HaeIII[GGmCC]	GGCCAN6TGG
ClaI[ATCGAT]	dam[GmATC]	GATCGAT
DdeI[CTNAG]	M. AluI[AGmCT]	AGCTNAG
EcoRV[GATATC]	M. TaqI[TCGmA]	TCGATATC
FnuDII[CGCG]	M. HhaI[GmCGC]	GCGCG
HinfI[GANTC]	M. HphI[TmCACC]	GANTCACC
HinfI[GANTC]	M. TaqI[TCGmA]	TCGANTC
HphI[GGTGA]	dam[GmATC]	GGTGATC
MboI[GATC]	M. ClaI[ATCGmAT]	ATCGATC
MboI[GATC]	M. TaqI[TCGmA]	TCGATC
MbolI[GAAGA]	dam[GmATC]	GAAGATC
MspI[CCGG]	M. BamHI[GGATmCC]	CCGGATCC
MspI[CCGG]	M. HaeIII[GGmCC]	GGCCGG
NaeI[GCCGGC]	M. HaeIII[GGmCC]	GGCCGGC
NcoI[CCATGG]	M. HaeIII[GGmCC]	GGCCATGG
NheI[GCTAGC]	M. AluI[AGmCT]	AGCTAGC
NruI[TCGCGA]	dam[GmATC]	GATCGCGA
PstI[CTGCAG]	M. AluI[AGmCT]	AGCTGCAG
SacII[CCGCGG]	M. HaeIII[GGmCC]	GGCCGCGG
Sau3AI[GATC]	M. HphI[TmCACC]	GATCACC
TaqI[TCGA]	M. ClaI[ATCGmAT]	ATCGAT
TaqI[TCGA]	dam[GmATC]	GATCGA
XbaI[TCTAGA]	dam[GmATC]	GATCTAGA
XmnI[GAAN4TTC]	M. TaqI[TCGmA]	TCGAAN4TTC

* DNA 经 M. HpaII 修饰之后,虽然 CCGGCGPyC 序列中重叠有 AhaII 识别序列,但不能再被 AhaII 切割。表中所列其他序列均以此类推。

参 考 文 献

- [1] Maniatis, T., Fritsch, E. F. and Sambrook, J., *Molecular Cloning, A Laboratory Manual*, pp 107—148, Cold Spring Harbor Laboratory, 1982.
- [2] Boyer, P. D., *The Enzymes*, vol. XIV, *Nucleic Acids, Part A*, 3rd edition, Academic Press, 1981.
- [3] Boyer, P. D., *The Enzymes*, vol. XV, *Nucleic Acids, Part B*, 3rd edition, Academic Press, 1982.
- [4] Sugino, A. *et al.*, Interaction of bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base-paired ends. *J. Biol. Chem.*, 252, 3987, 1977.
- [5] Higgins, N. P. and Cozzarelli, N. R., DNA-Joining Enzymes: A Review, In *Methods in Enzymology*, vol. 68, Ed. by R. Wu, 50—71, Academic Press, 1979.
- [6] Brennan, C. A. *et al.*, Using T4 RNA Ligase with DNA Substrates, in *Methods in Enzymology*, vol. 100, Ed. by R. Wu *et al.*, 38—52, Academic Press, 1983.
- [7] Weiss, B. A. *et al.*, Enzymatic breakage and joining of DNA. VI. further purification and properties of polynucleotide ligase from *E. coli* infected with Bacteriophage T4., *J. Biol. Chem.*,

243, 4543, 1968.

- [8] Richardson, C. C. *et al.*, Enzymatic synthesis of deoxyribonucleic acid, *J. Biol. Chem.*, **239**, 222, 1964.
- [9] Sanger, F. S. *et al.*, DNA sequencing with chain-terminating inhibitors, *Proc. Natl. Acad. Sci.*, **74**, 5463, 1977.
- [10] Verma, I. M., Reverse Transcriptase, in *The Enzymes*, vol. XIV, Ed. by Boyer, P. D., 87—103, Academic Press, 1981.
- [11] Ratliff, R. L., Terminal Deoxynucleotidyltransferase, in *The Enzymes*, vol. XIV, Ed. by Boyer, P. D., 105—118, Academic Press, 1981.
- [12] Richardson, C. C., Polynucleotide kinase from *E. coli* infected with bacteriophage T4, *Proc. Nucleic Acid Res.*, **2**, 815, 1971.
- [13] Rogers, S. G. and Weiss, B., Exonuclease III of *Escherichia coli* K-12, an AP Endonuclease, in *Methods in Enzymology*, vol. 65, Ed. by Grossman, L. and Moldave K., 201—211, Academic Press, 1980.
- [14] Vogt, V. M., Purification and further properties of single strand specific nuclease from *aspergillus oryzae*, *Eur. J. Biochem.* **33**, 192, 1973.
- [15] Laskowski, M. Sr., Purification and Properties of the Mung Bean Nuclease, in *Methods in Enzymology*, vol. 65, Ed. by Grossman, L., and Moldave, K., 263—276, Academic Press, 1980.
- [16] Lau, P. P. and Gray, H. B. Jr., Extracellular nuclease of *alteromonas espejiana* Bal 31. IV, The single strand specific deoxyribonuclease activity as a probe for regions of altered secondary structure in negatively and positively supercoiled closed circular DNA, *Nucleic Acids Res.*, **6**, 331, 1979.
- [17] Chase, J. W. and Richardson, C. C., Exonuclease VII of *Escherichia coli*, *J. Biol. Chem.*, **249**, 4545, 1964.
- [18] Little, J. W. *et al.*, An exonuclease induced by bacteriophage lambda, *J. Biol. Chem.*, **242**, 672, 1967.
- [19] Marinus, M. G. and Morris, N. R., Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12, *J. Bacteriol.*, **114**, 1143, 1973.
- [20] Roberts, R., Restriction and modification enzymes and their recognition sequences, *Nucleic Acids Res.*, **10**, R117, 1982.

第四章 分子克隆技术(一)

——基因与载体的剪切及重组

蔡良琬 李进 罗刚

第一节 重组技术概述

分子克隆基本技术包括四个方面:基因与载体的剪切;剪切的基因与载体的重组;重组载体的转化;对经过转化生长出来的菌株进行筛选。本章主要讨论前两个方面,后两个方面在下一章介绍。基因与载体的剪切和重组是分子克隆的核心,称为 DNA 的体外重组技术。即 DNA 体外重组技术包括 DNA 分子的剪切和 DNA 片段的连接。通过合理的剪切获得所需要的目的基因片段,再和适当的已剪切的载体片段连接,才能形成一个能够在相应的宿主细胞内自主复制的重组载体。由于基因和载体都是 DNA,故又称 DNA 重组技术。

DNA 体外重组技术能够建立起来,是因为发现了一系列不同的限制性内切酶和 DNA 连接酶。限制性内切酶主要来自微生物体内,可识别 DNA 序列中某些特定的短迴文序列(一般为四至六个核苷酸),并在识别部位催化 DNA 链的水解,形成带有粘末端(3'突出或 5'突出)或平末端的 DNA 片段。在分子克隆中,应用同一个限制性内切酶剪切基因或载体分子,可获得相同粘末端的 DNA 片段,连接方便。例如最早用于分子克隆的限制性内切酶有 Hind III、EcoR I。现在已有许多种限制性内切酶可供选用。DNA 连接酶可催化带有同一互补粘末端或任一种平末端的 DNA 片段之间的连接。各种基因与载体 DNA 体外重组技术的关键是如何选用合理的剪切方法,以及选用适当的剪切酶,剪切后再通过适当的条件把它们连接起来,获得理想的重组载体分子。重组技术主要包括剪切出特定的基因片段、载体的选择和剪切以及基因片段与载体的重组。

一、剪切获取特定的基因片段

这是指未分子克隆的基因而言。要从一个分子量很大的 DNA 分子中(如包括生物的各种染色体 DNA)分离得到某一特定的基因片段,首先需要用限制性内切酶对该 DNA 分子进行降解,得到较小的 DNA 片段。在含量较少的情况下要通过电泳等方法分离,观察究竟有多少片段,特定的基因是否有易回收的片段。象病毒的 DNA 就可用此法,小病毒甚至可采用完整基因组的全部克隆。细菌及哺乳动物细胞的染色体很大,而且目的基因又多,则常采用基因文库的方法建立(见本书第六及第七章)。在获得含有特定基因的 DNA 片段后,即可直接分离回收,用于分子克隆。再次分子克隆时,基因的剪切则应根据剪切的特点进行,如做基因表达则应按表达基因的需要进行剪切。

二、载体的选择

载体在分子克隆中是不可缺少的，目的基因片段只有与载体片段共价连接形成重组体后，才能进入合适的宿主细胞内进行复制和扩增。使用最多的载体是质粒(见本书第二章)。目前在分子克隆中常用的质粒载体如 pBR322、pAT153 等。它们都具有以下特点，在宿主细胞内能自主复制，可容纳较大量的外源 DNA 片段，有多个可供选择插入外源 DNA 片段的单一限制性内切酶位点，并带有合适的遗传学标记便于筛选等。此外，特别大的外源基因应选择 λ 噬菌体衍生 DNA 类的载体。

三、载体的剪切

为了顺利地达到分子克隆的目的，在第一次克隆(初级克隆)时，载体除了需要具备上述条件外，最好能用剪切基因时所用的限制性内切酶剪切，并使达到重组后插入灭活等容易识别的条件。对载体进行剪切的目的，是使其能够同特定的目的基因片段连接形成重组体，故在进行剪切时要尽量使产生的载体片段末端能与目的基因片段的末端互补。为了筛选方便插入灭活等要注意利用位于特殊基因标记内的限制性内切酶位点。

经过初级分子克隆后，再次剪切克隆以达到其他目的时，应按再次分子克隆的目的来剪切载体。

四、基因与载体的重组

带有互补粘末端的基因与载体片段之间，以及带有平末端的 DNA 片段之间，都可以用 T4 DNA 连接酶催化进行共价连接，形成重组载体分子。一般说来，连接时目的基因片段的量与载体片段的量之间的比例较为重要。在某些情况下，用载体末端脱磷酸的方法防止载体分子的自身连接也很重要，但这不是所有分子克隆中都要采用的。

有关载体与各种剪切、连接中所用的酶前面都有介绍，本章重点叙述应用分子克隆时所采用的各种剪切条件，连接条件及应注意的事项。

第二节 基因剪切常用方法

一、基因剪切的特点

在剪切基因时，由于不同的基因或 DNA 具有不同的组成和结构，含有限制性内切酶切点不同，以及分子克隆的目的不同，形成了不同的剪切特点。

所谓剪切特点主要是指一种基因或 DNA 曾用某些限制性内切酶剪切过并确定了酶切图谱(物理图谱)。这种基因或 DNA 经过剪切后很容易在电泳分离的区带上分辨出所需要的各片段部位。

酶谱分析中应用的剪切方法见上册第三部分的物理图谱分析。分子克隆中的剪切方法大致可分为以下几种：

1. 应用限制性内切酶水解基因与载体,使其形成粘末端

2. 有目的地形成带有平末端的基因

(1) 应用限制性内切酶剪切产生平末端。

(2) 从粘性末端改造成平末端。包括补齐法和削平法。

(3) 应用外切酶 III 与核酸酶 S1 剪切出长短不同的平末端片段。

(4) 应用核酸酶 Bal 31 处理基因以造成长短不同的平末端。

3. 从平末端改造成粘末端

(1) 加连接子或接头。

(2) 加尾法——在基因与载体剪切后加互补的核苷酸尾,形成粘末端关系。有关以上剪切特点将分别叙述如下。

二、应用限制性内切酶水解基因与载体形成粘末端

由限制性内切酶水解产生的粘末端有两种,即 5' 突出末端和 3' 突出末端。两者中以 5' 突出者最为实用。特别是在初级分子克隆中,从一个未知酶切图谱的 DNA 中,剪切成片段进行克隆时,往往选用能产生 5' 突出末端的限制性内切酶,因为这种末端最易连接,不致因连接不好而丢失应分子克隆的任一片段。这类酶中最常用的为识别 6 个核苷酸的酶,如 EcoRI、BamHI 等。有关这些酶的基本知识已在本书上册中详细介绍。为了达到容易分子克隆的目的,近年来人们应用的初级克隆载体上,往往都注明有关单一酶切点的部位,一些具有多酶连接子的载体,也都有一些适用的酶切位点,其中大部分为形成 5' 突出末端所用的酶。至于形成 3' 突出末端的酶如 PstI 等也可利用,特别是当没有形成 5' 突出末端的酶在剪切时可被利用,常用 PstI 酶来剪切。PstI 酶还有一个优点,即在载体 pBR322 质粒上,它恰好位于 Ap^r 基因的一个适合位点,在此酶切点处连接的外源基因不仅能达到分子克隆的目的,还常有表达的可能。

(一) 应用的范围

1. 在初级分子克隆中的应用

初级分子克隆是指生物体中一个基因或基因组的第一次克隆。首次剪切重组,一般尽量获得完全的基因组。象病毒类生物, DNA 一般都只有 10^3 — 10^4 水平的核苷酸对,这样长度的 DNA 有些可以切割成几段分别进行分子克隆,有些甚至不切割成片段,而采用全基因组的分子克隆。如果是环形 DNA,只要找到一个合适的单一酶切点即可。如果是线形的 DNA 则需要把 DNA 的两个末端处理,使它们也具有可连接用的末端。至于比较大的细菌类基因组一般有一 10^6 水平的核苷酸对,要直接用限制性内切酶剪切后做简单的分子克隆比较困难,因为剪切出来的片段太多,至少会有上千个不同片段。为了保证得到的全部基因组的 DNA 都进行分子克隆,必须制作基因文库。基因文库是简单分子克隆的发展,也同样需要剪切、重组、转化和筛选四个步骤。如何保证基因文库中有全部基因,这是个复杂的问题,将在第六章专题讨论。

病毒 DNA 比较容易纯化,也易得到一定数量的纯 DNA。至于 RNA 病毒,则还需

逆转录成为 cDNA, 有关内容将在第七章讨论。

初级分子克隆是在未知基因的 DNA 序列及酶切图谱的情况下进行的, 是否能在剪切中获得全部 DNA 序列, 而不至把小片段 DNA 丢失, 应先作电泳分析。如果只想把剪切的某一片段分子克隆, 则可在进行电泳后回收个别片段。回收的方法曾在上册 DNA 电泳一章中叙述。其中以透析袋回收法较好, 这样回收的片段在剪切时不受混杂物的影响。

2. 在次级克隆中的应用

次级分子克隆是指已经初级分子克隆的基因片段, 经过剪切等处理, 再次作新的分子克隆。分子克隆的操作主要可达到两个目的, 一是基因的纯化, 二是基因的扩增。在初级分子克隆时, 所用的基因或基因组有时是不够纯的, 但经分子克隆后, 利用了载体的不兼容性和细菌的无性繁殖, 任何一种重组载体都得到纯化。只要再从重组载体上将基因剪切下来, 则都是纯化后的目的基因。如果大量培养这些带有重组载体的菌株, 还可利用菌体的扩增和重组载体的扩增达到目的基因量的扩增。所以经过初级克隆后, 再次克隆时都是有其他目的的克隆, 如果原克隆的基因片段过大, 则可以剪切去除一些不需要的部分, 如利用这些基因做基因表达、基因结构分析等, 也必须对这些基因进行再次剪切。

次级克隆的剪切与初级克隆不同, 从 DNA 量上看, 比初级克隆前要多。初级克隆后, 往往已进行了所重组基因的酶谱分析, 甚至核苷酸序列的分析, 所以对可能利用的酶以及切点位置已清楚, 故酶解比较容易, 特别是当有适合的限制性内切酶可以应用时, 最为方便。其他剪切方法, 将在后面叙述。

3. 限制性内切酶应用的范围与根据

已知的限制性内切酶有上百种, 在分子克隆中究竟如何选择利用? 最简单的办法是先选择载体, 看所用的载体上有哪些最适用的单一酶切点。如质粒 pBR322 上的 BamHI 切点, 在此切点处剪切后, 很容易插入目的基因 (详细介绍见本书第二章), 这样就在可能的条件下运用此酶来水解欲做分子克隆的目的基因。应注意不要选择有多个酶切点的酶, 否则在剪切基因时会出现过多的片段, 在处理载体时也常不适用。不少载体经过改造, 在分子上带有多个单一酶切点的连接子, 这就更方便选择。在剪切基因时剪切出的片段最好不超过 2—3kb, 因为过大的分子会造成分子克隆困难。然而, 10kb 以下的 DNA 一次初级克隆成功也完全是可能的, 特别是一些环形质粒经过单一酶切点剪切形成线形, 与载体重组, 这样的例子是很多的。只要遵循这一原则, 即选择单一酶切点, 不论初级克隆还是次级克隆都容易成功。

对于分子较大的基因组, 经剪切后出现几个片段, 可同时进行分子克隆, 用筛选法找出全部片段, 分别分析, 重新组织绘出基因组的酶切图谱。

(二) 剪切时应注意的几个问题

分子克隆时, 剪切基因所用的酶解条件与一般酶解条件稍有不同。这些不同是因为酶解后还有分离、连接和鉴定等步骤。

1. DNA 的浓度

通过酶解产生的 DNA 片段通常都较小; 剪切是否完全或正确, 有时还需要进行电泳鉴定, 因此需增加 DNA 的用量。一般是增加 DNA 的浓度, 以免反应体积过大。如

果必须经过鉴定,甚至还需从电泳后的区带处回收一定的片段,常用的 DNA 量一般为 10—50 μg 。大量的 DNA 是为了在回收 DNA 片段后还可用于修饰加工, DNA 溶液的浓度可根据 DNA 的溶解度确定,一般在 1 $\mu\text{g}/2\mu\text{l}$ 至 1 $\mu\text{g}/4\mu\text{l}$ 。

2. 酶解反应体积

限制性内切酶是保存在 50% 的甘油中,为了减少甘油对酶活性的影响,通常需用单位体积中含酶单位较多的酶液,以减少加入的甘油量。如需增加酶液用量时,为了不致因加入甘油过多而影响酶解反应,可增加反应体积,有时可增加至 100—200 μl 。

3. 反应时间

由于用酶量常多于所需量,就要缩短反应时间,以防止因保温反应时间过长,而发生酶液中混有的杂酶对 DNA 的影响,通常保温时间不要超过三小时。

4. 鉴定的必要性

在分子克隆前可以先鉴定剪切是否能够成功。也可在剪切的同时取出少量酶解液进行鉴定,以保证剪切正确与完全。如果需要在剪切后分离出某一片段进行分子克隆,则以电泳鉴定与回收同时进行为好。但因为回收率往往达不到 100%,故 DNA 浓度与反应液量都应增加,以期获得足够的回收量。因为单位体积的凝胶中,总以 DNA 量愈大愈能回收完全,有关电泳中 DNA 的回收方法,已在上册中介绍,可根据不同实验室条件选择应用。

(三) 基因剪切实例

基因或载体经限制性内切酶剪切后,有些可以直接按一定比例混合,然后进行连接反应。有些则必须经过电泳回收基因的某一片段或载体的某一部分。有时基因和载体剪切后都必须经过电泳回收,才能进行连接反应。以下分别讨论。

1. 剪切后直接进行连接

在以下几种情况中,基因和载体经限制性内切酶剪切后可以不经电泳回收,直接进行连接反应。

(1) 当基因和载体用同一种限制性内切酶剪切,或用不同的限制性内切酶剪切,产生的末端相同时,并且基因在载体上的插入位置正好是载体上某一抗药性基因所处的位置,这时可通过基因的插入使抗药性基因失活对重组子进行筛选,而不必回收某一欲克隆的片段。

(2) 在进行初级分子克隆时,欲克隆基因的量通常很小,只有几十纳克到几百纳克水平,剪切后一般不经过回收而直接进行连接。

下面举一实例。由乙型肝炎病毒 (HBV) DNA 携带者血清中提取的含 HBV 基因的 DNA,其量非常少,全部 DNA 只有 40ng。HBV-DNA 是一个环形、双链并带有缺口的 DNA,约 3.2kb。已知在 HBV-DNA 上有 BamHI 的单一酶切点,在进行 HBV 的初级分子克隆时,选择的载体是质粒 pAT153。它是载体 pBR322 的衍生物,也是一环形 DNA,约 3.6kb,并有 Ap 和 Tc 抗性标记,在 Tc 抗性 (Tc^r) 部位有 BamHI 单一酶切点。因此将 HBV-DNA 和 pAT153 分别用 BamHI 剪切成线形,直接进行连接,挑选 Ap 抗性 Tc 敏感的重组子,即可得到有 HBV-DNA 插入的重组质粒 pHBV-NC1。由于 HBV-DNA 量少,为了使连接反应中总 DNA 量达到要求, pAT153 用量为 0.4 μg ,

而不能按一般连接反应中基因与载体的比例。反应物用量如下:

pAT153-DNA (线性)	0.4 μ g
HBV-DNA (线性)	40ng
连接酶缓冲液 $\times 10$	2 μ l
T4 DNA 连接酶	1 μ l
加 H ₂ O 至	20 μ l

反应液在 16 $^{\circ}$ C 保温 3 小时后在 4 $^{\circ}$ C 继续反应过夜。其中连接酶缓冲液的成分见本章第三节。整个剪切,连接过程见图 4.1。

2. 剪切后的回收

基因或载体经过限制性内切酶第一次剪切后,有些情况下需要进行回收以得到单纯的某片段,一般是回收剪切后的基因片段。

(1) 基因和载体剪切后得到不同的末端,可通过电泳回收基因或载体中的某一片段,再对其末端进行改造。

(2) 基因经限制性内切酶剪切后,产生若干片段,其末端与载体的末端可以互相结合。但欲克隆的基因片段在插入载体后没有明显的筛选标记,在这种情况下,通常要经过电泳分离,回收其中欲克隆的片段。

图 4.2 是一个基因经剪切、回收再进行连接的实例。载体质粒 pCQV2 (约 5.0kb) 上只有一个 Ap 抗性基因,对于外源基因的插入没有明显标记。欲克隆的基因是乙肝病毒克隆 pHBV-NC1 上的一部分,即约 400bp 的核心抗原 (cAg) 基因。pHBV-NC1 质粒用 BglII 剪切后产生两个片段: 6.4kb 和 0.4kb。cAg 基因位于 0.4kb 的片段上,用电泳回收此片段后再进行连接。pCQV2 用 BamHI 剪切成线形分子后,其末端和 BglII 产生的末端可以互补。所以将回收的 0.4kb 片段 (BglII 末端) 与剪切后的 pCQV2 (BamHI 末端) 进行连接,得到的重组子一般可以不经筛选,而认为是纯片段克隆。

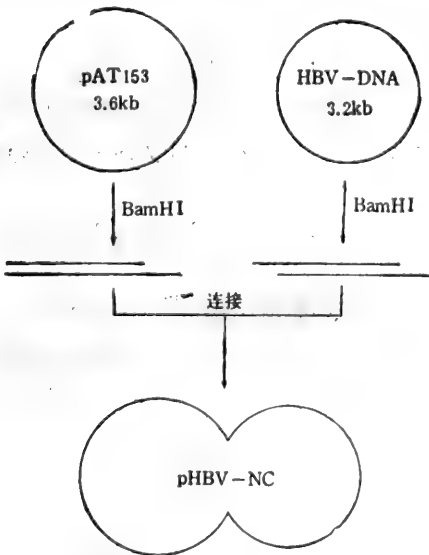


图 4.1 乙肝病毒 DNA 的分子克隆。

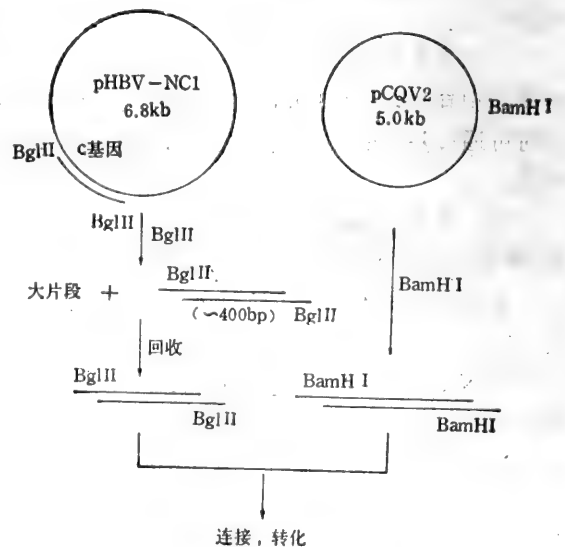


图 4.2 cAg 基因的分子克隆。

3. 剪切重组改造载体

在本书第二章介绍过质粒载体 pBR322，它是一个人工改造的质粒，其复制起点部分是属于 ColE1 多拷贝质粒。ColE1 在大肠杆菌中每个细胞约有 20 个拷贝。从 ColE1 衍生出来的质粒虽然都属于松弛型，但在大肠杆菌中的拷贝数却并不完全相同。这是因为控制 ColE1 质粒复制需要一种属于质粒的特异蛋白，在缺失这一基因的非主要区时，可以提高该质粒的拷贝数。这种质粒仍然很稳定，如恢复原有结构仍可恢复原有质粒的性质，pAT153 就属于这一类造成缺失的质粒。

pAT153 质粒只有 3.6kb 大小，它是由 pBR322 上除去了一个 HaeII-B 片段（约 0.62kb）重新复环而成。pBR322 上有 11 个 HaeII 的酶切点，如图 4.3 所示。按片段从大到小分别命名为 A、B、C 等片段，其中 B 片段是在 1727—2349 核苷酸之间，位于 pBR322 复制起点的右侧。将 pBR322 改造成 pAT153 的做法是应用 HaeII 部分酶解的方法，即在很短的时间内酶解 pBR322、得到许多不同的剪切片段，把这些剪切片段再

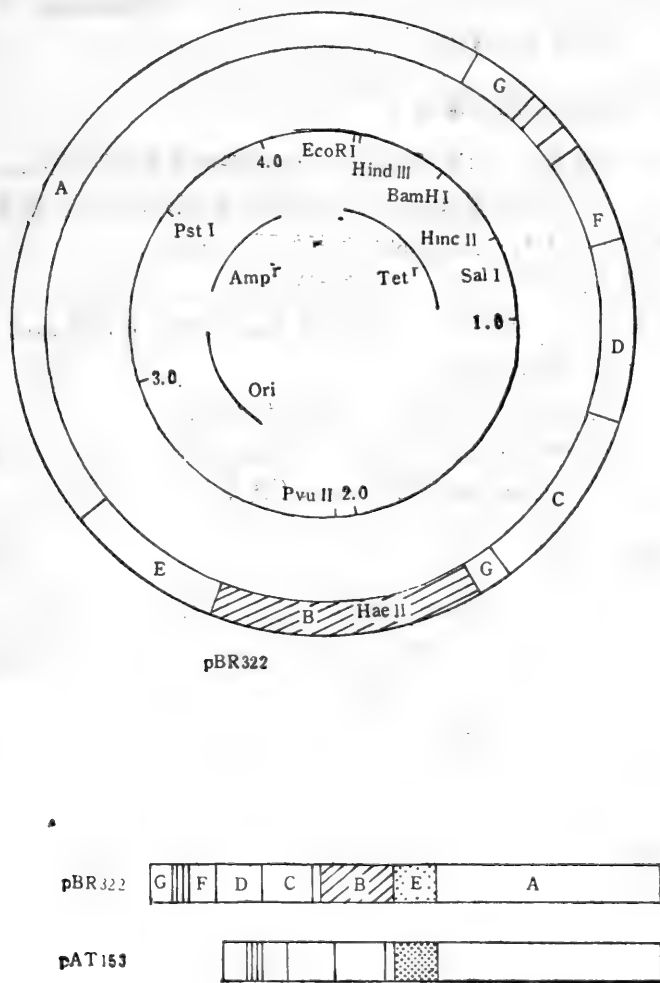


图 4.3 pBR322 与 pAT153 的 HaeII 切点示意图。

重新自身连接起来,用 Ap 及 Tc 抗性筛选,就能得到 pAT153。

pBR322 上的 11 个 HaeII 酶切点,剪切后得到的 11 种片段大部分是在 Tc 抗性基因上(有 7 个),其余的 1 个含有全部 Ap 抗性基因,1 个带有复制起点,只有在 1644—1727 和 1727—2349 之间的两个片段不影响 Ap 和 Tc 抗性,及质粒的复制。因此利用重新复环的质粒(带有复制起点)及 Ap 和 Tc 抗性筛选标记,可以获得 pAT153 质粒。在 pAT153 质粒上由于缺少 1727—2349 之间的 0.62kb 这一段,实验证明提高了原有 pBR322 的拷贝数,pAT153 拷贝数约为 pBR322 的 1.5—3 倍。它仍可用氯霉素进行质粒的扩增。pAT153 质粒除了提高拷贝数外,还因缺失了 nic 基因,比 pBR322 更稳定,不会因结合转移反应而发生丢失。实验证明它确实很稳定,拷贝数也高,还可增加 Ap 与 Tc 的抗性。这也是剪切后直接连接的一个例子。

三、形成带有平末端的基因

在初级分子克隆中很少用平末端剪切方法,但在次级克隆或组建一个特殊载体及表达质粒时,剪切成平末端则十分有用。

(一) 应用限制性内切酶剪切的平末端

有不少限制性内切酶如 HincII、HaeIII、SmaI 等可用于剪切平末端。平末端是比较难于连接的,与 5' 突出的粘末端相比,效率常不到其 1/50,故很少被直接用于初级分子克隆。只有特殊需要时才利用,应用的范围有如下几方面。

- (1) 剪切 5' 或 3' 突出末端时无适当的酶可被利用。
- (2) 一个质粒有两个以上的同一种酶的酶切点,可取消部分切点,再复环。
- (3) 为进一步修饰处理准备的剪切。
- (4) 为一定目的而形成的—粘末端与—平末端的片段。

平末端的剪切条件与一般应用其他限制性内切酶时一样。为了分子克隆等剪切的需要,剪切时条件与上面所述的粘末端剪切相一致,不再举例说明。

(二) 从粘末端改造成平末端

此种剪切修饰是次级克隆或基因表达研究中常用的方法,它适用于以下几种情况:

- (1) 目的基因与载体的剪切不能形成适用的粘末端,而二者又需要连接时。
- (2) 为减少一个酶切点或改造一个酶切点。
- (3) 在基因表达中为达到适合的框架连接。

有两种方法可改造粘末端成为平末端,即补齐法与削平法。分述如下。

1. 补齐法

应用的酶为 DNA 聚合酶 I(Klenow 大片段),此酶的性质在本册第三章中已介绍过。方法较简单,容易完成,现讨论如下。

- (1) 反应试剂。修复缓冲液(用于限制性内切酶和 DNA 聚合酶 I Klenow 大片段)

50mmol/L Tris-HCl (pH7.6)
50mmol/L KCl

10mmol/L MgCl₂

10mmol/L 二硫苏糖醇 (DTT)

四种脱氧核苷三磷酸 (dNTP)

0.2mol/L dATP

0.2mol/L dTTP

0.2mol/L dCTP

0.2mol/L dGTP

用 TE 缓冲液饱和过的酚

DNA 聚合酶 I(Klenow 大片段)

(2) 反应步骤。进行补齐反应时首先要选择脱氧核苷三磷酸。根据限制性内切酶产生的末端,其单链的突出部分的顺序,选择和这些顺序互补的脱氧核苷三磷酸。例如 EcoRI 产生的粘末端的突出部分是 5'AATT-, 在进行反应时只要加入 dATP 和 dTTP 即可;而 BamHI 产生的突出部分顺序是 5'GATC-, 则在反应中要加入四种脱氧核苷三磷酸。反应的具体步骤如下:

a. 以修复缓冲液作为限制性内切酶酶解缓冲液,将 DNA 用限制性内切酶剪切,得到 5' 突出或 3' 突出的 DNA 片段。补齐反应的最适体积是 50 μ l, 限制性内切酶的反应体积以保持在 30 μ l 为宜。

b. 酶解后,将限制性内切酶在 65 $^{\circ}$ C 水浴失活,再进行 DNA 聚合酶 I (Klenow 大片段)的补齐反应。若酶解反应体积为 30 μ l, 则在反应液中补加:

10 倍浓度修复缓冲液 2 μ l

0.2mol/L 脱氧核苷三磷酸各 4—5 μ l

加 H₂O 至 50 μ l

DNA 聚合酶 (Klenow 大片段) 0.1—1 单位

c. 反应在室温(15—25 $^{\circ}$ C)进行 15—30 分钟。

d. 补齐反应后在 65 $^{\circ}$ C 水浴中,将反应液保温 10 分钟,并用等体积饱和酚抽提反应液使 DNA 聚合酶 I(Klenow 大片段)失活。

e. 反应液经过失活后可直接用于连接反应,或用乙醇沉淀 DNA 后备用。

(3) 反应后的鉴定。

a. 若在补齐反应后还要进行其他剪切反应,可以在补齐反应完毕,取出一部分进行平末端连接,转化。如果补齐反应完全,可以得到转化子。另外还可以通过限制性内切酶酶解反应来鉴定补齐反应是否完全。方法是在连接、转化后,将提取的质粒再用原限制性内切酶剪切,如果补齐反应完全,则质粒不再被切开。此方法常用于为消除酶切点而进行补齐反应的鉴定(图 4.4)。

图 4.4 中最后一步,即在用 EcoRI 剪切时,质粒 pBR322 不被切开,说明已消除了 EcoRI 的切点。

b. 为了鉴定,可以在进行补齐反应的同时作一个对照实验,反应条件与补齐实验类似,只是用一个 ³²P 标记的核苷三磷酸取代反应中的一种核苷三磷酸。在四种 α -³²P-dNTP 中可以选择要进行补齐反应的粘末端最靠外的互补核苷酸。如 BamHI 酶解产生的粘末端顺序是 5'GATC-, 最靠外的一个是 G。对照实验可用 α -³²P-dCTP (与 G 互补)

标记末端,然后进行电泳分析及放射自显影,即可了解补齐反应是否完全。

(4) 补齐反应的应用。

a. 将粘末端改造成平末端。

b. 消除酶切点。

c. 在物理图谱分析中应用于识别剪切片段的位置。

d. 补齐反应容易进行,比削平法好,如果两种方法皆可选用,往往采用补齐法,以保证末端的完整。

e. 用于末端标记。但应注意,如果有两个以上核苷酸可被标记时,不能用于顺序分析,不然在电泳后会出现重复区带,不能阅读顺序。

2. 削平法

削平一个粘末端,常用核酸酶 S1 处理。此酶在第三章中虽已介绍过,为了合理使用,此处将再作详细说明。

核酸酶 S1 可以选择性地消化双链 DNA 分子中突出的单链末端部分,如果条件控制适当,不会产生明显的不平末端。其作用方法被描述为轻咬或轻啃。如果反应不适当,除完成上述反应外,还会在双链 DNA 中引入其他缺口。因此正确运用核酸酶 S1 十分重要。一般购买的商品核酸酶 S1 每毫克酶蛋白都含有 10^4-10^5 酶单位,也有以溶液状态保存出售的。在应用时必须稀释,但稀释后不宜再保存。

核酸酶 S1 还有以下几种用途,如分析 DNA-RNA 杂合体的结构,打开合成双链 cDNA (ds-cDNA) 时产生的发夹结构。

(1) 核酸酶 S1 酶解反应试剂。

酶储存缓冲液

20mmol/L Tris-HCl(pH7.5)

50mmol/L NaCl

0.1mmol/L ZnCl₂

50% 甘油

或为

10mmol/L 醋酸钠 (pH5.0)

0.1mmol/L ZnCl₂

1mmol/L 半胱氨酸 (cystein)

0.001% Triton X-100

50% 甘油

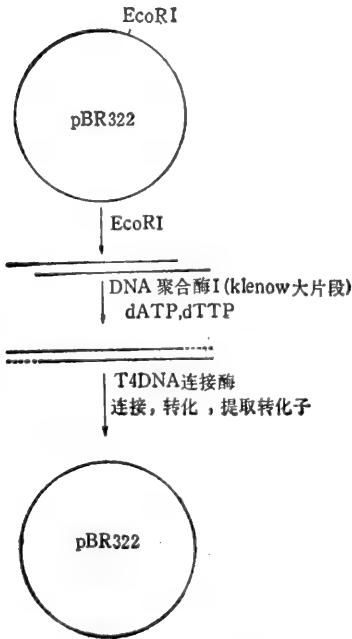


图 4.4 消除质粒 pBR322 上的 EcoRI 切点。

酶解缓冲液

300mmol/L 醋酸钠(pH5.0)

100mmol/L NaCl

2mmol/L ZnCl₂

10% 甘油

酶反应终止液

0.5mol/L EDTA (pH8.0)

2mol/L Tris-HCl (pH8.0)

(2) 核酸酶 S1 反应步骤。

a. 用 10 μ l TE 缓冲液或消毒蒸馏水溶解 1 μ g DNA 样品, 并与 150 μ l 2 倍浓度的核酸酶 S1 反应缓冲液混合, 然后加 H₂O 至 300 μ l, 再加入 0.5—1 u 的核酸酶 S1。

b. 将反应混合液在 15 $^{\circ}$ C 水浴中反应15分钟(不要超过 15 分钟)。

c. 立即在反应液中加入反应终止液 EDTA 和 Tris-HCl (分别为 10 μ l 和 35 μ l), 振荡以终止反应。

d. 反应液用酚抽提两次, 吸取水相。

e. 在上清水相中加入蒸馏水饱和的乙醚, 振荡后吸去上层的乙醚, 重复 2—3 次, 将残留的酚去除干净。

f. 用乙醇沉淀 DNA, 即可得到平末端的 DNA 片段或载体。

(3) 注意事项。

a. 核酸酶 S1 浓度一般较大, 1 μ l 酶液浓度可达上万单位, 应用时必须稀释。如果用核酸酶 S1 反应缓冲液稀释, 则反应必须及时进行。如果将稀释的核酸酶 S1 保存在冰箱, 可以用酶储存缓冲液将酶稀释成 10—20 u/ μ l, 在 -20 $^{\circ}$ C 保存。

b. 酶反应缓冲液中 NaCl 的浓度较大, 通常将溶液配成 2 倍浓度备用。

c. 核酸酶 S1 在 37 $^{\circ}$ C 时活性不易丧失, 故反应时要严格控制时间, 注意不要超过 15 分钟。反应完毕后必须立即加入终止液, 即改变溶液的 pH 和离子强度, 以停止反应。

d. 核酸酶 S1 和外切酶 III(exonuclease III) 结合使用, 可随机缩短 DNA 双链, 所得双链有时并不完全是平末端的。故在核酸酶 S1 反应终止后, 最好再用 DNA 聚合酶 I(Klenow 大片段)将末端补齐, 以保证缩短的 DNA 双链是齐头平端的。

(4) 核酸酶 S1 反应后产物的鉴定。鉴定核酸酶 S1 消化反应是否完全的方法很多, 这里介绍一种较简便的方法。

在用核酸酶 S1 消化 DNA 片段中单链的同时, 做一个空白对照实验, 在对照中不加核酸酶 S1。反应完毕后将两份样品都进行末端的同位素标记(方法见本书上册)。末端标记反应用 DNA 聚合酶 I(Klenow 大片段)催化。此方法是用核酸酶 S1 消化的样品, 因为已将 DNA 片段的单链部分消去, 而对照样品中没有加核酸酶 S1, 单链部分没有消去。在用 DNA 聚合酶 I(Klenow 大片段)进行末端标记时, 前者没有单链 DNA 为模板, 故不能加上任一核苷酸, 后者有单链部分可作为模板, 故可以加上标记的核苷酸。通过电泳及放射自显影, 就可知道核酸酶 S1 消化是否完全。

(三) 应用外切酶 III 与核酸酶 S1 剪切形成长短不同的平末端

当一个基因需要剪切成一个比较短的片段, 或把载体缩小时, 如果在基因或载体上无合适的限制性内切酶可用于剪切时, 可以用外切酶 III (exonuclease III, Exo III) 与核酸酶 S1 共同作用。剪切后获得的片段, 长度并非完全一致, 而是在一定范围内长短不同的片段。这种剪切是先用外切酶 III 剪切线形的基因或载体片段, 降解其中一条链达

到所需的长度后停止反应，再用核酸酶 S1 处理另一条链，至两条链达到一致的长度为止。

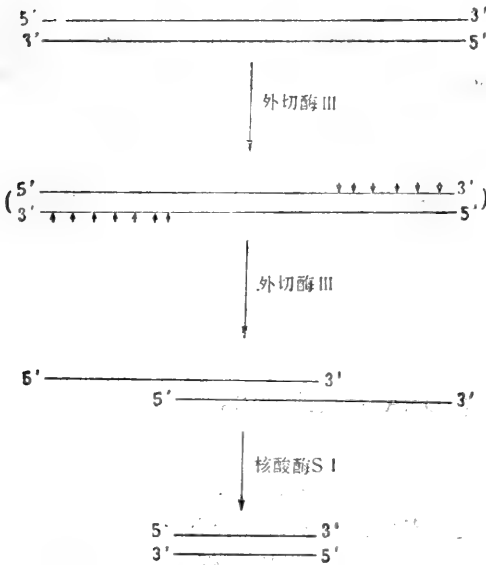


图 4.5 应用外切酶 III 和核酸酶 S1 的剪切反应。

5 × 核酸酶 S1 缓冲液

500mmol/L NaAc(pH4.0)

500mmol/L NaCl

60mmol/L ZnCl₂

10 × 修复缓冲液

500mmol/L Tris-HCl(pH7.6)

500mmol/L KCl

100mmol/L MgCl₂

100mmol/L DTT

反应终止液

100mmol/L EDTA

1.5mol/L NaAc

DNA 聚合酶 I(Klenow 大片段)

核酸酶 S1

1mol/L Tris-HCl(pH8.0) 饱和的酚

2. 降解反应

(1) DNA 以修复缓冲液为限制性内切酶酶解缓冲液，先剪切成线性 DNA 片段。如果剪切后的 DNA 片段为 3' 突出的粘末端，则应先按本节 (二) 中叙述的补齐反应将 DNA 补齐成为平末端的 DNA 片段，因为外切酶 III 对 3' 突出的 DNA 降解不完全。补齐反应后应在 68℃ 水浴中失活。

其过程如图 4.5 所示。

大肠杆菌外切酶 III 降解双链线性 DNA 时，只要反应条件适当，可以按同步作用速度消化掉一定数量的核苷酸。这是由于大肠杆菌外切酶 III 从 3' 末端消化时没有碱基的特异性。1982 年郭礼和等曾报道过使用外切酶 III 的条件，即在室温 23℃ 下反应时，外切酶 III 每分钟可在 DNA 链的每个 3' 末端去掉 10 个左右的核苷酸。现将其方法介绍如下。

1. 试剂与酶

10 × 外切酶 III 缓冲液

660mmol/L Tris-HCl(pH8.0)

770mmol/L NaCl

50mmol/L MgCl₂

100mmol/L DTT

(2) 线性平末端或 5' 突出末端的 DNA 片段反应液 10 μ l (含 1pmol DNA), 进行以下反应:

酶解反应液	10 μ l
补加: 10 \times 外切酶 III 缓冲液	4 μ l
外切酶 III	1—4u
加 H ₂ O	26 μ l

反应时间及所加的外切酶 III 量见表 4.1。通常反应时间应比计算的时间稍长几分钟。

(3) 上述反应在 37 $^{\circ}$ C 进行, 反应到时间立即加入 10 μ l 反应终止液, 以停止反应。

(4) 反应液在 65 $^{\circ}$ C 水浴中失活 10 分钟后, 用等体积饱和酚抽提反应液, 吸取水相后, 用乙醇沉淀 DNA 备用。

3. 外切酶 III 消化 DNA 的速率

外切酶 III 的用量及消化反应时间的选择可参考表 4.1。

表 4.1 外切酶 III 消化 DNA 的速率

从 DNA 每一端消化的核苷酸数	保温时间(分)	外切酶 III 量 (u/pmole DNA)
100—250	10—25	15—20
250—500	25—50	20—25
500—750	50—75	25—30
750—1,000	75—100	30—35
1,000—1,500	100—150	35—45

由上表可看出, 外切酶 III 每分钟从 DNA 的一端 (3 端) 约可消去 10 个核苷酸, 在双链 DNA 的两个 3' 端的消化是同步的。另外在做外切酶 III 实验前应注意先做预实验, 因为每一批酶的活性总是有差别的。

4. 外切酶 III 处理后的核酸酶 S1 消化反应

用限制性内切酶剪切产生的 DNA 线形片段, 经外切酶 III 消化反应后, 再用核酸酶 S1 消化 5' 端的单链部分, 形成平末端的缩短的 DNA 片段。具体方法与削平反应类似, 只是直接在外切酶 III 反应液中进行, 并补加 4.5 μ l 10 倍浓度的核酸酶 S1 缓冲液, 加入核酸酶 S1 的量是每 mg DNA 用 20 单位。图 4.6 是外切酶 III 和核酸酶 S1 消化反应的实例, 电泳图中所用的 DNA 是先用 BamHI 酶解后的线形片段 (约 2.7kb)。

(四) 应用核酸酶 Bal 31 处理基因造成长短不同的平末端

1. 核酸酶 Bal 31 介绍

核酸酶 Bal 31 具有以下两种主要功能: ①对双链 DNA 具有高度特异的双链 DNA 外切酶的活性。它可以同时对双链 DNA 分子的两条单链作用, 从 DNA 链的 5' 端和 3' 端开始降解, 产生单核苷酸和很小的寡核苷酸片段; 或者将一个较长的 DNA 双链缩短成一个较短的 DNA 双链分子。核酸酶 Bal 31 对两条链的酶解速度大致相等。另外核酸酶 Bal 31 还具有对单链 DNA 的特异内切酶的作用, 即可将超螺旋的质粒 DNA 变成线形 DNA。②对单链 DNA 具有特异的外切酶活性, 其作用方式类似于核酸酶

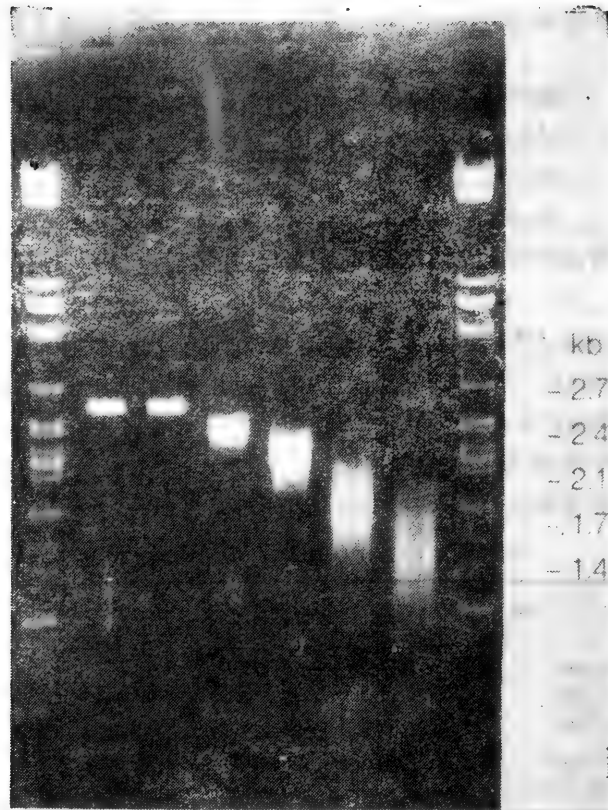


图 4.6 外切酶 III 和核酸酶 S1 反应电泳图。

图中 1.8: DNA 分子量标准 (λ -HindIII/EcoRI 和 λ -Xho I 混合物)。2. 对照(无外切酶 III 和核酸酶 S1)。3. 对照(无核酸酶 S1)。4,5,6,7. 分别为 1 μ g 片段用外切酶 III 各消化 20,40,60,80 分钟(23 $^{\circ}$ C)后,再用核酸酶 S1 消化 15 分钟。

S1, 即属于轻咬作用。

如果长时间用核酸酶 Bal 31 作用,可将双链 DNA 完全降解成为单核苷酸。

2. 核酸酶 Bal 31 的酶解反应体系及步骤

酶解缓冲液 I

600mmol/L	NaCl
12mmol/L	CaCl ₂
12mmol/L	MgCl ₂
20mmol/L	Tris-HCl(pH8.0)
1mmol/L	EDTA

酶解缓冲液 II

300mmol/L	NaCl
6mmol/L	CaCl ₂
6mmol/L	MgCl ₂

10mmol/L Tris-HCl(pH8.0)
1mmol/L EDTA

通常将反应体积控制在 30 μ l。将 1 μ g 线形 DNA 与 6 μ l 5 倍浓度的酶解缓冲液 II 混合,加蒸馏水至 30 μ l, 并加入 0.2 u 的核酸酶 Bal 31, 反应在 16 $^{\circ}$ C 水浴中进行。在此条件下, 酶解速度为每分钟 40bp, 即线形双链 DNA 分子的每一端的降解速度为每分钟约 20bp。用酚抽提终止反应。

3. 酶解实例及讨论

核酸酶 Bal 31 在保存过程中酶活性容易降低, 故在每次剪切实验前应先做预实验, 以确定酶反应的合适条件及反应时间。以下是核酸酶 Bal 31 预实验的一个实例。

反应条件

线形 DNA 片段	6.8kb
DNA 用量	0.5 μ g
核酸酶 Bal 31	0.2u
反应体积	30 μ l
反应温度	16 $^{\circ}$ C
反应时间	10—50 分钟
酶解缓冲液 II(5X)	6 μ l

共做六个预实验, 反应条件基本一样, 只是反应时间不同。每个反应到时间后立即加入二甲苯蓝和溴酚蓝染料 5 μ l, 以停止反应。然后一起上样于 1% 的琼脂糖凝胶电泳。在进行正式实验时, 因为 DNA 用核酸酶 Bal 31 酶解后还要进行其他剪切反应, 所以要用酚抽提来终止反应。以上例子的电泳图见图 4.7。

图中核酸酶 Bal 31 的酶解速度较慢, 这是由于酶不是新鲜制品, 活力有所下降, DNA 的一端降解不能达到每分钟 20bp, 而仅约 10bp。

根据上面的实验及其他一些实验, 已知核酸酶 Bal 31 酶解速度受以下因素影响。

(1) 缓冲体系。核酸酶 Bal 31 的酶活性需要钙离子 (Ca^{2+})。 Ca^{2+} 浓度愈大, 在其他条件相同的情况下, 酶解速度愈快。但是 Ca^{2+} 浓度大时, 酶解速度不容易控制。在酶解缓冲液 I 中, 核酸酶 Bal 31 每分钟可使线形双链 DNA 缩短 80bp, 即每一端分别缩短 40bp。在酶解缓冲液 II 中, 核酸酶 Bal 31 每分钟可缩短 40bp, 即每一端分别缩短 20bp。核酸酶 Bal 31 在两种缓冲液中的活性不同, 是因为缓冲液中 Ca^{2+} 浓度不同。当需要将 DNA 双链缩短较多时, 如缩短 2,000bp 以上, 可以选用酶解缓冲液 I。其他情况下一般以选择酶解缓冲液 II 较好。

(2) 反应温度。当核酸酶 Bal 31 酶解温度为 30 $^{\circ}$ C 时, 酶解速度比 16 $^{\circ}$ C 时提高 10

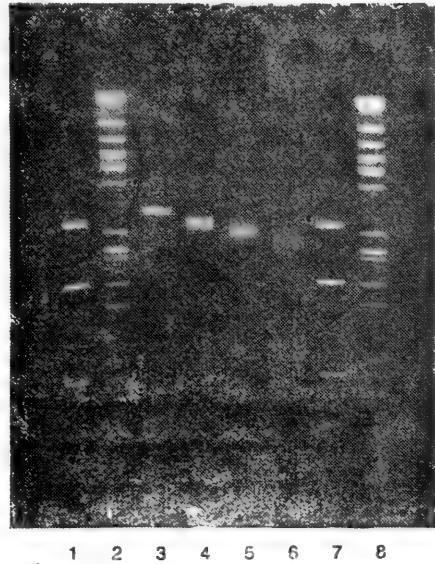


图 4.7 确定 Bal 31 核酸酶活性的电泳图谱。

1,7. 用 HaeIII 酶切 0.5 μ g 的 fdRF109DNA。2,8. 1 μ g DNA 分子量标准。

用 EcoRI 酶解的 2 μ g 的 pUR222 DNA 同 Bal 31 核酸酶保温: 3为 0 分钟。4为 5 分钟。5为 10 分钟。6为 15 分钟。1% 的琼脂糖凝胶。

倍,为了有效控制酶解反应速度,一般保持反应温度在 16℃ 左右。

(3) 反应时间。反应温度为 16℃ 时,在酶解缓冲液 II 中,核酸酶 Bal 31 的酶解速度大约每分钟 DNA 链的每一端消化掉 20bp。一般是根据预实验结果及所要消去的 DNA 片段长度控制酶解反应时间。

(4) DNA 的组成。核酸酶 Bal 31 对含 GC 较多的 DNA 的降解速度要比含 AT 较多的 DNA 快。如果 DNA 链中 GC 含量局部较多或很少时,会发生不等速现象。

(5) DNA 浓度。当 DNA 浓度较大时,不容易控制反应速度。若要进行较大量 DNA 的实验,可以将 DNA 分成几份,保持 30μl 的反应体积,其中含 DNA 0.5—1μg。在用酚终止反应时,再合并反应时间相同的 DNA 反应液。这样就能得到长短大约在一个范围的 DNA。

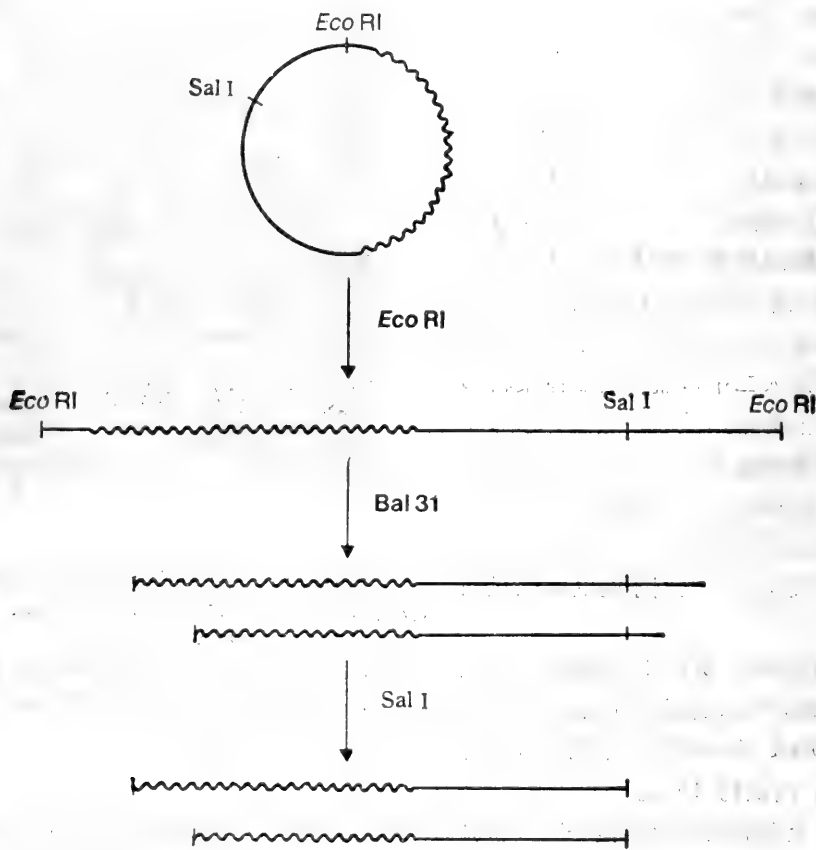


图 4.8 核酸酶 Bal 31 缩短链的降解作用。

温度对核酸酶 Bal 31 的反应速度影响较大。应用核酸酶 Bal 31 的优点在于能够通过控制反应时间来控制酶对 DNA 的降解速度。在较温和的条件下,即在较低的反应温度、较低的 Ca^{2+} 浓度(酶解缓冲液 II)下进行酶解反应,能够较好地控制酶解反应速度,因为这时核酸酶 Bal 31 的作用较慢,而且也比较均匀。如果酶反应速度太快,在较短的时间内酶对 DNA 的降解也较快,就不易控制酶解反应产生的 DNA 片段的大小,而且产生的片段大小差别也比较大。所以在进行核酸酶 Bal 31 的反应时,常选择比较

温和的反应条件,即在 16℃ 反应,用酶解缓冲液 II。

4. 核酸酶 Bal 31 的应用 (1:48:25) 酶解缓冲液 II

(1) 可以将一个较长的 DNA 分子降解缩短成为一个较短的 DNA 分子。降解产生的 DNA 分子的大小可通过控制反应时间来控制。

(2) 缩短 DNA 分子可得到较短的 DNA 分子,有时还可把所需要的基因一端引入片段的一端,以备作适当的连接。

(3) 对缩短的分子还可以增加另一个限制性内切酶的剪切,这样就得到两个不同的末端,即一个平末端与另一个粘末端的片段,如图 4.8 所示。图中质粒先用 EcoRI 切成线形片段,经过核酸酶 Bal 31 的消化,将 DNA 链缩短,再用 Sal I 剪切,这样既可以缩短 DNA 链,又可以形成一个一端平头一端粘头的 DNA 片段。

四、从平末端改造成粘末端

(一) 用加连接子或接头改成粘末端

1. 连接子介绍

连接子是一段含有一个或多个酶切点的序列已知的寡核苷酸。根据连接子中所含酶切点的多少及位置,可将连接子分为两种:连接子(linker)和接头(adapter sequences)。前者一般只含有一个酶切点,或有两个以上相关酶的切点,在酶切点的 5' 及 3' 端往往还有几个用于连接的不属于酶切点的核苷酸。接头一般都含有两个或两个以上不同的酶切位点,其 5' 或 3' 端不但无多余的核苷酸,而且酶切点应识别的顺序也往往呈剪切后状态。连接子为单链自身互补顺序,可复性为平端双链,接头即有单链的,又有双链的。

连接子有磷酸化和非磷酸化之分。接头一般都不必磷酸化。非磷酸化连接子在前都要用 T4 多核苷酸激酶(T4 polynucleotide kinase)处理,成为末端磷酸化的连接子。非磷酸化和磷酸化的连接子表示如下:

磷酸化的: d(pCGGATCCG)

非磷酸化的: d(CGGATCCG)

连接子的 5' 端只有带上磷酸才能和要连接的 DNA 片段的 3' 端羟基形成磷酸酯键。下面介绍连接子的磷酸化反应及常用的连接子。

试剂 2: 连接子的磷酸化

(1) 试剂。

a. T4 多核苷酸激酶

b. T4 多核苷酸激酶缓冲液($\times 10$)

700mmol/L Tris-HCl(pH7.6)

100mmol/L MgCl₂

50mmol/L DTT

c. 10mmol/L ATP

d. 500mmol/L EDTA

e. TE 缓冲液

10mmol/L Tris-HCl(pH7.8)

1mmol/L EDTA

f. 酚-氯仿-异戊醇混合液(25:24:1)

(2) 磷酸化反应步骤。

a. 混合以下反应物,在 37℃ 水浴保温 30 分钟。

非磷酸化连接子	1—5pmol
10 × 激酶缓冲液	5μl
10mmol/L ATP	5μl
T4 多核苷酸激酶	10u
加 H ₂ O 至	50μl

b. 在反应物中加入 2μl 500mmol/L EDTA, 并用等体积的酚-氯仿-异戊醇混合液抽提一次。

c. 用 95% 乙醇沉淀反应物,用 TE 缓冲液溶解后,即得到磷酸化的连接子。

在以上反应中,由于铵离子能抑制多核苷酸激酶的活性,所以全部反应试剂应避免接触含铵离子的溶液。另外反应中所用的 ATP 要新鲜配制,防止降解成 ADP。

3. 常用连接子

有关连接子与接头在本书上册限制性内切酶一章中已叙述,此处仅列出目前常用的连接子和接头,见表 4.2。

4. 连接子浓度的表示

连接子或接头是一段寡核苷酸。它的浓度常用 OD 值表示。1 OD 的连接子或接头约为 40—50μg, 即相当于 40—50μg 的连接子或接头溶于 1ml 水中,置于 1cm 宽的比色杯中,在 260nm 波长处的吸收值 $A = 1.0$ 。由于连接子或接头都很短,一般只由几个至几十个核苷酸组成,所以在和 DNA 片段或载体 DNA 连接时,用量很少,通常只用几十纳克。

5. 加连接子所形成的粘末端

这不是一种简单的剪切反应,首先是连接,然后再剪切成粘末端,即是两步反应。连接子和 DNA 片段的连接属于平末端连接,但它与 DNA 片段的平末端连接又不完全类似。因为连接子很小,连接时用量容易达到要求,所以连接效率比一般平末端连接反应高,容易完成。使用此法时应注意以下几点。

(1) DNA 片段的预处理。由限制性内切酶剪切产生的 DNA 片段如果是平末端的,可以直接和连接子连接。例如由 HincII、PvuII 产生的 DNA 片段,都是平末端的。如果 DNA 片段经限制性内切酶处理后,是粘末端的,则必须先将此粘末端处理成平末端,才能和连接子连接。具有 5' 或 3' 突出粘末端的 DNA 片段可以用两种方法处理成平末端,即削平法,用核酸酶 S1 削去突出的单链部分;或补齐法,用 DNA 聚合酶 I(Klenow 大片段)将缩进的部分补齐。这两种方法在本节前面都已详细介绍过,此处不再重复。

(2) 连接中的末端比例。与 DNA 片段的大小相比较,连接子则很小,只有几对至十几对核苷酸。当加入的连接子和 DNA 片段的量一样时,连接子的 pmol 数要比 DNA 片段的 pmol 数大很多,即连接子 DNA 分子的末端数比 DNA 片段的末端数大很多。在进行连接反应时,通常要保持连接子的末端数比 DNA 片段的末端数多一些,一般以

表 4.2 常见的连接子和接头

非磷酸化连接子	BamHI	d(CGGATCCG)
	BamHI	d(CGGGATCCCG)
	BamHI	d(CCGGATCCGCG)
	BclI	d(CTGATCAG)
	BglII	d(CAGATCTG)
	BglII(MbolI-MbolI)	d(GAAGATCTTC)
	BglII	d(GGAAGATCTTCC)
	ClaI	d(CATCGATG)
	ClaI	d(GATCGATC)
	EcoRI	d(GGAATTCC)
	EcoRI	d(CGGAATTCCG)
	EcoRI	d(CCGGAATTCCGG)
	HindIII	d(CAAGCTTG)
	HindIII	d(CCAAGCTTGG)
HindIII	d(CCCAAGCTTGGG)	
磷酸化连接子	BamHI	d(pCGGATCCG)
	BclI	d(pCTGATCAG)
	BglII	d(pCAGATCTG)
	ClaI	d(pCATCGATG)
	EcoRI	d(pGGAATTCC)
	HindIII	d(pCAAGCTTG)
	HindIII	d(pCCAAGCTTGG)
	KpnI	d(pGGGTACCC)
	NcoI	d(pCCCATGGG)
	NdeI	d(pCCATATGG)
	Sall	d(pGGTCCACC)
	SmaI	d(pCCCCGGGG)
	SphI	d(pGGCATGCC)
	XbaI	d(pCTCTAGAG)
	XhoI	d(pCCTCGAGG)
XmaIII	d(pCCGGCCGG)	
接头	BamHI-SmaI	d(GATCCCCGGG)
	EcoRI-SmaI	d(AATTCCTCCGGG)
	HindIII-SmaI	d(AGCTCCTCCGGG)
	Sall-SmaI	d(TCGACCTCCGGG)
	SmaI-hexamer	d(CCCGGG)

表中连接子和接头均为单链,复性后呈双链。

10:1 为宜,即保持连接反应体系中 10 个连接子 DNA 分子对 1 个 DNA 片段分子,这样能保证 DNA 片段和连接子充分连接。

可用以下公式计算连接子和 DNA 片段的 DNA 用量比(重量比):

$$\frac{\text{连接子用量}}{\text{DNA 片段用量}} = \frac{\text{连接子长度}}{\text{DNA 片段长度}} \times 10$$

确定了连接子和 DNA 片段的用量比,根据 DNA 连接反应要求的 DNA 总浓度(一般为 0.01—0.5mg/ml)就可以分别确定连接子的用量和 DNA 片段的用量。

例如,有一长为 800bp 的 DNA 片段需要加一个有 8 对核苷酸的连接子,如果连接

反应体积定为 20 μ l, 根据上面的公式:

$$\frac{\text{连接子用量}}{\text{DNA 片段用量}} = \frac{\text{连接子长度}}{\text{DNA 片段长度}} \times 10 = \frac{8}{800} \times 10 = 1:10$$

若连接反应中总 DNA 浓度为 0.01—0.5mg/ml (0.2—10 μ g/20 μ l) 时, 连接子用量应为 0.018—0.91 μ g, DNA 片段和连接子的连接条件与 DNA 的平末端连接条件一样。但是因为连接子很小, 一个 DNA 分子很容易接上多个连接子分子。不过在连接反应后, 总要用能切开连接子上酶切点的限制性内切酶切出粘末端, 这样很自然即可将多余的连接子去除, 整个程序有三部:

- a. 连接子磷酸化。
- b. 平末端的 DNA 片段与磷酸化的连接子连接, 形成带连接子的平末端片段。
- c. 利用相应的限制性内切酶酶解形成粘末端, 即可用于下一步的连接和反应。

6. 连接子的应用

(1) 连接子在分子克隆中的应用。在分子克隆中通过加连接子可使 DNA 片段和载体具有相同的粘末端。这样一方面使连接反应容易进行; 另外, 有时欲分子克隆的 DNA 片段上没有适合的带有粘末端的酶切点, 而选择的载体上有, 这时即可在 DNA 片段的平末端上先加上与载体上欲选用的限制性内切酶连接子。下面以一个平末端的 HaeIII 片段(即含 HaeIII 切点) 克隆到质粒 pBR322 的 PstI 切点处为例说明连接子的这一应用(图 4.9)。

(2) 连接子在基因表达中的应用。有不少连接子以几种不同长度的形式可供选择, 如果分别与 DNA 片段连接, 则可形成几种不同的阅读框架。一个密码子有三个核苷酸, 在 DNA 片段末端往往并非正好是一个密码子, 而是多一个或两个核苷酸。所以可通过加连接子来改变或保持某一种基因片段的阅读框架。以 BamHI 连接子为例。BamHI 连接子有三种, 即 8 个、10 个和 12 个核苷酸对, 它们都可以与一个能被表达的基因片段连接, 形成 BamHI 切点的粘末端。加上 8 个核苷酸时, 可补上两个核苷酸(即 6 + 2); 加上 10 个核苷酸时, 可补上一个核苷酸(即 9 + 1); 当加上 12 个核苷酸时则常用在完全适用于框架本来合适的片段连接。

(3) 连接子的加入可在载体中引进酶切点或消除酶切点。通过加连接子可以在载体上引入新的酶切点。如果一个载体上有两个以上同种限制性内切酶的切点, 可以通过部分降解再加连接子的方法消除某个酶切点, 并引进某个新的在连接子上的酶切点。

(二) 加尾法——在基因与载体剪切后加互补的核苷酸尾, 形成粘末端关系

在末端脱氧核苷转移酶(简称末端转移酶)催化下, 通过载体 DNA 和基因片段如 cDNA 末端分别加上互补的核苷酸尾, 可以使 DNA 具有互补的粘末端关系, 便于连接, 其应用与连接子类似。

1. 末端转移酶反应体系

(1) 末端加接多聚嘧啶 (dT 和 dc) 的条件。

a. 反应缓冲液

100mmol/L (pH7.5) 二甲砷酸钠 (sodium cacodylate)

1.0mmol/L CoCl₂

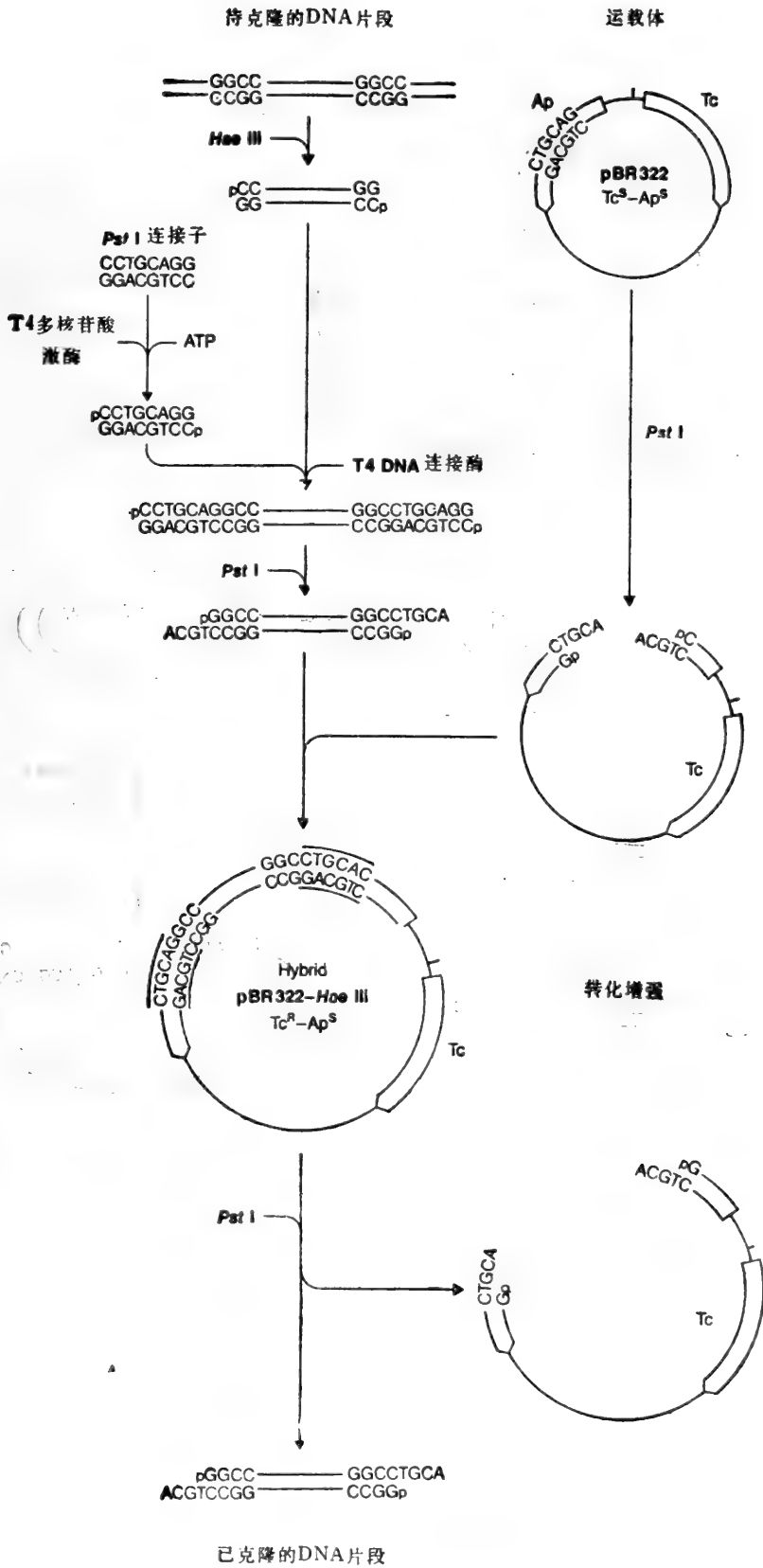


图 4.9 连接子在分子克隆中的应用。

0.1mmol/L DTT
 100 μ mol/L dTTP 或 dCTP

- b. 末端转移酶浓度: 400u/ml
 c. DNA 片段浓度: 1—5 μ g/ml
 (2) 加接多聚 dA 的条件。

a. 反应缓冲液

20mmol/L(pH7.5) 磷酸钾
 4mmol/L MgCl₂
 0.1mmol/L DTT
 100 μ mol/L dA

b. 末端转移酶浓度: 400u/ml

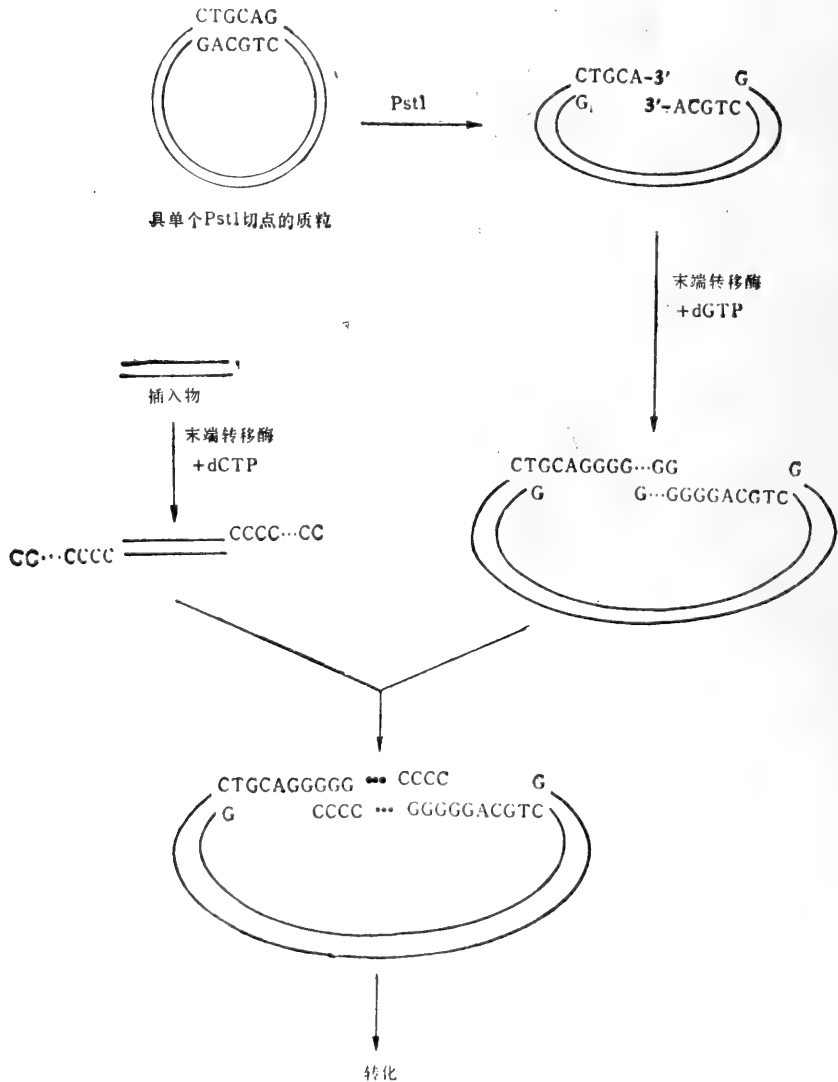


图 4.10 PstI 切点的恢复。

(3) 加接多聚 dG 的条件。

a. 反应缓冲液

140mmol/L

二甲磷酸钠

30mmol/L

Tris-HCl(pH6.9)

1mmol/L

CoCl₂

0.1mmol/L

DTT

b. 末端转移酶浓度: 1000 单位/ml

c. DNA 片段浓度: 1—5μg/ml

d. dGTP 的量为 DNA 片段的 100—150 倍

2. 反应步骤

(1) 在加入末端转移酶之前, 应先将反应体系加热到 37℃, 末端转移酶一经加入反应立即开始。

(2) 加接多聚 dG 时, 在上述的反应体系中, 在 37℃ 反应 30 分钟, 可在 DNA 末端加上 10 个左右核苷酸长度的多聚 dG。

3. 通过加接多聚 dG 和多聚 dC 恢复一些限制性内切酶的识别位点

在 PstI 剪切形成的末端 3' 突出部分加接多聚 dG 后, 再和带有多聚 dC 的 DNA 片段连接, 可以恢复 PstI 切点(如图 4.10)。

4. 多聚核苷酸与 DNA 的比例

对于具有 3' 突出末端的 DNA 片段, 在加接多聚核苷酸 (dNTP) 时最适比例为: dNTP:DNA = 20:1。对于具有 5' 突出或平末端的 DNA 片段, 最适比例为: dNTP:DNA = 100:1。

(三) 人工合成的寡核苷酸片段

近年来应用人工合成的方法已能很迅速地得到各种基因片段, 象连接子一类的基因片段寡核苷酸已有商品出售。其他人工合成的寡核苷酸有以下几类: 各种调节基因片段如各种启动子, 通用翻译终止位点, 核糖体结合点即 Shine-Dalgarno 片段, 携带着翻译起始位点, 转录终止位点等等。如图 4.11。

这类合成的寡核苷酸片段的特点, 除了有该基因的特殊顺序外, 不少片段还带有适当的能剪切成粘末端的酶切位点, 如图 4.11 (6) 中的 HindIII 及 BamHI。因此使用这类寡核苷酸片段比连接子更容易连接, 有的可以按连接子的应用方法连接(但不剪切), 有的直接按粘末端连接。这类顺序多应用于基因表达的研究中, 有关理论可见本册基因表达部分的内容。

第三节 剪切后基因与载体的连接

目前常用的 DNA 连接酶有两种, 一种来源于大肠杆菌, 一种由 T4 噬菌体诱导产生。其中以 T4 DNA 连接酶应用较多, 它不仅能催化带有互补粘末端的 DNA 片段之间的相互连接, 还能催化带有平末端的 DNA 片段之间的连接。这两种作用在 DNA 体外重组中十分有用。

一、带有粘末端的 DNA 片段之间的连接

大多数限制性内切酶在降解 DNA 分子后都产生粘末端。如果基因和载体用同一种限制性内切酶处理,或虽然用不同的限制性内切酶处理,都产生可以互补的粘末端,如 XhoI 和 S_{al}I 产生的粘末端都是 TCGA。其互补末端顺序经过退火,在 DNA 连接酶的催化下,也很容易共价连接,形成重组分子。重组分子因具备原载体的特性,经过转化到受体菌中,就可以达到分子克隆的目的。

粘末端的连接不仅可在同一酶切的片段中进行,也可在不同限制性内切酶剪切的片段中进行,只要这些限制性内切酶剪切后的片段突出末端一致即可。如 BamHI、BglII 及 MboI 酶剪切后的末端序列都有 -GATC- 的粘末端。在相互连接时反应与同一酶解片段的连接是一致的,只是连接后的片段有时不能再用其中一些酶剪切开,这是由于连接后的酶切顺序特点改变所致。

(一) 基因与载体的连接反应

1. 常用试剂

(1) T4 DNA 连接酶及其用量。此种连接常用的酶是 T4 DNA 连接酶。T4 DNA 连接酶的商品制剂的单位标准不同,如有的用 Weiss 单位,它是指在 37°C 反应 20 分钟时生成 mmol³²P 的催化酶单位。如以 1 μg λ-DNA-Hind III 片段连接 50% 计算,约含 0.006 Weiss 单位。像 Biolab 厂所出的 T4 DNA Ligase 则直接用另一种单位,即在 16°C 下 20 μl 连接液中,30 分钟内将 0.12 μmol (300 μg/ml) 的 5' λ DNA Hind III 片段末端连接 50% 的酶量为 1 个单位。显然,两种单位相差很大,所以使用时应注意商品的原有单位标准。在连接基因与载体片段时,所用酶量往往要超过原应用的酶量,因为长途运输与保存都会降低酶活性。如果酶比较好,稀释用时,也需注意此酶是不能长期保存的。只要酶液足够用,往往都以加入微升量来计算,例如以 0.5 μl 或 1 μl 量加入。

(2) 连接缓冲液。常用的缓冲液为 Tris-HCl 溶液, pH7.8, 使用时需加含 -SH 的化合物如巯基乙醇或二巯苏糖醇,还应有一定浓度的镁离子及新鲜未降解的 ATP,除此之外还常加入一定量的牛血清蛋白。下面举一种常用缓冲液:

50mmol/L	Tris-HCl (pH7.8)
10mmol/L	MgCl ₂
20mmol/L	巯基乙醇
1.0mmol/L	ATP
50 μg/ml	牛血清蛋白

为了保证 ATP 不被降解,每次要用新鲜的,也可直接配制在以上缓冲液中,因为在以上溶液中 ATP 是比较稳定的。此溶液应在 -20°C 冰箱中保存,储备液不可反复冻融,最好配制后随即分装保存。

2. 连接时基因与载体的用量

DNA 加入的总量应以含有 0.1—1.0 μmol 的 5' 末端为最适。根据所加 DNA 片段分子大小不同,总计算一下所有末端数,即一个片段有两个末端,1.0 μmol 即有 0.5 μmol

DNA 分子,再算出 DNA 量。如以 300bp 的基因片段及 3,000bp 的线形末端为例,二者的比例为 1:1 时,前者应加约 0.2 μ g,后者应加 2 μ g,在 10 μ l 缓冲液中为 0.1 μ mol 溶液。当然 DNA 浓度大会有利于连接,但连接还与 DNA 分子大小、可能的溶解度及两种片段的比例有关。在分子克隆的连接中注意基因片段与载体的比例往往很重要。特别是在载体剪切后不去末端磷酸的情况下,比例适合可以防止载体本身的复环,也防止基因本身的相互连接。

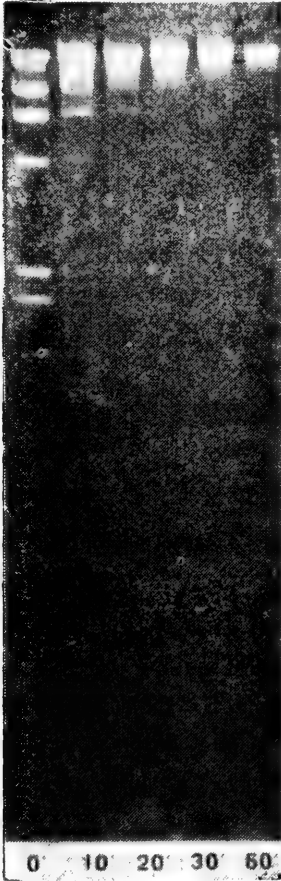


图 4.12 不同时间的连接反应(分钟)

3. 基因与载体在连接时常用的比例

通常以末端数比例计算,基因与载体的最适比例在 5:1 至 10:1 之间。如果二者分子大小接近可采用 2:1 至 3:1。如果分子大小十分悬殊,二者比例可达 100:1,

采用以上比例的理由是十分清楚的,基因分子小时多加一些以免不足。量多于载体也可防止载体形成二聚体及本身的复环。如果基因分子过多则又会造成基因片段本身的重复连接,这样不利于基因与载体二者的连接。以上比例经不少做分子克隆的工作者试用,都是比较合适的。

4. 连接反应的温度与时间

由于各产家生产的 DNA 连接酶活性不同,所定的活性单位也不同。常见的连接反应温度和时间有以下几种。

(1) 在 20 μ l 的反应体积中,当 DNA 总末端为 10 μ mol 时,用 0.01—0.1 单位的 T4 DNA 连接酶,在 22 $^{\circ}$ C 至少连接 1 小时。

(2) 在 20 μ l 的反应体积中,当总 DNA 量为 1 μ g 时,用 0.01—0.1 单位 T4 DNA 连接酶,在 22 $^{\circ}$ C 至少反应 1 小时。

(3) 在 20 μ l 的反应体积中,当 DNA 末端为 0.1—1.0 μ mol 时,用 0.01—0.1 单位 DNA 连接酶,在 16 $^{\circ}$ C 至少反应 30 分钟。

由于酶经过长途运输及长期保存会降低活性,在实际工作中常用的反应温度和时间如下:在 20 μ l 的反应体积中,当 DNA 量为 0.1—1 μ g 时,用 0.5—1.0 μ l T4 DNA 连接酶,在 16 $^{\circ}$ C 连接 3 小时后,继续在 4 $^{\circ}$ C 过夜。

(二) 基因与载体连接的具体实例

6 μ g λ DNA 被限制性内切酶 Hind III 剪切成 7 个片段,在 20 μ l 的反应体积中,用 1 μ l T4 DNA 连接酶再进行连接,结果见图 4.12。

二、平末端及 3' 突出末端的连接特点

(一) 平末端的连接

有许多限制性内切酶剪切 DNA 分子后产生带有平末端的 DNA 片段,如 HindII、

HpaI 等。5' 或 3' 突出末端的 DNA 片段经核酸酶 S1 削平, 或用 DNA 聚合酶 I (Klenow 大片段) 补齐, 都可使 DNA 末端成为平末端。T4 DNA 连接酶可以催化带有平末端的 DNA 片段之间的连接, 其反应条件与 T4 DNA 连接酶催化粘末端片段的连接条件类似, 但有以下几点应注意。

(1) T4 DNA 连接酶酶量。在进行平末端 DNA 片段的连接时, 通常要增加 T4 DNA 连接酶的量, 一般要比同样条件下的粘末端连接时的酶量多 2-5 倍。

(2) 当 T4 DNA 连接酶浓度较低时, 可加入少量 RNA 连接酶, 以促进连接。一般情况下都不加 RNA 连接酶。在 4-6% PEG 6000 存在条件下, 可增加连接效率。

(3) 平末端连接多用于基因表达中, 一般分子克隆中在遇到平末端时常加一连接子或同聚物尾, 使连接容易进行。

(4) 有人认为加大 DNA 用量也有利于连接。

(二) 3' 突出末端的连接

由于 DNA 合成的方向是 5' → 3', 5' 突出的粘末端在进行连接时, 很容易被 DNA 连接酶识别, 再加上粘末端互补顺序之间的退火, 连接比较容易进行。而 3' 突出粘末端之间的连接主要是靠末端互补顺序之间的退火, 所以连接效率较低。因此在剪切时以选择能产生 5' 突出末端的限制性内切酶为好。但有些酶切点由于位于载体的筛选标记部位, 因此也常被选用。如质粒 pBR322 Ap^r 基因上的 PstI 切点, PstI 剪切产生的末端为 3' 突出。具有 3' 突出末端的 DNA 片段在连接时也比 5' 突出末端的困难。一般采用的连接反应条件和平末端片段的连接条件相同。

(三) 平末端及 3' 突出末端的连接反应温度与时间

在 20 μl 反应体积中, 当 DNA 量为 0.1—1.0 μg 时, 用 1—2 μl T4 DNA 连接酶, 在 16℃ 连接过夜 (16 小时以上)。

三、基因片段大小与连接的关系

通常以小片段更易连接, 例如连接子片段只有几个或十几个核苷酸, 即使是平末端连接也容易成功。反之, 片段过大则连接困难。在本章开始曾讨论过载体选择问题, 而在载体一章中已经分别详述了质粒与 λ 噬菌体的特点, 其中也包括对被分子克隆基因大小的选择, 即 λ 噬菌体 DNA 通常适用于大片段基因, 特别是 10—20kb 的片段的分子克隆。质粒载体则只适用于 10kb 以下的基因分子克隆。在应用质粒来进行不同分子大小的基因克隆实验中, 1980 年 G. Drouin 曾做过一些统计。他将人线粒体 DNA 纯化后用不同的方法剪切成几类片段, 如用 E. coRI 和 BamHI 合并剪切, 用 HindIII 和 BamHI 合并剪切, 由于线粒体 DNA 较大, 也用 MboI 剪切, 分别分子克隆, MboI 剪切共得到 21 个大小不同片段, 其中最大的为 2,800bp, 最小的只有 25 及 33bp。这 21 个片段在剪切后不分别回收, 而是混在一起与 pBR322 连接, pBR322 是用 BamHI 剪切成线形分子。连接后转化到大肠杆菌 X-1776 中, 用 Ap^r 及 Tc^r 法筛选, 获得的有插入片段的阳性菌, 又经过电泳酶切分析, 在这一连接中每个片段与 pBR322 的分子比例都是一样的。筛选后除了第二个大片段 (2,400bp) 未得到分子克隆外, 其他均得到分子克隆菌株。克隆

株的成功率稍有不同,小于 200bp 的片段的成功率比较高,200—400bp 大小的基因片段的成功率约为 8%,400bp 以上至 2,800bp 者成功率为 4% 左右。其中第 5 片段(1,000bp)及第 8 片段(620bp)成功的菌落出现很少,在 705 个菌落中只出现一个。这一结果说明片段大的不易分子克隆。至于分子克隆中哪些步骤最为关键,一般认为是连接与转化,这两步骤中又以连接最为重要。因为连接反应的成功与难易,与所选用的载体、基因片段的大小、基因与载体使用的分子比例均有关。从以上实验结果看,在载体适用的范围内,如果不是把大小片段放在一起进行连接,并适当注意二者加入的分子比例,则各种大小的片段都可以成功,只是所获得的分子克隆数可能稍有差别,即成功率稍有不同,分子小的片段优于分子大的片段。

四、防止载体自身连接的脱末端磷酸基团

连接反应完成后,可直接转化合适的受体菌,挑选含有重组质粒的转化菌株进行大量培养,从中提取得到大量的重组质粒。可是往往会有这种情况,即从阳性转化菌株中得到的质粒,经限制性内切酶分析或分子杂交等方法检验,发现大部分并不带有连接上去的目的基因片段。这是因为用一种限制性内切酶剪切产生的载体片段,其两个末端的顺序是互补的,在连接反应中能够自身连接,恢复成环,重新形成原来的载体分子。为了防止这种情况出现,有人用碱性磷酸酶处理载体 DNA 片段的末端,以去除末端的磷酸基团,使载体 DNA 片段不能自身连接回复成环。同样,目的基因片段之间也可互相连接,不过这样形成的环形分子没有在宿主细胞内复制的能力,因此也不会形成可筛选的菌落。

DNA 片段末端用碱性磷酸酶处理后,在进行连接反应时所需要的 DNA 连接酶的量要比不用碱性磷酸酶处理时大得多。如一般用连接酶的浓度为 8u/ml,经过碱性磷酸酶处理后则有人加到 50u/ml 才能达到有效连接。

末端脱磷酸常用的碱性磷酸酶有两种,一种是细菌碱性磷酸酶,另一种是小牛肠碱性磷酸酶。一般多用小牛肠碱性磷酸酶(CIP),它在 SDS 中加热至 68℃ 时可完全失活。

(一) 脱磷酸的方法

1. 试剂

10 × CIP 缓冲液

0.5mol/L	Tris-HCl (pH 9.0)
10mmol/L	MgCl ₂
1mmol/L	ZnCl ₂
10mmol/L	亚精胺

10 × STE 缓冲液

100mmol/L	Tris-HCl (pH8.0)
1mol/L	NaCl
10mmol/L	EDTA

10%SDS

小牛肠碱性磷酸酶

一般 0.01 单位小牛肠碱性磷酸酶可去除 1pmol 5' 末端上的磷酸基(4kb 线性 DNA 1 μ g, 其 5' 末端是 1pmol)。

2. 方法

(1) 将酶切过的并经过乙醇沉淀纯化的 DNA 溶于 10mmol/L Tris-HCl (pH 8.0) 中, 使体积尽量小。

(2) 在 DNA 中加入 5 μ l 10 \times CIP 缓冲液, 并加水至 48 μ l, 再加小牛肠碱性磷酸酶适量。

(3) 如为 5' 突出的 DNA 片段, 则将反应物在 37 $^{\circ}$ C 保温反应 30 分钟, 然后加一次小牛肠碱性磷酸酶, 继续保温 30 分钟。如为平末端或 5' 缩入的末端则先在 37 $^{\circ}$ C 保温 15 分钟, 然后在 56 $^{\circ}$ C 保温 15 分钟, 再加一次小牛肠碱性磷酸酶, 并如上述两种温度下各保温 15 分钟。

(4) 反应到时间后立即加入 40 μ l 水, 10 μ l 10 \times STE 缓冲液, 5 μ l 10% SDS, 在 68 $^{\circ}$ C 加热 15 分钟, 用酚-氯仿抽提后, 乙醇沉淀 DNA。

(二) 脱磷酸的载体与基因的连接实例

把质粒 DNA 用 EcoRI 或 HindIII 剪切后, 再用碱性磷酸酶处理, 然后经过酚抽提、乙醇沉淀纯化, 即可将此除去磷酸的质粒 DNA 加到 cDNA 中 (此 cDNA 带有 Hind III 或 EcoR I 末端)。二者分子比为 3:1 (质粒:cDNA)。混合液在 60mmol/L Tris-HCl (pH7.6)、6.6mmol/L MgCl₂、10mmol/L DTT、1mmol/L ATP 液中连接, 加入 T4 DNA 连接酶, 浓度为 50 单位/ml。在 16 $^{\circ}$ C 反应 3 小时后再在 4 $^{\circ}$ C 放置过夜, 此连接的混合液反应后可直接转化到大肠杆菌受体细胞中。

(三) 脱末端磷酸在分子克隆中的使用范围

(1) 当目的基因 DNA 片段量非常小时(如只有几十纳克水平), 为了保持连接体系中总 DNA 的量, 就必须加大载体 DNA 的量, 这时可用碱性磷酸酶去除载体末端的磷酸, 防止载体的自身连接。

(2) 由于载体末端去磷酸, 不能自身连接, 同时也使基因 DNA 不容易与载体连接。

(3) 碱性磷酸酶的纯度及脱磷酸反应是否适当都影响连接。

(四) 用其他方法控制载体的自身连接

(1) 在用同一种限制性内切酶剪切载体和基因, 再进行片段之间的连接时, 如果基因片段的量大大超过载体 DNA 的量, 则载体片段末端之间相互退火的机会就大大降低, 自身连接就很少了。

(2) 如果用两种产生不同末端的限制性内切酶剪切载体, 则得到的载体片段的两个末端不能自身连接。如果基因片段也是用这两种限制性内切酶剪切得到, 则载体片段的两个末端可分别与基因片段的两个末端连接, 而且连接的方向是确定的。例如 pBR322 载体上有 EcoRI 和 BamHI 的两个切点, 用这两种限制性内切酶剪切得到的载体片段的两个末端分别是 EcoRI 的 AATTC $\overline{\text{G}}$ 和 BamHI 的 GATCC $\overline{\text{G}}$ 这两种顺序

的末端是不能互相连接的。如果基因片段也是用这两种限制性内切酶剪切得到, 则与载体片段连接时, 就不可能发生载体的自身连接, 但应注意防止载体二聚体的形成。

(3) 当载体片段的两个末端分别为平末端和粘末端时, 情况和上述类似, 只是需要把基因片段的两个末端也分别处理成一个平末端和一个能与载体的粘末端互补的粘末端。但这只有在基因表达质粒的重组中才常应用, 因为平末端的连接比粘末端困难得多, 剪切也不多用。

五、连接后限制性内切酶识别位点的恢复

如果载体片段和基因片段是用同一种限制性内切酶剪切得到的, 则在利用其粘末端进行连接后, 可以恢复原有的酶切位点, 即在连接的部位上仍具有原来的限制性内切酶的识别位点。

如果载体片段和基因片段是利用了不同的限制性内切酶剪切, 产生可互补的粘末端序列进行连接的, 形成重组分子后, 不会在连接部位上恢复原来的限制性内切酶的识别位点。例如 Xho I 和 Sal I 产生的粘末端顺序分别是 $\begin{array}{c} \text{TCGAG} \\ \text{C} \end{array}$ 和 $\begin{array}{c} \text{TCGAC} \\ \text{G} \end{array}$, 互补连接后, 产生的顺序是 $\begin{array}{c} \text{---GTCGAG---} \\ \text{---CAGCTC---} \end{array}$, 既不被 Sal I 识别也不被 Xho I 识别。

六、同聚物加尾后的连接

有时仅用限制性内切酶剪切, 无法使载体片段和基因片段具有可互补的粘末端, 所以连接比较困难。运用同聚物加尾的方法, 可在载体和基因片段上造成互补的末端顺序再进行连接, 方法与 5' 突出末端的连接类似。

参 考 文 献

- [1] Mamiatis, T. *et al.*, *Molecular Cloning*, Cold Spring Harbor Laboratory, 107-148, 1983.
- [2] Li-He Guo and Ray Wu, Exonuclease III use for DNA sequence analysis and in specific deletions of nucleotides, *Method in Enzymology*, **100**, 60-96, 1983.
- [3] Timothy nelson and Douglas Brutlag, Addition of homopolymers to the 3'-Ends of duplex DNA with terminal transferase, *Method in Enzymology*, **68**, 41-73, 1979.
- [4] Andrea, J., Twigs and David Sherratt, Trans-complementable copy-number mutants of plasmid ColEI, *Nature*, **283**, 216-220, 1980.
- [5] Jacques Drouin, Cloning of human mitochondrial DNA in *Escherichia coli*, *J. Mol. Biol.*, **140**, 15-34, 1980.
- [6] 马贤凯等译, 遗传工程, 人民卫生出版社, 91-119, 1978.
- [7] Li-He Guo and Ray Wu, New rapid methods for DNA sequencing B based on exonuclease III digestion followed by repair synthesis, *Nucleic Acids Research*, **10**(6), 2065-2082, 1982.

第五章 分子克隆技术(二)

——重组 DNA 的转化、筛选和鉴定

袁建刚 许彩民

目的基因和载体 DNA 在体外重组后,必须重新引入活细胞内才能进行增殖。外源 DNA 引入寄主细胞的过程一般称为转化。在正常生长条件下,大多数细菌并不发生转化作用。Mandel 等人发现,细菌细胞经过氯化钙处理之后,可促进 λ 噬菌体的转染作用。Cohen 等人后来证明, Ca^{2+} 同样可以介导质粒 DNA 对细菌细胞的转化。由于钙法操作简便有效,省时省力,目前已广泛应用于诱导重组 DNA 对细菌细胞的转化。

DNA 经过转化之后,就需要从大量的菌落中挑选细胞的克隆。从目前的技术水平看,转化的效率是比较低的,即使我们用转化效率比较高的细菌突变株作为受体菌,也不能使全部细胞转化成功。因此,筛选是基因工程研究工作中必不可少的过程。一般情况下,可通过以下三种主要途径进行筛选。①遗传学方法:即根据重组 DNA 转化后的细菌细胞具有某种特殊的遗传性状进行筛选。②菌落杂交方法:用体外标记的目的基因 DNA 或相应的 RNA 为探针同菌落进行原位杂交,筛选含重组 DNA 的克隆。这个方法可以快速地数百个菌落中挑选出含有重组 DNA 的菌落。③放射免疫方法:这种方法实际上是通过检测目的基因的表达产物,筛选含有重组 DNA 的克隆。此法需要制备 ^{125}I -标记的特异性抗体,然后根据抗原(表达产物)只能与特异性抗体形成复合物的原理进行筛选。

从菌落中筛选到所需要的含有重组 DNA 的克隆之后,必须对重组 DNA 或表达产物作进一步的分析和鉴定,这包括分析重组 DNA 的分子量以及目的基因插入载体的部位和方向是否正确。有时还需要对目的基因所表达的多肽的分子量和生物活性等进行全面分析和鉴定。只有通过这些鉴定工作,才能最终验证我们所设计的基因工程研究方案是否取得了圆满成功。

基因工程技术已广泛应用于与生物学和医学有关的科学领域,以及与工农业生产实践有关的研究工作中。由于选择的目的基因和研究的最终目标不同,往往需要选择不同的载体和寄主细胞。常用的基因载体有质粒、噬菌体、病毒等具有自我复制能力的小分子 DNA, 可以作为高等细胞载体的有线粒体和叶绿体 DNA 等。为了满足实际研究工作的需要,目前已有许多经过人工改造而组建的具有各种特殊功能和选择性标记的载体。对于不同载体需要选择相应的细菌、酵母、动物和植物等各种不同的细胞株作为受体细胞。然而不同的基因载体带有不同的选择性标记;不同的细胞株具有不同的遗传特性;不同的目的基因表达不同的多肽。我们不可能设计出一种对各种研究方案都适用的筛选和鉴定方法。因此,在本章我们只对一些基本的和通用性比较强的方法进行介绍。筛选和鉴定工作中需要使用的某些方法,例如 DNA 的限制性内切酶图谱分析、凝胶电泳、缺口翻译

法制备 DNA 探针等,在本书上册中已有介绍,本章就不再重复。

第一节 重组 DNA 对大肠杆菌的转化

重组 DNA 只有转入活细胞内进行扩增和表达才具有研究和实用意义。70年代以前,由于未能找到合适的方法和途径,外源 DNA 引入细胞的过程效率低、重复性差,不能满足日益发展的研究工作的需要。70年代初发现氯化钙处理后的细菌细胞可增加外源 DNA 的摄入,在此基础上相继建立了外源 DNA 进入细菌细胞的方法。目前大多数转化方法可达到每微克完整的质粒 DNA 产生 10^5 — 10^7 个转化子。有些方法可产生 10^7 — 10^8 个转化子,而且具有很好的重复性。这些方法的建立为基因工程工作奠定了坚实的基础。

不同种类的细胞甚至不同的细胞株,外源 DNA 的引入方法也不相同。大肠杆菌在目前分子克隆工作中应用最广泛,本节主要介绍几种常用的重组 DNA 对大肠杆菌的转化方法。

一、氯化钙法

(一) 试剂和材料

LB 肉汤

胰蛋白胨 10g/L
酵母浸膏 5g/L
氯化钠 5g/L
pH 7.5

平板培养基

胰蛋白胨 10g/L
酵母浸膏 5g/L
氯化钠 5g/L
琼脂 15g/L
pH 7.5

顶琼脂培养基

胰蛋白胨 10g/L
酵母浸膏 5g/L
氯化钠 5g/L
琼脂 6g/L

80mmol/L 氯化钙溶液

(二) 操作

- (1) 将受体菌种在平板培养基上划线, 37°C 过夜,进行活化。
- (2) 挑选单菌落,接入 5ml LB 肉汤, 37°C 震荡过夜。

(3) 从上述培养液中取 0.5ml, 转入 50ml LB 肉汤, 37℃ 震荡培养 2—4 小时, 到细菌密度大约为 5×10^7 /ml。

(4) 在水浴中将菌液冷却 10 分钟, 然后于 4℃ 下 3,500r/min 离心 15 分钟, 倾去上清液。

(5) 用 25ml 预先冰冷的高压灭菌氯化钙溶液 (80mmol/L) 重新悬浮细胞。

(6) 将细胞悬浮液在冰浴中冷 15 分钟, 再于 4℃ 下 3,500r/min 离心 10 分钟, 倾去上清液。

(7) 用 3ml 预先冰冷的高压灭菌氯化钙溶液 (80mmol/L) 分装在预先冰冷的高压灭菌 Eppendorf 离心管内, 每管 0.2ml, 4℃ 放置 12—24 小时。

(8) 加入待转化的重组 DNA 溶液, 和 0.2ml 菌液混匀, 在冰中放置 30 分钟。

(9) 在冰浴中放置 30 分钟后, 转入 42℃ 水浴中预热 2 分钟。

(10) 加入 1ml LB 肉汤, 于 37℃ 复壮 30 分钟到 1 小时, 不用震荡。这个步骤是为了使细胞复壮和抗性基因开始表达(用于四环素筛选一般需 1 小时)。

(11) 分别取 0.1ml、0.2ml、0.3ml 等合适的量, 涂布在含抗生素的平板培养基上。涂布时可使用扩散棒, 也可加入顶琼脂培养基铺布的方法。顶琼脂培养基铺布方法如下: 将融化的顶琼脂培养基在 43℃ 左右保温, 取合适量的转化细胞加入 2—3ml 顶琼脂培养基中, 混匀后, 迅速倒入固体平板培养基上, 转动平皿, 使之分布均匀。

(12) 将涂布好的平皿在室温放置 20 分钟到半小时, 使菌液被固体培养吸收或顶琼脂凝固。

(13) 将平皿放于 37℃ 培养箱中过夜, 一般在 12—14 小时菌落即可出现。

氯化钙法是用质粒 DNA 转化细菌细胞, 尤其是大肠杆菌细胞最常使用的方法。在转化时要注意以下问题:

(1) 低温对 Ca^{2+} 介导 DNA 的转化是必须的, 氯化钙处理时一定要在低温条件下进行。

(2) 用于转化的细菌细胞在对数生长期的转化能力最高, 静止生长期以后, 转化能力逐渐丧失。

二、氯化钙-氯化铷法

氯化钙和氯化铷协同处理受体菌, 可增加大肠杆菌许多菌株的转化能力。

(一) 试剂和材料

溶液 I:
10mmol/L MOPS (morpholinopropane sulfonic acid)
10mmol/L RbCl
pH 7.0

溶液 II:
0.1mol/L MOPS
50 mmol/L CaCl_2

10mmol/L RbCl

pH 6.5

二甲亚砜: 光谱纯

LB 肉汤

胰蛋白胨	10g/L
酵母浸膏	5g/L
氯化钠	5g/L
pH	7.5

平板培养基

胰蛋白胨	10g/L
酵母浸膏	5g/L
氯化钠	5g/L
琼脂	10g/L
pH	7.5

(二) 操作

- (1) 将受体菌种在平板培养基划线,于 37℃ 培养箱活化过夜。
 - (2) 挑选单菌落接入 LB 肉汤,37℃ 震荡过夜。
 - (3) 从上述培养液中取 0.5ml,转入 50ml LB 肉汤,37℃ 震荡 2—4 小时,到细菌密度大约为 5×10^7 /ml。
 - (4) 将细菌培养液于 4℃ 下 3,500r/min 离心 15 分钟。
 - (5) 倾去上清液,将细胞轻轻地重新悬浮于 25ml 高压灭菌的溶液 I 中。
 - (6) 再于 4℃ 下 3,500r/min 离心 15 分钟。
 - (7) 倾去上清液,将细胞轻轻地悬浮于 25ml 高压灭菌的溶液 II 中。
 - (8) 将细胞悬浮液于冰浴中放置 15 分钟。
 - (9) 每次转化取细胞悬浮液 1ml,4℃ 下 3,500r/min 离心 15 分钟。
 - (10) 尽量倾尽上清液,再将细胞重新悬浮于 0.2ml 高压灭菌的溶液 II 中。
 - (11) 在 0.2ml 细胞悬浮液中加入 $3\mu\text{l}$ 二甲亚砜, $10\mu\text{l}$ 以下体积的待转化重组 DNA (1—200ng)。
 - (12) 冰浴中放置 30 分钟。
 - (13) 再于 43—44℃ 水浴中保温 30 秒。
 - (14) 加入 5ml LB 肉汤,37℃ 保温 1 小时后即可涂布平板。
- 注意: 二甲亚砜的氯化物对转化有很强的抑制作用,需要将新鲜的光谱纯二甲亚砜分装成小包装,封好放于 -70℃ 保存,每个小包装启封后只能用一次。

三、大肠杆菌 X1776 的转化

大肠杆菌 X1776 用常规的钙处理方法进行转化时效果往往不太理想。Norgard 等人用质粒 pBR322 对大肠杆菌 X1776 转化的各种最适条件进行了摸索,在此基础上

Hanahan 制定了下述转化方案。

(一) 试剂和材料

X1776 培养基

胰蛋白胨	25g
酵母浸膏	7.5g
1mol/L Tris-HCl	20ml
pH 7.5	

定容到一升,高压灭菌后,冷到 50℃ 以下时加入以下成分:

1mol/L MgCl ₂	5ml
1% 二氨基庚二酸	10ml
0.4% 胸苷	10ml
20% 葡萄糖	25ml

除了 MgCl₂ 可用高压灭菌外,另外三种用消毒的硝酸纤维素膜(孔径 0.45 μm)过滤处理。

转化缓冲液

RbCl	100m mol/L
MnCl ₂	45m mol/L
CaCl ₂	10m mol/L
LiCl	0.5m mol/L
醋酸钾	35mmol/L (pH 6.2)
葡萄糖	15% (超纯)
pH 5.8	

缓冲液的制备:先溶解其他成分,再加入固体 RbCl 和 MnCl₂,并用稀醋酸仔细调节 pH 到 5.8。

(二) 操作

- (1) 将大肠杆菌在平板培养基上划线,37℃ 过夜进行活化。
- (2) 挑选单菌落,接入 5ml X1776 液体培养基 37℃ 震荡过夜。
- (3) 取上述培养液 0.5ml,转入 50ml 的 X1776 培养基中,震荡培养 3—4 小时,直到细菌密度大约为 5×10^7 /ml。
- (4) 于 4℃ 下 3,500r/min 离心 10 分钟。
- (5) 倾去上清液,用 10ml 冰浴预冷的转化缓冲液温和悬浮细胞,将细胞悬浮液在冰浴中放置 5 分钟。
- (6) 于 4℃ 下 3,500r/min 离心细胞悬浮液 5 分钟。
- (7) 倾去上清液,重新将细胞轻轻悬浮于在 2ml 冰浴中预冷的转化缓冲液中。分装于冷的高压灭菌 Eppendorf 离心管内,每管 0.2ml。冰浴中放置 5 分钟。
- (8) 每管中加入 7 μl 光谱纯二甲亚砜 (DMSO),再于冰浴中放置 15 分钟(二甲亚砜的用法见本节方法二)。

(9) 放入干冰-乙醇浴中, -55°C 冷冻 20 秒。

(10) 放入室温水浴中迅速融化冰冻细胞。融化后, 转到冰浴中放置 5 分钟。

(11) 加入 $7\mu\text{l}$ DMSO, 混匀后再于冰浴中放置 5—10 分钟。

(12) 加入 1—25ng 待转化的 DNA (总体积 1— $10\mu\text{l}$), 混匀后冰浴中放置 10—30 分钟。

(13) 在干冰-乙醇浴中冷冻 30—60 秒。

(14) 在室温水浴中迅速融化后, 放置到冰浴中。

(15) 将细胞转入 42°C 水浴中 1 分钟, 或 37°C 水浴中两分钟。

(16) 再加入 1ml X1776 培养基, 混匀后于 37°C 保温。如用于四环素筛选放置 30 分钟, 如用于氨苄青霉素或卡那霉素则放置 1 小时。

(17) 随后即可在选择性平板上涂布。

本方法不但适用于大肠杆菌 X1776 菌株, 略加改动后 (转化缓冲液中不加 LiCl 和 MgCl_2 并省去步骤(9)和(13)的冷冻), 同样可用于大肠杆菌 C600、DH1、MM294 等菌株。但不适用于 HB101 菌株。

大肠杆菌 X1776 菌株对去污剂非常敏感, 实验中所用器皿必须仔细冲洗干净。

为了提高转化效率, 在转化过程中要注意上述对寄主细胞的处理条件, 以保持寄主细胞的良好感受态。另外, 外源 DNA 的大小和结构也对转化效率有很大影响。一般说来, 分子越小, 转化效率越高; 环形分子的转化效率比线形分子的转化效率要高得多。共价闭环的超螺旋的质粒 DNA 比同种线形 DNA 分子的转化效率高上千倍。因此, 在转化前要尽量保证重组 DNA 为完整的环状分子。

四、重组 DNA 克隆的初步筛选

从转化后得到的大量菌落中初步筛选可能含有重组 DNA 的克隆, 最简便快速的方法是与微生物技术配合的遗传学方法。目前经常使用的基因载体都携带选择性遗传标记, 或具有一种可供选择的遗传学特性。例如常用的质粒载体一般具有抗药性标记或营养标记, 而噬菌体载体形成的噬菌斑本身就是可供挑选的遗传学特性。

利用载体 DNA 的遗传学特性进行筛选是初筛的第一步, 随后可应用插入失活的原理区分重组 DNA 和载体 DNA 自身连接所形成的菌落。以下用质粒 pBR322 和 M13 噬菌体 DNA 为例, 说明利用基因载体的选择性标记进行初筛的基本原理。

(一) 质粒 pBR322 为基因载体的筛选

质粒 pBR322 是分子克隆中使用最广泛的基因载体之一, 关于质粒 pBR322 的分子生物学特性, 第二章已有详细介绍。它具有两个抗药性标记, 一个是氨苄青霉素 (Ap) 抗性, 另一个是四环素 (Tc) 抗性。用质粒 pBR322 作基因载体时, 常常是将外源基因插入其中一个抗性基因中。外源基因插入后, 则引起这个抗性基因失活, 在转化受体菌细胞后, 可直接在平板培养基上筛选单抗性克隆。例如, 外源基因插入 Bam HI 位点后, 引起 Ap 抗性基因失活, 转化后的受体菌细胞直接涂布到含有 Ap 的平板培养基上 (一般含 Ap $40\mu\text{g}/\text{ml}$ 左右), 在 Ap 培养基上, 不具有 Ap 抗性的受体菌细胞不能生长。 37°C

培养过夜后,再将转化后的细胞形成的菌落按相同顺序接种到含有 Ap(40 μg/ml) 和 Tc (25 μg/ml) 的平板培养基上。在 Ap 培养基上生长而在 Tc 培养基上不能正常生长的单抗性菌落,即可能是含有重组 DNA 的克隆。在两种抗菌素培养基上都能生长的双抗性菌落,则是载体 pBR 322 DNA 自身连接后转化细胞所形成的菌落。这样可以方便地初步选出所期望的克隆。

(二) M13 噬菌体为基因载体的筛选

M13 噬菌体 DNA 是单链环状 DNA。但在细胞内的复制型 DNA 是双链 DNA。双链复制型 M13 DNA 可用作基因载体。由于分泌到细胞外的成熟 M13 噬菌体的 DNA 是单链的,因此在 DNA 序列测定中是常用的基因载体。M13 噬菌体产生的噬菌斑的大小与插入的外源 DNA 片段的大小成反比,这可作为一个方便的筛选标记。但是小噬菌斑的形成可受其他因素的影响,因此用这个性质筛选并不十分可靠。不过目前经过人工组建的 M13 噬菌体 DNA 的衍生物都易于筛选,这是因为它们分子中都插入了乳糖操纵子的一部分,含有半乳糖苷酶基因。在此基因产生的半乳糖苷酶的作用下,M13 噬菌体在含有异丙醇-β-D 硫代半乳糖苷 (IPTG) 和 5-溴-4-氯-3-吡啶-β-D 半乳糖苷 (X-gal) 的平板培养基上形成蓝色噬菌斑。当外源基因插入半乳糖苷酶基因中时,则不再产生有活性的半乳糖苷酶,使 IPTG 和 X-gal 不再产生蓝色反应,只能形成无色透明的噬菌斑。因此,可以根据颜色区分重组体 M13 DNA 和载体 M13 DNA 形成的噬菌斑。

根据基因载体的遗传学特性进行筛选常常可达到事半功倍的效果。因此,重组 DNA 转化受体细胞后,一般先利用基因载体的遗传学特性进行初筛,然后再进行进一步的筛选和鉴定。除去利用基因运载体的遗传学特性进行筛选外,有时也可利用外源基因的遗传学特性进行筛选。如果含有外源基因的重组体在受体细胞内得到了表达,并且这个外源基因的表达产物和受体细胞的营养缺陷株之间有互补作用,则含这种重组 DNA 的缺陷型受体细胞表现出非缺陷型的遗传学特性,从而可筛选出我们期望的克隆。但由于这种筛选方法限制因素很多,大多数情况下不易找到与目的基因相互补的营养缺陷型细胞株,因此,这种互补选择法远不如载体筛选法应用广泛。

第二节 重组 DNA 克隆的核酸杂交筛选方法

一、菌落杂交

菌落杂交方法在筛选工作中被广泛使用,它的最大优点是通用性比较强,可以用来检测任何一种插入的 DNA 序列。这种方法依据的是核酸分子杂交的原理,而不以插入的 DNA 序列能否实现基因表达为前提,因此,只要有可利用的标记探针就可检测重组 DNA。关于核酸探针的标记方法,在本书上册第一章已有介绍,这里只介绍菌落杂交的方法。

(一) 材料和试剂

变性溶液

NaOH 0.5mol/L

NaCl 1.5mol/L

中和溶液

Tris-HCl 0.5mol/L

NaCl 3mol/L

pH 7.0

预杂交溶液

6 × SSC

SDS 0.5%

(1 × SSC 为 0.015mol/L 柠檬酸钠, 0.15mol/L NaCl)

1 × Denhardt 溶液(含非可、聚乙烯吡咯烷酮和牛血清清蛋白各为 0.02%)

杂交溶液

2 × SSC

1 × Denhardt 溶液

SDS 0.5%

磷酸钾缓冲液 25mmol/L

EDTA 2mmol/L

pH 7.2

洗膜溶液

2 × SSC

SDS 0.5%

(二) 操作

(1) 硝酸纤维素膜的处理。把硝酸纤维素膜(孔径 0.45 μm)剪成直径 8cm 左右的圆片(略小于培养皿的内径),装入一干净培养皿中,高压消毒。也可用紫外灯在无菌橱中每面照射 30 分钟进行消毒。

(2) 滤膜菌落和参比菌落的制备。硝酸纤维素膜作好记号后,平铺于琼脂培养基平皿上,用扩散棒除去膜与培养基之间的气泡。大约一分钟后,培养基即可将滤膜均匀浸湿。然后用消毒牙签在滤膜上接种单菌落。37℃ 培养 2—3 小时,待菌落直径大约为 1 mm,使硝酸纤维素膜菌落面向上,移到另一块含氯霉素(170 μg/ml)的琼脂培养基平皿上,37℃ 培养 15—20 小时以扩增质粒。在滤膜上扩增质粒的同时,在另一块琼脂培养基平皿上按相同顺序和位置转种,作好顺序记号,作为菌落参照系,37℃ 培养过夜,2—4℃ 密闭保存。

(3) 菌体裂解和 DNA 的变性及固定。氯霉素扩增后,把滤膜从平皿中取出,菌落面朝上,放置于滴有一薄层变性溶液的塑料膜上。5—10 分钟后移到干滤纸上吸干碱液。再使菌落面向上放置于滴有一薄层中和溶液的塑料膜上,中和两次,每次 15 分钟。室温

晾干后, 80℃ 真空干燥 2 小时。

为确保细菌细胞或 DNA 在操作时不发生移位, 也可用以下方法进行。滤膜取出后, 菌落面向上, 放于 4 层用变性溶液浸透的 Whatman 3MM 滤纸上。5 分钟后把滤膜放到一合适的抽滤器上, 用水泵抽至菌落接近干燥 (1—2 分钟)。然后用 120ml 中和液分 4 次抽过滤膜, 室温晾干后, 70℃ 真空干燥两小时。

(4) 杂交。用同位素标记的 DNA 为标记探针时, 把滤膜浸入预杂交液中, 于 60℃ 预杂交 6 小时。DNA 探针变性后加入到杂交溶液中, 然后把滤膜浸入含变性 DNA 探针的杂交溶液中, 55℃ 杂交 16—24 小时。探针的用量一般为 10^5 — 10^6 cpm/ml。为了节约探针, 也可用以下方法: 预杂交后, 将膜取出, 于 60℃ 干燥 15 分钟, 用 1—2ml 含变性 DNA 探针的杂交液浸湿滤膜, 在矿物油中 55℃ 杂交 16—24 小时。

杂交后, 滤膜用洗膜溶液于 60℃ 洗 3 次, 每次 45 分钟。

如果用标记的 RNA 作为杂交探针, 则可省去预杂交步骤, 直接将膜浸入含 RNA 探针的 $5 \times \text{SSC}$ 溶液中进行杂交。55℃ 杂交 16—24 小时后, 滤膜用 $2 \times \text{SSC}$ 洗两次, 每次 30 分钟, 再用 50ml 含 $20 \mu\text{g/ml}$ RNase A 的 $2 \times \text{SSC}$ 洗 30 分钟。

(5) 放射自显影。膜洗好后, 用滤纸吸干, 夹于两层塑料薄膜中包好, 放入暗盒, 压好 X 光片和增感屏, 于 -60℃ 冰箱曝光。

菌落杂交与 DNA 的吸印转膜杂交和斑点杂交不同, 后者一般只有 DNA, 而菌落杂交时, 许多细胞成分与 DNA 一起形成痕斑, 即使经过固定步骤之后, DNA 与膜的结合仍然不很牢固。因此, 菌落杂交与纯核酸的杂交最适温度不同, 一般温度略低一些可防止 DNA 从膜上脱落。

菌落杂交时, 硝酸纤维素膜要经过碱溶液处理。膜经碱处理之后变脆, 因此, 随后的杂交和压 X 光片时, 操作必须小心, 以免膜破碎。

二、重组噬菌体 DNA 的杂交筛选

菌落杂交方法经过适当修改即可用于噬菌斑的筛选。最初 Kramer 等人用和菌落杂交相似的方法筛选含重组 DNA 的噬菌斑, 即挑取单个噬菌斑, 按顺序点到生长有寄主细胞的硝酸纤维素膜上, 培养一段时间后, 将生长有噬菌斑的硝酸纤维素膜碱变性和 DNA 固定, 再用标记的核酸探针进行杂交和放射自显影。这个方法也具有很高的灵敏性, 但需要逐个挑取噬菌斑, 比较费时。为了加快筛选速度, Benton 和 Davis 设计了另一种方法。即把硝酸纤维素膜放置在经过转化和培养的含噬菌斑的琼脂平板表面上, 使滤膜和噬菌斑直接接触。部分噬菌体可扩散到滤膜上, 随后进行变性、固定和杂交。由于省去了挑取噬菌斑, 大大加快了筛选速度。但是, 由于膜上的噬菌体 DNA 比较少, 大大降低了灵敏性。Savio 将两种方法的优点结合起来, 建立了灵敏、快速的下述筛选方法。

(一) 试剂

见本节菌落杂交。

(二) 操作

(1) 噬菌体的转移和增殖。硝酸纤维素膜剪成大小合适的圆片(直径略小于平皿内径),用笔标好方向记号后,放入一干燥的平皿内,高压灭菌 20 分钟。

将培养好的受体细菌用液体培养基稀释 10 倍,把经过消毒的硝酸纤维素膜浸入稀释后的细胞悬浮液。取出已包上一层细胞的硝酸纤维素膜,放到一张经过高压灭菌的 Whatman 3MM 滤纸上,再把膜转到第二张灭菌的干燥 Whatman 3MM 滤纸上,室温干燥 30—60 分钟。然后将受体菌的硝酸纤维素膜平铺到待筛选的长有噬菌斑的琼脂培养基平板上(整个过程要在无菌橱中进行),随后在 4℃ 冰箱放置 10 到 30 分钟使噬菌体扩散转移。

噬菌体转移后,按硝酸纤维素膜上的记号,将平板标好方向。再将硝酸纤维素膜从平板培养基上轻轻揭下来(一般涂布含噬菌体的平板时,最上面是一层 0.6% 的顶琼脂培养基,操作时一定要小心,防止把平板表层揭坏)。揭下的滤膜铺到一块新鲜的平板培养基上(注意使接触噬菌斑的一面向上),于 37℃ 保温过夜。

在此过程中,膜上的受体菌生长繁殖,而从原平板上转到膜上的噬菌体也同时进行增殖,从而在膜上形成一个个噬菌斑,并且每个噬菌斑出现的位置也和原平板一致。原来的平板放 4℃ 冰箱保存。硝酸纤维素膜上的噬菌斑即可做原位杂交。

(2) 噬菌体 DNA 的变性和固定。把长有噬菌斑的硝酸纤维素膜从平皿上取下,放到用变性溶液浸湿的 Whatman 3MM 滤纸上,室温变性 10—15 分钟。然后放到用中和溶液浸湿的 Whatman 3MM 滤纸上中和两次,每次 30 分钟。再将膜放在干净的滤纸上室温晾干后,于 70℃ 真空干燥 2 小时。

(3) 杂交和放射自显影见本节菌落杂交。

(三) 噬菌体转移时应注意的问题

本方法省去了挑取单个噬菌斑的步骤,经过噬菌体的转移和扩增又保证了有足够的噬菌体 DNA 用于随后的原位杂交,从而可进行快速灵敏的筛选。整个操作过程中噬菌体的转移是最关键的一步。只有转移到硝酸纤维素膜上的噬菌斑在数目和位置上与原来的平板一致,才能进行成功的筛选。为了做到这一点,硝酸纤维素膜用受体菌悬浮液浸过之后,必须放到干燥的 Whatman 3MM 滤纸上吸干,充分吸干后的膜铺到含噬菌斑的平板上后,在平板培养基中的溶液慢慢浸湿膜的同时,噬菌体颗粒按原位向上扩散到硝酸纤维素膜上,就可得到与原平板完全一致的复制品。如果将很湿的膜铺到含噬菌斑的平板上,膜上的溶液则向下扩散,不但影响噬菌体向膜上转移,而且会导致原平板上的噬菌斑向四周扩散,从而使转移失败。

第三节 免疫化学筛选方法

抗原与对应的抗体之间可发生特异反应,根据这个原理可设计免疫化学的筛选方法。免疫化学方法的优点是它不依赖外源基因和载体为寄主细胞提供选择性遗传标记,也不依赖于外源基因与寄主细胞之间的功能互补,甚至也不依赖于外源基因所编码的多肽得

到完整的功能活性的表达。只要产生了含有抗原决定簇的多肽片段，原则上就可以用这种方法进行筛选。随着基因工程研究的发展，目前已经建立了几种比较灵敏快速的放射性抗体筛选方法。

一、抗体的制备

被免疫的动物放血后，取血 10—20ml，在室温下放置 2 小时，血块用玻璃棒捣松散，4℃ 放置 16 小时，凝血块经离心（10,000g）15 分钟去掉。上清液部分加入固体硫酸铵到 50% 饱和度（每毫升血清加 0.313g 硫酸铵），在 4℃ 轻轻摇动使硫酸铵溶解。于 4℃ 离心（9,000g）15 分钟后去上清液，将沉淀溶于原体积的 10mmol/L 磷酸钾缓冲液（pH 6.8）中。再加硫酸铵到 50% 饱和度，同样条件离心，收集沉淀用原体积的 PBS 溶液（10mmol/L 磷酸钠，0.14mol/L NaCl，pH 7.0）溶解，并对 PBS 透析。

二、抗体的纯化

用于同位素标记的抗体还需要作进一步的纯化，比较好的纯化方法是亲和层析。一般是将抗原蛋白分子共价结合到用 CNBr 活化的琼脂糖颗粒上，然后将抗体粗制品通过抗原琼脂糖柱，进行亲和层析，即可分离得到比较纯净的抗体制品。

（一）琼脂糖 4B 的活化

15ml 的琼脂糖 4B 用蒸馏水充分洗涤，在显微镜下检查有无细菌污染，并去除破碎颗粒。在冰浴中用等体积的冷蒸馏水悬浮，并通过搅拌在整个反应过程中保持悬浮状态。用 10% 的 NaOH 将悬浮液调到 pH 大约为 11。每 30ml 悬浮液加入 4g CNBr，在反应中随时监测 pH，并通过滴加 10% NaOH 以保持 pH 在 10.8—11.2。反应终止倾去上清液，将琼脂糖 4B 颗粒倒入玻璃沙芯漏斗内，迅速用 5 倍体积预冷的 PBS 溶液（10mmol/L 磷酸钠，pH9.5，0.142mol/L NaCl）充分洗涤，为了加快冲洗速度可用水泵温和地抽滤。

（二）亲和层析

将 1g 制备好的活化琼脂糖 4B 加到 4ml 含有 10—20mg 抗原蛋白质的 0.1mol/L 碳酸钠溶液中（pH9.0）。反应混合物温和震荡 1 小时后，于 4℃ 放置过夜。离心去上清液，将琼脂糖 4B 倒入 50ml 乙醇胺溶液中（1mol/L 乙醇胺，0.1mol/L 碳酸钠，pH8.0），于 4℃ 放置 4—6 小时，以封闭活化的琼脂糖 4B 的剩余的活性基因。离心去上清液，用 PBS 溶液（10mmol/L 磷酸钠，pH7.2，15mmol/L NaCl）洗两次。将洗好的抗原-琼脂糖 4B 装入小层析柱内（1.5 × 4cm）。为了去除未能共价结合但已经吸附到琼脂糖 4B 上的抗原蛋白质分子，可用以下两种溶液交替地连续洗柱：① 醋酸缓冲液（0.1mol/L 醋酸钠，1mol/L NaCl，pH4.8）50ml；② 碳酸缓冲液（0.1mol/L 碳酸钠，1mol/L NaCl，pH7.6）50ml。洗完三个循环之后，用上述的 PBS 溶液 50ml 平衡层析柱，然后将抗体粗提物上样。样品进入柱床后，至少先要保持 30 分钟以上，以使抗体和抗原结合。

然后用 50ml PBS 溶液洗柱,再用 50ml 上述的醋酸和碳酸缓冲液交替洗柱三次。最后用 4.5mol/L 的 $MgCl_2$ 或用 0.1mol/L 的甘氨酸 (pH2.8) 洗脱抗体蛋白质。

三、抗体蛋白质的 ^{125}I 标记

在氯胺 T 等氧化剂存在时,可用 $Na^{125}I$ 标记蛋白质,但反应条件比较剧烈,易引起蛋白质变性。用乳糖过氧化物酶为催化剂,以过氧化氢为氧化剂,进行标记反应时,条件温和不会导致标记蛋白质的变性。在溶液中进行标记时酶蛋白会造成污染,因此,David 等人将乳糖过氧化物酶共价连接到琼脂糖 4B 上,再用固相酶催化标记反应。这种方法避免了酶蛋白对标记蛋白质的污染,而且蛋白质溶液中有尿素和 SDS 等变性剂存在时也能进行有效的标记。

(一) 固相乳糖过氧化物酶的制备

用上述方法以 CNBr 活化琼脂糖 4B。取 3g 活化好的琼脂糖 4B 倒入 5ml 含 1mg 左右的乳糖过氧化物酶的 PBS 溶液中 (10mmol/L 碳酸钠, 0.14mol/L NaCl, pH 7.0), 于 4°C 震荡过夜,再用冷 PBS 溶液充分洗涤。然后重新悬浮于磷酸-甘氨酸缓冲液中 (2.0mol/L 甘氨酸, 10mol/L 磷酸钠, pH7.5)。4°C 放置 5 小时,充分洗涤后存放于含 10^{-3} mol/L 硫酸盐的 PBS 溶液中。

(二) ^{125}I 标记

固相乳糖过氧化物酶用前需用 PBS 溶液充分洗涤。在 Eppendorf 管内依次加入下述反应组分: 50 μ l 的 $10 \times$ PBS 溶液,非标记 KI 溶液 (到终浓度 10^3 mol/L), 0.5mCi 放射性 $Na^{125}I$ (一般 $Na^{125}I$ 是溶于 0.1mol/L 的 NaOH 溶液中,所以随后应加入等体积 0.1mol/L 的 HCl 溶液中和), 10 μ l 的固相乳糖过氧化物酶,最后加入终浓度为 10^{-3} — 10^{-4} mol/L 的 H_2O_2 (反应体积为 0.5ml 左右)。立刻封口进行反应。反应过程中不停地震荡,以保持固相酶的悬浮状态,室温反应 10 分钟后加入 1/100 体积的 2.5mol/L NaN_3 终止反应。反应终止后,加少量固体非放射性 KI。为了减少随后操作的损失,也可加少量载体蛋白。标记混合物用琼脂糖 G-25 柱使标记蛋白质和游离 ^{125}I 分离,然后分管记数,第一个洗脱峰为标记的蛋白质。

四、溴化氰活化滤纸法

此方法是先将抗体与溴化氰活化的滤纸结合,然后将活化滤纸同琼脂平板上的裂解克隆保温,表达克隆所产生的抗原将与活化滤纸上的抗体结合,再经过与 ^{125}I 标记的抗体保温后,就可以进行放射自显影检测。这种方法使用 ^{125}I 标记的抗体,因此有比较高的灵敏性。可检测出 50pg 的抗原。

(一) 溴化氰活化滤纸的制备

将 20 张直径为 9cm 的 Whatman 40 号滤纸用 0.1mol/L 的碳酸氢钠浸泡 15 分钟, 用蒸馏水漂洗后备用。在 1L 烧杯中将 500ml 2.5mol/L 的碱性磷酸钾 (pH 11) 预热至 50℃, 加入 CNBr 12g, 边加边摇至全部溶解, 在 20—25℃ 时, 放入漂洗后的滤纸, 反应中不时轻轻摇动烧杯, 反应进行 8 分钟 (CNBr 有毒, 操作要在通风橱中进行)。倾出 CNBr 溶液使反应停止, 用 500ml 的 0.1mol/L NaHCO₃ 洗两次, 每次 5 张滤纸依次通过 250ml 0.1mol/L NaHCO₃, 250ml 水, 200ml 50% 丙酮(水溶液), 50ml 丙酮抽滤, 滤纸放入真空干燥器中干燥, 最后在干燥器中 4℃ 保存。活化滤纸比较稳定, 可保存几个月。

(二) 抗体的结合

30—60 μ l 抗血清在培养皿中同 3ml 的 20mmol/L 磷酸钠缓冲液 (pH 6.8) 混合, 如果抗体的效价低, 也可多用些抗血清。将每张活化的滤纸放入一个培养皿中, 使纸和平皿之间不要有气泡, 并且保持平皿水平。用石蜡密封后在 4℃ 下至少 18 小时。滤纸取出后放入含有 0.1mol/L 的甘氨酸、0.5% 牛血清白蛋白、0.14mol/L NaCl 和 10mmol/L 磷酸钠 (pH 7.0) 的溶液中, 每张滤纸用 5ml, 置 37℃ 下保温 6 小时, 此时滤纸即可使用。如不立刻使用, 于 4℃ 下过夜后, 再放在 -20℃ 冰箱备用。使用前浸入含 5% 牛血清、0.14mol/L NaCl、10mmol/L 磷酸钠 (pH 7.0) 的溶液中, 每张滤纸用 5ml。浸泡 30 分钟后, 再用 25ml 的此溶液抽滤冲洗。最后按每张滤纸 4ml 的量浸泡到使用。

(三) 克隆的裂解和抗体滤纸的应用

将用重组质粒转化后的克隆在 S 琼脂平板培养基 (每升含胰蛋白胨 32g、NaCl 5g、酵母浸膏 20g、琼脂 15g, pH7.5) 上生长 1—2 天。将平板倒扣在预先滴有 1ml 氯仿的直径为 7cm 的 Whatman 3MM 滤纸上, 30 分钟后拿开平板, 放置 1 小时。然后在平板上加一层顶琼脂, 此顶琼脂层中含 1.5ml 顶琼脂 (每 100ml 含 1g 胰蛋白胨、0.8g Difco 纯琼脂、0.8g NaCl, pH 6.8)、100 μ l 0.5% SDS 和 100 μ l 溶菌酶 (12.5 mg/ml), 于 45℃ 加到平板上。顶琼脂凝固后, 将溶于 PBS 溶液 (0.14 mol/L NaCl、10mmol/L 磷酸钠缓冲液, pH 7.0) 中的纯化抗原按 1 μ l 含 10ng、1ng 和 0.1ng 的量点在平板上作阳性对照。以后琼脂板上的滤纸方向就通过这些对照点和所刺的位置记号来确定。再将共价结合了抗血清的滤纸平铺于琼脂板上, 要保证纸和平板之间不能有气泡, 上面再加 5—10cm 的粗滤纸, 于室温反应 5—9 小时, 使滤纸上的抗体和菌落中可能含有的特定抗原结合。然后每次取 4 张滤纸放在含有 5% 牛血清的 100ml PBS 溶液中轻轻摇动洗 1—2 小时。每张滤纸再用 200ml 的 PBS 溶液抽滤冲洗, 最后每 4 张纸滤一组放到含 5% 牛血清的 PBS 溶液中。

(四) ¹²⁵I 标记抗体的结合和放射自显影

将上述浸泡滤纸的 PBS 溶液倒掉, 再加入 10ml 含 4 × 10⁶cpm 的 ¹²⁵I 标记的抗体溶液 (此溶液含 10mmol/L Tris-HCl, pH 7.5, 0.14 mol/L NaCl, 25% 牛血清),

将容器盖好,室温下轻轻摇动 5—7 小时,然后将放射性溶液倒出,换上 50ml 含 5% 牛血清的 PBS 溶液,30 分钟后,用 200ml PBS 抽滤冲洗,最后浸泡于 PBS 中 15 分钟后取出,使标记面向上晾干,压好 X 光片后,于低温冰箱放射自显影。

五、固相筛选法

抗体能牢固结合到塑料面上,并且保持与抗原结合的能力,因此可将抗体预先结合到合适的塑料制品(如带小井的塑料盘、小塑料珠等)的表面上。加入待测的重组 DNA 克隆的粗抽提物后共同保温,特定的抗原和抗体将结合。冲洗后再加入用 ^{125}I 标记的同种抗体保温,则放射性抗体就会和已结合到冷抗体上的抗原结合。经充分冲洗后进行计数,根据放射性强度就可以检测出所需的克隆。

(一) 细胞粗抽提物的制备

将待检测的克隆转移到 S 琼脂平板上(每升含胰蛋白胨 25g、NaCl 5g、酵母浸膏 20g,琼脂 15g, pH7.5),每个平板可移植 50 个克隆左右。生长过夜后,每个克隆用无菌的刮勺刮下放入 1.5ml 的 0.05mol/L Tris-HCl (pH7.5) 并含有 0.5 毫克溶菌酶的小离心管内,在室温下保温 30 分钟后,置 0℃ 下 1 小时,然后,用液氮冻融三次。最后,加入苄甲基磺酰氟到终浓度为 10 nmol/L,在 37℃ 处理 10 分钟,离心收集上清,即可用于测定。

(二) 塑料井检测

7 μg 抗体溶于 1ml 的 10mol/L Tris-HCl (pH 9.2) 溶液中,然后取抗体溶液加到塑料微滴度盘的每个井内,一个井加入 150 μl 。在室温放置过夜,然后用蒸馏水冲洗并使其干燥。在每个井内加入 150 μl 的待检测的克隆细胞抽提物,37℃ 保温 4—6 小时,并轻轻震荡,保温后用蒸馏水冲洗,然后每个井内再加入 150 μl 的 ^{125}I 标记的同种抗体(标记抗体溶于 10mmol/L Tris-HCl pH 7.5、0.14 mol/L NaCl、25% 小牛血清溶液中,每个井大约加 10^5cpm 计数的抗体),在室温保温过夜。倾去溶液,用 0.8% NaCl 溶液充分冲洗后干燥。将每个井切下后分别记数,如同时有阴性对照和阳性对照,即可根据放射性强度筛选出所需的克隆。这种方法可以检测出 1ng 以下的蛋白质。

(三) 塑料珠检测

塑料井法检测的缺点是在测定计数时要分别切下每个塑料小井。后来,人们又把抗体包被在小塑料珠上,然后把小塑料珠放到塑料微滴度盘的小井内与抗原和标记的抗体反应。反应结束后取出每个小塑料珠分别进行记数。此方法的操作步骤和原理基本上和塑料井法相同。目前已有测定乙型肝炎表面抗原和 e 抗原的包被抗体的塑料小珠和成套试剂的商品药盒出售,此药盒即可用于临床检测,也可用于乙肝疫苗的基因工程工作中。下面以乙型肝炎抗体塑料小珠的应用为例,说明塑料珠检测方法的具体操作。

在每个塑料盘的小井内加入一个包被有抗体的小塑料珠。再于每个井内加入待测的样品 0.2ml (病人血清或乙肝疫苗基因工程的待检测克隆细胞抽提物)。同批作阴性对

照 3 个,阳性对照 2 个。塑料盘封好胶纸轻轻震动后,置室温过夜。吸出每个井内的溶液后用蒸馏水冲洗干净。于每个井内加入 0.2ml 的 ^{125}I 标记的同种抗体。轻轻震动塑料盘,于 45°C 水浴中保温 1 小时,吸出反应液,各孔用蒸馏水冲洗干净。将塑料小珠移到计数管内,用 γ 计数器分管计数。如果阳性对照平均计数和阴性对照平均计数之比 (P/N 值) 大于 5, 则实验成立。如果样品计数大于 2.1 倍的阴性对照平均计数阈值, 则为阳性结果。

第四节 重组质粒和 λ 噬菌体 DNA 的小量快速制备方法

一般经过筛选之后,得到的克隆数目仍然比较多,这就需要建立小量快速的分离重组质粒和噬菌体的方法,以便尽快完成分析和鉴定。以下介绍几种质粒 DNA 和 λ 噬菌体 DNA 的快速分离制备方法。

一、碱裂解法制备质粒 DNA

(一) 试剂

LB 肉汤培养基

胰蛋白胨 10g/L

酵母浸膏 5g/L

NaCl 10g/L

pH7.5

TEG 缓冲液

Tris-HCl 25mmol/L

EDTA 10mmol/L

葡萄糖 50mmol/L

pH8.0

TE 缓冲液

Tris-HCl 10mmol/L

EDTA 1mmol/L

pH8.0

碱裂解液

NaOH 0.2mol/L

SDS 1%

醋酸钠溶液

醋酸钠 3mol/L

冰醋酸 2mol/L

pH4.8

(二) 操作

筛选出的单菌落接种到 5ml LB 肉汤培养基,并根据重组质粒的抗性加入合适的抗菌素,37℃ 震荡过夜。取 1.5ml 培养液放到 Eppendorf 管内(其余培养液放 4℃ 冰箱保存),离心后吸出上清液。将离心管倒扣在干滤纸上,尽量空干溶液。将 0.4mg 溶菌酶溶于 100 μ l 冰冷的 TEG 缓冲液中,加入离心管,旋转混匀后悬浮细胞,室温放置 5 分钟后,加入新配制并预冷的碱裂解液,盖好盖,上下翻转两三次混匀。将离心管在冰水中放置 5 分钟,加入 150 μ l 醋酸钠溶液,盖好盖,用旋转混匀器混匀并旋混 30 秒钟,冰水中放置 5 分钟。于 4℃ 下 18,000r/min 离心 5 分钟,上清液转移到另一个干净的 Eppendorf 管内,加入等体积的酚或氯仿,涡旋混匀后,5,000r/min 离心 5 分钟,再把上层水相转移到一个干净 Eppendorf 管内。加入 2 倍体积乙醇混匀,于 -70℃ 放置 5 分钟后 18,000r/min 离心,倾去上清液,将 Eppendorf 管倒扣在干滤纸上,空干溶液。再加入 1ml 70% 冷乙醇,混旋后,18,000r/min 离心,倾去上清液,真空干燥(不要过干)。加入 50 μ l 含有 RNase A (20 μ g/ml) 的 TE 缓冲液(为了去除 RNase 中可能的 DNase 污染,使用前可将 RNase 于 80℃ 处理 5 分钟)。

处理后的质粒 DNA 可直接用于限制性内切酶分析。取 10 μ l DNA 溶液加入 1.2 μ l 合适的 10 \times 酶解缓冲液,加 1—2 单位所需的限制性内切酶,37℃ 保温 1—2 小时,即可进行凝胶电泳分析。从电泳图谱可以确定是否有插入片段。剩余的质粒 DNA 溶液放 -20℃ 冰箱保存。

二、煮沸法制备质粒 DNA

(一) 试剂

TETS 溶液

Tris-HCl 10mmol/L

EDTA 50mmol/L

Triton X-100 0.5%

蔗糖 8%

溶菌酶溶液

Tris-HCl 10mmol/L pH 8.0

溶菌酶 10mg/ml

RNase A 溶液

Tris-HCl 10mmol/L

EDTA 1mmol/L pH 8.0

RNase A 50 μ g/ml

用前 80℃ 处理 10 分钟。

(二) 操作

将每个单菌落接种于 5ml LB 肉汤培养基,并根据重组质粒的抗性加入合适的抗菌

素,于37℃震荡过夜。取1.5ml到Eppendorf管内(剩余培养液放4℃冰箱保存),8,000r/min离心5分钟,弃上清液,倒扣Eppendorf管在干燥滤纸上,空干溶液。将细胞重新悬浮于0.35ml TETS溶液,加入25 μ l新配制的溶菌酶溶液。旋转混匀3秒钟后将Eppendorf管放沸水浴中煮40秒钟,立刻用离心机于18,000r/min离心10分钟。将上清液转入另一干净Eppendorf离心管内,加入420 μ l异丙醇和40 μ l的2.5mol/L醋酸钠溶液。混匀并于乙醇干冰浴中放置15分钟后,在4℃下,18,000r/min离心15分钟。弃上清液,空干异丙醇溶液。加入50 μ l RNase A溶液,37℃保温10分钟。

取10 μ l制品加入干净的Eppendorf管内,加入1.2 μ l合适的10 \times 酶解缓冲液,再加入1—2单位所需的限制性内切酶,37℃反应1—2小时后即可进行凝胶电泳,从电泳图谱可确定是否有插入片段。剩余样品放-20℃低温冰箱保存。

三、从单菌落中制备质粒 DNA

有时候在载体DNA中插入的外源DNA片段比较大,因此重组质粒DNA和载体质粒DNA的分子量相差很多。在这种情况下,不用限制性内切酶消化,经过凝胶电泳就可以判断那个菌落中含重组质粒。同时,无需扩大培养,从一个单菌落就可制备足够的质粒DNA,进行一次琼脂糖凝胶电泳。此方法的操作过程如下。

在含有合适抗菌素的琼脂平板培养基上将菌落养到直径3mm左右。用消毒牙签逐个地从菌落上挑取少量细胞到另一个含抗菌素的平板培养基上保种,用刮勺从原平板上逐个把剩下的细胞转到Eppendorf管内,加入25 μ l裂解溶液(50m mol/L NaOH、5m mol/L EDTA、0.5% SDS、0.025% 溴甲酚绿),用牙签温和搅匀裂解液。盖好盖,68℃保温45—60分钟。加入2.5 μ l的25% 菲可溶液,上样到不含溴乙锭的0.7% 琼脂糖凝胶进行电泳。电泳后再浸入溴乙锭溶液(0.5 μ g/ml)中染色45分钟。电泳时不加溴乙锭,能更准确地反映分子量大小。

四、平板裂解法制备 λ 噬菌体 DNA

(一) 试剂

SM 培养基

白明胶	0.01%
NaCl	5.8 g/L
MgSO ₄ · 7H ₂ O	2g/L
Tris-HCl	50mmol/L
pH	7.5

NZCYM 培养基

NZ 胺	10g/L
NaCl	5g/L
酵母浸膏	5g/L
酪蛋白酸	1g/L

MgSO₄ · 7H₂O 2g/L

加入 1.5% 琼脂糖,作平板培养基;

加入 0.7% 琼脂糖,作顶平板培养基。

PEG 溶液

聚乙二醇 20% (MW = 6000)

2 mol/L NaCl

溶于 SM 培养基。

TE 缓冲液

10mmol/L Tris-HCl

1mmol/L EDTA

pH 8.0

(二) 操作

挑取噬菌斑转入 1ml SM 培养基中,加一滴氯仿。于 4℃ 放置 4—6 小时,使噬菌体颗粒从顶琼脂糖培养基中扩散出来。取 50—100 μ l 噬菌体悬浮液,加 100 μ l 受体菌悬浮液,37℃ 保温 20 分钟。再加入 2.5ml 的融化并于 45℃ 保温的 0.7% 顶层琼脂糖培养基,混匀后,铺于 NZYCM 平板培养基上(制作平板培养基时,最好用琼脂糖而不用琼脂,因琼脂中含有抑制限制性内切酶的物质,影响随后的酶切分析)。倒放平板,于 37℃ 培养 9—14 小时。加入 5ml SM 培养基在室温温和摇荡 1—2 小时后,将上层培养液转入离心管中,于 4℃ 下 8,000g 离心 10 分钟。取上清液,加入 RNase A 和 DNase I 至终浓度各为 1 μ g/ml,于 37℃ 保温 30 分钟。加入等体积的 PEG 溶液,在 0℃ 放置 1 小时,沉淀噬菌体颗粒。于 4℃ 下,10,000g 离心 20 分钟后,吸出上清液,倒扣离心管在干滤纸上,空干液体。加入 0.5ml SM 培养液,用涡旋混匀器悬浮噬菌体颗粒,于 4℃ 下 8,000 g 离心 2 分钟去细胞碎片。上清液转移到一干净 Eppendorf 离心管内,加入 10% SDS 和 0.5mol/L EDTA 各 5 μ l;于 68℃ 保温 15 分钟。依次用酚、酚和氯仿(1:1)、氯仿各抽提一次。每次都要将水相转入一个干净 Eppendorf 管内再进行下次抽提。最后将水相加入等体积的异丙醇,于 -70℃ 放置 20 分钟,或液氮中冷冻 5 分钟,融化后于 4℃ 下 18,000r/min 离心 15 分钟,用 70% 乙醇洗沉淀一次,18,000r/min 离心后去上清液,抽干后把沉淀溶于 50 μ l。

取 10 μ l 样品,加入 1.2 μ l 合适的 10 \times 酶解缓冲液,加入 1—2 单位所需的限制性内切酶,于 37℃ 保温 1—2 小时后进行凝胶电泳分析,以确定是否有插入 DNA 片段。剩余部分于 -20℃ 下保存。

五、液体培养基裂解法制备 λ 噬菌体 DNA

(一) 试剂

见方法四。

(二) 操作

挑取噬菌斑接种于 1ml SM 液体培养基内,加入一滴氯仿,4℃ 放置 4—6 小时,使噬菌颗粒从琼脂糖培养基中扩散出来。取 0.5ml 噬菌体悬浮液(大约 3×10^6 个噬菌体),加入处于静止生长期受体细胞悬浮液(大约加入 1.6×10^8 个细胞),于 37℃ 保温 15 分钟,加入 4ml NZCYM 液体培养基,37℃ 下震荡 9 小时。再加入 0.1ml 氯仿,37℃ 下震荡培养 15 分钟,将裂解物转到离心管中,4℃ 下 8,000g 离心 10 分钟,取上清液,加入 RNase 和 DNase 到终浓度各为 $1 \mu\text{g}/\text{ml}$,32℃ 保温 30 分钟,加入等体积的 PEG 溶液,0℃ 放置 1 小时,沉淀噬菌体。4℃ 下 10,000g 离心 20 分钟,吸出上清液,将离心管倒扣在干滤纸上,空干液体。加入 0.5ml SM 培养液,旋转混匀悬浮噬菌体颗粒,4℃ 下 8,000g 离心 2 分钟,上清液转到干净 Eppendorf 管内,加入 10% SDS 和 0.5mol/L EDTA 各 $5 \mu\text{l}$,68℃ 保温 15 分钟,依次用酚、酚和氯仿(1:1)、氯仿各抽提一次,每次离心后将水相转入一干净的 Eppendorf 管内再进行下次抽提。最后将水相加入等体积的异丙醇,于 -70℃ 放置 20 分钟,或液氮中 5 分钟,于 4℃ 下 18,000r/min 离心 20 分钟。再用 70% 乙醇洗沉淀一次,再经过 18,000r/min 离心后去上清液,抽干沉淀并溶于 $50 \mu\text{l}$ 的 TE 缓冲液中。

取 $10 \mu\text{l}$ 作限制性内切酶和凝胶电泳分析,剩余部分放 -20℃ 冰箱保存。

第五节 重组 DNA 的酶切分析和鉴定

用快速提取方法从筛选出的克隆中制备了重组质粒或重组噬菌体 DNA 之后,随后要对重组 DNA 进行分析和鉴定。鉴定的第一步是用限制性内切酶和凝胶电泳确定重组 DNA 的物理图谱。利用限制性内切酶绘制 DNA 物理图谱的一般方法,本书上册的有关章节中已有详细描述。在这一节中,只针对重组 DNA 的特点简单介绍有关的限制性内切酶的分析方法。另外,重组 DNA 经过酶切和凝胶电泳分析后,经常需要进行转膜杂交鉴定,有时还要进行核苷酸序列分析。有关杂交和序列分析的具体方法请参阅上册的有关章节。

一、单酶解分析

为了简化重组 DNA 的鉴定工作,一般在设计重组方案时,往往选择载体 DNA 的单切口限制性内切酶酶切位点作为目的基因的插入部位。如果目的基因的 DNA 片段也是用同一种酶切所获得的,则和载体 DNA 连接之后,会在目的基因 DNA 片段两侧产生两个相同的酶切位点。这样制备重组 DNA 后,只要使用一种相同的限制性内切酶消化,就可从重组 DNA 上切下目的基因 DNA 片段,再经过凝胶电泳,就可得到目的基因 DNA 片段的电泳区带,具体作法如下。

根据插入的 DNA 片段在整个重组 DNA 分子中所占的比例,取 $0.5-2 \mu\text{g}$ 重组 DNA,加入 $1-5 \text{ u}$ 合适的限制性内切酶和缓冲溶液,于 37℃ 保温 2 小时。然后上样到 1% 的琼脂糖凝胶上进行电泳,同时还要上样另外四份样品作对照,一份不经过酶切的

重组 DNA, 一份不经过酶切的载体 DNA, 一份经过同一种酶切的载体 DNA 和一组大小合适的 DNA 片段作分子量标准。

电泳结束后, 用溴乙锭染色, 在紫外灯下观察和照相, 即可得到电泳图谱。

从电泳图谱中, 一般只有在经过酶切的重组 DNA 样品中才可看到: 一条插入的外源基因 DNA 的区带。根据分子量标准可测得它的分子量, 根据分子量可初步判断它是否为原来的插入片段。

为了进一步鉴定, 照相后可将凝胶进行变性和中和, 再将 DNA 转移到硝酸纤维素膜上进行杂交(具体方法见本书上册第六部分)。杂交后可在放射自显影图谱上看到未经过酶切的重组 DNA 有杂交区带, 而经过酶切的重组 DNA 则只有插入的 DNA 片段, 部分有杂交区带。载体 DNA 部分不再出现杂交区带。其他作为对照的样品则没有任何杂交区带。

二、双酶解分析

有时为了定向连接, 常选用两种不同的限制性内切酶。进行酶切分析时, 自然还需要用两种限制性内切酶对重组 DNA 进行消化。但有时目的基因 DNA 片段和载体 DNA 连接后, 接口处不再保留任何合适的酶切位点, 在作酶切分析时, 往往要在两个接口的附近各找一种合适的酶切位点。同样也就需要用两种酶作酶切分析。具体的酶切和电泳方法和单酶切法相同。但要注意酶切缓冲液要合适。

如果两种酶所要求的缓冲系统相同或相近, 则可在一种缓冲系统加入两种酶同时消化, 如果两种酶所要求的缓冲系统相差比较大, 则需要先用一种限制性内切酶进行消化, 用等体积饱和酚和氯仿抽提后, 再加入 2 倍体积的乙醇沉淀 DNA。将沉淀的 DNA 溶于 5--10 μ l 水后, 再进行另一种限制性内切酶消化。

三、多次酶解和电泳分析

有时为了对重组 DNA 作详尽的酶切分析, 在进行第一次酶切后, 还要从凝胶中回收酶切片段, 然后再对回收的 DNA 片段进行酶切和电泳分析。酶切片段的回收本书上册已介绍了一些方法。回收步骤比较复杂, 而且回收过程中 DNA 片段有丢失。以下介绍一种可用于多次酶切和电泳的低熔点琼脂糖凝胶电泳方法, 电泳后可直接把含 DNA 区带的凝胶切下来, 凝胶薄片于 65 $^{\circ}$ C 融化后, 于 37 $^{\circ}$ C 保温, 在此温度下, 低熔点琼脂糖不发生凝聚, 可直接加入第二种酶消化。这种方法既快速又可减少样品丢失, 步骤如下:

(1) 将低熔点琼脂糖加入电泳缓冲液中(不要使用含磷酸盐的缓冲体系, 磷酸盐改变低熔点琼脂糖的融解性质), 加热融化后用常规方法铺胶。

(2) 凝胶制好后即可将酶解后的 DNA 样品上样进行电泳, 因为低熔点琼脂糖凝胶的强度比普通琼脂糖凝胶小, 电泳应尽可能在低温条件下进行。

(3) 电泳后, 将凝胶浸入 0.5 μ g/ml 的溴乙锭溶液中染色 30 分钟, 紫外灯下照相后, 浸入水中 10 分钟洗去多余溴乙锭, 再用刀片切下需要进一步分析的 DNA 区带, 凝胶要尽量切得薄一些。

(4) 将含 DNA 片段的凝胶薄片加到小试管内, 加入 10 倍体积的冷水, 4 $^{\circ}$ C 下浸泡

(6) 选用合适的第二种单切点限制性内切酶将标记的 DNA 切成两段,电泳后分别回收两个标记片段,随后即可分别进行部分降解酶切分析。

(二) 部分降解酶切分析

(1) 取 10^4 cpm 的末端标记 DNA, 加入 $1\mu\text{g}$ 载体 DNA, $1\mu\text{l}$ 10X 酶解缓冲液, 1—2 单位限制性内切酶, 补水到 $10\mu\text{l}$, 混匀后于 37°C 保温。

(2) 保温到 2、5、10、15 和 30 分钟时, 分别取出 $1.8\mu\text{l}$, 转入同一个含有 $1\mu\text{l}$ 0.5 mol/L 的 EDTA 溶液的 Eppendorf 管中终止反应。最后将几批反应产物混匀。

(3) 加入 $2\mu\text{l}$ 指示染料后, 分成两份, 一份上样到 5% 的聚丙烯酰胺凝胶, 一份上样到 1.4% 的琼脂糖凝胶, 即分别可进行电泳。如需要分子量标准, 可取相同计数的一组已知分子量的标记 DNA 片段同时上样。

(4) 指示染料溴酚蓝走到凝胶长度的 $3/2$ 处时, 停止电泳, 用塑料薄膜包好凝胶, 压好 X 光片后, 置 -70°C 冰箱放射自显影。

冲洗 X 光片之后, 从最小片段的区带读起, 再从小到大顺序读出各片段之差, 即为彻底降解时应得到的各片段的排列顺序, 从而可方便地确定各酶切点的位置。

五、人工合成探针在鉴定中的应用

分子杂交是克隆筛选和鉴定中应用最广泛的技术, 经常遇到的困难是不容易得到纯化的探针。近年来, 多聚核苷酸人工合成技术的发展为解决这个难题作出了贡献。目前, 多聚核苷酸人工合成的反应程序已自动化, 并已有被称为基因合成仪的 DNA 合成仪器出售, 它可以根据已知的基因序列快速合成所需的探针。目前有许多具有重要生物功能的酶或多肽的氨基酸顺序已分析清楚。在不知道基因的核苷酸序列的情况下, 还可根据基因所编码的蛋白质的氨基酸顺序推算出基因的核苷酸顺序。一般作为探针用的 DNA 片段并不需要很长, 只要合成十几个核苷酸的寡核苷酸就可以作为杂交探针。

人工合成的单链 DNA 片段一般采用末端标记的方法进行标记。末端标记的方法可参照本书上册第五部分中的方法进行。合成的探针没有 5' 磷, 可省去脱磷一步。也可以在合成一个十几个核苷酸片段的同时, 按照此片段 3' 末端的核苷酸顺序合成一个 5—6 个核苷酸的互补顺序作为引物。二者复性后, 再用大片段大肠杆菌 DNA 聚合酶 I 按本节中第四部分的步骤(3)—(5)介绍的方法进行合成标记。为了提高探针的比活性, 要选用合适的 $[\alpha\text{-}^{32}\text{P}]$ dNTP, 即此种核苷酸在其余的互补顺序中出现次数最多, 再补加其他的非标记 dNTP。例如, 假如有下面带有引物的合成探针:



则应选用 $[\alpha\text{-}^{32}\text{P}]$ dATP (与 dT 互补), 再补加另外三种非标记 dNTP 即可得到高比活性的探针。这种方法可称为引物延伸法。用这种方法算记小片段的 DNA 探针, 比活性可达到 1×10^9 cpm 以上。如想得到更高比活性的探针, 也可用一种以上的 $[\alpha\text{-}^{32}\text{P}]$ dNTP。

人工合成寡核苷酸探针在克隆筛选和鉴定以及其他一些核酸研究中用途越来越广

泛。尤其在 cDNA 克隆的筛选和鉴定工作中有重要作用。使用人工合成探针不但可以克服纯化探针的困难,而且通过合成不同部位的核苷酸顺序作为探针,可以鉴定 cDNA 克隆中的重组 DNA 是否包含完整的 mRNA 的基因顺序。

一般杂交反应的温度要低于杂交双链的 T_m 值。而用小片段的合成探针时,形成的杂交双链很短,因此一定要根据探针的长度控制杂交反应的温度。小片段探针形成的杂交双链的 T_m 值可根据下式计算:

$$T_m = 16.6 \log [\text{Na}^+] + 0.41(\%G + C) + 81.5 - \frac{500}{\text{片段长度}}$$

Na^+ 浓度用克离子浓度表示,片段长度用碱基数表示。

第六节 SDS 聚丙烯酰胺凝胶电泳和 Western 技术在克隆鉴定中的应用

限制性内切酶分析、凝胶电泳、Southern 杂交等 DNA 分析方法在分子克隆鉴定中具有重要作用。但是,在许多情况下,还需要对表达产物做一定的分析才能最终确定我们所期望的克隆。对表达产物的分析包括测定目的基因表达的多肽的分子量、N 末端等是否正确,有时甚至还要测定是否具有特定的生物活性以及测定全部氨基酸顺序。在目前分子克隆的鉴定工作中,多肽的 SDS 聚丙烯酰胺凝胶电泳和 Western 技术应用比较广泛,因此,本节将对这两种多肽分析技术的原理和方法作简单的介绍。

聚丙烯酰胺凝胶电泳是分离蛋白质、核酸等生物大分子应用最广泛的技术之一。凝胶电泳能把一个成分复杂的大分子混合物分离开,主要依据样品中各种分子的分子量和所带电荷不同。尤其是在蛋白质的凝胶电泳时,电荷效应对蛋白质的迁移率有很大影响,这在分离蛋白质混合物时是一个有利因素。但是,如果要利用凝胶电泳测定某种蛋白质的分子量,就必须将电荷差异而引起的效应消除或减少到可以忽略不计的程度,使蛋白质迁移率的大小完全取决于分子量,这样才能根据标准蛋白质的分子量的对数和迁移率所做的标准曲线得出待测蛋白质的分子量。在基因工程中,对表达产物的鉴定,很重要的一项工作就是测定表达产物的分子量,因此,必须选择合适的电泳体系,使电荷效应降低到最小程度。目前,测定蛋白质分子量最常用的电泳技术是 SDS 聚丙烯酰胺凝胶电泳,即在电泳体系中加入一定量的 SDS。SDS 是一种表面活性剂,它以一定比例和蛋白质分子结合成复合物,使蛋白质分子带大量负电荷。SDS 的负电荷大大超过蛋白质分子自身的电荷差异,从而可消除或大大降低蛋白质自身的电荷影响。而且 SDS 和蛋白质分子的结合是与蛋白质分子的质量成正比的,所以各种 SDS-蛋白质复合物电泳时的迁移率可反映分子量的不同。

对表达产物粗抽提物进行聚丙烯酰胺凝胶电泳分离后,分析结果时常遇到的一个困难是如何确定哪条区带是外源基因的表达产物。一般情况下,未经纯化的粗提物中含有许多寄主细胞的蛋白质组分,如果有的组分和外源基因的蛋白质分子的分子量相近,就会给判断造成困难。1979 年 Towbin 等人建立了将蛋白质从聚丙烯酰胺电泳中转移到硝酸纤维素膜上的方法。这个方法实际上是 Southern 技术的改良和发展,因此,又称为

Western 转膜技术。利用此方法,可在合适条件下通过电泳使分离后的蛋白质分子从聚丙烯酰胺凝胶中按原位转移到硝酸纤维素膜上,从而可在硝酸纤维素膜上得到与原蛋白质凝胶电泳分离图谱完全一样的复制品。随后可在硝酸纤维素膜上进行免疫反应。如果使用 ^{125}I 或酶标的特异性抗体,则可得到待测蛋白质的放射自显影或显色区带。这种方法把聚丙烯酰胺凝胶电泳、免疫化学、同位素或酶标技术有机地结合在一起,具有灵敏、微量、分辨率高、特异性强等优点。是从多组分蛋白质混合物中检测和鉴定某一种蛋白质的非常有效的手段。这种方法和核酸的硝酸纤维素杂交方法相似,因此,又可称为蛋白质杂交。

一、SDS 聚丙烯酰胺凝胶电泳

(一) 试剂

聚丙烯酰胺凝胶贮存液

丙烯酰胺 40g

甲叉双丙烯酰胺 1.5g

溶于 100ml 去离子水,过滤后装入棕色瓶中,4℃ 下保存。

4×分离胶缓冲液

Tris-HCl 1.5 mol/L pH 8.8

4×大孔胶缓冲液

Tris-HCl 0.5 mol/L pH 6.8

SDS 溶液

SDS 20%

过硫酸铵溶液

过硫酸铵 1.5%

用前配制。

TEMED 溶液

TEMED 0.5% (V/V)

用前配制。

5×样品处理缓冲液

Tris-HCl 0.25 mol/L

DTT 0.1 mol/L

SDS 5%

蔗糖 1.25mol/L

pH 7.5

电泳缓冲液

甘氨酸 0.36mol/L

Tris 5m mol/L

SDS 0.1%

pH 8.3

电泳时用蒸馏水稀释一倍。

染色液

考马斯蓝 R250 0.05%

异丙醇 25%

醋酸 10%

脱色液

水:乙醇:冰醋酸为 8:3:1 (V/V)

(二) 操作

(1) 制胶。垂直板和水平板凝胶的制备方法是相似的,下面以垂直板为例说明聚丙烯酰胺凝胶的制备过程。首先选择两块 2—3mm 厚的表面平整的玻璃板(大小根据需要而定)用去污剂洗净,再分别用自来水和蒸馏水冲洗干净。移动洗净后的玻璃板时最好带医用乳胶手套,不要用手接触玻璃板表面。玻璃板直立空干后,在两块板之间的左、右、下三面夹上塑料条或玻璃条,厚度可根据需要和上样量而定,一般在 1—4mm。用夹子夹好后,再用石蜡或 1% 的琼脂糖凝胶封边,垂直放好。然后在上口先按所需的胶浓度加入分离胶,到离顶部 3—4cm 时,再轻轻加入高 0.2cm 的无离子水以隔离空气。等分离胶聚合后,倒掉上层的水,再加入大孔胶,放样品梳子,梳齿下沿离分离胶 1cm。封水隔离空气。等大孔胶聚合后倒掉水,小心取掉梳子和下边的边条,安装在电泳槽上,装好电泳缓冲液,随后上样。

几种常用凝胶浓度的配方如下:

分离胶

胶浓度(ml)	6.5%	7.5%	9%	12%
丙烯酰胺贮存液	7.3	5.36	6.8	13.5
4×分离胶缓冲液	11.25	7.5	7.5	11.25
20% SDS	0.22	0.15	0.15	0.22
1.5% 过硫酸铵	2.25	1.5	1.5	2.25
0.5% TEMED	2.25	1.5	1.5	2.25
H ₂ O	21.75	13.73	12.6	15.53
总体积	45	30	30	45

大孔胶

丙烯酰胺贮存液 0.67ml

4×大孔胶缓冲液 2.5ml

20% SDS 0.05ml

1.5% 过硫酸铵 1.03ml

0.5% TEMED 1.03ml

H₂O 4.62ml

总体积 10ml

(2) 上样。先在加样孔加满电泳缓冲液。将样品和样品处理缓冲液按 4:1 比例混合

好，沸水浴煮 1 分钟，冷却后，加少量示踪染料（碱性缓冲液一般可用 0.1% 溴酚蓝），5—10 μ l，即可用微量加样器上样。加样时将加样器下口伸入到加样孔下部小心加入。聚丙烯酰胺凝胶电泳所需样品的量不大。一个加样孔只需 5—100 μ l 的样品溶液（1—100 μ g 蛋白质样品）。由于样品组分不同，用量可在一定范围内调整，组分越少的样品用量也越少。

（3）电泳。用电泳缓冲液浸湿的三层滤纸将上槽缓冲液与凝胶板上口的缓冲液沟通，形成缓冲液-滤纸桥。要尽量缩短上槽和凝胶板之间的距离，以减短纸桥，降低电阻。接通电源，起始用 30mA 电流，示踪染料通过大孔胶后，用 70mA 的恒定电流。电泳 5—6 小时。

（4）染色和脱色。电泳结束后，取出凝胶，浸入到考马斯蓝 R250 染色液中过夜后，再浸入脱色液中脱色，换 2—3 次脱色液，直到把本底洗脱干净。

如在聚丙烯酰胺电泳后进行电泳转膜和蛋白质杂交方法检测，则不用染色。

（5）电泳图谱分析。为了检测外源基因在寄主细胞内的表达产物，一般要将重组 DNA 克隆的蛋白质抽提物和不含重组 DNA 的寄主细胞的蛋白质抽提物同时上样，不含重组 DNA 的寄主细胞抽提物作为阴性对照。同时还要上样一组合适的已知分子量的蛋白质作为标准，以便绘制分子量标准曲线。如果有该外源基因编码的蛋白质纯化样品作为阳性对照，则更有利于对电泳结果进行分析。电泳分离和染色后，如果重组 DNA 克隆蛋白质抽提物的电泳图谱中含有比阴性对照多余的蛋白区带，而此蛋白区带的迁移率和阳性对照蛋白质的迁移率相同，则可初步判断此蛋白区带可能是外源基因的表达产物。如果没有纯化样品作为阳性对照，则可根据此多余蛋白质区带的迁移率和分子量标准曲线计算此蛋白质的分子量。如果和期望的外源基因编码的蛋白质分子量相同，也可初步判定为阳性结果。为了得到确切结果，可用以下蛋白质杂交方法做进一步鉴定。

二、电泳转移

剪切与电泳后的蛋白质聚丙烯酰胺凝胶同样大小的一张硝酸纤维素膜和两张 Whatman 3MM 滤纸。在硝酸纤维素膜上用软铅笔标好方向和记号，将膜和滤纸用电泳缓冲液浸湿。把一张滤纸放到一块海绵上，海绵下边是一块布满微孔的有机玻璃板。把硝酸纤维素膜放在滤纸上，然后小心地把蛋白质聚丙烯酰胺凝胶铺到膜上，用玻璃棒赶出膜和胶间的气泡，再放一张滤纸在胶上，然后盖一块海绵，压一块带微孔的有机玻璃板。将上、下两块微孔有机玻璃板压紧后插入电泳槽壁上的沟槽内，在电泳槽内加满电泳缓冲液，接通电源就可进行电泳转移。接电极时要注意，从凝胶到硝酸纤维素膜的方向要和电泳方向一致。电压用 6V/cm，一般 1—2 小时就可转移完全。电泳装置示意图见本书上册图 4.1.1。

电泳转移后的硝酸纤维素可以染色（染色液为 0.1% 氨基黑、40% 甲醇和 10% 醋酸）和脱色（脱色液为 90% 甲醇和 2% 醋酸）。如果随后进行免疫学方法测定，则不用染色。

三、抗体的放射性检测法

电泳转移后的硝酸纤维素膜浸入含牛血清白蛋白的缓冲液中 (3% 牛血清白蛋白、0.9% NaCl 和 10 mmol/L Tris-HCl, pH 7.4), 于 40℃ 保温 1 小时, 封闭膜对蛋白质的非特异结合。将 ^{125}I 标记抗体加到含 3% 牛血清白蛋白和 10% 载体血清的盐溶液中 (0.9% NaCl 和 10 mmol/L Tris-HCl pH 7.4), 每毫升中含抗体的量大约为 10^6c/m , 再加入 0.01% 的 NaN_3 。然后把膜浸入上述反应液中, 100cm^2 膜用 3ml 左右的反应液, 在室温反应 6—7 小时, 反应结束后用磷酸缓冲的生理盐水 (10 mmol/L 磷酸钠和 0.9% NaCl) 充分洗涤 (一般洗 4—5 次), 洗好后用吹风机吹干, 压好 X 光片, 置低温冰箱自显影。

第四章和第五章讲述了分子克隆技术的基本原理和方法。其主要过程可用图 5.1 的简图表示。

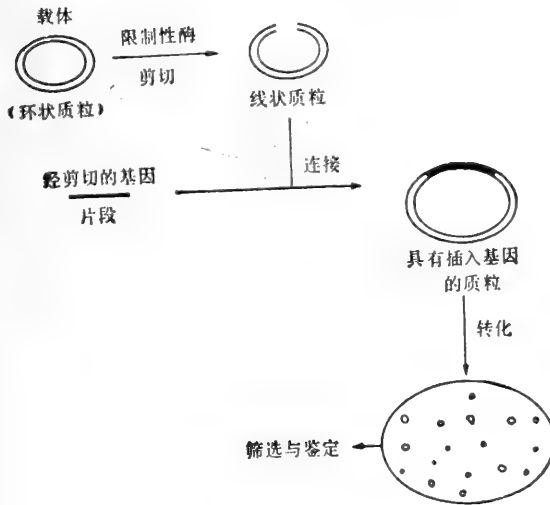


图 5.1 基因克隆简图。

参 考 文 献

- [1] Norhard, M. V. *et al.*, Factors affecting the transformation of *Escherichia coli*, strain X1776 by pBR322 plasmid DNA, *Gene*, **3**, 279, 1978.
- [2] Maniatis, T. *et al.*, Introduction of plasmid and Bacteriophage λ DNA into *Escherichia coli*, *Molecular cloning*, 247—254, 1982.
- [3] Michael, G. and John, W., Colony hybridization, *Methods in enzymology*, **63**, 379, 1979.
- [4] Benton, W. D. and Davis, R. W., Screening λ gt recombinant clones by hybridization to single plaques in situ, *Science*, **196**, 180, 1977.
- [5] David, G. S. *et al.*, Protein iodination with solid state lactoperoxidase, *Biochemistry*, **13**(5), 1014, 1974.
- [6] Louise, C. and John, C. A., Sensitive and rapid method for recombinant phage screening, *Methods in enzymology* **68**, 389, 1979.
- [7] David, J. S. *et al.*, Isolation of hen ovikuct ovalbumin and rat liver albumin polysomes by indirect immunoprecipitation, *The Journal of Biological Chemistry*, **249**(12), 3665, 1974.

- [8] Louise, C. *et al.*, Selection of specific clones from colong banks by screening with radioactive antibody, *Methods in enzymology*, **68**, 436, 1979.
- [9] Maniatis, T. *et al.*, Rapid isolation of Plasmid or Bacteriophage λ DNA, *Molecular cloning*, 365-371, 1982.
- [10] Towbin, H. *et al.*, Electrophretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and applications, *Proc. Natl. Acad. Sci. USA*, **76(9)**, 4350, 1979.

01 10 1982

第六章 基因文库

强 伯 勤

第一节 引言

分子克隆技术是基因工程的核心,同时,它为从分子水平研究生物体基因的结构、表达和调控提供了必要的技术。进行分子克隆的首要工作是分离所需的特定基因片段。

一、基因的分离

基因实质上就是染色体 DNA 上一段特定的核苷酸序列,这段序列包含了细胞中某种蛋白质或酶或 RNA 的完整的遗传密码。分离基因存在许多困难。首先是基因种类繁多。简单生命物质——DNA 病毒就含有近十个至数百个基因,原核生物有几千个,真核细胞的基因数目则有上万至几万个,而人类染色体中可达 10^6 个左右的基因。如此众多的基因,其化学结构均由腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)和胸腺嘧啶(T)四种核苷酸组成,使彼此在理化性质上十分相近,因此,采用一般的物理、化学方法不易将它们分开。其次是基因本身结构的复杂性。对于真核细胞来说,结构基因不仅包含了全部编码区,还包括了间隔于编码区之间的非编码区(插入序列,即内含子),一个结构基因内可存在几个以上的内含子,从而使基因结构变得复杂、长度增大。除结构基因外,还有控制和调节基因表达的序列。造成基因分离困难的第三个因素是细胞内每种基因的含量极少。不论是原核还是真核细胞,大多数结构基因都是单拷贝,例如珠蛋白 β 链有146个氨基酸,其基因密码相当于约500bp长度,它在细胞染色体DNA的 2.9×10^9 bp中所占比例极小(1.7×10^{-7})。最后的困难则是如何检测被分离的基因,证明它确实是所需要的基因。

尽管如此,近年来还是提出了一些分离某些特殊基因的方法。例如,利用氯化铯密度梯度平衡超速离心法把G-C配对含量高的基因和G-C含量低的基因片段分开,有人曾用此法分离了爪蟾中的5S DNA。再如,大肠杆菌乳糖操纵子的分离是用携带乳糖操纵子方向相反的两特定噬菌体完成的。当把来自此两种噬菌体的DNA单链进行退火时,两条链上的操纵子部分的核苷酸重新以氢键组成双链,而DNA链的其余部分仍为松散的单链,它们可用单链核酸酶切除,留下的双链部分即是乳糖操纵子的基因。在真核细胞染色体DNA中,rDNA(核糖核蛋白体RNA的基因有数百次重复)、5SDNA(5S RNA基因)、tRNA基因和组蛋白基因均为多拷贝基因,分离就相对容易些。

为提供基因工程用的目的基因,真核细胞基因分离主要选用反转录的方法,即本书第七章介绍的cDNA的制备。首先分离真核细胞的mRNA,借反转录酶合成互补于mRNA的DNA拷贝(称之cDNA)。早期,人们从总mRNA中分离纯化或富集某特异

mRNA, 再借反转录酶合成 cDNA, 可是, 特异性 mRNA 含量往往很低, 极难分离纯化到足够的量, 因此, 现在大多采用 cDNA 库的方法, 即用总的 mRNA 反转录成 cDNA, 把双链的 cDNA 克隆到合适载体之中组成 cDNA 库, 然后从 cDNA 库中再筛选出所需特异基因。此方法分离到的真核细胞特异基因就不含有原来结构基因中的非编码序列, 所以可直接提供基因工程表达研究, 生产相应的蛋白产物。

人工合成基因是为分子克隆提供外源基因的又一重要途径。DNA 序列测定揭示了基因的一级结构; 蛋白质纯化和氨基酸顺序分析技术的发展, 可测出某蛋白质肽链的氨基酸序列, 这些均为基因的人工合成提供了必要的信息。尤其是 DNA 化学合成技术的迅速发展、固相合成方法的出现和 DNA 合成仪的诞生, 使之有可能在短时间内合成某特定的基因片段。在一些多肽类激素的基因工程研制中, 不少外源基因是由化学合成或部分化学合成得到的。譬如生长激素释放抑制因子 (somatostatin) 在大肠杆菌中克隆和表达用的基因, 就是由八段人工合成的寡核苷酸经连接酶作用拼接成功的。人生长激素克隆用的基因则由合成片段(与激素中 1—24 位氨基酸相应的核苷酸片段)和双链 cDNA (相应于 24—191 位氨基酸序列)连接组成的。

基因都定位于细胞的染色体上, 所以染色体 DNA 是分离基因的主要材料。早期, 采用限制性内切酶或者机械方法(例如超声波处理)将染色体 DNA 打断, 通过 T4 DNA 连接酶把片段与质粒 DNA 重组, 并把重组体 DNA 转入相应受体菌中进行无性繁殖, 得到的是一大群含不同外源 DNA 片段的克隆, 这就是所谓的鸟枪法 (shot gun)。然后以放射性标记的纯 mRNA 或 DNA 片段作探针, 与菌落进行原位杂交选出所需的特异基因。或者可利用特定基因营养缺陷型受体菌的遗传标记筛选方法分离相应基因。可是质粒 DNA 可容纳的外源 DNA 的长度有限, 超过 7—10kb 的基因片段的重组与克隆就较困难。随着噬菌体载体的应用, 使得可插入的外源 DNA 片段长度大为增加, 最大限量在 22kb。尤其是考斯质粒 (cosmid) 载体, 它可携带长度为 45kb 的外源 DNA。在此基础上发展了构建基因文库的技术, 从基因文库中分离和筛选同一生物体内的各种基因片段。

二、什么是基因文库

基因文库是指含有某种生物基因组中不同基因片段的 DNA 重组体的一群克隆。把原核或真核细胞的染色体 DNA (即基因组, genome) 切割成一定大小的片段, 将它们和合适的载体(例如质粒或 λ 噬菌体)进行重组, 可得到一系列的 DNA 重组体。当以质粒为载体时, 重组的 DNA 分子可被引入宿主细胞中, 形成一大群各含不同 DNA 重组分子的转化菌; 若采用 λ 噬菌体 DNA 作载体, 重组的分子可在体外包装成一群重组的噬菌体颗粒。虽然一个克隆只含基因组内的一个片段, 但只要切割基因组是随机的, 而且得到的重组体克隆数目足够大, 那么构建成的克隆群体就可能包含基因组内全部基因片段, 成为基因组的代表。因此, 这样的克隆群体犹如一个贮存有基因组内各种基因的库, 故称为基因文库, 英文为 genomic library, 也称 gene library 或 gene bank。

对于原核细胞和低等真核细胞来说, 基因结构比较简单, 它们的基因文库就可以作为直接提供基因工程所需的各种有兴趣基因的重要来源。而较高等的真核细胞, 尤其是人

和哺乳动物细胞,基因结构复杂。一个结构基因往往可被几个以上的插入序列所间隔,因此由基因文库中分离得到的基因片段很难直接用于体外的表达研究。但是它们对真核细胞基因结构的分析、基因表达和调控的研究有着重要作用。尤其在染色体步移研究中,基因文库更为重要。

一个理想的基因文库应该是: ① 构建成的克隆群体包含完整基因组的全部 DNA 序列。这就要求在打断基因组 DNA 时尽可能做到随机切割,并且提供克隆用的片段要有足够的长度,使其含有整个基因以及它们的侧翼序列。通过对外源基因序列有部分重叠的某些克隆株的分析,可呈显出基因的全部序列以及邻近区域,从而有助于选择合适克隆,切出感兴趣基因的全部序列,同时群体的克隆数目要足够,保证基因组中各基因序列全部被重组于克隆群体中。② 构建基因文库所需的起始材料用量要少。这样就使那些不易取得的材料如肿瘤组织等有可能被建成基因文库,以进行其基因结构和功能的研究。③ 建成的基因文库可被扩增,并在扩增过程中不会导致重组体的丢失或产生不正常的新重组体。基因文库要稳定,便于贮存,不因保存几年而引起克隆的明显丢失。只要在操作中载体和宿主细胞选择恰当,便可以达到这一目的。④ 基因文库的克隆数目要在一个便于筛选的数量级范围内。克隆量过大将会使筛选增加很多工作量,又要花费过多的财力、物力。噬菌体 λ 的替代型载体或考斯质粒的选用,使插入载体的 DNA 片段长达 22kb, 以至 45kb, 这样相应地减少了构建一个完整的基因文库所需达到的克隆数目。考斯质粒的最大载量虽为 45kb 左右,但由于使用时它比 λ 噬菌体的替代型载体困难,所以构建基因文库时更多地选用噬菌体的替代型载体。

三、基因文库构建的基本步骤

目前较多地选用 λ 噬菌体的替代型载体来构建基因文库,图 6.1 介绍了使用这样的载体构建基因文库的基本步骤:

(1) 首先,制备染色体(基因组)DNA,并用合适的限制酶(如 *Sau* 3A1)进行部分降解和分离合适大小的插入片段(18—22kb),以提供重组用的外源 DNA 片段;当然也可采用机械切割的方法(如超声波破碎)来降解染色体 DNA,而且这种方法切割 DNA 随机的程度更高,但是产物要经过末端修饰,使之含有与载体片段相对应的粘末端结构才能与载体重组;

(2) 其次,选择和处理载体。图中介绍的是经过人工改建过的 λ 噬菌体的替代型载体(如 *EMBL3*),它有两个 *Bam*HI 的切点,用酶切割之后就分为三段:左臂、右臂和中央片段。左、右两臂包含了感染的噬菌体颗粒繁殖所必需的全部基因。中央片段又称 *Stuffer*,含有的基因并非噬菌体生存所必需的,它在一定程度上起着支撑载体 DNA 分子特定长度的作用,可被相当长度的外源 DNA 所取代。线性载体 DNA 分子的左、右两端的 5' 位各有长度为 12 个核苷酸的单链突出,两单链的核苷酸序列是互补的,因此可借氢键把线性末端粘合一起,此末端结构称之粘末端(*cohesive end*),因此该部位序列被命名为 *cos* 位点。它对于 λ 噬菌体是很重要的,能使左、右两臂互相粘合而形成串联状的 DNA 线性分子,提供了重组 λ DNA 分子在体外包装成噬菌体颗粒所需的分子结构。噬菌体载体的结构已在第二章第二节中详细介绍。在使用替代型载体时,重要的是除去中

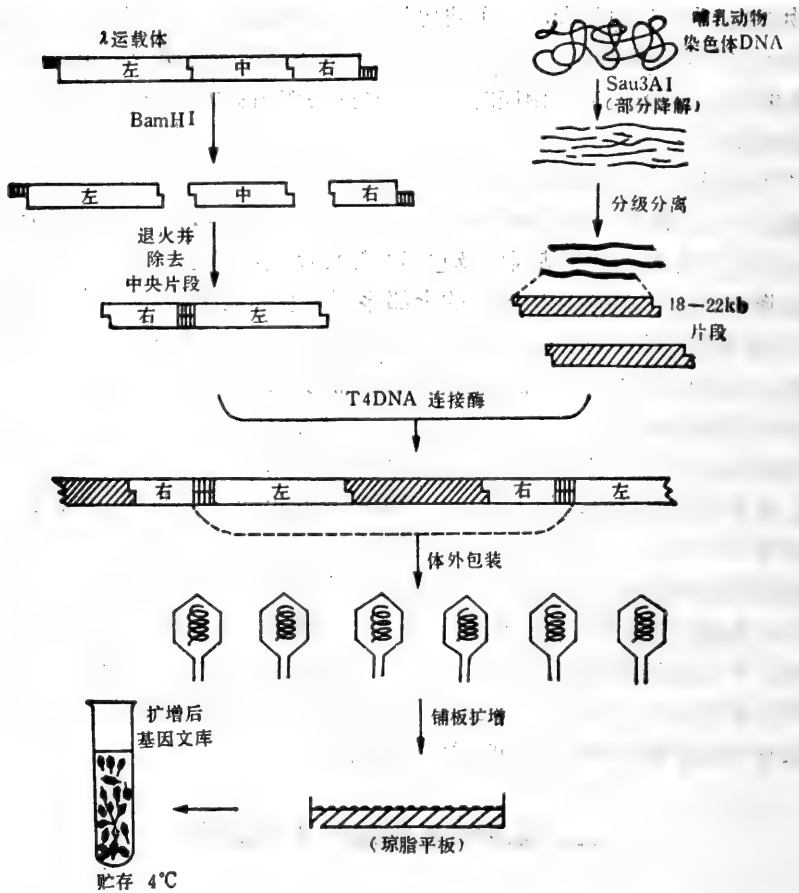


图 6.1 基因文库构建示意图。载体中标记“左、右和中”分别表示 λ 噬菌体 DNA 的左臂、右臂和中央片段。

央片段，或者经酶促反应使它不能与两臂自身相连，以提高两臂与外源 DNA 重组的效率；

(3) 第三步是将载体两臂与外源 DNA 片段重组。这主要靠 T4 DNA 连接酶的作用。在连接反应中，适当的底物浓度以及插入片段和载体的适宜比例是得到高产量重组体和创造下一步体外包装时最适底物的重要条件。

(4) 重组 DNA 分子在体外包装成噬菌体颗粒，便可直接用于筛选感兴趣的基因。重组的 DNA 还可转染大肠杆菌，在细菌体内繁殖噬菌体颗粒。但转染的效率比较低，以每微克插入 DNA 的量计算在 10^4 pfu (噬菌斑形成单位) 左右。这样的数目对于组建基因组分子量较小的细菌或酵母类生物的基因文库尚可满足需要，但对构建含重组体百万种以上的高等真核细胞的基因文库来说，效率实在太低。1975 年 Becker 和 Gold 制备了噬菌体体外包装用抽提物，并成功地实现了 λ DNA 在体外包装成噬菌体颗粒的实验。实验表明，每微克 λ DNA (野生型) 体外包装效率为 10^7-10^8 pfu。重组的 DNA 分子包装效率也可达每微克 DNA 10^5-10^6 pfu，这比噬菌体 DNA 转染宿主细胞的效率大大提高了。这样，使真核细胞基因文库的构建既减少了工作量，又节省了大量的起始材

料。

(5) 基因文库的扩增和贮存。把体外包装得到的重组噬菌体感染合适的大肠杆菌宿主菌,可将构建的基因文库扩增。一般认为琼脂板的固体培养比液体培养的扩增要好。把在琼脂板上扩增的噬菌斑用噬菌体缓冲液洗脱,并低速离心除去未被感染的细菌,取出上清液,加入少许氯仿防止菌的繁殖,噬菌体就可于 4℃ 贮存。在扩增时要注意防止噬菌体克隆的丢失,这将在以后章节中讨论。

关于基因文库的构建, Maniatis, Fritsch 和 Sambrook 编著的《分子克隆——实验室手册》(参考文献 2)一书已有详细的叙述。读者可以参照手册中介绍的方法操作。本章仅将逐一讨论构建基因文库各步骤中的有关问题,以供读者参考。

第二节 供体 DNA

如上所述,构建基因文库的一般步骤中,首要的一步是从细胞或组织中抽提供体 DNA 以及制备克隆用的 DNA 片段。抽提的供体 DNA 的质量,以及制备方法,将影响所构建基因文库的代表性。

一、染色体 DNA 的抽提

病毒、细菌、低等真核生物以及哺乳动物和人的基因组的大小差别甚大,但是,不论什么来源的供体 DNA,在构建基因文库时都必须满足一个基本的要求:要尽可能保持抽提染色体 DNA 的高分子量。从细胞中分离染色体 DNA 时,由于机械切割作用不可避免地会有所断裂,但要尽可能地把断裂作用减到最低程度。建立真核细胞的基因文库,主要是采用替代型的 λ 噬菌体载体,插入的基因片段一般在 18—22kb 大小范围(14kb 以上片段也符合需要);若用考斯质粒作载体,插入的片段要求更长(30—45kb)。按我们的经验,抽提的染色体 DNA 制剂的片段长度必须超过 100kb。如果抽提的染色体 DNA 断裂成比较小的碎片(小于 50—100kb),当用限制酶(Sau 3AI)进行部分水解时,产生的合适大小范围的酶解片段时会有相当一部分片段的一端不持有 Sau 3AI 粘末端。这样的片段借 BamHI 位点和 λ 载体的两臂相连时,往往只有一端可与载体臂相连,另一臂不能连接。所以,不合格的供体 DNA 片段,犹如重组链的“中止者”,抑制着与载体的重组连接以及重组 DNA 的体外包装效率。所以说,高分子量染色体 DNA 的提供是构建理想的基因文库的重要基础,尤其是对于复杂结构的真核细胞的基因文库,因为它的完整文库需要包含有百万个以上的克隆。

(一) 酚法抽提染色体 DNA

这是制备 DNA 最常使用的方法。在细胞裂解之前,可以先向组织中加入裂解缓冲液并匀浆。缓冲液含有强的去污剂(SDS、Triton X-100 或 Sarcosyl),有益于 DNA 从细胞核中释放,同时抑制着内源性 DNase 活性。裂解后溶液变得粘稠,此后要尽可能减少机械切割。这种切割作用往往出现于制剂的剧烈混和、振摇的过程中。当通过注射器针头或者用吸管转移裂解液时,也会导致 DNA 分子的机械断裂。所以在进行酚或氯

仿-异戊醇抽提时, 要注意: ①裂解体积适当大些, 不要使裂解液太粘稠; ②酚抽提最好在一种耐酚和有机溶剂的 50ml 带盖塑料管中进行, 裂解液加等体积饱和酚后, 拧紧盖子, 靠慢慢地上下颠倒管子混匀两相, 切忌剧烈振摇; ③吸取粘稠的 DNA 溶液时, 选用口粗而光滑的吸管, 避免用口径细小的滴管; ④酚抽提后, 以透析法除酚为好, 尽可能不采用乙醇沉淀。

从培养细胞中制备高分子量 DNA 的方法基本上包括下述几步:

(1) 裂解细胞。把经过洗涤后的培养细胞悬浮于 TE 缓冲液中 (10mmol/L Tris-HCl, 1mmol/L EDTA, pH8.0), 使浓度为每毫升约含 10^8 细胞, 加入 10 倍体积的裂解溶液 (0.5mol/L EDTA, pH8.0), 100 μ g/ml 蛋白酶 K, 0.5% Sarcosyl)。混匀后, 于 50 $^{\circ}$ C 水浴裂解 3 小时。

(2) 酚抽提三次, 每次借离心分相。弃有机相和中间层, 取水相。在高盐浓度时, 往往有机相在上层, 必须注意区分。若采用含 8-羟基喹啉的酚液, 可由相的颜色 (酚相为黄色) 来区分有机相和水相。在两相的中间层主要是变性蛋白, 但常常夹有 DNA, 一般可用适量 TE 缓冲液反抽来收回 DNA。

(3) 透析除酚。透析液可用 50m mol/L Tris-HCl (pH8.0)、10m mol/L EDTA、10m mol/L NaCl。透析过程中会伴随样品体积的急剧增加 (为原来的 2—3 倍), 所以透析袋要留有一定空间。

(4) 去除 RNA。透析后样品用无 DNase 活力的 RNase 处理 (10—100 μ g 酶/ml 样品), 37 $^{\circ}$ C 保温 3 小时。然后用酚抽提两次, 灭活 RNase, 并对 TE 缓冲液透析。除尽酚后, 样品于 4 $^{\circ}$ C 贮存。

(二) 氯化铯平衡密度梯度离心法制备 DNA

这是借氯化铯密度梯度的平衡超速离心作用来分离高分子量的染色体 DNA, 方法简便, 容易操作, 减少了抽提 DNA 步骤, 也无需酚抽提和乙醇沉淀。它已被成功地应用于由细菌、真菌、果蝇、人工培养的真核细胞以及高等真核生物的组织中分离高分子量的 DNA。

(1) 裂解细胞。对于培养细胞可同酚抽提法。如果从组织中制备 DNA, 可先分离细胞核, 再裂解细胞核, 释放 DNA。譬如从肝脏组织中提取 DNA, 先把组织用液氮快速冰冻, 并在液氮存在下将组织磨成细粉, 再悬浮于冷的缓冲液中 [10m mol/L Tris-HCl (pH7.4)、60m mol/L NaCl, 10m mol/L EDTA, 0.15mmol/L 精胺, 0.15m mol/L 亚精胺; 0.5% (V/V) Triton-X100], 并匀浆。匀浆液经细纱布过滤后, 滤液在 7,000 r/min 速度下 4 $^{\circ}$ C 离心 7—10 分钟, 收集细胞核再悬浮于上述缓冲液中 (1—2g 组织在 10ml 体积左右), 并转入带盖的塑料管内, 加 10% Sarkosyl 至终浓度为 2%, 50 $^{\circ}$ C 水浴保温 60 分钟, 可完全裂解细胞核。

(2) 氯化铯密度梯度离心。将裂解液称重, 按每克裂解液加 1.25g 固体氯化铯加入氯化铯。加盖后, 轻轻来回颠倒管子使固体盐全部溶解, 于离心管中室温 (20 $^{\circ}$ C) 超速离心。离心条件为:

Beckman SW41 离心头	39,000 r/min	48 小时, 或者
Beckman 65 型离心头	55,000 r/min	24 小时

Beckman VTi50 离心头 45,000 r/min 16 小时

离心毕, RNA 沉于管底, DNA 区带处于梯度的中间部位。为避开 RNA, 在管底靠上一点的位置插一针头, 刺破管壁, 使溶液流出, 以 0.5—1.0ml 分部收集样品。当有 DNA 流出时, 溶液会变粘稠。分别从每管中取 2—5 μ l 样品, 用 0.3—0.5% 琼脂糖电泳以确定 DNA 峰管, 且合并一起。

一般经一次离心, 所得 DNA 已有足够纯度, 如果需要更纯更浓样品, 可把合并的 DNA 样品稀释到含氯化铯 5.5% 浓度, 再重复超速离心一次。

(3) 透析。样品经 4 $^{\circ}$ C 下对 TE 缓冲液充分透析后, 可贮存在消毒试管中于 4 $^{\circ}$ C 保存。若透析后 DNA 样品太稀, 还可用等体积正丁醇抽提, 得到合适浓度后, 再透析除去残存的丁醇。

按上述步骤制备的 DNA, 分子量很高, 制剂完全可以满足构建基因文库的要求。如果认为还需进行蛋白酶 K 处理和酚抽提, 则所用蛋白酶 K 必须是无 DNase 活性的, 而且酚抽提时一定要操作缓和, 以防 DNA 断裂。

(三) DNA 制剂的分子大小检测

样品 (50—100ng 左右) 在 0.3—0.5% 琼脂糖凝胶板上, 低电压 (2—5V/cm) 下电泳, 同时用 λ DNA (48.5kb) 和 T4DNA (165kb 左右) 作标准, 进行比较。注意电泳时要控制 DNA 的量 (50—100ng 左右为好), 过多的 DNA 使电泳后荧光区带变宽, 难以相互比较。

二、染色体 DNA 的切割

在基因文库构建中, 为了保证生物体基因组的全部序列均可分段地克隆于 DNA 重组体的群体之中, 在制备插入的 DNA 片段时, 要尽可能做到随机切割基因组 DNA。一般认为, 机械性切割随机率最大, 但机械力打断的 DNA 片段, 其末端往往带有一个短的单链尾巴。此单链的长短并不均一, 其核苷酸序列也不能互补, 因此被切割的片段不宜直接与载体重组连接, 需要把末端处理成平末端, 并加上合适的寡核苷酸连接子, 提供特定的限制性内切酶粘末端后才行。否则, 会严重影响连接效率。较多的人主张采用限制性内切酶的部分降解来制备插入片段。限制性内切酶有特定的识别位点, 但用四核苷酸识别特异性的酶 (例 AluI、HaeIII、MboI、和 Sau 3AI 等) 切割随机排列的 DNA 分子时, 理论上为每 256 个碱基即有一个切口 (1/4⁴)。哺乳动物细胞染色体 DNA 含有 3 \times 10⁹bp, 要从中切割出大小为 2 \times 10⁴bp 的一整套有部分重叠序列的片段, 采用这种四核苷酸识别特异性的酶基本上满足了切割的随机性要求。而且选用象 Sau 3AI 或 MboI 酶的部分降解供体 DNA 产生的片段持有 pGATC 的 5' 粘末端, 因此可直接插入载体的 BamHI 位点, 而无需其他修饰或处理。下面将介绍用 Sau3AI 部分切割供体 DNA 以及降解产物片段大小的分级方法。

(一) 用 Sau 3AI 部分降解 DNA

在应用 λ 噬菌体改建的载体中, 采用比较多的是带有 BamHI 插入位点的载体, 例

Charon 28、 λ 1059, 新近改建成的 λ EMBL3、L4 和 Charon 34、35 等等。因此, 一些能产生与 BamHI 相同粘末端结构的四核苷酸识别序列的酶如 MboI、Sau3AI 可被选用于降解供体 DNA。Sau3AI 识别序列和切割的末端与 MboI 相同 (GATC); 但当 GATC 序列中的 A 甲基化后, MboI 不能再识别, 而 Sau3AI 则不受影响。因此用 Sau3AI 替代 MboI 进行 DNA 部分降解, 随机率就更大, 故目前更多选用 Sau3AI 进行部分降解, 其步骤如下。

1. 部分降解反应条件的测试

(1) 取 20 μ g 染色体 DNA, 加 40 μ l 5 \times sau3AI 缓冲液, 并补水至终体积 200 μ l。充分混合, 并在台式高速离心机离心 5 秒钟, 使 DNA 溶液全部集中于管底。然后取 0.5 ml 或 1.5ml 尖底塑料管 (Eppendorf 管) 9 支, 并编号从 1 至 9。在第 1 管中加入 40 μ l DNA 溶液, 其余各为 20 μ l。剩余部分转入第 9 号管, 作为不加酶的对照管。反应管置于冰浴中。

(2) 在第 1 号管中加入 4u 的 Sau3AI (1u/ μ g DNA), 轻轻充分混匀。从中取出 20 μ l 快速转移至第 2 管中, 依次转移并混合均匀, 直至第 8 管, 其终体积将为 40 μ l。第 9 管则不加酶。如此转移使每反应管中酶量对倍下降, 第 1 管为 1u/ μ g DNA/20 μ l, 第 2 管稀释为 0.5u, 至第 8 管中酶量将为原浓度的 1/256。

转移稀释酶量后, 同时将 1 至 9 管置 37 $^{\circ}$ C 水浴保温 60 分钟, 并立即转至冰浴中, 向各管内加入 5 μ l 100m mol/L EDTA, 以停止反应。

(3) 用 0.5% 琼脂糖凝胶电泳检查酶解反应。取各反应产物 10 μ l, 分别加入甘油-溴酚蓝染液 3 μ l 后, 电泳 (2V/cm)。电泳时用 DNA 分子量标准作对照。常用的分子量标准为 λ DNA/HindIII 降解物, 其片段分子量为 23.1、9.4、6.5、4.3、2.3、2.0 和 0.56kb。如果再选用 λ DNA/sacII 降解物 (有片段 20.3、18.8、8.1、1.1 和 0.2kb) 和 λ DNA/pvuII 降解物 (片段为 14.3、12.7、11.9 和 9.5kb), 可提供 14—20kb 的更精确范围。凝胶经溴化乙锭染色后, 即可在紫外灯 (254nm) 下观察结果并照相。

(4) 适宜酶浓度的选择。由所得照相结果, 遮盖 >22kb 和 <18kb 的区域, 观察 22—18kb 区域间产量最大的反应管号, 得到相应的酶浓度, 然后选择此浓度的二分之一量为大量制备时的酶量。另一种选择酶量方法是, 在确定 22—18kb 片段产生量最大的反应管号后, 选用此管号及其前后相连管号的浓度。用这三种酶浓度分别水解 DNA, 然后合并产物再进行分级, 这样可避免观察时造成的误差。

同时, 我们知道 DNA 链上酶识别位点周围的结构会影响酶解速度, 即在相同的反应条件下, 酶切割各位点的速度不一致。从上述测试酶浓度的实验中, 将发现此现象。在较大的酶量时, 总有一小部分产物处于较大分子量的范围内, 这是属于不易被酶降解的部分; 同样在很小酶量下, 极大部分 DNA 尚未被水解, 但已有小部分被水解, 这一部分称之为 DNA 链中对酶敏感的部分。为了保证降解作用尽可能地随机, 即保证在所取的一定大小片段 (18—22kb) 中包括 DNA 链上的这些不同反应速率的序列, 建议在供体 DNA 量足够的条件下, 多取几种酶浓度来进行部分降解。至少可以把在做测试酶浓度实验时剩余的反应物合并于大量级制备的产物中, 一起进行片段的分级, 以提高酶解作用的随机率。

(5) 大量级的部分降解。在保持与测试实验所选择的酶浓度、反应时间、温度以及

DNA 量尽量一致的条件下,进行大量级的酶解反应。在 1.5ml 带盖的尖底塑料管中含 DNA 50—80 μg ,反应体积相应地在 500—800 μl ,共取三份,每份加入预先选择好的酶量进行部分降解(37 $^{\circ}\text{C}$ 60 分钟),准时加 0.5mol/L EDTA(pH8.0) 至终浓度 20mmol/L 以终止反应。从每反应管中各取 10 μl 反应液进行琼脂糖凝胶电泳检查酶解程度,其余的反应产物可放冰浴中存放。

(6) 合并酶解产物。用等体积酚和氯仿-异戊醇(24:1V/V)的混合物(1:1V/V)抽提两次。水相加 NaCl 至 0.3mol/L 后,加入两倍体积的乙醇沉淀。回收沉淀后,再悬浮于 200—400 μl 的 TE 缓冲液中。

在合并酶解产物时,可把做测试酶浓度时剩余的反应产物一起合并回收。

(二) 反应产物的大小分级

常用的是蔗糖密度梯度离心法,也可以选用氯化钠梯度离心法。后者的优点是可省去去除蔗糖的步骤,而且离心时间也大为缩短。

1. 蔗糖密度梯度离心法

参考文献 2 的第 275 页介绍了蔗糖密度梯度离心法的操作。下面简述另一方法。

在适用于 Beckman SW50.1 离心头的离心管中制备总体积为 5ml 的蔗糖密度梯度,即 5—25% 蔗糖的线性梯度,缓冲液为 0.1mol/L NaCl、20 mmol/L Tris-HCl (pH 7.5)、2 mmol/L EDTA。加入 DNA 的 *Sau3AI* 部分降解产物 50—100 μg ,体积不超过 200 μl 。加样后,4 $^{\circ}\text{C}$ 下 20,000r/min 离心 18 小时。小心取出离心管,借管底扎洞或从管底开始的虹吸方法,分部收集(100 μl /管)。

每隔三管,取 5 μl 样品,加 10 μl TE 缓冲液稀释后进行电泳(0.5% 琼脂糖凝胶),并用 λ DNA/*HindIII* 降解物作分子量标准,如条件可能,还可用 λ DNA/*pvuI* 和 λ DNA/*sacII* 两种酶解物作标准,以提供 18—22kb 或 14—20kb 的更精确部位。由电泳结果把合适大小 DNA 片段(18—22kb 或 14—20kb)合并,加等体积 TE 缓冲液(稀释蔗糖浓度)后,再加 1/10 体积 3mol/L NaAc 和 2.5 倍体积的 95% 乙醇, -20 $^{\circ}\text{C}$ 2 小时(或 -70 $^{\circ}\text{C}$ 30 分钟)后,离心,沉淀经用 70% 乙醇洗涤,然后减压抽干,并悬浮于 TE 缓冲液(10—20 μl 左右)。取 0.5 μl 样品,加 10 μl TE 稀释,再电泳,进一步观察合并产物的分子量大小。若合适,即可以供构建基因文库,样品于 -20 $^{\circ}\text{C}$ 保存、备用。

2. 氯化钠密度梯度离心法

当采用 Beckman SW41 离心头时,借梯度形成仪在 12.5ml 离心管中制备容积为 11ml 的 5—25%(W/V) NaCl 梯度。NaCl 中含有 3mmol/L EDTA, pH8.0。注意在形成梯度时速度应维持在每分钟 1ml。也可用 5ml 的管子制备 4.5ml 的梯度,用 Beckman SW50.1 头离心。

11ml 梯度上铺 200—250 μl 的 DNA 样品(含 DNA 约 150 μg),注意样品体积和 DNA 量不宜过大,否则影响分辨率。将梯度装入 SW41 头,室温离心 37,000 r/min, 4.5 小时。然后分部收集离心溶液,每 0.25ml 一管。每隔 2—3 管取样品 20 μl ,经 2—3 倍稀释后,用 0.5% 琼脂糖凝胶电泳检测分级的样品,由分子量标准来衡量 DNA 片段大小。其余样品管于 -20 $^{\circ}\text{C}$ 冰冻保存。在以上各步操作中,要注意清洁,尽可能无菌,以防止杂酶污染。

合并分子量在 18—22kb(或者 14—20kb 的部分降解片段)范围的样品,加等体积无菌水稀释浓度后,加两倍体积乙醇沉淀 DNA,沉淀用 70% 冷乙醇洗两次,尽可能除去过多的盐,每步均需伴随离心收集沉淀。最后,减压抽气除乙醇,样品溶于 20—50 μ l 消毒的 TE 缓冲液中。

(三) DNA 含量测定

在将供体 DNA 片段与载体重组时,载体和供体 DNA 分子的合适比例以及反应浓度可提高重组效率。所以在进行重组前要得知 DNA 的含量。由于制备得到的插入用 DNA 片段量很少,无法测得绝对量,但可按下述方法测出大致含量,供重组时参考。

一种方法是取 1 μ l 制得的 DNA 样品,用 TE 稀释至 20 μ l,加于 0.5% 琼脂糖凝胶中电泳。同时用相同体积 (20 μ l) 含不同量 λ DNA (例 1、0.75、0.5、0.25 和 0.1 μ g) 标准作对照,在含溴乙锭 (0.5 μ g/ml) 缓冲液下经短时间电泳 (1 小时) 后,进行紫外照相,照相时选用 2 至 3 种不同曝光时间 (注意避免曝光时间过长)。通过相片中样品区带与标准 DNA 区带比较,可估计供体 DNA 的含量。如果电泳时间过长,供体 DNA 样品区带变宽,就不易估计了。

另一种方法是用含溴乙锭 (1 μ g/ml) 的缓冲液稀释样品,配制每 20 μ l 体积中含 1、0.75、0.5、0.2 和 0.1 μ g λ DNA 的标准溶液,点于蜡膜 (parafilm) 上,同时取 1 μ l 供体 DNA 样品,稀释到同样体积 (20 μ l),并排点于纸上,在 254nm 波长的紫外灯下照相,取二或三种不同时间的曝光条件,由所得照片来比较供体 DNA 和标准 DNA 所呈现的斑点亮度,可大致测出样品的 DNA 含量。

测定了含量后的样品,贮存于 1.5ml 或 0.5ml 的带盖尖底塑料管中,用蜡膜封口,于 -20 $^{\circ}$ C 保存,防止染菌和蒸发。若无菌条件控制得好,可在 4 $^{\circ}$ C 保存。

(四) 插入考斯质粒的外源 DNA 片段的制备

载体考斯质粒的一个优点是容纳外源 DNA 片段的能力比替代型的 λ 噬菌体载体还大,其最大的容纳量可达 45kb。因此,在构建基因文库时,选用考斯质粒作载体,全文库的总克隆数目可以大为减少,有益于重组和筛选。但是,在构建文库中,对供体 DNA 制剂的要求甚高,其 DNA 的分子量必须大于 100kb,而且越大越好。这就需要在从细胞或组织中抽提 DNA 时特别小心,尽可能减少染色体 DNA 链的机械性断裂或其他因素导致的降解。制备插入载体用 DNA 片段与上面介绍的 sau3A1 或 MboI 的部分降解法相同,但在进行降解产物的分级时,要收集的片段大小为 30—45kb。测量产物大小所用分子量标准可选用野生型 λ 噬菌体 DNA (48.5kb) 及其合适限制酶的降解产物,如 λ DNA 的 NarI 降解物(最大片段为 45.7kb),XhoI 降解物(最大片段 33.5kb)。也可用噬菌体 T7DNA (39.9kb)、腺病毒-2 型 (Ad2) DNA (35.9kb),或者 pBR322 合适大小的聚合体作分子量标准。酶解片段的分级,较多采用琼脂糖凝胶电泳的方法。

第三节 载体 DNA

供 DNA 重组用的克隆载体种类较多,有质粒、噬菌体和考斯质粒,还可选用 DNA

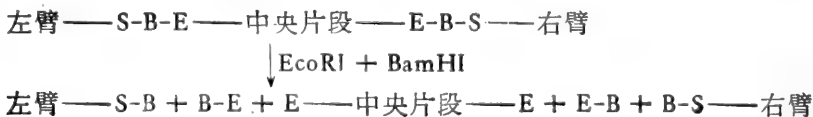
病毒。当构建基因文库时，尤其是真核细胞的文库，要求载体可容纳大的 DNA 片段，18—22kb，甚至 30—45kb 的外源 DNA 片段。由于考斯质粒使用上的困难，它的应用不够普遍，因此，经常选用替代型 λ 噬菌体作载体。本节将重点介绍此类载体的使用。

一、 λ 替代型载体的结构特点

本部分第二章介绍基因工程载体系列时，已就 λ 噬菌体的生物学及作载体用的噬菌体 DNA 的人工改建作了简要叙述。野生型 λ 噬菌体基因组是双链线性 DNA 分子，全长为 48,502bp，其中约 40% 的核苷酸序列是噬菌体生存非必需的，可被外源 DNA 片段取代，这是 λ 噬菌体 DNA 可以人工改建成替代型载体的基础。图 6.2 以 EMBL3 为例展示了 λ 噬菌体替代型载体分子的基本结构。载体包括了左、右两臂和中央片段（或称 *stuffer*）三个部分。左臂有密码噬菌体头部和尾部蛋白的全部基因；右臂包含 DNA 链复制起始点、主要的转录启动子和其他必需基因。两臂的末端称为 *cos* 位点，即各带有十二核苷酸长度的单链末端，两单链序列的碱基是互补的，末端单链粘合一起时， λ DNA 就呈环形分子。在两臂之间为中央片段。它不含噬菌体颗粒存活所必需的基因，但维持了基因组被包装成噬菌体颗粒时所需的足够长度。这部分片段可被合适长度的外源 DNA 片段所替代。

在载体 λ EMBL3 中，两臂和中央片段的相连是借两个序列相同而方向相反、长度短于 40bp 的寡核苷酸片段（图 6.2b），此片段含有 *Sall*、*Bam*HI 及 *Eco*RI 的识别序列。而在构建 λ EMBL3 时，已把噬菌体 DNA 两臂中的这三种酶的识别序列都除去了，只在中央片段还保留两个 *sall* 切口（图 6.2a），因此可以借助这三种酶的位点，除去中央片段，插入外源 DNA。其中，普遍采纳的是 *Bam*HI 的位点。因为识别序列为六核苷酸（GGATCC）的 *Bam*HI 酶切割双链 DNA 分子中的识别序列后产生 5' 位粘末端 GATC，这与能识别四核苷酸 GATC 的 *Sau*3AI 或 *Mbo*I 产生的粘末端结构完全一样，因此 *Sau*3AI 或 *Mbo*I 的降解作用产生的片段可以直接插入 *Bam*HI 位点。虽然选用其他酶的插入位点（*Sall*、*Eco*RI）同样可以构建重组 DNA 克隆，但是当用同样的酶（六核苷酸识别序列）去切割供体 DNA 时，从理论上分析，对随机排列的 DNA 分子来说，每 4,096bp 长度才出现一个切口，因此六核苷酸识别序列的酶切割 DNA 的随机率远低于四核苷酸的酶，这对构建基因文库来说是不适宜的，若只要克隆包含某些特定基因的 DNA 片段，并非构建完整的基因文库，则可以随研究者需要来选择合适的酶插入位点。在 λ EMBL3 载体出现之前，已有各种类型的替代型载体（见第二章第二节介绍的），例如 λ gtWes- λ B(*Eco*RI 位点)，Charon 4A(*Eco*RI 位点)，Charon 17(*Hind* III 位点)，Charon 28(*Bam* HI 位点)， λ 1059(*Bam*HI 位点)等等。

λ EMBL3 载体构建时，在两臂与中央片段相连处加入的多连接子（polylinker）序列，使同一部位具有三种限制酶的识别序列，它们是 *Sall*(S)、*Bam*HI(B) 和 *Eco*RI(E)，其结构如下：



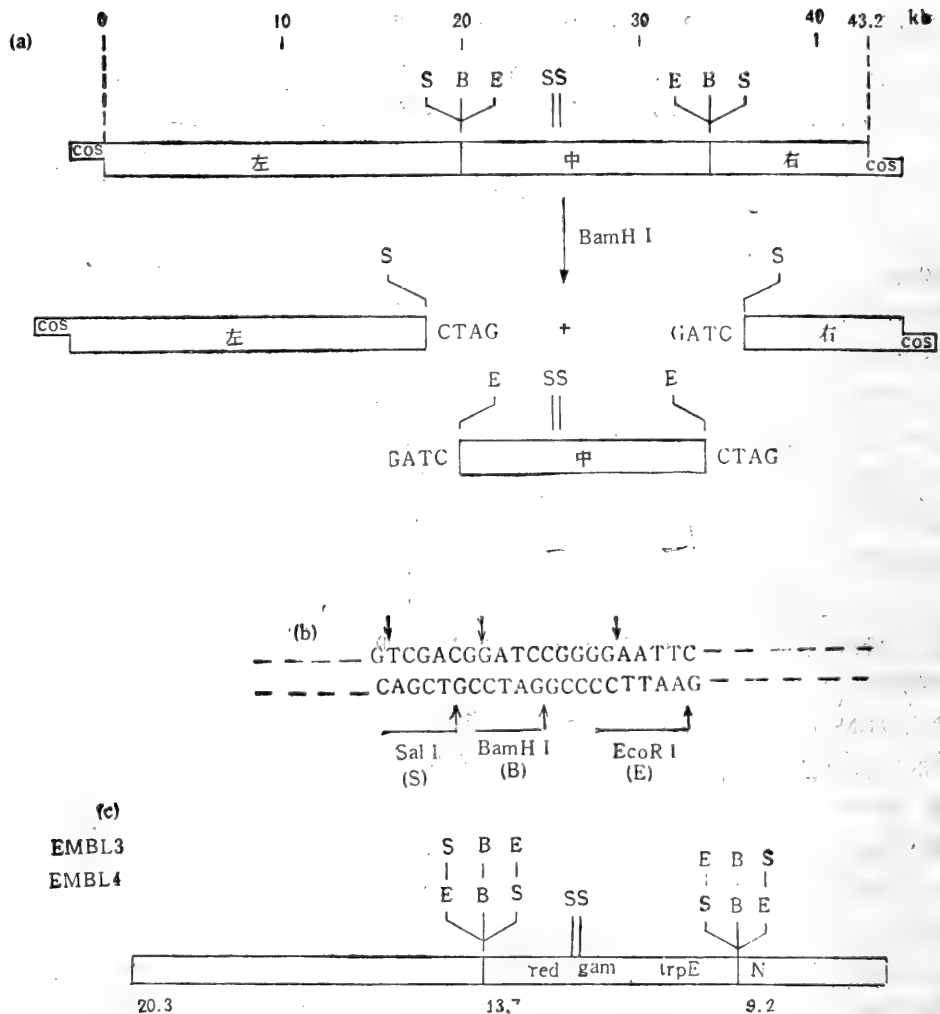


图 6.2 替代型载体 λ EMBL3 DNA。

(a) 为 λ EMBL3 DNA 结构示意图, 分子长度为 43.2kb, “左”、“右”分别表示噬菌体 DNA 的左臂和右臂, “中”为中央片段。S、B、E 分别代表 SalI、BamHI、EcoRI 的酶切位点。用 BamHI 降解载体分为三段: “左”、“右”和 “中”。(b) 展示了含有 S、B、E 三种酶识别序列的多连接器 (polylinker) 的核苷酸序列。(c) λ EMBL 3 和 L4 两种载体中多连接子的酶切点序列。red、gam 和 N 是 λ DNA 的三个基因, trpE 基因来自大肠杆菌。

其优点是①当选用 BamHI 位点插入外源 DNA 片段时, 为制备噬菌体的两臂, 可选用 BamHI 和 EcoRI 的双酶降解, 结果是两臂持有 BamHI 粘末端, 而中央片段则为 EcoRI 末端, 同时还出现两个长度相等、序列一致的寡核苷酸片段 (B-E), 它们一端为 BamHI 粘末端, 另一端是 EcoRI, 只要采用简便的异丙醇沉淀方法 (见下述), 即可把寡核苷酸产物除去, 这样噬菌体 DNA 的两臂就不可能和中央片段相连, 所以在噬菌体两臂与外源 DNA 重组时, 中央片段干扰就少, 无需将片段除去, 而可直接使载体与外源 DNA 重组。往常, 选用一种酶 (BamHI) 切割载体, 结果两臂的一端与中央片段有相同的粘末端, 若在重组前, 不先分离除去中央片段, 那末在与外源 DNA 进行重组时, 由于两臂与中央片段的自身连接作用而大大降低两臂与外源 DNA 的重组效率。②由重组子制

备插入片段时,可以选用 Sall 切割重组 DNA 分子,以保证切下的酶插入片段中不含噬菌体两臂的任何序列,这对于进一步的次级克隆或用插入片段作探针时,均去除了载体序列的干扰。如果采用 BamHI 切割重组体 DNA,往往不能把插入片段全部切割下来,因为重组时外源片段由 Sma3AI 降解获得,其末端结构为 5'pGATCN-3'。就理论上讲,
N-5'

对随机排列 DNA 分子,仅有四分之一机率会出现 BamHI 的末端结构: 5'pGATCC-3'
G-5',

也只有出现此结构时, BamHI 才能把重组的位点切开。若选用其它酶时,常常在载体两臂的序列中也有切口存在,酶解后得到的插入片段就会携带一段载体上的序列。当载体如 λ EMBL3 在分子中含有多酶连接子序列,就可避免上述情况。与 λ EMBL3 载体配对的是 λ EMBL4,两者具有相同的多酶连接子序列,差别仅在于多酶连接子序列的方向正好相反。新组建成的 Charon 34 和 35 的连接子含有八种限制酶序列(见第二章第二节),应用时更方便。

二、载体 DNA 的制备

替代型 λ 噬菌体载体 DNA 的制备与裂解型的 λ DNA 制备方法(见本书上册第一章)相同。首要的是选取适合于该载体 DNA 的宿主菌,例如 EMBL3 可选用大肠杆菌 C600。在感染宿主菌进行 λ 噬菌体裂解性繁殖时,选择合适的宿主菌和噬菌体的比例,将会获得较满意的噬菌体产量。载体 DNA 量一般需要制备 100 μ g 以上,尤其对初次开展工作,制备较多量的载体 DNA 是有益的。得到噬菌体裂解液后,借 PEG6000 收集噬菌体颗粒,并采用 CsCl 阶段梯度超速离心法纯化噬菌体,最后裂解噬菌体颗粒,从中抽提 DNA。酚抽提后的 DNA 最好不用乙醇法沉淀,因为沉淀后的 DNA 极难完全溶于缓冲液中。所以,采用直接透析除酚,较为理想。制得的载体 DNA 在用前要进行检测。

首先取 0.5—1 μ g DNA (估计量)进行琼脂糖凝胶(0.5% 浓度)电泳,观察 DNA 区带以及有无 RNA。如果混杂的 RNA 量较大,还需再用 RNaseA 处理。少量 RNA 则不必除去。制剂中 DNA 含量的测定,最好选用紫外分光光度法,1 OD₂₆₀ 按 50 μ g DNA 计算。同时通过 OD₂₆₀/OD₂₈₀ 比值来检测 DNA 纯度,一般需达到 1.8 至 2.0,比值过低表明蛋白等杂质含量高。当制备的 DNA 量不多时,只能采用与不同标准含量 λ DNA 一起电泳,通过区带比较,可估计大致含量。

其次,载体 DNA 制剂的酶切鉴定。选择制备载体两臂用的限制酶,例如用 BamHI 和 EcoRI 切 λ EMBL3 DNA。在 50 μ l 反应体系中,以 1—3 u 的酶量切 1 μ g DNA。37 $^{\circ}$ C 保温 1—2 小时,并用 λ DNA 作对照。若电泳表现能完全降解,说明 DNA 制剂合格。当酶量大至 5u 以上还不能降解完全,表示 DNA 制剂含有抑制酶活力杂质,需要再纯化。用 λ DNA 进行酶解的对照实验,可排除因酶活力降低而造成的不完全降解。

此外,需要用 DNA 的体外包装试验来检验制备的载体。一般实验室制备的 λ 噬菌体 DNA,体外包装抽提物的包装效率可达 1—2 $\times 10^3$ pfu。若载体 DNA 可达此种包装效率,说明制备的载体 DNA 符合构建文库的要求。

经检验合格的载体 DNA,可按需要分装保存。分装用的器皿要经过消毒处理,尽可能不带入细菌。DNA 可以在 4 $^{\circ}$ C 保存,也可在 -20 $^{\circ}$ C 贮存,但要避免反复冻化。

三、制备载体的两臂和去除中央片段

在将外源基因片段与载体进行重组之前，需用酶切割载体 DNA，并分去载体的中央片段，以利于载体两臂与供体 DNA 片段的重组连接。

1. 载体 DNA 的酶解

当选用 Sau3AI 或 MboI 的部分降解作用处理供体 DNA 时，载体 DNA 要用 BamHI 酶切；如用 λ EMBL3 时，常用 BamHI 和 EcoRI 双酶降解。

(1) BamHI 酶切。反应缓冲液为 150mmol/L NaCl、6mmol/L Tris-HCl (pH7.9)、6mmol/L MgCl₂、100 μ g/ml 的牛血清清蛋白(若没有无 DNase 级别的 BSA，可以不加，切不可用一般 BSA 替代)。DNA 浓度为 100—150 μ g/ml，反应总体积最好不超过 1.0ml，通常以 0.5ml 为宜。酶量为 2—3 单位/ μ gDNA，37 $^{\circ}$ C 水浴保温 1—2 小时，不宜采用过长的酶解时间。反应管置 0 $^{\circ}$ C 冰浴终止反应。取少于 0.5 μ g 的 DNA，先加适量能终止反应的含 EDTA 和甘油的溴酚蓝染液，于 65—68 $^{\circ}$ C 水浴加热 10 分钟，使粘合的 cos 位点解开，再在 0.5% 琼脂糖凝胶上电泳，检查酶解反应是否完全。电泳时用不经酶切的 DNA 作对照。BamHI 在 65—68 $^{\circ}$ C 条件不会很快失活，因此在借电泳检查酶解产物时，必须保证酶反应完全停止了，再在 65—68 $^{\circ}$ C 加热，否则，电泳显示的完全酶解并不真正反映酶解时的状况。还需提醒在反应时要控制反应体系中随加酶而带人的甘油浓度。若甘油的终浓度在等于和超过 5% 时，反应会出现二级酶活性，又称星号活性 (star activity)。为避免此副反应，要选用高浓度的 BamHI。另外低离子强度也会出现这样的活性。

(2) BamHI 和 EcoRI 的联合酶切。这是 λ EMBL3 载体选用的酶解条件。此时，一般可以选用 EcoRI 的酶反应条件：100mmol/L Tris-HCl (pH7.5)，50mmol/L NaCl，5mmol/L MgCl₂，100 μ g/ml BSA。反应体系中离子强度比较高，BamHI 呈现较好的活性。DNA 浓度保持在 100—150 μ g/ml。每种酶的用量为 2—3u/ μ g。为防止甘油的终浓度超过 5% 而出现星号活性，必须选用高浓度的酶制剂 (10—20u/ μ l)。

对于其他含多酶连接子 (polylinker) 序列的载体，例如 λ EMBL4，Charon 34 或 35，也可选用合适的双酶降解。降解时，两酶可灵活运用，或同时降解，或先用一酶降解，随后再进行第二个酶的切割，这取决于所用酶的反应条件。

(3) 酶解产物的处理。电泳鉴定证实酶反应完全后，产物需用等体积的酚或者酚与氯仿-异戊醇的混合液抽提两次，使酶彻底灭活。抽提后的水相(含 DNA 部分)，经乙醚除酚和氯仿，再乙醇沉淀(加 NaCl 至终浓度为 0.3mol/L，两倍体积 95% 乙醇，-20 $^{\circ}$ C 1—2 小时)。DNA 沉淀经 70% 乙醇洗涤后，减压抽干，并溶于 TE 缓冲液中，终浓度为 200—500 μ g/ml。4 $^{\circ}$ C 下贮存，有待进一步去除中央片段，以纯化载体两臂。

对于用 BamHI 和 EcoRI 双酶解的 λ EMBL3 产物或经类似处理的其他载体产物，因无需分离中央片段，它们在酚抽提后，可按下述方法沉淀 DNA，以选择性地去除双酶解产生的寡核苷酸产物：水相 DNA 溶液中加入 1/10 体积的 3mol/L NaAc，pH6.0 和 0.6 体积的异丙醇，混匀后在冰浴中放置 15 分钟，室温下于 Eppendorf 离心机离心 5 分钟，弃上清液，沉淀用 0.5 体积含 0.3mol/L NaAc，pH6.0 的 40% 异丙醇洗，并离心。

倾去上清液后,沉淀被减压抽干(注意不宜过干,否则 DNA 不易溶解),并溶于 TE 缓冲液中,使浓度为 200—500 $\mu\text{g}/\text{ml}$ 。借电泳法测定 DNA 含量后,样品于 4 $^{\circ}\text{C}$ 保存备用,也可于 -20 $^{\circ}\text{C}$ 长期贮存。

2. 载体两臂的纯化

对于 λ EMBL3 一类载体,经双酶解处理,所得中央片段的末端结构与两臂末端的粘性结构不相同,当将反应产物中的寡核苷酸经选择性沉淀法去除之后,在载体与供体 DNA 片段重组时,中央片段不会与载体两臂相连而影响与供体片段的重组效率。因此,载体酶解物不必先除去中央片段,即可直接进行重组连接。但是,当载体仅用一种限制酶切割,产生的两臂与中央片段将具有相同的粘末端。在载体与供体 DNA 片段重组时,中央片段将会参与连接反应而严重影响两臂与供体片段的重组。因此为提高重组效率,必须除去中央片段,纯化载体两臂。方法可采用蔗糖密度梯度离心(见本章第二节)或醋酸钾梯度离心(见参考文献 2);也可选用低熔点琼脂糖凝胶的电泳分离法。无论采用上述何种方法,载体的酶解产物均需先经过“退火”处理,把左、右两臂借 *cos* 位点粘合在一起(见图 6.1)。

(1) “退火”(annealing)处理。向酶解后的产物(在 TE 缓冲液中,浓度为 200 $\mu\text{g}/\text{ml}$)中加入 1mol/L Tris-HCl (pH8.0),使其浓度达 100mmol/L, MgCl₂ 达 10mmol/L。混匀后,42 $^{\circ}\text{C}$ 放置 1 小时,使 λ 载体两臂粘合。粘合程度可借低温(10—4 $^{\circ}\text{C}$)、低电压(2V/cm)的 0.5% 琼脂糖凝胶电泳检查。样品共两份,每份 0.2 μg 左右,用 100mmol/L Tris-HCl, 10mmol/L MgCl₂ 溶液稀释至 20 μl , 其中一份加热 68 $^{\circ}\text{C}$ 10 分钟;另一份不作其他处理,分别进行电泳。

为了提高两臂的粘合率,也有人把载体 DNA 经“退火”后,先借 T4 DNA 连接酶连接两臂,然后用限制酶切割。产物主要是连接在一起的噬菌体 DNA 两臂和一个中央片段,这样有益于两者的分离。

(2) 低熔点琼脂糖凝胶电泳法制备纯化载体两臂。低熔点琼脂糖(low melting agarose)可为美国 BRL 公司(现改名为 Life Technology Inc; LTI)或 FMC 公司的 Sea Plaque 产品(Rockland ME 04841)。电泳缓冲液为 TAE (0.04mol/L Tris-HAc, 0.001mol/L EDTA, pH8.0)。制备用凝胶浓度为 0.8—1.0%,胶厚度为 0.8—1cm,胶的大小可由实验者持有的电泳槽规格而定,一般为 15(宽)×20(长)cm。样品槽面积为 0.3×10cm。以上凝胶可适用于分离 50—100 μg 的 DNA 酶解物。为了保证胶凝固完全,制胶时以 4 $^{\circ}\text{C}$ 下进行为好。若想节省低熔点的琼脂糖,可先用普通琼脂糖铺制相同规格的 0.7—1.0% 凝胶,待完全凝固后,在离加样槽前沿 1 和 11 厘米处,用薄刀片切开凝胶(必须注意保持胶切面的整齐平滑),挖去中段胶,于 4 $^{\circ}\text{C}$ 下把冷却至 40 $^{\circ}\text{C}$ 左右的低熔点凝胶填补其中,待充分凝固后即可使用。

加样槽中加入 0.25—0.5ml 载体 DNA 的酶解产物(含 DNA 50—100 μg 左右),酶解物中预先已加有 1/10 体积的甘油-溴酚蓝染液。加样后即可电泳。起始用 5—8V/cm 电压条件进样,20—30 分钟后,将电压调至 2V/cm, 4 $^{\circ}\text{C}$ 下电泳,直至酶解片段的区带能彼此清晰地分离(约需 18 小时或更长)。电泳缓冲液中含有 0.2 $\mu\text{g}/\text{ml}$ 的溴乙锭,电泳过程中可随时于长波紫外灯下观察分离情况。电泳最好在避光条件下进行。

得到满意的分离之后,在长波紫外灯下把粘合一起的两臂区带,以及单独的左臂与右

臂区带全部切下,切时必须注意勿混入中央片段的区带,并尽可能切去不含样品的空白凝胶。然后,按下述步骤从凝胶中抽提回收 DNA 片段。

把待抽提 DNA 的凝胶条放入可耐酚和有机溶剂的 50ml 带盖塑料管中,加入 1—1.5 倍体积的 0.5mol/L NH_4Ac -TE 溶液 (0.5mol/L NH_4Ac , 10mol/L Tris-HCl, 1mmol/L EDTA, pH8.0), 加盖后在 68—70°C 水浴中融化凝胶 (20—30 分钟), 间歇地轻摇以助熔融。

加入等体积酚 (Tris 缓冲液饱和的酚液), 缓和而充分的混合。室温下离心 3,000—5,000r/min \times 10 分钟。将水相转移到新管中, 酚相和中间层再用 0.5mol/L NH_4Ac -TE 溶液 (原体积的 1/4) 反抽提一次, 离心后, 上清液与第一次的合并, 再用 1/2 体积酚液抽提两次。

若上清液体积过大, 可用正丁醇抽提法浓缩, 即加入等体积的正丁醇 (A. R.) 充分混匀并离心, 除去有机相, 反复数次直至 DNA 样品浓缩为 50—100 $\mu\text{g}/\text{ml}$ 的浓度。浓缩后的 DNA 溶液, 再用等体积酚-氯仿混合物 [1/2 体积酚, $\frac{1}{2}$ 体积的氯仿-异戊醇, 24:1(V/V)] 抽提一次, 氯仿-异戊醇抽提两次, 最后用等体积醚抽提三次。

向所得 DNA 溶液中加入两倍体积的 95% 乙醇, 按常规沉淀 DNA。DNA 沉淀物必须用 70% 乙醇洗两次, 并离心去乙醇, 减压抽干沉淀物后溶于 TE 缓冲液中, 使 DNA 浓度约在 500 $\mu\text{g}/\text{ml}$ 。4°C 下保存, 也可分装后于 -20°C 保存。

制得的载体两臂, 必须测定其浓度及质量。DNA 浓度可通过电泳与标准的 λ DNA 区带比较, 而测其近似值。至于两臂的质量检测, 则取 0.5 μg 的两臂 DNA 两份, 一份加入约等摩尔量的同法抽提制备的中心片段, 一份不加, 分别进行连接和体外包装。通过感染宿主细胞来观察噬菌斑的形成数目 (pfu), 计算出每微克载体臂的 pfu 数值。若单纯两臂的 DNA 片段自身连接, 得到噬斑数目小于 10³pfu/ μg , 而加入中央片段后的连接产物, 其形成噬斑能力可提高至 10⁵—10⁶pfu/ μg , 说明制备的载体臂符合要求。

当用低熔点琼脂糖的凝胶电泳法纯化载体两臂时, 低熔点琼脂糖的质量对能否制备合格的两臂有很大影响。但是尚不清楚是何种杂质干扰。实验经验表明, 应当注意琼脂糖批号, 因为经常发现有的批号合格, 换另一批就不一定符合要求。

第四节 重组连接、体外包装和感染方法

构建一个理想的基因文库, 取决于提供重组用的载体和供体 DNA 制剂的质量, 也取决于连接反应的条件以及重组的噬菌体 DNA 分子的体外包装效率。此外, 还和宿主细胞和感染的方法有关。本节主要介绍供体 DNA 片段和载体间的重组连接、重组 DNA 的体外包装和重组噬菌体的感染等有关方法。

一、重组连接

将供体 DNA 片段插入到 λ 载体两臂之间构成重组 DNA 分子, 需依靠 T4 DNA 连接酶, 它是从 T4 噬菌体感染的大肠杆菌中纯化得到的。它能催化 DNA 片段间的 3'OH 和 5' 磷酸单酯的相互作用, 形成 3', 5' 磷酸二酯键, 但底物分子必须为双链 DNA (见本

书上册第三章介绍)。因此带相同粘末端的双链 DNA 片段之间的相互连接是 T4DNA 连接酶催化的典型反应。两个带平末端的 DNA 片段,也可用 T4 DNA 连接酶催化形成 3', 5' 磷酸二酯键,但连接效率远不如两个相同粘末端 DNA 片段间的相互连接。在基因文库的构建中,一般都选粘末端连接方式。但在进行载体和外源 DNA 片段间的连接时,要得到理想的重组效率,还需考虑以下因素:

第一,载体和外源 DNA 的摩尔比例。在采用 λ 噬菌体的替代型载体构建基因文库时,只有按左臂-插入的外源 DNA 片段-右臂的顺序连接,才会建成有效的重组分子。要做到如此连接,反应中各底物的分子比例必须是 2:1:2 (左臂:外源片段:右臂),因为左臂和右臂只含一个由限制酶降解提供的粘末端,它们的另一末端是 *cos* 位点结构,而插入的外源片段的两个末端均是限制酶的粘末端。当参加连接反应的 λ 载体与外源片段的摩尔数为 2:1 时,即达到了上述的 2:1:2 的比例要求。

第二,载体和外源 DNA 的总浓度。由噬菌体 DNA 两臂和外源 DNA 片段组成的一个典型分子,如果借 *cos* 位点自身形成环状分子,它就不能被包装成噬菌体颗粒。只有许多典型的重组分子互相以 *cos* 位点连接成一个长串排列的线性多联体 (concatemer) 时,才能被有效地包装。因此,连接反应体系中 DNA 的绝对浓度要适当地高,使之主要产生重组分子之间的连接,减少自身连接。

(一) 理想连接条件的理论值估算

Dugaiczyk 等曾经从理论上讨论了片段的大小和 DNA 浓度对连接产物特性的影响 (*J. Mol. Biol.*, 96, 171, 1975)。提出了与连接反应密切有关的两个参数: j 和 i 。 j 是指同一 DNA 分子中两末端的有效浓度。分子内两个末端碰撞一起的机率可由下列公式来计算

$$j = \left(\frac{3}{2\pi lb} \right)^{3/2} \text{ 末端数/ml} \quad (1)$$

l 是 DNA 长度(以厘米为衡量单位), λ DNA 分子长度 $l = 13.2 \times 10^{-6}$ cm; b 为 DNA 分子杂乱线团的长度,其数值取决于缓冲液的离子强度,它影响了 DNA 的刚性,溶于连接反应缓冲液中的 λ DNA 的 b 值为 7.7×10^{-6} cm。 j 值与 DNA 分子的长度成反比关系。分子越大,末端间的相互作用就越弱。 j 值还可用下面公式计算

$$\begin{aligned} j &= j_{\lambda} \left(\frac{M_{w_{\lambda}}}{M_w} \right)^{3/2} \text{ 末端数/ml} \\ &= \frac{5.5 \times 10^{22}}{M_w^{3/2}} \text{ 末端数/ml} \end{aligned} \quad (2)$$

$j_{\lambda} = 3.22 \times 10^{11}$ 末端数/ml (λ DNA 分子的有效末端数由公式 (1) 计算得到), $M_{w_{\lambda}} = 30.8 \times 10^6$ 。由公式(2)可知 j 值仅与其自身分子大小有关,而和 DNA 浓度没有关系。 i 是指可互补的末端总浓度。对于自身互补的双链线性 DNA, 粘末端的浓度

$$i = 2N_0M \times 10^{-3} \text{ 末端数/ml} \quad (3)$$

N_0 是阿佛加德罗 (Avogadro's) 常数, M 是 DNA 摩尔浓度。

如果 DNA 分子所带末端不能自身互补, 那末每一种末端的浓度 $i = N_0 M \times 10^{-1}$ 末端数/ml。

从理论上, $j = i$ 表示 DNA 分子的一个末端与同分子内的其他末端或与其他分子的末端间的接触机会相等, 也就是说, 连接反应中自身环化和分子间的串联状连接的速度相同; 当 $j > i$, 主要形成自身环化的分子; $i > j$ 形成串联状线性多联体。在基因文库的构建中, 希望载体两臂和插入片段之间的连接能提供体外包装成噬菌体颗粒所需的有效的 DNA 重组分子, 即形成分子间呈串联状连接的线性多联体。因此, 在连接重组时, 需选择合适的载体和插入片段浓度, 造成 $j \ll i$ 的反应条件。

假如载体两臂淬灭粘连后的分子量为 31kb (1.9×10^7), 插入片段的平均大小是 20 kb (1.25×10^7), 两者的 j 值可由公式(2)计算得出:

$$j_{\text{两臂}} = 6.6 \times 10^{11} \text{ 末端数/ml}$$

$$j_{\text{插入片段}} = 1.25 \times 10^{12} \text{ 末端数/ml}$$

i 表示连接反应体系中所有粘末端的总浓度, 为了提供 $j \ll i$ 的反应条件, Maniatis 等人曾提出, 使 $i = 10 \times j_{\text{插入片段}}$ 。因此由 $j_{\text{插入片段}}$ 的值可以求出

$$i = 10 \times 1.25 \times 10^{12} = 1.25 \times 10^{13} \text{ 末端数/ml} \quad (4)$$

同时 i 又是 $j_{\text{两臂}}$ 和 $j_{\text{插入片段}}$ 之和。故按公式(3)可以得知 $i = (2N_0 M_{\text{两臂}} + 2N_0 M_{\text{插入片段}})$ 为了造成“左臂-插入片段-右臂”的有效重组分子, 上述曾提出了载体两臂和插入片段间的摩尔数应为 2:1, 即选择摩尔浓度 $M_{\text{两臂}}$ 应为 $M_{\text{插入片段}}$ 的 2 倍。依此换算:

$$\begin{aligned} i &= (2N_0 \times M_{\text{插入片段}} + 2N_0 \times 2M_{\text{插入片段}}) \times 10^{-3} \text{ 末端数/ml} \\ &= 6N_0 \cdot M_{\text{插入片段}} \times 10^{-3} \text{ 末端数/ml} \end{aligned} \quad (5)$$

因此, 把 (4) 式中 i 值和 N_0 代入 (5) 式中, 即可得到插入片段和两臂的摩尔浓度值 ($M_{\text{插入片段}}, M_{\text{两臂}}$)

$$M_{\text{插入片段}} = 3.5 \times 10^{-9} \text{ mol/L (即 } 43 \mu\text{g/ml)}$$

$M_{\text{两臂}} = 2 \times M_{\text{插入片段}} = 7.0 \times 10^{-9} \text{ mol/L (相当于 } 135 \mu\text{g/ml)}$ 这两个数值的含意是, 要把分子大小为 31kb 的载体两臂和平均为 20kb 大小的插入片段重组连接时, 两底物的浓度应是 $43 \mu\text{g/ml}$ (插入片段) 和 $135 \mu\text{g/ml}$ (载体两臂)。它们在理论上提供了连接反应中两底物的适宜浓度。可供实验者参考。

实际上, 由于制备所得底物 DNA 的含量难以确切测得, 以及每次制备时带入的抑制连接反应的杂质不等, 按理论计算的两种底物浓度往往不能得到满意的重组效率, 一般都经过测试实验, 求得较适宜的底物浓度。

(二) 连接条件的测试以及正式实验

在测试合适的底物反应浓度之前, 为保证连接反应的成功, 先要检测载体两臂和插入的外源 DNA 片段的质量, 然而再开始测试合适的连接反应条件, 以及采用测得的合适条件进行重组连接实验。

1. 噬菌体载体两臂的质量检测

在制备噬菌体 λ 载体两臂时, 无论用物理方法还是酶促方法, 通常不能完全地除去噬菌体 DNA 的中心片段。用载体两臂的 DNA 进行自身连接(反应中不加插入片段), 再体外包装之后, 滴定其噬菌体形成的噬菌斑数目。每 μg DNA 的噬菌斑总数不能超过

10^3 pfu (噬菌斑形成单位), 而向载体两臂加入适量的、合适大小的中央片段 (来自相同载体) 后, 重组效率应提高 10^2 左右, 如此的载体两臂 DNA 才为合格制剂。

2. 外源 DNA 片段的大小

野生型 λ DNA 分子的长度为 48.5 kb, 当其被包装成噬菌体颗粒时, DNA 的最小长度不得小于全链长的 78%, 最长者不超过全长的 105%。即可被包装成噬菌体颗粒的 λ DNA 的长度范围是 38 至 51 kb。当插入的外源 DNA 片段大小适宜, 可以提高重组的 λ DNA 的包装效率, 而靠近极限值范围, 包装效率较低, 外源 DNA 片段的大小范围取决于载体两臂的长度。若两臂长度之和为 30 kb, 插入片段的大小可为 9 至 21 kb, 适宜的平均长度是 18—19 kb。因此, 在进行 *Sau* 3A I 或 *Mbo* I 部分酶解产物的分级时, 一定要严格控制长度, 不让过长或过短的片段混杂。

3. 连接反应的条件

T4 DNA 连接酶的反应缓冲液通常用的是:

- 50 mmol/L Tris-HCl, pH 7.4
- 10 mmol/L $MgCl_2$
- 10 mmol/L DTT
- 1 mmol/L 亚精胺 (spermidine)
- 1 mmol/L ATP

加入 BSA (牛血清白蛋白) 至浓度 $100 \mu\text{g}/\text{ml}$, 若没有合格的 BSA (不含 DNase 级试剂), 则以不加为好。缓冲液一般配制成 10 倍或 5 倍于上述的浓度, 分装成 0.5 ml 的小包装, 贮存于 -70°C 或 -20°C 。

反应中 T4 DNA 连接酶用量为每 $10 \mu\text{l}$ 反应体积含 2—10 Weiss u。水浴保温, 反应温度为 $12-16^\circ\text{C}$, 时间从几小时到过夜 (18 小时), 一般多采用保温过夜。反应体系中 DNA 浓度常保持在 $150-200 \mu\text{g}/\text{ml}$ 左右。

合适的连接酶反应条件主要是指载体和插入 DNA 片段的合适的摩尔比。测试时, 一种方法是固定载体 DNA 的用量为 1 或 $1.5 \mu\text{g}$ DNA/ $10 \mu\text{l}$ 反应体积, 而改变插入片段的 DNA 量, 使之达到约相当于载体摩尔数的 0、0.5、1 和 2 倍 (分别标以第 1、2、3、4 四管)。插入片段为“0”是对照, 以观察载体自身的本底。除上述四种反应管外, 有时增设第 5 反应管。其插入片段量同第 3 管, 而载体 DNA 量比原来增加一倍。若为节省载体 DNA, 也可以把反应体积缩小为 $5 \mu\text{l}$, 用 $1-5 \mu\text{l}$ 标示量的微量毛细管转移缓冲液、DNA 和酶, 并使反应在密封的毛细管内进行。

另一种测试方法是固定插入片段的用量, 改变载体两臂的量, 使两者摩尔比值为 4:1、2:1、1:1 和 0.5:1 (载体:插入片段)。

反应停止后, 定量地取 $3 \mu\text{l}$ 连接产物进行体外包装 (见本节第二部分), 将包装的重组噬菌体悬浮于 $50 \mu\text{l}$ 的 SM 缓冲液 [含 $100 \text{mmol}/\text{L}$ NaCl, $10 \text{mmol}/\text{L}$ $MgSO_4 \cdot 7H_2O$, $50 \text{mmol}/\text{L}$ Tris-HCl (pH 7.5) 和白明胶 0.01%], 取 1 或 $10 \mu\text{l}$ 感染宿主细胞, 由平板上所得噬菌斑数目计算出每 μg 插入片段 DNA 所得噬菌斑数 (pfu); 也可以每 μg 载体 DNA 来计算, 然后根据实验数值, 选择最佳的连接反应条件。

有人还用琼脂糖凝胶电泳方法检测连接效果, 即比较连接反应前后的电泳谱, 不过实际意义不大。因为常在电泳谱上可见到连接区带, 而进行体外包装时, 却效率很低。我们

从实践中体会到,合适的连接条件选择主要依据重组产物被包装成噬菌体颗粒的效率。

当换用不同批号制备的载体和插入片段时,需要重新测试。因此,在测试之前,最好制备有一定数量的载体和插入片段,以确保在选到合适条件后,有足够量的同批样品来完成整个基因文库的构建。

一般测试实验剩余的连接产物不要丢弃,一个完整的基因文库需要包含有一定数量的重组噬菌体。在得到了质量好的载体和插入片段制剂,以及有高效率的噬菌体包装抽提物时,通过上述测试已获得完整基因文库所需要的重组体数目或其二分之一,只要按测试的合适条件再重复1—2次,即可完成基因文库的构建。如与所需数量相差甚远,则可用放大体积的连接反应或采用与测试实验相同反应体积的数个反应管平行连接。

二、重组噬菌体 DNA 的体外包装

按理论计算,每微克野生型 λ DNA 可形成 1.8×10^{10} 个噬菌斑。当用 λ DNA 直接感染大肠杆菌,效率仅为 10^4 — 10^6 pfu/ μ g DNA。1974年, B. Hohn 和 T. Hohn 提出了可在试管中将 λ DNA 包装成噬菌体颗粒。Becker 和 Gold(1975年)报道了从感染的宿主细胞中分离噬菌体 A 基因蛋白方法,以及用它进行 λ DNA 体外包装的步骤。Sternberg 等和 Faber 等都曾介绍了包装抽提物的制备,操作较简便的是 Maniatis 等主编的《分子克隆》一书中介绍的方法(见参考文献2)。 λ DNA 体外包装抽提物已有商品,例如美国 Stratagene 公司的产品,包装效率可高达 $1-2 \times 10^9$ pfu/ μ g λ DNA。高效率包装抽提物的提供,不仅减少了实验者的工作量,而更重要的在于节省了构建基因文库所需的材料,便于不易获得的组织和细胞构建基因文库。

(一) 制备包装抽提物用的菌株

A. BHB2688 [N205 recA⁻(λ imm434 cIts b2 red3 Eam4 Sam7)/ λ]

B. BHB2690 [N205 recA⁻(λ imm434 cIts b2 red3 Dam15, Sam7)/ λ]

这两株菌都是含有 λ 前噬菌体的溶源菌 (lysogen), 它们的宿主细胞是大肠杆菌 N205 recA⁻ (来自 N. Sternberg)。整合于细菌染色体 DNA 链上的 λ 前噬菌体基因分别是两种 λ 突变体的重组产物。菌株 BHB2688 中的前噬菌体基因来自 λ cIts857 red3 b2 Eam4 Sam7 和 λ imm 434cIts Eam4 两种 λ 突变体的重组体。BHB2690 是由 λ cIts 857 red3 b2 Eam4 Sam7 和 λ imm434 cIts Dam15 重组而来的。两种前噬菌体基因表型的差别在于 BHB2688 株中为 E 蛋白基因的琥珀型突变 (amber mutant), 2690 株是 D 蛋白基因的琥珀型突变。E 和 D 两种蛋白是构建 λ 噬菌体头部的重要成分。D 蛋白定位于噬菌体头部的外侧,它的功能涉及 λ DNA 插入到头部的头前身物(即前头体)中,以及与头部的成熟有关。当 D 蛋白基因突变 (Dam15) 时,导致大量前头体的堆积,而 λ DNA 又不能插入到前头体中。E 蛋白是噬菌体头部的主要成分,是构建最早期的头部前身物不可缺少的蛋白。E 蛋白基因的突变 (Eam4) 导致了构成头部的其他蛋白呈可溶状态,不能组装成头部外壳的形态。总之,这两株溶源菌 (BHB2688 和 2690) 诱导后均不会产生噬菌体颗粒。但是,当把这两株溶源菌的裂解液合并一起,由于互补而可以提供形成成熟的噬菌体颗粒所需要的全部成分。当有 ATP、亚精胺和腐胺存在下, λ DNA 就可被

包装成具有感染性的成熟的噬菌体颗粒。

这两种溶源菌在前噬菌体基因上的突变还有：① *imm434cIts*。该基因产物是温度敏感型的阻遏物。当溶源菌在 32℃ 下繁殖时， λ 基因处于溶源状态，即被整合在宿主菌染色体中，随染色体 DNA 一起复制；当培养温度提高到 42—45℃ 时，前噬菌体基因组被有效地诱导，由染色体 DNA 链上释放出来，进入细胞质中自行复制，并繁殖噬菌体；② *Sam7* 突变，这是 S 基因的琥珀型突变。S 基因产物是噬菌体裂解宿主细胞所需要的，*Sam7* 的突变，使溶源菌诱导后，噬菌体可在细胞内繁殖 2—3 小时，而不引起宿主细胞裂解；③ b 区的缺失 (*b2*)，可减少来自包装抽提物内源性 λ DNA 的体外包装；④ *red3* 突变使 λ DNA 重组体系失活。此外是宿主菌的 *recA* 突变，*recA*⁻ 是宿主细胞染色体的重组体系失活。*recA*⁻ 的宿主菌，加上噬菌体的 *red3* 突变和 *b2* 区缺失，使重组噬菌体 DNA 体外包装时，大大减少存在于包装抽提物中内源性 DNA 的包装。

要制备满意的包装抽提物菌株是很重要的，每次制备前一定要认真检测细菌的溶源性能。检测方法为：把贮存的菌种（甘油保存菌或冷冻干燥菌株）接种于 LB 平板上，32℃ 培养过夜，长出单菌落之后，取长得小的单菌落同时接两块 LB 平板（一次可接几个单菌落），一块板于 32℃ 培养过夜，另一块在 42℃ 培养，挑取在 32℃ 生长而 42℃ 不长的菌作制备抽提物。另外还应用紫外线辐射检测 *recA*⁻ 菌株，作制备包装抽提物用的菌株应是 *recA*⁻，即在紫外线照射后不能再生长。

（二）包装抽提物的制备

制备包装抽提物至少有两种以上的方法（见参考文献 [2]）。由 BHB2688 和 2690 分别制备抽提物的方法简述如下：

1. 菌体的培养

菌体培养基采用 NZC 肉汤（即每升培养基中含 N-Zamine 10g，NaCl 5g，MgCl₂·6H₂O 2g 和酪蛋白氨基酸 1g，pH7.5）。N-Zamine 为美国 Humko-Sheffield 公司产品，可用水解酪蛋白替代。有的实验室（中国科学院动物研究所沈孝宙，私人通信）采用 TBY 培养基，也得到良好结果。其成分为：每升培养基中含蛋白胨 1g、牛肉膏（上海产品）5g、酵母抽提物（Difco）2g 和 NaCl 10g，pH7.5。另外配制 1mol/L MgSO₄，分开消毒，用前按 1/100 体积加量。

菌株 BHB2688 和 2690 经温度敏感性及 *recA*⁻ 检测之后，将合格的菌株制备过夜培养液 50—100ml（32℃ 振荡培养）。次日接至 500ml NZC 培养基（装于 2 升容积的三角瓶）中，使其 OD₆₀₀ 读数为 0.08—0.1，随后 32℃ 振荡培养至 OD₆₀₀ 为 0.3。45℃ 水浴热诱导 15—18 分钟后，移至 37—38℃ 剧烈振荡培养 3 小时。取少许培养液（1—2 ml），借氯仿检测裂解程度。即加数滴氯仿并用手振荡后，培养液在静置几分钟后需变得清亮。

由 BHB2690 制备“超声”抽提物，一般培养 500ml 即可；BHB2688 菌用于“冻-融”抽提物的制备，则需 3 × 500ml 培养液。若制备量不大，培养液可减半。

2. BHB2690 超声抽提物的制备

取 500ml BHB2690 培养液，于低温（4℃）离心后，收集菌体，弃尽上清液。菌体悬浮于 2.5ml 预冷的超声缓冲液 [20mmol/L Tris-HCl (pH8.0)、1mmol/L EDTA、3m

mol/L $MgCl_2$ 、5mmol/L 巯基乙醇], 在冰浴中用超声波发生器的微型探头间歇地打碎细胞 (5—10 秒/次, 20 次以上), 直至裂解液变清亮, 不见粘稠。超声破碎时, 注意温度勿超过 $4^{\circ}C$, 防止出现泡沫。裂解液在 $4^{\circ}C$ 下以 10,000r/min 速度离心 10 分钟, 取上清液 2 ml, 加入等体积预冷超声缓冲液及 1/6 体积的新鲜配制的包装缓冲液 [6mmol/L Tris-HCl (pH8.0)、50mmol/L 亚精胺、50mmol/L 腐胺、20mmol/L $MgCl_2$ 、30mmol/L ATP、30mmol/L 巯基乙醇], 轻轻混合均匀, 分装于 1.5ml 容积的尖底塑料管中, 加 15 μ l 抽提物于每管底部, 加盖后立即浸于液氮中速冻, 并在 $-70^{\circ}C$ 贮存。全部操作需在冷室中进行, 分装用塑料管要先预冷。

3. BHB2688 “冻-融”抽提物的制备

按上述培养的 3×500 ml BHB2688 菌培养液, $4^{\circ}C$ 离心收集菌体, 弃尽全部上清液, 移去全部残留液。菌体用 2ml 预冷的蔗糖缓冲液 [10% 蔗糖、50mmol/L Tris-HCl (pH7.5)] 悬浮, 并转至离心管中, 向管中再加入 100 μ l 新鲜配制的溶菌酶溶液 [2mg 酶/ml、0.25mol/L Tris-HCl (pH7.5)], 混匀后, 浸没于液氮中速冻, 然后置于 $0^{\circ}C$ 冰浴中融化。向粘稠的裂解液中加入 100 μ l 包装缓冲液 (新鲜配制, 配方见超声抽提物的制备一段), 充分混匀。经超速离心 (35,000r/min, 60 分钟, $4^{\circ}C$), 吸出上清液, 分装于 1.5 ml 尖底塑料管中, 每管 10 μ l, 立即置于液氮中速冻, 并于 $-70^{\circ}C$ 保存。按我们的实验经验, 若在裂解液融化之后, 用超声波发生器微型探头打 5 秒钟 \times 3, 可减低裂解液粘稠度, 提高“冻-融”抽提物产量。

(三) 抽提物包装效率的检测

制备的抽提物应先检测它们包装 λ DNA 的效率, 然后才能使用。

1. 内源性 DNA 的检查

制备包装抽提物的两种菌株 (BHB2688 和 2690) 都是含 λ 前噬菌体的溶源菌, 若内源性的噬菌体 DNA 可被包装成噬菌体颗粒, 将干扰重组噬菌体的筛选, 因此制备的包装抽提物要首先检测内源性 DNA 的被包装程度。检测方法是加任何外源性 DNA, 把两种包装抽提物按正规的包装操作步骤进行空白包装, 包装后感染宿主菌 (相同于重组噬菌体 DNA 包装用受体菌) 并铺板检测有无噬菌斑形成及形成噬斑的数量。理想的结果应无噬菌斑可见, 或者整个包装物形成的噬菌斑不能超过 10^2 个。

2. 天然 λ DNA 包装效率的测试

包装用样品采用经过连接或淬火粘合的串联状线性多聚体的天然 λ DNA 或 λ 载体 DNA (相同于重组实验用的 λ 载体) 制剂。包装步骤为: 从 $-70^{\circ}C$ 取出“超声”和“冻融”两种抽提物 (每次取量不能超过 3 对), 放于冰浴中, 先用手指融化“冻融”抽提物, 融化后, 即刻用微量可调式加样器把 10 μ l 抽提物转移到未融化的“超声”抽提物管中, 并迅速加入 DNA 样品 (含 DNA 0.3—0.5 μ g, 体积不超过 5 μ l), 用加样器上的塑料头 (tip) 或封顶的玻璃毛细管轻轻搅混样品和抽提物。此时“超声”抽提物才全部融化。注意尽量不要产生气泡, 如有少许气泡, 可立刻离心 (Eppendorf 离心机) 2—3 秒钟, 除去气泡。加盖后, 反应管于室温 ($20-25^{\circ}C$) 放置 60 分钟。最后加入 0.5ml SM 缓冲液 [0.1mol/L NaCl、10mmol/L $MgCl_2$ 、50mmol/L Tris-HCl (pH7.5)、0.01% 白明胶] 和氯仿 20 μ l, 轻轻混匀, 即可滴定被成功包装的噬菌体颗粒数目。

在滴定被包装的噬菌体颗粒数目时,要先把噬菌体悬液稀释。按常规每 μg 野生型 λDNA 或载体的包装效率为 $10^8\text{pfu}/\mu\text{g DNA}$, $500\ \mu\text{l}$ 包装物中含 10^8pfu , 则每 μl 约含 10^5 数量级 pfu。因此将包装的噬菌体原液稀释 10^{-3} 和 10^{-4} , 再每取 $10\ \mu\text{l}$ 噬菌体去感染细胞、铺板检测噬菌斑数目,计算出每 $\mu\text{g DNA}$ 可被包装成噬菌体颗粒(即噬菌斑)的数量。感染和铺板的操作步骤见本节第三部分所述。由实验室制备的抽提物,其包装效率一般可达 $1-2 \times 10^8\text{pfu}$ 以上。

3. 合适的 DNA 用量及其样品体积的检测

噬菌体 DNA 的体外包装效率,除了与抽提物及 λDNA 样品质量密切有关外,包装时所用 DNA 的量和样品体积也对包装效率有影响。在上述的体外包装操作中,当 DNA 用量在 $1\ \mu\text{g}$ 之内,包装成噬菌体的数目会随 DNA 量的加大而增多,超过 $1\ \mu\text{g}$ 之后,再加大 DNA 量,可包装成噬菌体的数目不会增加。当固定 DNA 用量(例如用量均为 $0.5\ \mu\text{g}$ 左右),则包装效率往往随体积的加大而降低。体积由 $3-5\ \mu\text{l}$ 加大至 $10\ \mu\text{l}$, 包装效率下降不明显,而样品体积加大到 $20-50\ \mu\text{l}$, 效率会急剧降低。因此在进行重组体 DNA 包装之前,应该测定 DNA 用量和样品体积对包装效率的影响曲线,以求得合适的样品量和体积。取每 $5\ \mu\text{l}$ 体积中含 $0.1, 0.3, 0.5, 0.8, 1.0, 2.0\ \mu\text{g DNA}$ 分别进行体外包装,检测最佳 DNA 用量(一般为 $0.3-0.5\ \mu\text{g}$),然后把 DNA 固定为 $0.5\ \mu\text{g}$, 使其分别溶于不同体积的 SM buffer ($1, 3, 5, 10, 20, 50\ \mu\text{l}$) 中,测定包装效率(一般体积用 $2-5\ \mu\text{l}$ 为多)。

(四) 重组 DNA 分子的体外包装

鉴于重组时连接反应的 DNA 浓度较高(一般为 $150-200\ \mu\text{g}/\text{ml}$), 因此在进行重组 DNA 分子的体外包装实验中,只要取 $2-3\ \mu\text{l}$ 的连接产物就行。重组 DNA 的包装过程与上述介绍的野生型 λDNA 包装相同。先把“冻-融”的抽提物($10\ \mu\text{l}$) 转移到超声抽提物管内,再加 $2-3\ \mu\text{l}$ 连接产物,在超声抽提物融化过程中将 DNA 与它们混匀,在室温中保温 60 分钟,再加入 $500\ \mu\text{l}$ SM 缓冲液及一滴氯仿摇匀即可。

在滴定重组噬菌体的数目时,毋需把噬菌体悬液稀释,可以直接进行滴定,因为,一般说来,重组的 λ 噬菌体 DNA 的包装效率比野生型的 λ 噬菌体 DNA 要低 10^2-10^3 。

包装后,测得噬菌体颗粒若达到 $2-5 \times 10^5\text{pfu}/\text{ml}$ 的滴度,可直接用于克隆的筛选实验。由于 SM 缓冲液中含有 Mg^{2+} , 使被包装的噬菌体颗粒很稳定。若在噬菌体悬液中加入少许氯仿后,可在 4°C 保存一年以上,滴度不会明显下降。不过没有扩增的基因文库不宜保存过长时间,以防重组体丢失。

三、重组噬菌体的感染

λ 噬菌体感染大肠杆菌,首先由噬菌体尾部纤维的顶端吸附于细胞膜上,然后将 λDNA 注入到细胞内。噬菌体的吸附作用发生于细胞外膜的特定受体部位。这种受体是由大肠杆菌基因组中 lam B 基因所密码的。它的形成需要有麦芽糖存在。另外,噬菌体吸附作用还需要有镁离子的帮助,而温度对吸附作用无明显影响。 λ 的宿主细胞需要在含有麦芽糖作碳源的培养基中生长,但是培养基中不能含有葡萄糖,以防止分解代谢物的

阻遏作用。至于 λ DNA 的转染作用,即将 λ DNA 直接导入宿主细胞的过程,和细胞内膜的成分有关。温度对转染作用有明显影响,37°C 时的转染效率高,降低温度,转染作用就减慢。

不论是噬菌体的感染,还是用 DNA 直接转染宿主细胞,引入细胞内的线性 DNA 分子首先要形成共价闭合环形分子。这种环化作用是借其自身的 *cos* 位点及宿主细胞内的连接酶(大肠杆菌 DNA 连接酶)而实现的。环化的 DNA 分子随即开始其自身复制以及转录。

λ 噬菌体是温和型噬菌体,它的繁殖有裂殖型和溶源型两种方式。感染开始的 15 分钟内,呈共价闭合环形的 λ DNA 分子进行双向复制(θ 型复制),产生环形 DNA 拷贝。以后进入滚筒式复制,产生串联状线形多聚分子。同时,随着 DNA 的转录作用由早期进入晚期;大量的噬菌体外壳蛋白被合成,线形的多聚分子即可在宿主细胞内包装成成熟的噬菌体颗粒,大量噬菌体的形成最终导致宿主细胞的裂解。这就是 λ 噬菌体的裂解型繁殖。当裂解繁殖在液体培养基中进行,原来有大量宿主菌的培养液会渐渐变得清亮,并出现了细胞裂解的残片;琼脂平板上的感染作用表现为形成透亮的噬菌斑。一个噬菌斑含有大量的噬菌体颗粒(一般认为有 10^6 噬菌体),正象一个单菌落包含有许多细菌数一样。同一噬菌斑内各噬菌体 DNA 分子的组分结构是一致的,因为它们都属同一祖先的后代。由重组噬菌体感染在平板上形成的噬菌斑,同一噬菌斑内噬菌体含有相同的重组 DNA 分子,不同的噬菌斑表示含有不同的重组 DNA 分子。噬菌体裂解细菌的现象与宿主细胞有关。

(一) 宿主细胞的选择

大多数的 λ 噬菌体宿主菌是由大肠杆菌 K12 衍生来的。K12 菌通常能产生限制性内切酶 *EcoK*。如果在噬菌体 DNA 分子上含有 *EcoK* 的识别序列,而且此序列未被 *EcoK* 甲基化酶修饰,那末,当它感染 K12,进入细胞内的噬菌体 DNA 就会被 *EcoK* 酶切割。然而,在具有修饰体系(即有 *EcoK* 甲基化酶)的宿主中繁殖的噬菌体,再去感染含 *EcoK* 酶活性的宿主细胞,噬菌体 DNA 不会被 *EcoK* 降解。对于重组噬菌体 DNA,即使载体两臂是来自经过修饰后的噬菌体 DNA,而插入的外源 DNA 片段中含有 *EcoK* 的识别序列,重组 DNA 分子同样会受到降解。含有 *EcoK* 限制酶活性的菌株以 rk^+ 表示,丧失此酶基因为 rk^- 菌株。采用 rk^- 菌株作宿主细胞,重组噬菌体 DNA 将不受影响。缺失了限制体系而保留有修饰体系的菌株是以 $hsdR^-$ 表示, $hsdS^-$ 则是限制和修饰体系均为缺陷的菌株 (rk^- , mk^-)。

宿主细胞的 rk^- 表型比较容易鉴别。其表现为对甲基化和非甲基化的两种噬菌体的感染效率不显差异。而 rk^+ 的细菌,非甲基化的 λ 形成噬菌斑效率会降低三个数量级 (10^3)。在基因文库构建中,选用的宿主菌都是 rk^- 的基因型。

下面介绍宿主细胞的遗传重组体系。

大肠杆菌 K12 可以发生 DNA 的同源重组 (homologous recombination) 现象。这样的现象可以有不同的重组通路(包括 *RecF*, *RecBC*, *RecE*) 来实现。在野生型 K12 菌中,主要的重组通路是 *RecBC*,它依赖于基因 *recBC^+* 的产物。基因型 *recB^-*, *recC^-* 或者 *recB^- recC^-* (也可写成 *recBC^-*) 的细菌都是属于缺失 *RecBC* 重组通路的突变株。

sbc 是 *recB* 和 *recC* 的校正基因, *sbcB⁺* 的表型表示 *RecBC* 和 *RecF* 两条重组通路是通的, *sbc⁻* 是表明上述两条重组通路之间的通道被阻断。使宿主细胞的重组作用只能由 *recF* 基因产物来催化。依赖于 *recBC⁺* 和 *recF⁺* 产物的两条重组通路 (*RecBC*, *RecF*) 同样取决于 *recA* 基因产物, 因此 *recA⁻* 突变株同时丧失上述两条重组通路。*RecE* 是宿主细胞的另外一条独立的重组通路, 从生化和遗传角度看, 在许多方面此重组通路和噬菌体仅有的 *Red* 重组通路相类似。

噬菌体 DNA 在宿主细胞内出现的重组现象, 除了与 λ 噬菌体本身的 *Int* 和 *Red* 重组系统有关外, 同时受到宿主细胞重组系统的影响。噬菌体 DNA 的复制既与其自身的分子重组密切相关, 也与宿主细胞的重组系统有关系。 λ 噬菌体能在宿主细胞内进行裂殖型生长, 必要的一步是 DNA 要由 θ 型复制进入滚筒式复制, 合成串联状的线形多聚分子, 以提供体内噬菌体颗粒包装所需的 DNA。宿主细胞的外切酶 *V* 能阻止 λ DNA 进入滚筒式复制。大肠杆菌外切酶 *V* 是其细胞 *recB* 和 *recC* 基因的产物, 因此在 *recBC⁺* 表型的大肠杆菌中, 噬菌体 λ DNA 难于进入滚筒式复制, 噬菌体颗粒的繁殖较困难。通常情况下, 野生型 λ 噬菌体具有 *gam* 基因, 在被感染的宿主细胞中, λ 的 *gam* 基因可被表达 (即为 *gam⁺* 表型), 其产物可与宿主产生的外切酶 *V* 相结合, 使酶失活, 保证了 λ DNA 可从 θ 型复制进入滚筒式复制, 噬菌体可以正常繁殖。这表明噬菌体为 *gam⁺* 表型时, 它可在 *recBC⁺* 宿主细胞中繁殖。但对于 *recBC⁻* 宿主细胞, 噬菌体的 *gam⁺* 表型并非必要, *gam⁻* 的噬菌体 DNA 可以正常复制。*gam⁻* 的噬菌体在 *recBC⁺* 的宿主中繁殖能力很低, 而其能存活, 靠的是单分子环形 DNA 重组成二聚和多聚体的功能, 因为多聚的环形分子也可以在体内被包装。

宿主的重组体系, 除了对噬菌体 DNA 的复制以及噬菌体的繁殖有影响外, 还会引起噬菌体 DNA 的同源重组。后者对已构建的基因文库中重组噬菌体的稳定性相当不利。此种同源重组会影响到插入的外源 DNA 的原有核苷酸序列, 给基因结构分析造成假象和错误。因此在进行基因文库的扩增和基因筛选时, 以及在大量制备重组体 DNA 时, 要考虑选用合适的宿主细胞。表 6.1 列举了一些常用的宿主细胞供参考。在用 λ 为载体构建基因文库中, 为减少同源重组, 应选择 *Rec⁻* 的宿主细胞。不过采用这样的宿主, 噬菌体的产量较低。其原因, 一是由于噬菌体颗粒释放减少, 再是在 *Rec⁻* 细胞的培养液中死细胞的比例高, 噬菌体吸附于死细胞上后就丢失。但为了使重组体在繁殖中能维持稳定不发生重组现象, 往往只能如此。

λ EMBL3 类的载体的特性是不能在含 P2 的溶源菌中繁殖, 这种表型称之为 *Spi⁺*, 它是因为噬菌体载体中 *gam* 和 *red* 基因表达产物所致。 λ EMBL3 是替代型载体, 当它和外源 DNA 重组后, 重组体为 *gam⁻* 和 *red⁻*, 因为 *gam* 和 *red* 基因定位于载体的中央片段, 重组时外源 DNA 替代了中央片段, 重组体即失去 *gam* 和 *red* 基因。它呈显 *Spi⁻* 表现型, 能在 *recA⁺* 的 P2 溶源菌 (例如 359) 中生长。因此, 在基因文库构建中选用 *recA⁺* 的 P2 溶源菌为宿主菌, 可以选择性繁殖 *Spi⁻* 的重组噬菌体, 除去 *Spi⁺* 表型的原载体噬菌体。但是 *Spi⁻* 重组体在 *recA⁺* 细胞中繁殖时, 由于宿主重组系统的影响, 会引起 DNA 的重排, 并且形成的噬菌斑很小。*gam⁻red⁻* 的重组噬菌体也可在 *recBC⁻* 的菌株中生长, 可允许噬菌体 DNA 进入滚筒式复制, 形成大的噬菌斑。但噬菌体生长较慢, 而且不能选用 P2 溶源菌的选择特性。为分离 *Spi⁻* 重组体 DNA, 可以选用由

表 6.1 常用宿主细胞的基因型及其用途

菌株	表型	有关的基因型	用途
C600		supE (tonA lacY)	滴定和繁殖 λ 噬菌体用
ED8654 ^a	rk ⁻ mk ⁺	supE supF hsdR	同 C600, 并是良好的转染用宿主菌
WA803	rk ⁻ mk ⁻	supE hsdS	同 C600, 并是良好的转染用宿主菌
NM531	rk ⁻ Rec ⁻	supE supF hsdR recA	同于 ED8654, 但为 Rec ⁻ 菌株
K802	rk ⁻ mk ⁺	supE hsdR	su ⁺ 菌株, 用于 λ 和重组噬菌体的繁殖
Q538	rk ⁻	supE hsdR	su ⁺ 菌株, 适合于 gam ⁺ 噬菌体 (λ 1059)
Q539	rk ⁻ (P2)	supE hsdR (P2)	su ⁺ 菌株, 为选择 Spi ⁻ 重组噬菌体用
Kro		recA	适合于繁殖 λ BF101 载体
546	rk ⁻ mk ⁻ Rec ⁻	supE hsdR recBC	RecBC ⁻ 宿主, 为 Spi ⁻ 噬菌体用
GC508	rk ⁻ Rec ⁻	supE supF hsdR recA/pgam	RecA ⁻ 宿主, 为 Spi ⁻ 重组噬菌体用
CES201	rk ⁻ Rec ⁻	hsdR recBC sbc15 recA	RecA ⁻ 宿主, 为 Spi ⁻ 重组噬菌体用
CES200	rk ⁻ RecF ⁺	hsdR recBC sbcB15	RecF ⁺ , 为 Spi ⁻ 重组噬菌体用
NM519	rk ⁻ RecE ⁺	hsdR recBC sbcA23	RecE ⁺ , 为 Spi ⁻ 重组噬菌体用

a. LE392, EQ82, BHB2600, BNN45 都是 ED8654 株的后代(本表资料主要来自参考文献 [1])。

recBC⁻, sbcA⁻ 或 recBC⁻, sbcB⁻ 衍生得到的 Rec⁺ 的大肠杆菌作宿主菌, 不论在液体培养液或琼脂板上繁殖, 都可得到高滴度的裂解液。

由 Charon 34 和 35 载体得到的 red⁻ gam⁺ 重组体能在严格的 Rec⁻ (recA⁻) 条件下繁殖。red⁻ gam⁻ 重组体也可在 Rec⁻ 条件下繁殖, 省去了遗传性选择。若使宿主带有含 λ gam 基因的质粒, 当引入噬菌体, 即可激活插入质粒上的 gam 基因表达, 则 recA⁻ recB⁻ 的宿主细胞同样可以采用。因为 recBC 的产物 (外切酶 V) 可被质粒上 gam 基因产物所灭活。另一种是用 recBC⁻ 或 recA⁻ recBC⁻ 的宿主, 而其不含带 gam 基因的质粒。必须注意, 在采用由质粒 DNA 提供放射性探针筛选基因文库中克隆时, 不能选择含 λ gam 基因片段质粒的宿主细胞。因为作探针的质粒 DNA 与宿主细胞内含 gam 基因质粒往往有同源性, 从而出现假阳性。

(二) 噬菌体的感染实验

正如前面介绍的, 为噬菌体感染用的宿主细胞, 需要在含麦芽糖作碳源的培养基 (例含 0.2% 麦芽糖的 LB 或 NZC 培养基) 中生长。在进行噬菌体感染实验之前, 先要培养和处理宿主细胞。由单菌落培养 5 至 50 毫升过夜菌, 测 OD₆₀₀, 培养液 4℃ 离心, 弃上清液, 菌体悬浮于 10mmol/L MgSO₄ 中, 使 OD₆₀₀ = 2.0OD (相当于 1.5 × 10⁹ 细胞/ml)。

取菌液 0.1 或 0.2ml⁻¹, 加于 10ml 带塞无菌管中, 再加入已包装的噬菌体 5—10 μ l (含适量噬菌体), 在 37℃ 或室温下保温 15—20 分钟, 以让噬菌体吸附于菌体表面。然后加入预热至 47℃ 左右的上层琼脂 (0.6% 琼脂的培养基) 3ml, 快速混匀后倒入琼脂板 (1.5% 琼脂), 室温下凝固上层琼脂, 然而在 37℃ 下培养过夜。一般 7—8 小时即可出现噬菌斑。

用琼脂板滴定噬菌体数目时, 先得把噬菌体适当稀释, 再行感染, 以获得可分辨易数的噬菌斑 (100 至几百个单一噬菌斑)。

噬菌体感染也可以在液体培养基进行 (见参考文献 [2]), 尤其是要大量制备重组体 DNA 时。

四、基因文库的构建规模

构成基因文库重组群体中的每一重组噬菌体就是一个克隆。它包含着该生物体基因组中某一特定的片段,那么要得到多少个不同的克隆才能具有该基因组的真正代表性,即把各克隆的插入片段借重复的部分叠加起来正好包含了基因组的全部序列。所需克隆总数取决于外源基因组的大小以及切割成插入片段的大小。Clark 和 Carbon (*Cell*, 9: 91, 1976) 提出了一个计算公式:

$$N = \frac{\ln(1 - P)}{\ln(1 - f)}$$

P 为所希望的具有代表全基因组的可能性, f 是插入片段的大小对全基因组大小的比值, N 为必需的重重组体克隆的数目。例如,要达到基因文库对所克隆的基因组有 99% 的代表性,此时 $P = 0.99$, 待克隆的外源基因组片段的平均大小为 17kb, 哺乳动物基因组长度为 3×10^9 bp, 此时

$$N = \frac{\ln(1 - 0.99)}{\ln\left[1 - \frac{1.7 \times 10^4}{3.0 \times 10^9}\right]} = 8.1 \times 10^5$$

$N = 8.1 \times 10^5$, 意思是对于长度为 3×10^9 bp 的基因组,当切割成平均大小为 17kb 片段用于构建基因文库时,至少要得到 8.1×10^5 数目以上的重组体克隆,才能保证基因文库的代表性为 99%。按照此公式,可以计算出用不同来源的基因组构建基因文库时需要的最少克隆数目的理论值(表 6.2)。表中列举的数值仅供参考,此理论值是按供体 DNA

表 6.2 不同来源基因组在构建基因文库时需克隆数目(理论值)

基因组来源	单倍体基因组的大小 (bp)	插入片段为 20kb ^a 时需克隆数目	插入片段为 45kb ^b 时需克隆数目
大肠杆菌(细菌) (<i>Escherichia coli</i>)	4.2×10^9	9.2×10^5	4.3×10^5
酵母(低等真核细胞) (<i>Saccharomyces cerevisiae</i>)	1.4×10^7	3.2×10^5	1.4×10^5
果蝇(真核细胞) (<i>Drosophila melanogaster</i>)	1.4×10^8	3.2×10^5	1.4×10^5
海胆(真核细胞) (<i>Strongylocentrotus purpuratus</i>)	8.6×10^8	2.0×10^6	8.8×10^5
人类 (<i>Homo sapiens</i>)	3.3×10^9	7.6×10^5	3.4×10^5
六倍体小麦 (<i>Triticum aestivum</i>)	1.7×10^{10}	3.9×10^6	1.7×10^6

a. 片段大小适用于插入 λ 替代型载体。

b. 片段适用于插入 cosmid 载体(例 pJB-8)。

本表数值是按 Clark^a 和 Carbon 提出的公式计算得出(见正文)。

可被随机切割的前提下计算出来的。实际上,采用 Sau3AI 或 MboI 的部分降解作用并不能达到理论上的完全随机切割。为了使基因文库具有真实代表性,一般构建基因文库的实际克隆数应比理论值大,可以二倍或三倍于理论值,以至更高。目前由于技术上的成熟,已能获得优质的载体和外源插入片段,保证了有效的重组,再加上高效率包装抽提物

的提供,不需要更多地取材就可以构建出理想的基因文库。

五、基因文库的扩增和贮存

按上述公式构建的基因文库可以直接由体外包装得到的重组噬菌体来进行筛选,用放射性标记的 DNA 片段或寡核苷酸作探针,借噬斑原位杂交技术找出有兴趣的基因。构建成的基因文库也可以先进行扩增。所谓扩增,是将基因文库中的每一个重组噬菌体通过感染繁殖获得数万至百万倍数量的增加。例如原来为 1×10^5 重组噬菌体的基因文库,经扩增后噬菌体的总数可增加到 10^{11} — 10^{12} 个。扩增的好处是既可以长期保存,重复使用,又可提供更多种基因供筛选,尤其适用于染色体步移研究。但是基因文库的扩增也遇到某些问题。由于插入到载体的外源 DNA 序列的影响,使得不同重组噬菌体的生长繁殖速度不一致。生长慢的重组噬菌体形成的噬菌斑小,往往在扩增后被筛选得到的机率会低于那些生长正常的呈显大型噬斑的重组体,甚至生长过慢的噬菌体在扩增过程中就丢失了。这意味着扩增会减低基因文库的完整性和代表性。其次是在扩增过程中会出现重组体 DNA 分子内及分子间的重排,它是由同源重复序列间的重排作用引起的。分子间的重组是出现在高密度的噬菌体对宿主细胞的感染作用,其使同一宿主细胞被 2 个以上不同的重组噬菌体感染,如果不同的重组体外源插入片段之间存在有相同重复序列,它们就因同源重组作用而导致分子间重排。但是只要在扩增时注意了以下几方面,还是可以使不足之处减到最小程度。

(1) 选用 Rec^- 的宿主细胞。它可减少由重组体内的外源插入片段特有的重复序列引起的同源重组现象;

(2) 在扩增基因文库时,避免采用液体培养,减少噬菌体的竞争性生长。一般选用在琼脂平板上扩增,并且注意减小感染时的噬菌体密度,防止两个以上噬菌体感染同一宿主细胞。另外,铺于琼脂平板上的噬菌体数目不宜过大,在 15 厘米直径的平板上控制噬菌斑数目在 $2-5 \times 10^4$ 范围内;

(3) 严格掌握噬菌体在琼脂板上生长的时间, 37°C 培养 8—10 小时,不得超过。以避免生长速度不一而引起的噬菌斑大小差别过大; 同时也可减少两种以上不同重组噬菌体感染同一细胞的机率,防止分子间重组;

(4) 铺板时注意保持水平,否则会人为造成噬菌斑的大小不均;

(5) 采用新鲜培养的宿主细胞进行噬菌体的预吸收,以减少因噬菌体吸附于死细胞上而导致的丢失机会。

基因文库的扩增方法是:

(1) 选择适宜的宿主菌,在 LB 或 NZCYM 琼脂板上划线, 37°C 培养过夜,长出菌落。

(2) 接种一单菌落于 5ml 液体培养基中。培养基为 LB 或 NZCYM, 并含麦芽糖 0.2%。 37°C 振荡培养过夜。

LB 培养基(每升液体):

蛋白胨

10g

酵母抽提物	5g
NaCl	10g

溶于蒸馏水中,用 NaOH 调 pH 至 7.5,并加水至终体积为 1L。

NZCYM 培养基(每升液体):

NZamine (或酪蛋白水解物)	10g
NaCl	5g
酵母抽提物	5g
酪蛋白氨基酸	1g
MgSO ₄ · 7H ₂ O	2g

溶于蒸馏水中,用 NaOH 调 pH 至 7.5,并加水至终体积 1L。

LB 或 NZCYM 琼脂板为含琼脂 1.5% 的上述培养基,消毒后铺于培养皿中得到。

LB 或 NZCYM 的顶层琼脂为含 0.7% 琼脂的相应培养基。

(3) 按 1/100 接种量,在 100ml 液体培养基 (LB 或 NZCYM, 并含麦芽糖 0.2%) 中接入 1ml 过夜菌液,37°C 振荡培养至 OD₆₀₀ 读数为 0.5 左右,然后冷却至 4°C,进行低温离心(4°C),收集菌体,弃去上清液后,菌体悬浮于 1/10 原体积的 10mmol/L MgSO₄ 溶液中。

(4) 于 15--20ml 容积消毒试管的底部加入 0.2ml 上述菌液,并加噬菌体 2—5 × 10¹⁰ 重组噬菌体,摇匀后 37°C 水浴预吸附 20 分钟。按基因文库拥有噬菌体总数确定预吸收的管数。

(5) 预先将顶层琼脂融化并恒温于 47°C 水浴中保持熔融状态。把隔夜铺制的新鲜琼脂平板 (15cm 直径的大培养皿) 在 37°C 恒温箱中预热。然后在预吸附噬菌体 20 分钟后的试管中准时加入 6ml 顶层琼脂 (47°C 温度左右),混匀后倒入琼脂平板上,置培养皿于较水平的桌面上,待顶层琼脂凝固。30 分钟后,再把培养皿移至 37°C 恒温箱,培养 8—10 小时。不让噬菌斑长得过大。

(6) 把培养皿移至 4°C 预冷 30—60 分钟,然后再向每块琼脂板加 SM 缓冲液 10ml,于室温下轻摇 1 小时,或于 4°C 放置过夜。吸取上清液,4°C 离心弃菌体,取上清液至带塞玻璃三角瓶中,加氯仿使含量为 0.3%,即可在 4°C 贮存。若加二甲基亚砷达 7% 的终浓度,可贮存在 -70°C。

(7) 若扩增后噬菌体浓度达不到 0.5—1 × 10¹¹ pfu/ml, 为避免扩增噬菌体悬液体积过大,可以用直接超速离心的方法将噬菌体浓集。离心条件为 22,000r/min, 3 小时 (4°C) (Beckman 离心机 SW28 转头或者 SW41 转头)。噬菌体颗粒将浓集于离心管底部,逐层小心移去离心管内上部溶液,留下 1/5 至 1/6 原体积的底层噬菌体悬液,缓和而又充分地混匀,回收噬菌体,一般可浓缩五倍以上。上层溶液经检测,几乎无噬菌体颗粒。此外,也可用 CsCl 离心方法浓集。

一般说来,扩增后的基因文库在几年内可保持高的噬菌体滴度,而且大部分重组体克隆是稳定的。

在构建基因文库时,若发现体外包装的重组噬菌体在琼脂板上生长不佳时,最好不要进行扩增。可直接借噬菌斑原位杂交法进行筛选,因为此筛选法比较灵敏,即使阳性克隆为生长慢的小噬菌斑,同样可被检测。

第五节 考斯质粒为载体的基因文库

考斯质粒 (cosmid) 又称为装配型质粒, 是质粒载体的衍生物。它具有质粒载体的一般特征: 分子量小, 含有 DNA 复制起始点和选择标记, 可以在大肠杆菌中自行复制。同时此类质粒都含有 λ 噬菌体 DNA 的粘末端 (cos 位点) 的核苷酸片段, 提供了噬菌体 DNA 可体外包装成噬菌体颗粒所必须具备的核苷酸序列 (cos 位点结构)。因此, 当考斯质粒在合适的限制酶位点插入了足够大的外源 DNA 片段, 并连接成近似 λ DNA 分子大小的重组体时, 它可借载体上特异的 cos 位点结构在体外被包装成类似于噬菌体样颗粒, 这种颗粒会象 λ 噬菌体一样, 具有感染细菌的能力。但是重组体 DNA 进入大肠杆菌宿主细胞后, 由于分子本身不具有 λ DNA 复制时所必需的基因, 并不能按 λ DNA 的复制方式繁殖, 仍只能按质粒 DNA 的方式进行复制。重组 DNA 存在于转化细胞中, 按其质粒上的选择性标记来挑选含重组子的转化菌落。

一、考斯质粒作载体的优缺点

用考斯质粒作载体构建的基因文库, 其优点在于: ①可容纳更大的外源 DNA 片段 (30—45kb), 使建成完整基因文库所需的克隆数目减少, 随之带来的益处是可减少筛选时的工作量。尤其对染色体步移的研究, 可节省许多步骤; ②载体本身的分子很小, 一般在 5kb 左右, 这在分析插入片段的限制性内切酶图谱时, 减少了来自载体的干扰。而用 λ DNA 作载体, 其自身片段约占 30kb 左右, 在重组子酶谱分析时, 干扰必然大, 除非把插入的片段分离出来, 单独分析酶谱。

然而, 采用考斯质粒构建基因文库时也有缺点和困难之处: ①应用常规的原位杂交技术筛选克隆时, 阳性检出率低。因为菌落原位杂交的敏感度低于噬菌斑的原位杂交法, 尤其是携带了巨大插入片段的重组质粒, 复制困难, 拷贝数较低; ②重组效率低。 λ 载体的重组效率可达 10^5 — 10^6 克隆数/ μ g DNA, 而考斯质粒只有 10^1 — 10^2 克隆数/ μ g DNA。导致效率低的主要原因是很难提供极大分子量 (>100 kb) 的供体 DNA。若得不到高质量的供体 DNA, 在制备 Sau3AI 或 MboI 部分降解产物和分离 30—45kb 大小的片段时, 相当一部分长度合格的酶解片段的两端不具备相同的粘末端结构, 它们的一个末端保留了原来的不均一的末端形式, 从而使重组时有效的连接效率降低。鉴于技术上的难度, 用考斯质粒来构建基因文库不及以 λ 噬菌体作载体那么普遍。但是随着技术和方法学上的改进, 以及新克隆体系的产生, 考斯质粒将成为构建真核细胞基因文库的良好载体。

Dahl 等介绍了某些考斯载体的特性 (表 6.3)。其中 pJB-8 是一个分子量比较小 (仅 5kb) 的载体, 有氨基苄青霉素抗性 (Ap^r) 的选择标记, 常用插入位点是 BamHI, 可以与 Sau3AI 或 MboI 部分降解的外源 DNA 片段 (35—45kb) 相重组。

二、基因文库的构建过程

图 6.3 示意了用考斯质粒 (pJB-8) 构建基因文库的过程。质粒用 BamHI 切开, 并

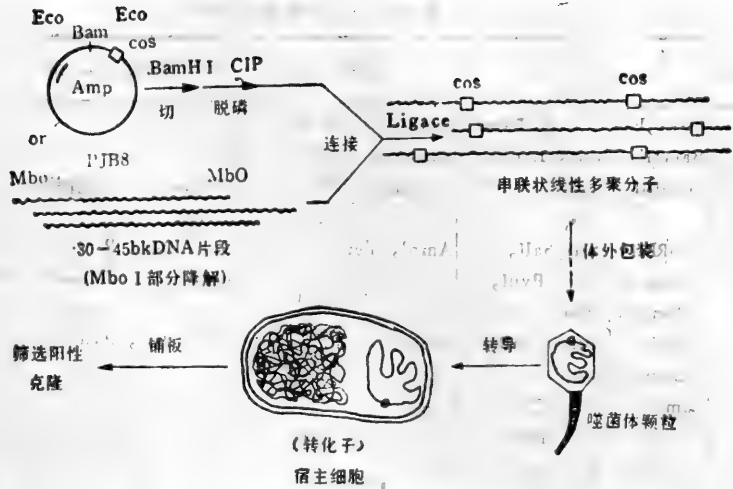


图 6.3 用考斯质粒构建基因文库。

考斯质粒 pJB8。Eco = EcoR I, Bam = BamH I, CIP 为牛肠碱性磷酸酶, cos, λ DNA 粘末端序列。

借碱性磷酸酶处理去除 5' 端磷酸单酯。外源 DNA 是借 MboI 酶的部分降解作用制备插入片段。详细操作步骤可参考文献[2]第 295—307 页的方法进行。构建基因文库的过程与用 λ 载体的步骤大致相似。所不同的,首先是载体的处理上比较简单。考斯质粒只需用限制性内切酶(例 BamHI)切成线性,毋需象 λ 载体那样进一步分离酶解产物,除去中央片段。为了减轻质粒的自身连接,酶切后质粒仍需用碱性磷酸酶(例如牛肠碱性磷酸酶, CIP)处理,去除粘末端 5' 磷酸单酯。在正式用质粒进行重组实验前,还应检测去磷酸单酯的效果。即取一定量去除了磷酸单酯的质粒 (0.5 μ g) 进行自身连接;另取同样量上述质粒,加入插入片段(例 λ DNA BamHI 降解产物)于相同条件下连接,用连接产物的转化效率检测质粒去磷酸单酯的效果。上述实验的对照为相同量的碱性磷酸酶处理之前的质粒限制酶解产物,观察自身连接效率。

其次,在插入片段的制备和分离上也稍有差别。制备供体 DNA(例如真核细胞的染色体 DNA)需特别小心,必须保证 DNA 分子尽可能大。因此,抽提 DNA 的条件要非常温和,减少大分子的机械性断裂;在用限制酶(Sau3AI 或 MboI)进行部分酶解作用时,酶反应要选择能产生较多 30—45kb 片段的条件;并在对这些片段进行分级时要谨慎操作,以减少片段的断裂。

第三,在连接重组时,质粒 DNA 的摩尔数要比插入片段高 10 倍以上,使插入片段最大限度地与质粒相连。同时,连接反应中 DNA 的总浓度要高于 200 μ g/ml,以便促其形成串联状线性多聚分子。连接产物的包装步骤与包装 λ 载体的重组 DNA 分子相同,但所用的包装抽提物中不能含有腐胺,以保证被包装的重组 DNA 分子中都连接有足够大的插入片段。因为不含有腐胺的包装抽提物,在包装时将对被包装的重组分子大小有明显的选择性,一旦重组 DNA 分子的长度为野生型 λ DNA 总长度的 80% 时,其包装效率将只有野生型 λ DNA 的 1/200 左右。

包装后的重组体颗粒可悬浮于 SM 缓冲液中,在 4 $^{\circ}$ C 保存。进行感染时,宿主细胞要选择 recA⁻ 的菌株,如 ED8767 或 DH1 的大肠杆菌株,以减少插入片段中因存在重

表 6.3 几种考斯质粒的特性

名称	克隆位点	选择标记	标记可被插入灭活的位点	分子大小 (kb)	克隆容量 (kb)
pJC 74	BglII, EcoRI, Sall, BamHI	Amp ^r , El imm	—	16	19—36
pJC75-58	BamHI, BclI, EcoRI	Amp ^r	—	12	23—40
pJC720	HindIII, XmaI	El imm, Rif ^r	Colicin (大肠杆菌素)产生: XmaI	24	11—28
pHC 79	EcoRI, HindIII, Sall, BamHI PstI, PvuI, XmaI, SphI	Amp ^r , Tet ^r	Amp: PstI, PvuI Tet: HindIII, Sall, BamHI, SphI	6	29—46
pJB-8	EcoRI, HindIII, Sall, BamHI, SphI	Amp ^r	Amp: PstI, PvuI	5	30—47
pOPF1	BamHI, KpnI, ClaI PvuI, BglII, HpaI	Amp ^r TK (在哺乳细胞中)	Amp: PvuI TK: BglII	8	27—44
MUA-3	PstI, EcoRI, AvalI, BclI PvuI, PvuII, XmaIII	Tet ^r	—	4.5	31—48

Amp^r, 氨基苄青霉素抗性; El imm, 大肠杆菌素的产生; Rif^r, 利福平抗性; Tet^r, 四环素抗性; TK, 胸腺嘧啶核苷激酶。

复序列而引起分子内的重排现象。感染宿主的过程基本上与 λ 噬菌体的感染相同, 仅有的差别是在 37°C 预吸附 20 分钟之后, 要加入 1ml 培养液 (例如 LB), 继续培养 45 分钟, 然后再铺板, 琼脂板应含有可供选择转化子的抗菌素 (例如 pJB-8 是用氨基苄青霉素)。过夜培养之后, 在平板上出现的是转化菌落, 不是噬菌斑。克隆的筛选和鉴定可参照菌落原位杂交的方法以及重组质粒的鉴定技术。

结 语

随着有效的 λ 噬菌体和考斯质粒新载体的构建、高效率包装噬菌体体系的出现以及分子克隆技术的改进和发展, 要构建某一生物体组织的基因文库已经不甚困难, 即使对于那些难以获得的材料如病变组织, 也可以在少量取材的情况下构建成满意的基因文库。

基因文库提供了构成基因组序列的全部克隆, 它们是研究基因结构及其表达的主要材料, 并通过染色体巡视分析, 提供了染色体基因组结构的重要信息。基因文库还为从生物进化角度研究基因结构以及生物体不同发育阶段的基因表达提供有意义的资料。对原核细胞和低等真核细胞来说, 基因结构比较简单, 它们的基因文库可为基因工程直接提供结构基因。而高等真核细胞, 尤其是人和哺乳动物细胞, 基因结构复杂, 一个结构基因往往可被几个插入序列 (内含子) 分割, 因此基因文库不能象 cDNA 库那样, 可直接为基因工程提供研究表达产物用的结构基因片段, 但是, 通过对其基因结构的分析, 可为该基因的体外表达和调控研究提供重要信息。

1985 年, 美国又出现了一种新的体外特异合成 DNA 的方法, 叫聚合酶链反应 (polymerase chain reaction, 简称 PCR)。该方法可以部分地代替基因文库。它是利用两个与待合成的 DNA 片段的 5' 端和 3' 端互补的短多聚核苷酸引物, 以全 DNA 为模板, 经过体外的 PCR 扩增方法, 可特异地合成一段 DNA。也有人将此方法称作无细胞体系的分子克隆 (cell-free molecular cloning)。该方法简化了分子克隆及筛选过程, 所以可快速

得到特异的 DNA 片段,目前在基因诊断上应用较广。有关此方法的文献,请参考 *Methods in Enzymology*. Vol 155, 335—350, 1987。

该方法虽有上述优点,但也有两个缺点:①它要求待合成的 DNA 片段的序列必须是已知的,至少要求 DNA 片段两端长约 20pb. 的序列是已知的,由此才可设计出 3'端及 5'端的引物序列,进行 PCR 反应;②该方法到目前为止最长只可合成长约 6 kb 的 DNA 片段,而用经典的基因文库的方法,我们最长可以得到长 20 kb 的片段。因此,该方法还需研究改进。不久的将来可望得以改进,用以取代基因文库的制作。应用 PCR 方法很重要的是需用一种耐热的 Taq DNA 多聚酶;另外,合成的引物必须很纯,否则不能得到特异的 DNA 链。

参 考 文 献

- [1] Kaiser, K. and Murray, N. E., The use of phage Lambda replacement vectors in the construction of representative genomic DNA libraries, in *DNA Cloning, a Practical Approach*, vol. 1, Ed. by Glover, D. M., 1—47, IRL Press, 1985.
- [2] Maniatis, T., Fritsch, E. F. and Sambrook, J., *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory, New York, 1982.
- [3] Hendrix, R. W., Roberts, J. W., Stahl, E. and Weisberg, R. A., *Lambda II* Cold Spring Harbor Laboratory, New York, 1983.
- [4] Dahl, H. H., Flavell, R. A. and Grosveld, F. G., The use of genomic libraries for the isolation and study of eukaryotic genes, in *Genetic Engineering*, vol. 2, Ed. by Williamson, R., 49—127, Academic Press, 1981.
- [5] Karn, J., Breaner, S. and Barnett, L., New bacteriophage Lambda vectors with positive selection for cloned inserts, in *Methods in Enzymology*, vol. 101, Ed. by Wu, R., Grossman, L. and Moldave, K., 3—19 Academic Press, 1983.
- [6] Hohn, B., In vitro packaging of Lambda and cosmid DNA, in *Methods in Enzymology*, vol. 68, Ed. by Wu, R., 299—309 Academic Press, 1979.
- [7] Collins, J., Escherichia coli plasmids packageable in vitro in Lambda bacteriophage particles, in *Methods in Enzymology*, vol. 68, Ed. by Wu, R., 309—326 Academic Press, 1979.
- [8] Davis, R. W., Botstein, D. and Roth, J. R., *Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory, New York, 1980.

第七章 cDNA 克隆

吴石君

cDNA 是互补于 RNA 的 DNA, 又称反向转录 DNA。cDNA 是指以 RNA (多为 mRNA) 为模板, 在逆转录酶作用下合成的互补 DNA 链。通过 DNA 克隆的方法, 把互补合成的 cDNA 链与载体 DNA 重组, 引入寄主细菌, 进行扩增的过程称为 cDNA 克隆。

由于 cDNA 是严格互补于模板 mRNA 核苷酸序列的。因此 cDNA 序列只反映基因表达的转录及加工后产物所携带的信息。也就是说 cDNA 序列只与基因的编码序列有关, 而不反映内含子 (intron)。它与相应的基因组比较, 提供了一种分析真核基因结构、组织和表达的强有力工具, 被广泛地用于以下几方面的研究。

(1) 真核细胞 mRNA 的结构和功能。一种特异的 mRNA 在细胞 RNA 中往往只占很小比例, 难以研究其序列、结构和功能。相应的 cDNA 则可以很方便地进行序列分析, 用这种途径已完成了许多特异 mRNA 的序列分析, 确定了 mRNA 起始、编码、转录与翻译的终止序列。

(2) 真核基因的结构和功能。cDNA 克隆与相应基因克隆的比较, 已确定了不少蛋白质, 如珠蛋白、胰岛素、生长激素的基因结构和含有插入序列的基因转录产物的加工, 以及免疫球蛋白基因的重排等。

(3) 分析了 RNA 病毒基因组的结构。如疱疹性口炎病毒、脊髓灰质炎病毒和流感病毒, 研究了这些病毒的结构、复制和表达, 并据此提出了发展抗病毒疫苗的方法。

(4) 研究真核细胞的染色质结构、组织和 DNA 的转录, 由于染色质结构在细胞分化、发育及受到激素等诱导时表现出不同的状态, 对脱氧核糖核酸酶 (如 DNase I、DNase II 及微球菌核酸酶) 表现出不同的敏感性。因而以 cDNA 为探针可检查经酶解的染色质 DNA 中的特异基因片段, 通过核小体结构变化, 了解染色质上某特异基因的活跃表达及不活跃状态, 并进而认识染色质结构变化对基因表达的调节作用。

(5) 从基因文库中分离特异的基因。在基因工程操作中, 常以 cDNA 为探针从基因文库中分离相应的基因组基因克隆。因而 cDNA 克隆常常是基因分离和结构分析的第一步。获得 cDNA 探针常常是许多分子生物学研究及基因工程操作的手头。

cDNA 克隆技术在现代生物学研究中占有显而易见的重要位置, 近十年来许多实验室致力于 cDNA 克隆技术的研究, 使此项技术日臻完善, 迅速简便。到目前为止已报道有数十个 cDNA 克隆成功, 为真核基因结构和表达的研究提供了一系列的探针。本章将介绍 cDNA 克隆的几种方法, 详细叙述具体的程序, 并对 cDNA 克隆中的诸因素进行讨论。

第一节 cDNA 克隆的一般方法和步骤

一、cDNA 克隆方法

所有 cDNA 克隆都包括 cDNA 合成、异源 DNA (cDNA 与载体 DNA) 重组、转化细菌和克隆筛选等步骤。目前广泛采用的方法有以下几种。

1. 双链 cDNA 合成-同源多聚物接尾重组法

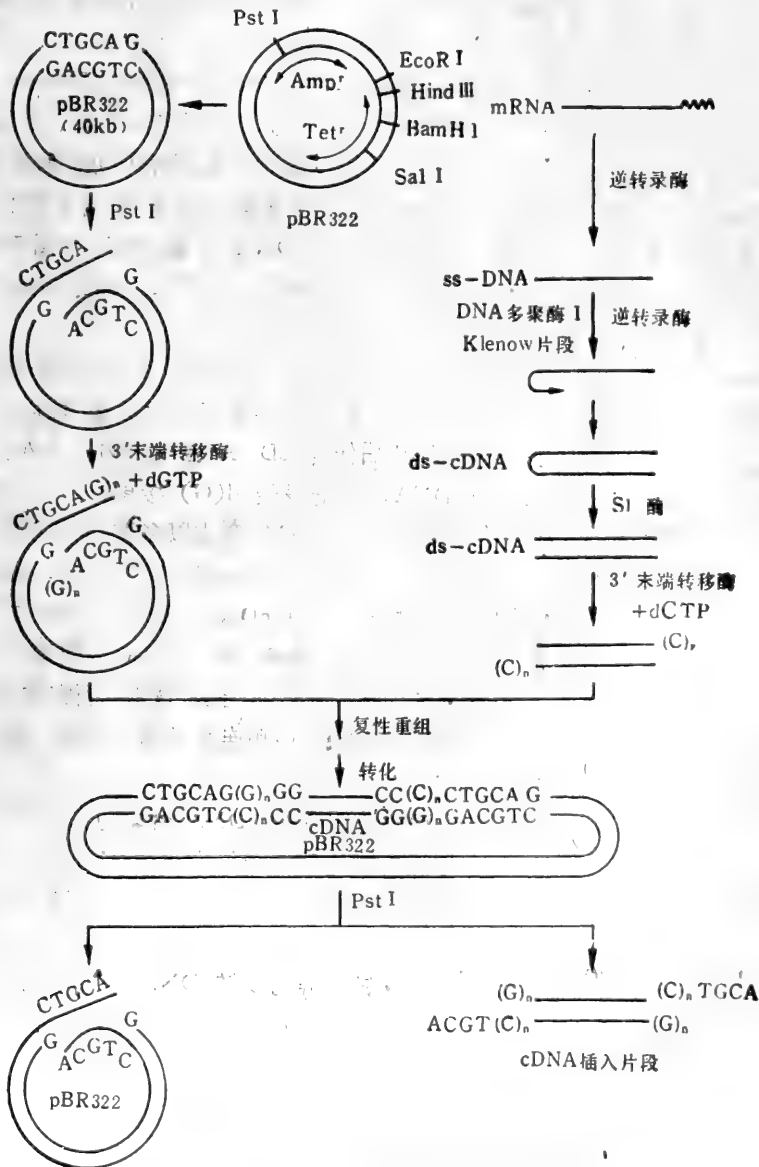


图 7.1 双链 cDNA 合成-同源多聚物接尾法示意图。

这是目前广泛采用的方法。它是多聚(A)RNA 为模板,经 AMV 逆转录酶催化,合成以寡聚 d(T) 为引物的单链 cDNA。然后由 DNA 多聚酶 I 或逆转录酶催化合成以第一条 cDNA 单链为模板的双链 cDNA, cDNA 的 3' 末端连接同源多聚 d(C), 而载体 DNA 由同源多聚 d(G) 接尾, 两种 DNA 由多聚 d(G) 与多聚 d(C) 的互补关系, 形成的粘末端重组, 转化细菌得到 cDNA 克隆。过程如图 7.1 所示。

2. 采用特异引物的 cDNA 合成和克隆法

此法是在单链 cDNA 合成中, 加入由人工合成的特异序列的寡聚脱氧核苷引物。引物的序列由蛋白质的氨基酸序列确定, 一般应首先确定某种蛋白质 C 端约 5-7 个氨基酸顺序, 然后合成相应引物。cDNA 合成时, 加入此引物, 使单链 cDNA 合成只能以具有与引物相应序列的 mRNA 作为模板, 而不需要分离特异 mRNA, 其他步骤仍可采用方法 1。

3. 人工合成连接子的 cDNA 与载体 DNA 连接方法

合成的双链 cDNA 先与第一个连接子连接, 由 DNA 多聚酶 I Klenow 片段催化; 然后用 S1 酶去除双链 cDNA 另一端的环形单链部分, 继而连接上第二个连接子。经

过第一、第二连接子的相应限制性内切酶消化后, cDNA 与载体 DNA (两端带有与 cDNA 两端连接子一致的粘末端) 连接重组建立 cDNA 克隆。程序如图 7.2 所示。

4. 用质粒作为引物的 cDNA 合成方法 (Okayama Berg)

如图 7.3 所示, 首先以限制性内切酶 KpnI 酶切质粒 DNA, 并以 3' 端接尾法连接上多聚 d(T), HpaI 酶切去一端经纯化后作为 cDNA 合成引物。另外, 以 PstI 酶切质粒 DNA, 3' 端多聚 d(G) 接尾, Hind III 酶切, 纯化后得到连接子。cDNA 合成以多聚 (A)-RNA 为模板, 加入质粒引物得到 cDNA-mRNA 杂交链, 再以 Hind III 酶切, 并通过连接子使 cDNA-mRNA 与质粒 DNA 连接成环形双链, 然后以 RNase H、DNA 多聚酶及大肠杆菌连接酶 I 去除 RNA 和置换合成 DNA 而完成重组。

目前已有质粒引物和连接子商品出售, 使这种方法简便易行。

5. 一种较为简单的 cDNA 合成方法

在方法 1 的基础上进行修改: 完成单链 cDNA 合成后, 不需去除 mRNA 链, 而在双链 cDNA 合成中同时加入 DNA 多聚酶 I、RNase H 和大肠杆菌 DNA 连接酶。反应中 RNA 链降解与双链 DNA 合成同时进行。cDNA 与载体 DNA 的连接重组、转化细菌步骤同方法 1。

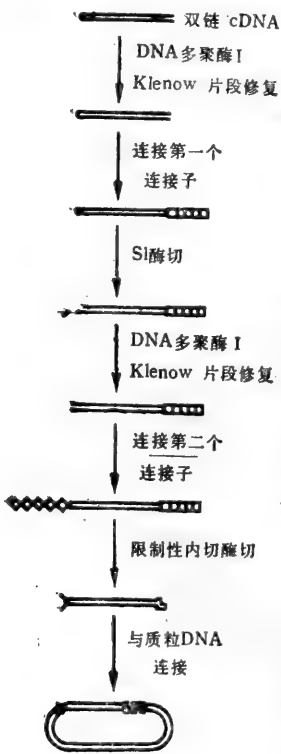


图 7.2 人工合成连接子的 cDNA 克隆方法示意图。

二、cDNA 克隆的具体步骤

下面以方法 1 为例叙述 cDNA 克隆的各具体步骤和操作。

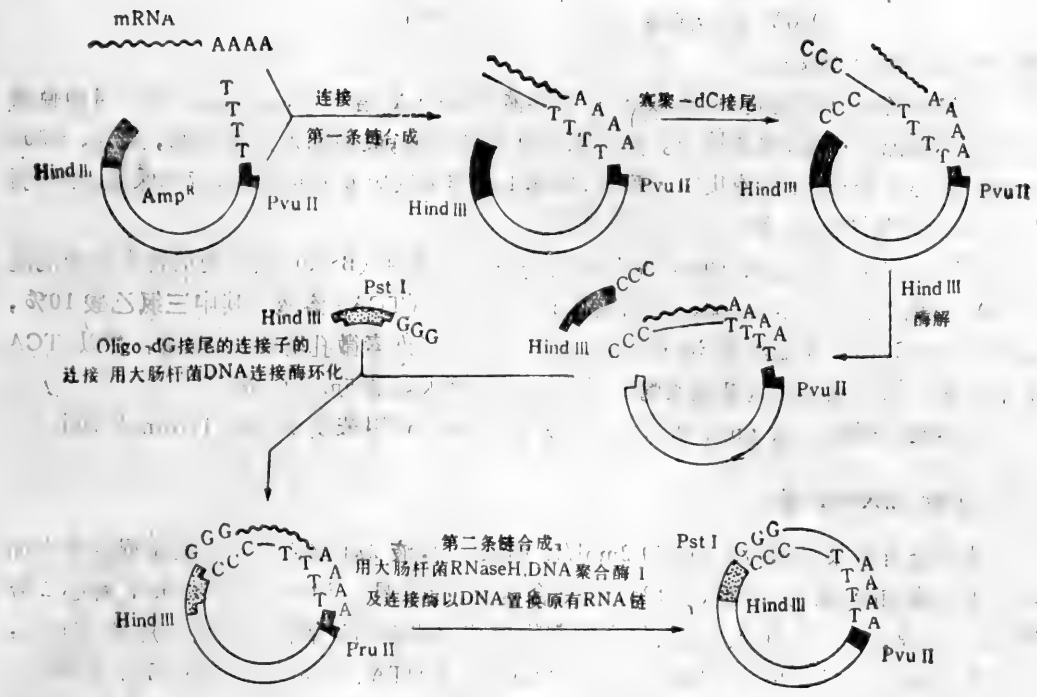


图 7.3 完整序列的 cDNA 克隆步骤(参考 Okayama 和 Berg)。

(一) 单链 cDNA 合成

1. 试剂

AMV 逆转录酶: 选用高比活性酶制剂, 其运输与保存应在 -20°C , 有效使用期为三个月。

低聚 d(T)₁₂₋₁₈: cDNA 合成引物, 结合于 mRNA 3' 端多聚(A)区域, 长度为 12—18 脱氧胸腺核苷。如使用合成的特异序列引物则不必加入此引物。

放线菌素 D: 抑制可能混杂在 mRNA 样中的 DNA 模板的活性。

标记化合物: 通常采用 ³H、³²P、³⁵S 标记 α 位的 dCTP。

2. 操作

反应混合物包括:

多聚 (A)-mRNA	1 μg
dATP、dGTP、dTTP	各 1m mol/L
dCTP	0.5mmol/L
Tris-HCl (pH8.3)	100mmol/L
MgCl ₂	10mmol/L
KCl	50mmol/L
DTT	10mmol/L
放线菌素D	50 $\mu\text{g/ml}$
低聚 d(T) ₁₂₋₁₈	200 $\mu\text{g/ml}$

^3H -dCTP	20 μCi
AMV 逆转录酶	100u

反应物总体积为 25 μl 。

事先将各成分配制成 10 \times 或 5 \times 贮存液。mRNA 溶于水,取 1 μg ,加入四种脱氧核苷酸及标记化合物,冰冻抽干。然后加入其他各溶液,最后加入逆转录酶,混匀,3,000 r/min 离心 1 分钟,使溶液集中于管底,抽样 1 μl 作为反应 0 分钟样品,37 $^{\circ}\text{C}$ 保温,并间隔 10—15 分钟取样 1 μl ,测定 ^3H -dCTP 掺入。

测定方法是:首先取 100 μl 0.4% 的牛血清白蛋白 (BSA) 置于测定管中作为沉淀 cDNA 的载体,加入样品 1 μl ,再加入 2ml 三氯乙酸 (TCA) 溶液 (其中三氯乙酸 10%, 焦磷酸钠 0.2%)。冰浴中置 30 分钟后,滴加于硝酸纤维素微孔滤膜上,抽滤,并以 TCA 溶液洗涤。加入水溶性闪烁液或烘干后加入脂溶性闪烁液 5ml,计数。

水溶性闪烁液: 4g ppo、0.05g popop 溶于 667ml 甲苯及 333ml Triton X-100。

(二) 去除 mRNA 链

选择恰当的保温时间,加入 0.2mol/L EDTA 溶液 2 μl 终止反应。反应混合物中加入等体积饱和酚-氯仿-辛醇 (25:24:1) 混合溶液,旋转混合 30 秒,5,000r/min 离心 2 分钟,取上清液,弃去蛋白质。为去除合成产物溶液中残余的脱氧核苷酸及小片段 cDNA,通过葡聚糖凝胶 G-100 柱层析。柱以 0.02mol/L EDTA、0.1mol/L Tris-HCl (pH7.1) 溶液平衡并洗脱,每管收集组分 0.2ml 左右。各收集组分取 5 μl 加入 5ml 水溶性闪烁液计数。第一洗脱峰为 cDNA,第二洗脱峰(滞留组分)是未掺入的核苷酸。将第一洗脱峰的各组分合并,得 cDNA-mRNA 杂交链。为去除 mRNA 链,在溶液中加入 3mol/L NaOH 至最终浓度为 0.3mol/L,60 $^{\circ}\text{C}$ 保温 1 小时,加入冰醋酸使溶液 pH7.0 (由于溶液体积很小,为避免损失和同位素污染,一般不用 pH 计测定反应溶液,而是通过一个同 cDNA-mRNA 溶液一致的空白对照管,加入 NaOH 至终浓度为 0.3mol/L,然后以冰醋酸调 pH7.0,记下加入冰醋酸的量,也就是应加入反应溶液中冰醋酸量)。随后加入大肠杆菌 tRNA 至 10 $\mu\text{g}/\text{ml}$,作为载体,加入 2 倍体积预冷乙醇沉淀过夜,16,000r/min 离心 30 分钟后,真空泵抽干,用于双链 cDNA 合成。

(三) 双链 cDNA 合成

可采用两种不同的反应体系,一为 DNA 多聚酶 I 体系,另一为逆转录酶体系。后一反应体系为:

单链 cDNA	
dATP、dCTP、dGTP、dTTP	各 1mmol/L
Tris-HCl (pH8.3)	100mmol/L
MgCl ₂	10mmol/L
KCl	50mmol/L
DTT	10mmol/L
逆转录酶	100u

反应体积为 50 μl , 37 $^{\circ}\text{C}$ 保温 2 小时,每间隔 30 分钟取样 1 μl ,用 S1 酶消化以测定双

链化程度。S1 酶反应体系为 (50 μ l 中):

NaAC-HAC (pH4.5)	30mmol/L
NaCl	300mmol/L
ZnCl ₂	7.5mmol/L
载体 tRNA	1 μ g
S1 酶	10u

室温下放置 30 分钟,然后加 100 μ l 0.4% BSA, 2ml 10% TCA 溶液, 硝酸纤维素滤膜过滤,液闪计数。保温结束时 cpm 数与 0 时 cpm 的比值即为双链 cDNA 合成效率。

双链 cDNA 反应混合物加入 2 μ l 0.2mol/L EDTA 终止反应。酚法去除蛋白,同样以 S1 酶反应体系消化去除单链区及两端环形区域, S1 酶消化后再以酚法去除蛋白,双链 cDNA 溶液加入 2 倍体积冷乙醇沉淀过夜。

(四) cDNA 合成的产量计算

计算 cDNA 合成的产量及效率是检定 cDNA 合成过程,并为单链 cDNA 合成后进行双链合成、3'端接尾和 DNA 重组提供计算依据的主要步骤。计算中已知的参数有: 1Ci ³H 标记化合物的 dpm (2.22 \times 10¹²); 液闪计数器效率,通过标准源测定得知; 标记化合物比活性; 脱氧核苷酸平均分子量 (按 350 计算); cDNA 中四种核苷酸掺入按各四分之一计算。

首先推知 25 μ l 或 50 μ l 反应混合物中 ³H-dCTP 的实际放射比活性。0.5mmol/L dCTP 在反应混合物中 mmol 数为:

$$\frac{0.5\text{mmol} \times 25\mu\text{l}}{10^6\mu\text{l}} = 1.25 \times 10^{-5}\text{mmol}$$

假定 ³H-dCTP 标记化合物比活性为 18Ci/mmol, 反应中加入 20 μ Ci, 标记化合物在总 dCTP 中所占 mmol 数忽略不计, ³H-dCTP 实际放射比活性为:

$$\frac{20\mu\text{Ci}}{1.25 \times 10^{-5}\text{mmol}} = 1.6\text{Ci/mmol}$$

因此新合成的 cDNA 中,每 1 μ g cDNA 应显示的 cpm 数为:

$$\frac{1.6\text{Ci/mmol} \times 2.22 \times 10^{12}\text{dpm/Ci}}{4 \times 350000\mu\text{g/mmol}} \times 40\% \frac{\text{cpm}}{\text{dpm}} = 2.54 \times 10^3\text{cpm}/\mu\text{g}$$

式中 40% 为液闪计数器的计数效率,如用 ³⁵S 或 ³²P 标记化合物则应改用相应数值。

根据新合成 cDNA 的计数即可推知所合成 cDNA 的微克数。

$$\frac{\text{cDNA 的 } \mu\text{g 数}}{\text{mRNA 的 } \mu\text{g 数}} \times 100\%$$

即为 cDNA 合成效率。

双链 cDNA 合成的有关计算原理与此相同。

(五) 3' 末端同源多聚物接尾及 DNA 重组

将含有双链 cDNA 的乙醇溶液离心去除乙醇,抽干,加入接尾反应体系,反应体系中含有:

二甲砷酸钾	0.2mol/L
Tris-HCl (pH6.9)	25mmol/L
DTT	2mmol/L
CoCl ₂	0.5mmol/L
BSA	250μg/ml
dCTP	1mmol/L
³ H-dCTP	10μCi
3' 末端脱氧核苷转移酶	10u

反应体积为 50μl, 37°C 保温 30 分钟, 0 分及 30 分各取样 1μl, 加入 BSA 及 TCA 测定酸不溶性成分 cpm。保温结束后, 加入 2μl 0.2mol/L EDTA 终止反应, 酚法去除蛋白, 葡聚糖凝胶 G-100 柱层析去除未掺入核苷酸, 加入 tRNA 至 1OD/ml 及 2 倍体积冷乙醇沉淀过夜。

载体 pBR322 DNA(经 PstI 酶切) 1μg, 接尾程序同 cDNA, 但 dCTP 及 ³H-dCTP 改为 dGTP 及 ³H-dGTP。

3' 端同源多聚物接尾 d(C), d(G) 数的计算如下。³H-dCTP 实际放射比活性为:

$$\frac{10\mu\text{Ci} \times 10^4\mu\text{l}}{1\text{mmol/L} \times 50\mu\text{l}} = 0.2\text{Ci/mmol}$$

首先根据总掺入的 cpm 推算 dCTP 掺入到 3' 端的 mmol 数, 每掺入 1 mmol dCTP 的 cpm 为:

$$0.2\text{Ci/mmol} \times 2.22 \times 10^{12}\text{dpm/Ci} \times 40\% \frac{\text{cpm}}{\text{dpm}} = 1.78 \times 10^{11}\text{cpm/mmol}$$

根据实际 cpm 计数即可推知掺入到 3' 端的 dCTP mmol 数。

然后根据 cDNA 的 μg 数及其分子量计算出 cDNA 的 mmol 数。

$$\frac{\text{dCTP 掺入的 mmol 数} \times \frac{1}{2}}{\text{cDNA mmol 数}}$$

即为每个 cDNA 分子 3' 端 dCTP 接尾的平均数值, 式中乘以 $\frac{1}{2}$ 是由于 cDNA 分子有两个 3' 末端。pBR322 DNA 3' 端 dGTP 接尾的计算类同。一般以 3' 端各接尾 20—30 d(C) 或 d(G) 为宜。

多聚 d(C)-cDNA 和多聚 d(G)-pBR322 DNA 的重组采用退火复性法。cDNA 和 pBR322 DNA 按 mmol 比值 1:4 的比例混合, 溶于 100μl 含有 0.1 mol/L NaCl、10mmol/L Tris-HCl (pH7.8)、0.1mmol/L EDTA 的溶液。65°C 保温 30 分钟, 并在 12 小时或更长时间内降至室温, 用于转化细菌。

(六) 细菌转化及克隆筛选

取受体菌—接种环, 接种于 LB 液体培养基 (1,000ml 中含有 10g NaCl、10g 水解酪蛋白、5g 酵母提取物, pH7.5), 37°C 培养过夜。取 5ml 培养物接种于 100ml LB 培养基中, 37°C 培养直到 600nm 光吸收为 0.6—0.7OD。取 6ml 新鲜培养物, 5,000r/min 离心 10 分钟, 菌体悬浮于 3ml 预冷的 50mmol/L CaCl₂ 中, 冰浴振荡 20 分钟, 再离心, 菌体悬浮于 0.2ml CaCl₂ 溶液中, 加入复性重组的 DNA 20μl。冰浴振荡 60 分钟, 转入 37°C 保

温5分钟,加入0.5ml LB培养基,37℃保温30分钟。涂布于含有15μg/ml四环素和1%琼脂的LB平板培养基,培养过夜。抗四环素菌落再以含50μg/ml氨苄青霉素的培养基筛选。挑出只在含四环素培养基生长而不在含氨苄青霉素培养基生长的菌落,进一步转入液体LB培养基扩大培养,以提取质粒。

(七) 重组质粒 DNA 的抽提和鉴定

经筛选得到的抗氨苄青霉素基因插入失活的细菌株,接种于10ml LB培养基,培养过夜。按如下程序提取质粒DNA:培养物8,000r/min离心10分钟收集菌体,菌体悬浮于100μl 25%蔗糖、50mmol/L Tris-HCl (pH8.0)、1mmol/L EDTA溶液中。再加100μl含有5mg/ml溶菌酶的上述溶液,室温放置5分钟,加入200μl 0.2mol/L NaOH、1% SDS溶液,冰浴放置10分钟,加入200μl 5mol/L NaAC (pH4.8),水浴置10分钟,离心10,000r/min 10分钟去沉淀。取上清,加入等体积饱和酚-氯仿-辛醇(25:24:1)溶液,提取二次去除蛋白质,取上清液加入2倍体积冷乙醇,沉淀过夜。

8,000r/min离心10分钟,去除乙醇,抽干,将DNA溶于含有50mmol/L NaCl、10mmol/L Tris-HCl (pH7.5)、10mmol/L MgCl₂、1mmol/L DTT的酶解缓冲液中,加入PstI限制性内切酶10u,37℃保温30分钟,置沸水浴中2分钟,用于琼脂糖凝胶电泳分析,以确定cDNA插入片段的大小。

第二节 特异 cDNA 的克隆

为获得与特异 mRNA 分子互补的 cDNA 克隆,常采取的方法有:①分离特异的 mRNA,并以此 mRNA 分子为模板,得到特异的 cDNA 克隆;②从某种组织总 mRNA 为模板得到的 cDNA 文库中分离特异的 cDNA 克隆;③首先分离纯化蛋白质,测定其氨基酸序列,根据氨基酸序列推知15—20个相应的脱氧核苷酸序列,以此合成一段寡聚脱氧核苷序列作为引物,加入到提取的总 mRNA 中合成 cDNA,得到特异的 cDNA 克隆。下面详细介绍由 cDNA 文库分离特异 cDNA 克隆的方法。

一、插入到表达载体中的 cDNA 克隆的筛选

在建立 cDNA 文库时,选用可使插入 DNA 表达的载体(如 λgt11 等)。在建立 cDNA 文库后,分组逐一鉴定各 cDNA 的表达产物(多肽)。常用的鉴定方法是免疫反应鉴定法,即以某种蛋白质的抗体寻找相应的特异 cDNA 克隆。此方法的细节已在有关章节中叙述。

二、选择杂交法筛选特异 cDNA 克隆

将 cDNA 文库的各菌落每十个一组,混合培养于10ml LB培养基,37℃过夜。采用前面叙述的质粒提取方法得到带有 cDNA 插入片段的质粒 DNA。DNA 溶于20μl 水,100℃煮沸10分钟,置-20℃速冷变性。加入20μl 1mol/L NaOH,室温放置20分钟,使

其中混有的 RNA 降解。加入 $10\mu\text{l}$ 1mol/L NaCl、 0.3mol/L 柠檬酸钠、 0.5mol/L Tris-HCl (pH8.0) 溶液及 $10\mu\text{l}$ 1mol/L HCl, 置冰水中。取 3mm^2 大小的硝酸纤维素滤膜, 置于蜡膜上, 滴加 $5\mu\text{l}$ DNA 溶液, 吹干, 再滴加, 再吹干, 直至加完所有样品。滤膜空气干燥, $6\times\text{SSC}$ 溶液洗两次, 室温干燥后, 80°C 真空干燥 2 小时。滤膜置于 100°C 沸水中 1 分钟, 再置于冰水中振荡摇动 3 分钟, 以除去未吸附成分。然后将滤膜置于 $50\mu\text{l}$ 多聚 A-RNA 杂交液中, 其中包含:

多聚 A-RNA	500 $\mu\text{g/ml}$
甲酰胺(去离子)	65%
PIPES pH6.4	20mmol/L
SDS	0.2%
NaCl	0.4mol/L
牛肝 tRNA	100 $\mu\text{g/ml}$

50°C 保温 3 小时, 吸干杂交液, 加入 1ml 预热至 65°C 的漂洗液, 内含:

Tris-HCl (pH7.6)	10mmol/L
NaCl	0.15 mol/L
EDTA	1mmol/L
SDS	0.5%

旋转振摇 30 秒, 吸去漂洗液, 如此重复 10 次, 再用不含 SDS 的上述漂洗液洗两次。滤膜置于含有 $30\mu\text{g}$ 的 $300\mu\text{l}$ 牛肝 tRNA 中, 煮沸 1 分钟, -20°C 速冷, 从溶液中取出膜, 弃去。在溶液中加入 $6\mu\text{l}$ 2mol/L NaAC (pH5.2) 及 1ml 冷乙醇, -20°C 沉淀过夜, 离心去乙醇, 抽干, 溶于水。各组质粒 DNA 经与多聚 A-RNA 杂交所得 mRNA, 用于体外翻译产物鉴定。以确定哪一组含有所需的特异 cDNA 克隆。然后再将这一组的十个菌落逐一进行选择杂交, 以得到特异 cDNA 克隆。

三、cDNA 克隆操作实例

氨甲酰磷酸合成酶 I (CPS_I) 是尿素合成的关键酶, 在生物体氮代谢中具有重要作用, 主要存在于肝脏, 是组织特异性酶。为研究 CPS_I 基因的表达和调节, 中国医学科学院基础医学研究所生物化学及分子生物学研究室完成了 CPS_I cDNA 的克隆。下面叙述建立 CPS_I cDNA 克隆的具体操作, 作为特异 cDNA 克隆的一个实例。

(一) CPS_I mRNA 的分离

取大鼠肝组织 30g, 生理盐水洗二次, 剪碎, 置聚四氟乙烯玻璃匀浆器中, 加入预冷的提取缓冲液 150ml, 其中包含 50mmol/L Tris-HCl (pH7.5)、 25mmol/L NaCl、 5mmol/L MgCl₂、 0.25mol/L 蔗糖、 0.2mg/ml 肝素及放线菌酮 $1\mu\text{g/ml}$ 。电动匀浆上下六次, $14,500\text{r/min}$ 离心 10 分钟, 于上清液中加入 1/10 体积的 10% 脱氧胆酸钠和 10% Triton X-100 溶液, 摇匀后铺于预先制备好的蔗糖阶梯梯度溶液表面 [梯度为 5ml 2.5mol/L 蔗糖、 1.5ml 2.0mol/L 蔗糖, 其中还含有 2.5mmol/L Tris-HCl (pH7.5)、 150mmol/L NaCl、 5mmol/L MgCl₂。选用 12.5ml 离心管, 每管铺加肝匀浆液 8ml]。 $20,000\text{r/min}$

(SW-28 转头)离心 16 小时, 离心后取出 2.5mol/L 蔗糖溶液, 其中含有细胞质多聚核蛋白体, 加入等体积的 2.5mmol/L Tris-HCl (pH7.5)、150mmol/L NaCl、5mmol/L MgCl₂、0.1% NP-40、1μg/ml 放线菌酮及 0.2mg/ml 肝素溶液, 并在此溶液中透析 20 小时。透析后贮存于 -80℃。

将多聚核蛋白体于冰水中解冻, 12,000r/min 离心 10 分钟, 去除不溶物, 加入抗 CPS₁ IgG (每克多聚核蛋白体中加 160OD₂₈₀IgG), 于 4℃ 反应 1 小时, 然后通过蛋白质 A-琼脂糖柱层析分离, 平衡及洗脱缓冲液为 2.5 mmol/L Tris-HCl (pH7.5)、150mmol/L NaCl、5mmol/L MgCl₂。流速为 4ml/小时。洗脱柱滞留成分的溶液是 25mmol/L Tris-HCl (pH7.5)、20mmol/L EDTA、0.2mg/ml 肝素。流速为 10ml/小时。收集滞留洗脱组分, 加入等体积 1mol/L NaCl 及 1% SDS 溶液, 然后进行寡聚脱氧胸腺核苷纤维素柱亲和层析。柱以 0.5 mol/L NaCl、0.5% SDS、10mmol/L Tris-HCl (pH7.5) 溶液平衡并洗脱非吸附成分。用 10 mmol/L Tris-HCl (pH7.5)、0.5% SDS、1 mmol/L EDTA 溶液洗脱多聚 A-RNA。收集含有寡聚 A-RNA 组分, 加入 NaCl 至 2% 及 2 倍体积冷乙醇, 沉淀过夜。12,000r/min 离心 30 分钟, 去除乙醇, 抽干, 溶于少量水, 即得到富含 CPS₁ mRNA 的溶液, 用作 cDNA 合成的模板。

(二) CPS₁ cDNA 克隆

取 CPS₁ mRNA 1μg, 按本章所叙述的单链 cDNA 合成方法进行, cDNA 合成保温曲线见图 7.4。

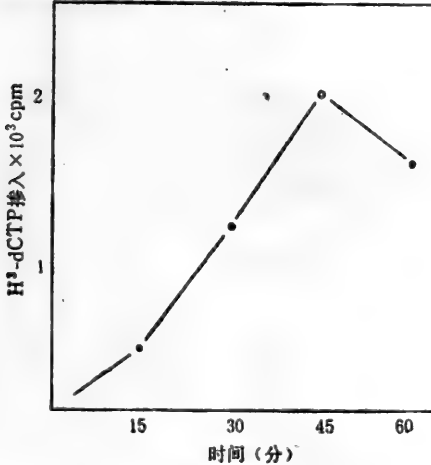


图 7.4 单链 cDNA 合成保温曲线。

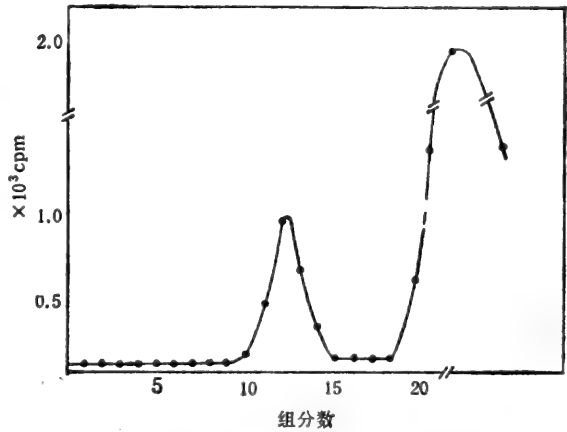


图 7.5 cDNA 的葡聚糖凝胶 G-100 柱层析分离。

如图所示, 保温 45 分钟可达到 dCTP 掺入高峰。单链 cDNA 的葡聚糖凝胶 G-100 柱层析分离结果见图 7.5。前一洗脱峰为 cDNA, 后一滞留洗脱峰为未掺入的脱氧核苷酸。

双链 cDNA 合成采用 AMV 逆转录酶催化, 合成保温曲线见图 7.6。

双链 cDNA 经 S1 酶解, 3' 端多聚 d(C) 接尾与 pBR322 质粒 DNA (经 PstI 酶切, 3' 端多聚 d(G) 接尾) 重组, 转化细菌 HB101。1μg mRNA 经此过程可得约 120 个转化细菌株。分成 12 组, 每组 10 株进行选择杂交, 筛选出 6 组菌株的质粒带有 CPS₁ cDNA 插

入片段。对其中一组的菌株逐一筛选得到一株菌落的质粒带有 0.8 kb CPS_1 cDNA 插入片段。图 7.7 是选择杂交的筛选结果。图左侧是各菌株质粒与多聚 A-RNA 杂交，并洗脱下 mRNA 进行体外翻译(在兔网织红细胞体外翻译体系中)的产物的电泳图谱。右侧是 ^{35}S -甲硫氨酸掺入到体外翻译产物中的放射自显影。

电泳采用 6% 聚丙烯酰胺凝胶、Tris-甘氨酸、SDS 缓冲液。样品 1—4 为与各质粒杂交并经洗脱下的 mRNA 体外翻译产物。5 为 CPS_1 (箭头所示)。右图为翻译产物电泳后的荧光自显影。

筛选得到的 CPS_1 cDNA 克隆，经 1% 琼脂糖凝胶电泳分析其 PstI 酶切图谱，得知插入片段为 0.8kb 左右(图 7.8)。

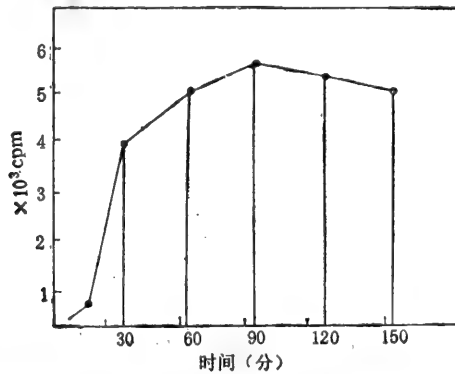


图 7.6 双链 cDNA 合成保温曲线。

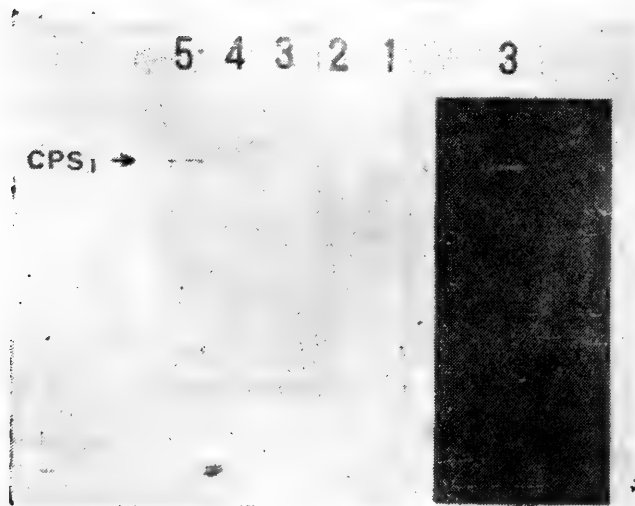


图 7.7 杂交选择法筛选 CPS_1 cDNA 克隆。

第三节 cDNA 克隆应注意的问题

一、cDNA 克隆所使用的酶制剂

在 cDNA 克隆的制备中使用了多种酶制剂，在运输、贮存和使用过程中都应注意保

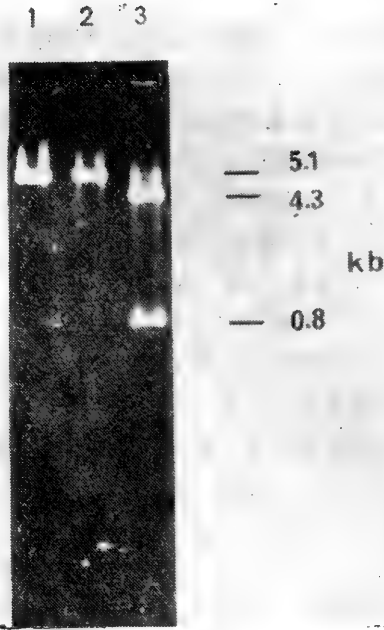


图 7.8 带有 CPS cDNA 插入片段质粒的确切图谱。

1. EcoRI 酶切。 2. Hind III 酶切。 3. Pst I 酶切。插入片段为 0.8kb。

持酶活性。一般应保存在 -20°C 条件下,实验中除保温进行酶促反应外,应在 0°C 冰水中进行。反复的冻融和其他溶液进入引起的污染都会导致酶活性的损失。其中几种关键的酶制剂及其保存、使用注意事项如下。

(1) AMV 逆转录酶。运输和贮存严格保持低温,有效使用时期为三个月。如贮存时间较长应在使用前测定酶活性。测定条件是加入多聚 r(A)-寡聚 d(T) 作为模板, cDNA 合成的反应系统, ^3H dTTP 为底物, 37°C 保温 30 分钟,测定酸不溶性成分 cpm,并计算酶活性单位。

(2) 3' 末端脱氧核苷转移酶。低温 (-20°C) 贮存,能在半年或更长时间内保持酶活性。3' 末端转移反应的缓冲液成分对保证接尾掺入脱氧核苷的数量有重要作用。可在订购酶制剂的同时订购相应缓冲液。

二、作为 cDNA 合成模板的 mRNA

mRNA 的纯度及分子的完整性直接影响 cDNA 合成的效率和 cDNA 的长度。采用任何一种成熟的分离纯化方法,只要操作正确都可得到用于 cDNA 克隆的 mRNA。在进行 cDNA 克隆操作之前,应对 mRNA 制剂进行分析鉴定。

1. mRNA 的分离及注意事项

用于分离 mRNA 的所有器械和溶液都应经过去除 RNase 的处理。器械处理方法是:洗液沉浸过夜,冲洗干净,高压蒸气灭菌或 180°C 烘烤 4 小时以上。所用溶液应尽量高压消毒(不能采用高压蒸气消毒的生物制剂,应以消毒重蒸水配制)。mRNA 分离操作

中动作要迅速,保持低温,配带手套。

为避免 RNase 的降解作用,常在溶液中加入 RNase 的抑制剂,如 RNasin、氧碲-核苷复合物等。RNasin 是一种由大鼠肝脏或人胎盘分离得到的小分子蛋白质,分子量为 40,000 左右。氧碲-核苷复合物是氧碲离子和四种核苷酸中的任意一种结合而成的复合物,可与 RNase 结合以抑制其活性。分离 mRNA 中还可采用高浓度的 (4mol/L) 异硫氰酸胍或其他胍盐溶液以抑制 RNase。下面介绍一种用异硫氰酸胍溶液分离 mRNA 的方法,此方法操作简单、能迅速获得较完整的 mRNA。

异硫氰酸胍溶液的制备方法是: 取 100g 异硫氰酸胍,加入 100ml 重蒸水、10.6ml 1mol/L Tris-HCl (pH7.6)、10.6ml 0.2mol/L EDTA, 室温搅拌过夜。加温至 60—70℃ 10 分钟,至完全溶解,或 5,000r/min 20℃ 离心 10 分钟去除不溶物。加入 21.2ml 20% 的 Sarkosyl 和 2.1ml β -巯基乙醇,并加入重蒸水到最终体积 212 ml, 备用。取 10g 组织或 10^8 培养细胞加入异硫氰酸胍溶液 10ml (组织电动匀浆 2 分钟),升温到 60℃,并在此温度下用 18 号针头的注射器反复吸推,直到悬浮液粘度降低。加入 10ml 60℃ 的饱和酚溶液,旋转振荡 3 分钟。加入 5ml 0.1mol/L NaAC (pH5.2)、10mmol/L Tris-HCl (pH7.4)、1mmol/L EDTA 和 10ml 氯仿-辛醇 (24:1), 60℃ 振荡 15 分钟。冰浴中冷却,5,000r/min 离心 10 分钟,取出上清液,用酚-氯仿-辛醇 (25:24:1) 提取两次,加入 2 倍体积冷乙醇,沉淀过夜。15,000r/min 离心 20 分钟,去除乙醇,沉淀溶于 10ml 0.1mol/L Tris-HCl (pH7.4)、50mmol/L NaCl、10mmol/L EDTA、0.2% SDS。加入蛋白酶 K 至 200 μ g/ml, 37℃ 保温 1 小时。升温至 60℃,加入 5ml 60℃ 的饱和酚溶液,旋转振荡,加入 5ml 氯仿-辛醇, 60℃ 保温 10 分钟。冰浴冷却, 5,000r/min 离心 10 分钟,取上清液,酚-氯仿-辛醇重复提取一次。取上清液加入 2 倍体积冷乙醇即可得 RNA 沉淀。一般 10g 组织或 10^8 培养细胞可得 RNA 1mg。mRNA 的分离可采用前面介绍的寡聚胸腺核苷纤维素柱亲和层析法。

2. 特异 mRNA 的富集

为提高某种特异 mRNA 在总 mRNA (多聚 (A)-RNA) 中的含量,常常需经过富集步骤。除了上述的通过抗体结合多聚核糖体的方法外,还可根据特异 mRNA 的特定沉降系数 (S 值),采用蔗糖连续密度梯度离心分离的方法。预先配制 15% 及 40% 两种蔗糖溶液,其中还含有 1mmol/L EDTA、0.1mol/L KCl、0.01mol/L Tris-HCl (pH 7.4) 和 RNasin 1u/ml。以梯度形成器制备 15—40% (W/V) 蔗糖连续密度梯度, 4℃ 静置 4 小时以扩散均匀。梯度表面铺加 RNA 溶液 0.2ml, 用 SW41 转头 36,000 r/min 离心 20 小时,离心管底部打洞,逐滴收集, 0.2ml 为 1 收集组分。离心时以 3 H-rRNA 或不标记 rRNA (23S + 16S + 5S) 为对照。在 260nm 测定各收集组分的 OD 值,对照管测定 cpm 或 OD₂₆₀, 同时测定各组分蔗糖浓度(通过折光率推算)。半对数座标纸同时标出各组分的 260nm 吸收及蔗糖浓度,根据对照 rRNA 的 S 值选取所需 S 值的 mRNA。此方法可使特异 mRNA 富集数倍至数十倍。

3. mRNA 的分析和鉴定

用作 cDNA 合成模板的 mRNA 需经过分析、鉴定。一般采用下述方法。

- (1) 通过聚丙烯酰胺或琼脂糖凝胶电泳,以确定其纯度、分子完整性如何。
- (2) 用体外翻译系统,确定其是否具有体外指导蛋白质合成的能力。

(3) 进行3'末端分析,确定多聚(A)的平均长度。

高质量的 mRNA 制剂应具备以下特征:纯度高,不混有其他种类 rRNA,体外翻译中模板活性高,3'端多聚(A)平均长度超过100个腺苷。

三、载体 DNA

用于 cDNA 克隆的载体有两类:质粒 DNA (如 pUC、pBR322 等)及噬菌体 DNA (如 λ gt 10、 λ gt11 等)。根据重组后插入的 cDNA 是否能够表达,经转录和翻译成蛋白质,又可将载体分为表达型载体及非表达型载体。pUC 及 λ gt11 即为表达型载体,而 pBR322 及 λ gt10 等为非表达型载体。表达型载体 DNA 的 cDNA 插入位置的上游具有启动基因顺序。

cDNA 克隆操作中应根据不同的需要选择适当的载体。cDNA 插入片段不太长(小于 10kb),可选用质粒载体,如大于 10kb 则应以噬菌体 DNA 为载体。选择载体 DNA 时另一个要考虑的因素是,pBR322 比 pUC、 λ gt10 比 λ gt11 更容易获得具有完整序列的 cDNA 克隆。

四、器械的消毒处理

cDNA 合成过程中所使用的所有器械及溶液都应经过高温消毒处理,以去除 RNase。处理方法是:玻璃器械应洗净、烤干,浸泡于 10% 硅油的氯仿溶液中 5 分钟,取出,蒸馏水洗涤多次,180°C 烘烤过夜。塑料器械及溶液应湿热蒸气消毒,15 磅压力 1 小时。

另外在实验操作中,应带手套,并防止皮肤直接接触器械。

在上一章最后曾介绍近一年的聚合酶链反应 (PCR) 方法的新进展,它有希望取代基因文库的工作。接着,人们又将它与逆转录方法结合起来,得到一种新的合成 cDNA 的方法,即逆转录-聚合酶链反应法 (RT-PCR 法)。该方法也是先用两个分别与待合成 cDNA 片段两端互补的多聚核苷酸链为引物,以全 RNA 为模板,经逆转录首先得到 cDNA 第一条链,然后再用 PCR 方法进行扩增,特异地合成目的 cDNA 链。它与 PCR 方法一样,有类似的优缺点,即可以部分代替 cDNA 文库的工作,但又不能完全代替 cDNA 文库。因为 RT-PCR 方法合成的 cDNA 也长度有限,故也有待进一步完善。有关该方法的文献,请参考 *Proc. Natl. Acad. Sci. USA* Vol. 85, 8998—9002, 1988。RT-PCR 方法除需用耐热的 Taq DNA 多聚酶外,还必须有比较好的 AMV 逆转录酶,否则不易得到相应的特异 cDNA。

参 考 文 献

- [1] Efstratiadis, A. and Villa-Komaroff, Cloning of double-stranded DNA, in *Genetic Engineering* (J. K. Stelow and A. Hollaender eds.), Vol. 1, 15, Plenum Press, New York, 1979.
- [2] Okayama, H. and Berg, P., High-efficiency cloning of full-length cDNA, *Mol. Cell. Biol.*, 2, 161, 1982.
- [3] Gubler, U. and Hoffman, J., A simple very efficient method for generating cDNA libraries, *Gene*, 25, 263, 1983.
- [4] Kraus, J. P. and Rosenberg, L. E., Purification of low-abundance messenger RNAs from rat

- liver by polysome immunoadsorption, *Proc. Natl. Acad. Sci. USA.*, **79**, 4015, 1982.
- [5] Ricciardi, R. P. *et al.*, Purification and mapping of specific mRNAs by hybridization-selection and cellfree translation, *Proc. Natl. Acad. Sci. USA.*, **76**, 4927, 1976.
- [6] Wood, K. O. and Lee, J. C., Integration of synthetic globin genes into an *E. Coli* plasmid, *Nucleic Acids Res.*, **3** 1961, 1976.
- [7] Maxam, A. M. and Gilber, W., A new method for sequencing DNA, *Proc. Natl. Acad. Sci.*, **74**, 560, 1977.
- [8] Goldberg, M. L. *et al.*, Isolation of specific RNAs using DNA covalently linked to diazobenzoyloxymethylcellulose on paper, *Methods Enzymol.*, **68**, 206, 1979.
- [9] Korman, A. J. *et al.*, cDNA clones for the heavy chain of HLA-DR antigens obtained after immunopurification of polysomes by monoclonal antibody, *Proc. Natl. Acad. Sci. USA.*, **79**, 1844, 1982.

第二部分 基因表达

第八章 基因表达与几种表达体系

蔡 良 琬

基因表达系指结构基因在生物体中的转录、翻译以及所有加工过程。在原核生物中基因表达是以操纵子形式进行的。当一个操纵子的调节基因与细胞内以 DNA 为模板的 RNA 聚合酶(简称 RNA 聚合酶)作用时,此操纵子的蛋白结构基因则开始转录,其中有意义的一条链转录形成相应的 mRNA。生成的 mRNA 又立即进入核糖体中,翻译出相应的多肽或蛋白质。生成的蛋白质都具有生物活性。这就是原核生物基因表达的总过程。在有细胞的生物体中都有类似的基因表达过程,此过程是简单还是复杂,视生物体的简繁而定。原核生物细胞没有核膜,其表达的基因中没有插入序列或内含子,因此表达过程中的转录与翻译往往同时进行,即转录完毕翻译也完成,随着 mRNA 也被水解掉。真核细胞不同,细胞内有核膜,转录是在核内进行,形成的核内不均一 RNA(hnRNA),还需加工,如去掉内含子、修饰、加工 5' 末端与 3' 末端,才能形成翻译用的 mRNA; mRNA 又只能在胞浆中进入核糖体,并翻译成多肽或蛋白质,再加上多肽或蛋白质的加工,如形成高级结构、糖化等,因此整个基因表达是复杂过程,这是原核生物与真核细胞在基因表达上的主要区别。

生物细胞中有许多可供表达的基因,但并非经常都在进行表达,有些基因在某种情况下或某种细胞中甚至根本不表达,可见基因表达是受调节与控制的。原核生物基因表达的调控比较简单,但对多细胞生物,特别是高等动植物,基因表达的调控有非常严密复杂的机制,称为多级调控。目前在基因工程中应用的调控手段与原理都是在细胞内的,及在基因水平的,不包括高等生物的整体调控。目前的基因工程生产方式也都是在单一细胞体系中完成的。

基因工程中基因表达的研究,是指所谓外源基因在某一种细胞中的表达活动,即剪切下一个外源基因片段,拼接到另一基因表达体系中,使其能获得既有原生物活性又是高产的表达产物,称为高效表达。最早应用的表达体系是大肠杆菌,即把欲表达的任一结构基因合理地拼接在大肠杆菌的质粒上,转化到适合的大肠杆菌受体菌中,使拼接的基因获得表达。近年来为了适应大部分真核基因不适用原核细胞生产表达产物的特点,又发展了真核细胞基因表达体系,如酵母与哺乳动物细胞体系,此三种基因表达体系已成为当前基因工程工业性生产的核心,由此也带动了这三大基因表达体系的研究。

为了满足各基因表达体系的工业生产需要,最重要的是必须研究这三类细胞的基因表达特点、基因表达调控的一般规律以及欲表达的基因的结构、特点及其在原存在细胞中的表达过程等。在具备这些必要知识后,基因的剪切、重组才能合理,也才能使欲表达基因在新宿主细胞中获得有生物活性的表达产物。

分子克隆技术是基因表达研究中极重要的部分,所有基因表达的实践都需分子克隆技术,表达理论也已应用于组建新表达质粒中。本章将从介绍简单基因表达理论开始,然后叙述表达调控的重要性与调控内容,并简单介绍这些理论在基因工程、分子克隆技术上的应用。

第一节 基因表达的基本原理

大肠杆菌中的乳糖操纵子和色氨酸操纵子是研究得最清楚的基因表达体系,而且它们的表达理论与转录的重要元件都已应用到基因工程中。本节将以这两个操纵子的特点为例,介绍基因表达的原理。真核生物的基因表达原理也将作简单叙述。

一、乳糖和色氨酸操纵子的特点及其转录翻译元件

乳糖(lac)操纵子是大肠杆菌分解和利用糖代谢的一个操纵子,可作为各种分解代谢酶的基因表达的一个代表。在大肠杆菌利用乳糖为碳源时,此操纵子即增强活动,产生有助于利用乳糖的一些酶。此操纵子中有三个酶蛋白基因,即 β 半乳糖苷酶、渗透酶(适用于乳糖等一些双糖)和巯基半乳糖转乙酰酶。整个操纵子有6,500bp,这三种酶蛋白结构基因(编码基因)约占5,000bp,属于调节控制基因表达的部分基因只有1,200bp,在此调控部分中还包括一个阻遏蛋白的编码基因,占1,080bp。此阻遏蛋白实质上只作用于乳糖操纵子,起调控作用。真正作为调控的基因只有112bp,包括一个P区和一个O区,如图8.1所示。

从图8.1中可以看到三个酶蛋白结构基因,简称为Z、Y和A。*lacI*为阻遏蛋白,在调控基因部位的*lacP*约84bp,*lacO*约28bp,它们是乳糖操纵子的启动基因(或称启动子)和操纵基因。整个操纵子中的三个酶蛋白基因是否表达,受Z基因前的调控部分*lacI*、*lacP*和*lacO*所控制。转录时生成的mRNA是一个整体,进一步翻译时将产生三种上述酶蛋白。

任何一个操纵子的基因转录都是由RNA聚合酶作用于启动子的部分开始的。此RNA聚合酶是大肠杆菌中唯一的转录酶,对任何一个启动子作用都是一样的,但各操纵子转录的速率则不相同。一个操纵子的基因转录速率的高低实际上取决于启动子的结构。现在已知原核生物操纵子中的启动子都是类似结构,即都带有两个能与RNA聚合酶结合的核苷酸序列区,如图8.2中所示,一个在-35左右,另一个在-10左右。编号1为转录的起始点,其上游即5'末端为负数,-35与-10区都是RNA聚合酶的结合点,它们附近的核苷酸序列大同小异,被称为是一种有同感作用的序列(consensus sequence)。不过-35与-10区的作用略有不同,一般认为前者是开始结合区,后者则不仅结合而且可使基因DNA的两链分开,启动转录反应。经过几年的研究,目前已知一个启动子的强弱以及所起动的转录速率的高低,除了决定于这两个区的核苷酸序列外,还与两区之间的距离有关。由于-10区的结构特点首先是由Pinnow D.发现并加以比较、确定的,故又称为Pinnow框架(Pinnow box)。关于转录过程在此不多叙述。

色氨酸操纵子是大肠杆菌中利用氮源合成氨基酸有关基因表达的一个代表,此操纵

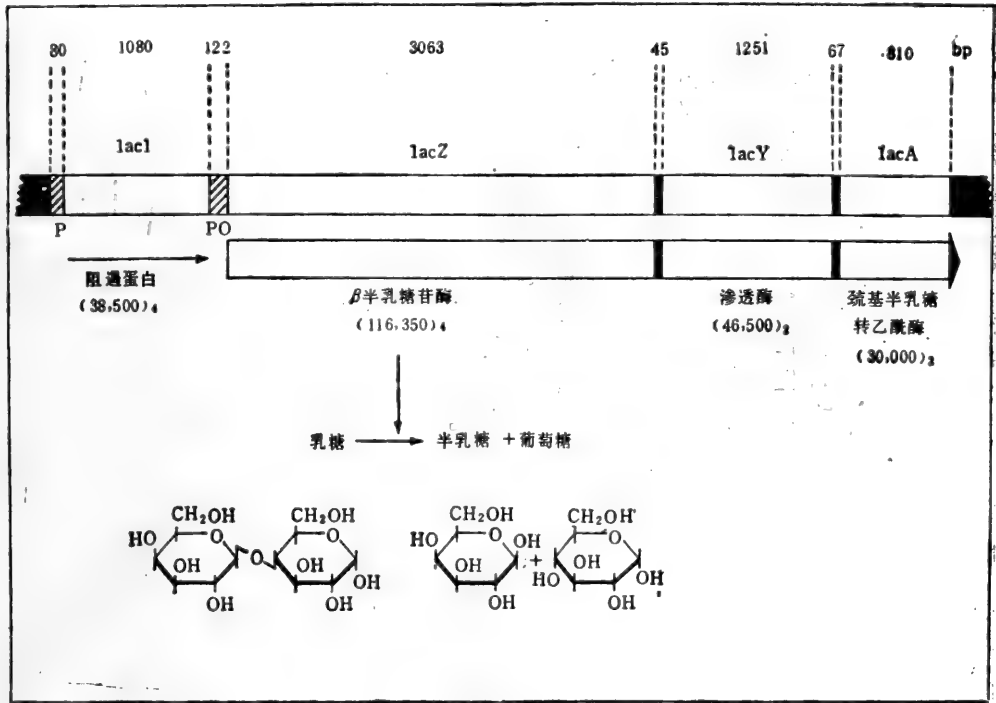


图 8.1 乳糖操纵子。

子中有 5 个与合成色氨酸有关的酶蛋白基因,全长约有 7,200bp, 其中 5 个结构基因约占 7,000bp, 其他为调控用基因。此操纵子中还应包括一个阻遏蛋白基因,但因这个基因不直接连在此操纵子任一侧,故未计算在内,整个操纵子的简图如图 8.3 所示。

图中标有 5 种酶的名称和作用点, 5 种酶的简称为 E、D、C、B 和 A。色氨酸操纵子的阻遏蛋白基因简称为 trpR, 其他调控基因有 trpP₁、trpO 与 trpP₂, 还有一个 trpL。由于 trpP₂ 的作用很弱, 只有 trpP₁ 的 3%, 实际上此操纵子的基因表达是受 5' 末端的 trpP (即 trpP₁) 和 trpO 等控制。trpP、O 控制区与 lac 的大体相似, 在 trpP 区中也有与 RNA 聚合酶的一 35 区和 -10 区同感序列, 从这两个操纵子的转录速率比较, trpP 明显强于 lacP。也有人曾将大肠杆菌用紫外线处理, 得到了一个突变株 UV5 型, 此株的 lacP 强于野生型。从此变异株的 lacP 分析中得知, 其 -35 区未变异, 但 -10 区核苷酸序列已有改变。现在公认的一 35 区与 -10 区中出现最多也是最适的核苷酸序列为:



任何一个基因转录并不是全部 mRNA 都翻译成蛋白质或多肽。mRNA 形成后可看出从转录起始至翻译密码出现, 有大约十几个核苷酸不等的序列是不翻译的, 但这段序列是翻译所必需, 即它是 mRNA 能否进入核糖体的关键。通过序列分析已发现, 一般在翻译起始密码前 10 个左右的核苷酸处有一段序列恰与大肠杆菌中核蛋白体中的小亚单位 16S rRNA 羧端的序列有互补关系, mRNA 能否进入核蛋白体就靠这一序列的作

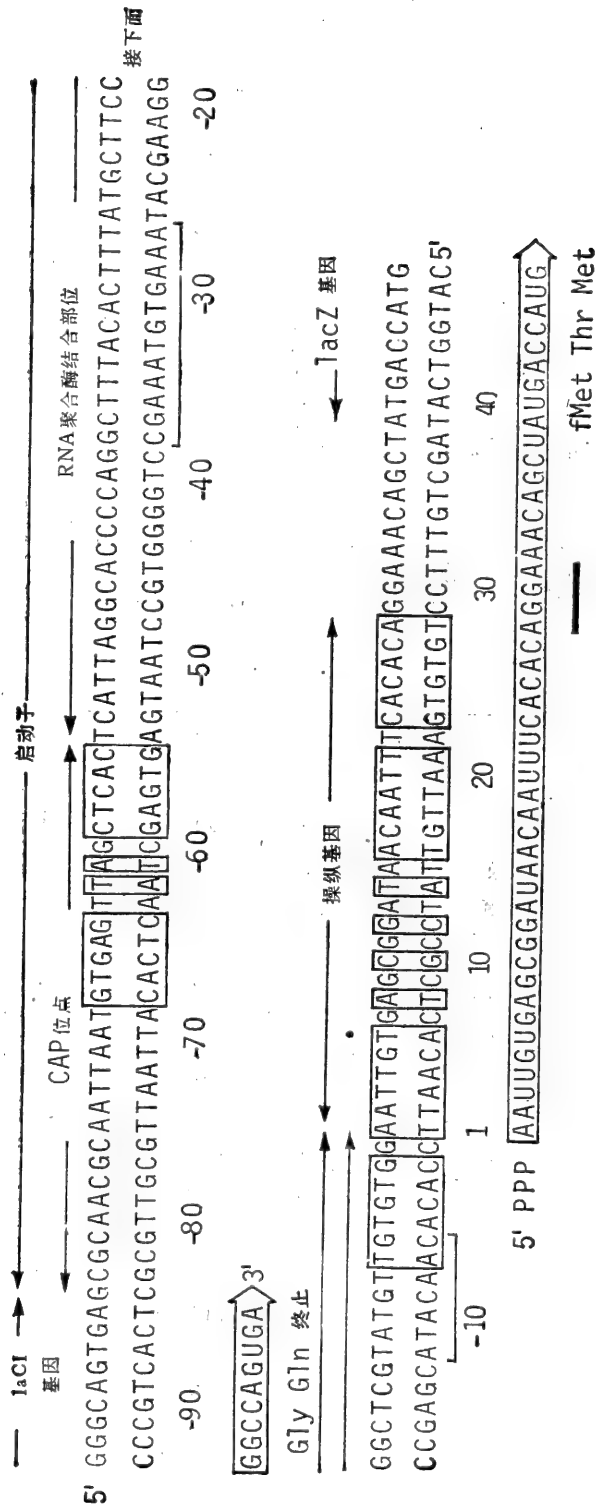


图 8.2 乳糖操纵子中的调节区。

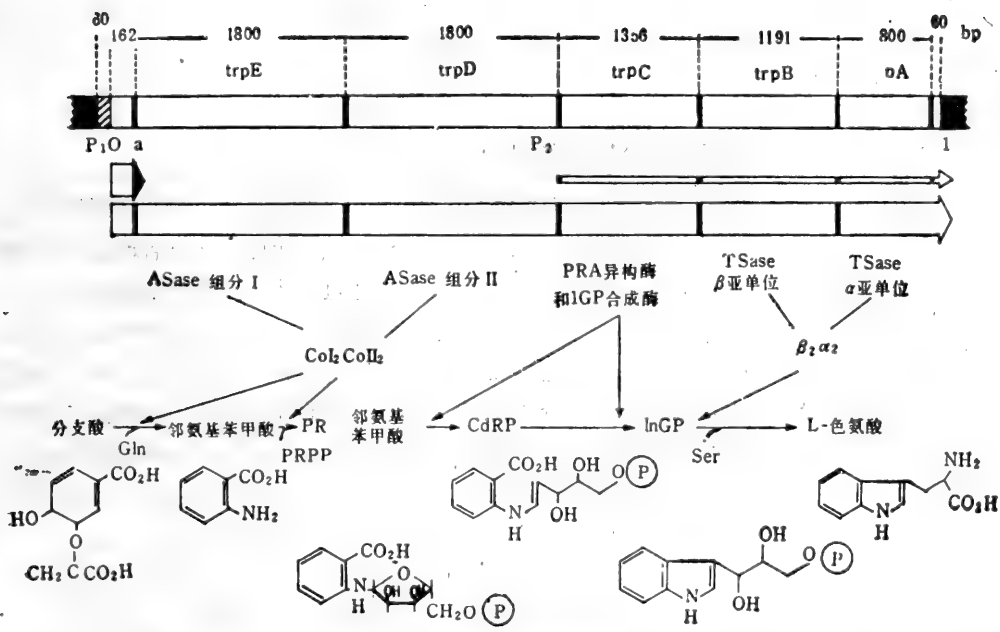


图 8.3 色氨酸操纵子。

用。这段序列是由 Shine G. 与 Dalgarno L. 发现的,故常称 Shine-Dalgarno 序列,或简称 SD 序列。SD 序列虽然是 mRNA 分子中的一个序列,但它仍是由 DNA 上转录而来,通常是以 RNA 序列表示。在不同基因中这一序列虽都存在,但不完全相似,只是一段大致相似而作用相同的同感序列,其最多出现的序列为:



SD 序列对基因的转录似无关,但对翻译十分重要,翻译是否顺利与这段序列及其与起始密码的距离都有关。任何一个基因都有它特异的 SD 序列。在生物体内各 SD 序列虽不相同,但在保持生物体的统一性上则有协调作用,这一点与每个启动子之间的差别有类似的作用。在基因工程中,欲达到使一个外源基因表达获得高产,研究这些同感序列就十分重要。有人把这些同感序列,即 P 区的 -35 与 -10 处的序列以及 SD 序列特点的研究,统称为基因表达单元的研究。对它不但简单加以利用,而且可改造后利用,以提高基因表达的速率。

二、lac 与 trp 操纵子的基因表达调节

生物细胞中所有的基因表达都受一定的调节控制,只有极少数组成型的不受控制,但这类基因表达的速率通常是比较缓慢的。调节控制有两方面,一是减慢转录速率至很低,称为负调节或阻遏作用;另一方面是加速转录,即正调节。对于一些有强启动子的基因表

达,调节控制尤其重要,这是任一细胞保持正常生长的必要条件。*lac* 与 *trp* 操纵子的基因表达皆受一定的调节控制,但两者的调节控制基因与机制不完全相同,它们代表了最简单的两类调控机制。

(一) *lac* 与 *trp* 操纵子的调控基因与作用物

lac 与 *trp* 操纵子都属于经常处在负控制或抑制状态下的操纵子,即当大肠杆菌没有利用乳糖类物质,或在色氨酸含量充足的情况下,*lac* 与 *trp* 操纵子很少活动。这种负控制在 *lac* 操纵子中是由 *lacO*,即 O 区与 *lacI* 基因表达产物阻遏蛋白结合而实现的,在 *trp* 操纵子中则是由 *trpO* 区与 *trpR* 表达产物经作用而产生的,不过 *trpL* 也有相当重要的作用。与此相反,一旦 *lac* 与 *trp* 操纵子基因表达开始,又会出现新的正调节,即增强转录水平的调节。在图 8.2 中 *lac* 操纵子的 *lacP* 区 5' 末端处,亦即在 *lacI* 基因后有一段序列可与 cAMP-CAP 受体结合,一旦结合就可出现这类正调节,使表达产物成倍增加。*trp* 操纵子则稍有不同,它不是通过以上的作用,而是通过减少色氨酸水平,改变 *trpL* 转录出的 mRNA 构型而加强转录速率的。

因此,正负调节的实现需有以下基因及作用物:

负控制 (<i>lac</i>)	<i>lacO</i> 基因	<i>lacI</i> 基因表达产物
(<i>trp</i>)	<i>trpO</i> 基因	<i>trpR</i> 基因表达产物+辅因子
正调节 (<i>lac</i>)		cAMP 及其受体
去阻遏调节 (<i>trp</i>)		减少环境中色氨酸含量

(二) *lac* 操纵子的去阻遏和正调节

lacO 是此操纵子的操纵基因,它的核苷酸序列见图 8.2。若以 11 核苷酸部位为中心,其两侧的核苷酸是处于旋转对称状态。说明 *lacO* 区是个容易改变构象的区域,即可由直线形双链变为链内互补的十字架结构。一般认为这种结构容易与某特异蛋白结合,即 *lacO* 易与 *lacI* 产生的阻遏蛋白结合,结合后影响了 RNA 聚合酶从 *lacP* 区向 3' 末端滑动,因此减缓转录反应至最低值。如遇适当的诱导物,加乳糖或目前常用的异丙基硫基半乳糖苷 (IPTG),则可改变阻遏蛋白的构型,使其不能与 *lacO* 区结合,此时 RNA 聚合酶则可通过,转录速率开始增加,这种调控为去阻遏。去阻遏反应在基因工程中是经常应用的一种方法,*lacZ* 表达生成的 β 半乳糖苷酶可应用于某些显色反应,用于筛选重组菌株。

去阻遏虽能增加转录速率,但并非达到最高(比抑制状态的速率增强约 1,000 倍),此时半乳糖苷酶产量可达细菌蛋白的 2% 左右。如果有适量的 cAMP 及其受体加入,或该菌能生成足够的 cAMP 及其受体, β 半乳糖苷酶的产量,还可再增加 2.5 倍。因此 cAMP 及其受体是 *lac* 操纵子的强有力正调节作用物。

(三) *trp* 操纵子的去阻遏与色氨酸调节

色氨酸操纵子经常处于负控制状态下,*trpR* 是其阻遏蛋白基因,此蛋白与 *trp* 结合后能与 *trpO* 区相结合。*trpO* 区也有一旋转对称的核苷酸序列,故与 *lacO* 类似,能通过变构形成十字架构型与 *trpR* 的产物结合,影响 RNA 聚合酶进行 *trpE*、D、C、B、A

酶基因的转录。不过阻遏作用不是通过诱导物,通常采用以下两种方式,一是遗传性去阻遏,即通过突变使 $trpR$ 变为 $trpR^-$; 二是生理性去阻遏, 因为 $trpR$ 基因不与 trp 操纵子直接相连, 故在基因工程中常常不是利用加 $trpR$ 的操纵子, 而是利用带有 $trpR$ 的受体菌来调节。

trp 操纵子的正调节实际上也是一种去阻遏作用, 因此也属于负调节的部分 (在许多合成氨基酸的操纵子中都有类似调节形式)。在 $trpP$ 与 O 以后, 首先有一段 $trpL$ 序列, 它不属于此操纵子的生成酶基因, 但可以转录成为前导序列。此序列中带有-一个转录终止的特有序列, 即像一个转录终止子 (transcriptional terminator), 简称为 $trp a$ (弱化基因, attenuation)。此 $trpL$ 约含 140 个核苷酸, 只转录而不翻译。当遇到大量色氨酸出现时, 能部分翻译出 14 肽并能阻止 90% 的 RNA 链继续延长而到此处终止。故只有在色氨酸量减少后, 这段类似转录终止子的弱化基因才改变构型, 使所有转录出的 RNA 链延长自 $trpE$ 到 $trpA$ 。

前面已经提到, $trpP_2$ 的作用只有 $trpP_1$ (即 $trpP$) 的 3%, 因为没有 $trpP_2$ 前的 $trpE$ 与 $trpD$ 基因表达; 尽管有足够量的 $trpC$ 、 B 、 A 基因表达, 所生成的酶蛋白还是不能进行色氨酸合成。 $trpR$ 产物的阻遏作用不影响 $trpP_2$ 处的转录速率, 只影响 $trpP$ 的转录速率。当应用去阻遏方法, 去掉 trp , $trpR$ 表达产物的作用, 一般可提高转录效率 70 倍, 如再用低色氨酸含量改变弱化基因的作用, 还可提高转录效率 10 倍, 总计为 700 倍。 $trpP$ 是个比较强的启动子, 在基因工程原核体系的基因表达中常被利用, 故以上的调控机制与方法也同时被应用。人工组建启动子如 $trp-lac$ 启动子(即 tac 启动子), 其调控部分用的是 lac 操纵子中的正负调节机制。有关人工强启动子的组建将在后面叙述。现已证明, 起始密码的特点、转录终止子与翻译终止子的正确应用等, 在提高基因表达速率上是很重要的。

三、真核细胞中的基因表达与多级调控

真核细胞远比大肠杆菌复杂, 所含基因很多, 其表达与调控特点目前了解还不像原核的清楚。在基因工程中常用的表达体系如酵母与哺乳细胞, 是前了解得比较清楚的。在真核细胞中一般都有三种 RNA 聚合酶, 在基因工程应用上主要是研究 RNA 聚合酶 II, 它能转录出各种 hnRNA。从核内不匀一的 RNA 到 mRNA 有个加工过程, 以后还有一系列翻译与翻译后加工等修饰蛋白的过程, 以及糖化、蛋白高级结构的形成等。目前研究的重点也仍在基因水平的转录与转录调控。

真核细胞中的基因表达单元不是由几个有关功能的蛋白组成一个操纵子, 而是一个蛋白质, 或以蛋白质中的一个多肽为单位, 其复杂性首先表现在几乎任何一种蛋白质或多肽结构基因中都含有大小不等的内含子, 因此转录出的 hnRNA 都比较大。通常把这种

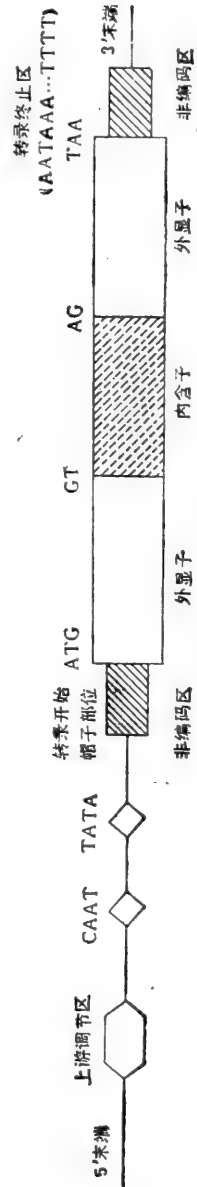


图 8.4 应用 RNA 聚合酶 II 转录的某些基因的结构与调节部位简图。

转录出来的 hnRNA 叫做转录本(transcript)。它包括类似于大肠杆菌操纵子的结构,有 5' 端的转录调节控制区,整个多肽的结构基因与内含子,还需加上 3' 端部分序列,如图 8.4。

5' 端又被称为该基因的上游,3' 端则为下游,真核细胞的基因转录调控主要与上游关系密切,3' 端也有影响。一般多肽结构基因前也有相应的启动子,启动子的核苷酸序列比较长,但不像原核中的启动子其长短已很清楚。在转录起始点前的 -30 区左右是一个 RNA 聚合酶结合区,其中有 TATA 序列,故又叫 TATA 框架,在其上游 -70 至 -80 处还有一个结合区叫 CAAT 框架。此两结合区的序列特点不十分一致,甚至有些很不相似。近年来又发现上游还有影响转录速率的序列,尚无统一的名称,一般在 -100 至 -1000 核苷酸处,因此在基因工程利用真核启动子时往往需要剪切出相当长的序列,以免影响转录速率。在上游与下游远端对基因表达有影响的序列中,已发现有增强子(enhancer)及上游活化序列(upstream activation sequences, UAS),二者的作用类似但又有差别,增强了序列不论放在上游或下游,顺向或反向都有增强转录速率的作用,UAS 序列则只在上游有作用。至于它们是怎样起增强活化作用的,目前还不十分清楚。

第二节 常用的基因表达质粒

应用上一节的基因表达理论与采用转录所需要的主要元件,组建一些应用于基因工程技术与生产的表达用质粒还在不断增多,有些已成为商品出售。本节将介绍几个大肠杆菌体系中表达质粒的特点与应用。可以看到,在组建一个表达质粒时,不仅需要组建相应启动子的核苷酸序列,还必须保留转录成 mRNA 时的 SD 序列,以使转录后能顺利翻译,还必须有表达的全部元件,调控基因也最好完整,当然也可以利用其他方法保持其完整。

一、lac 启动子的应用

lac 启动子的应用十分广泛,已成为基因工程常用的一种组建质粒,其他载体如噬菌体中筛选用的启动子。它有以下几个优点,在 lac 操纵子中 lacZ 的表达产物 β 半乳糖苷酶可作用于某些显色剂,产生颜色反应,容易识别,可应用于筛选菌落或噬菌斑。二是 β 半乳糖苷酶的生成还可用其他分析方法鉴定,故可帮助了解表达产物量的多少,特别是用于融合蛋白的生成。有关融合蛋白将在下一节介绍。

在本书上册末端终止法一章中曾介绍了 M13 噬菌体应用于测定核苷酸序列。M13 噬菌体的衍生物,一系列 mp1-19 都带有 lac 操纵子的部分,只是没有 lacY 与 lacA。另外一些常用质粒如 pUR222 也有以上部分 lac 操纵子基因,不过 lacZ 基因大多很短,有的只有 N 端少数氨基酸编码部分。其他含 lac 操纵子的载体还很多,它们的差别在于 lacZ 基因的长短和是否完整。lacZ 基因不论长短或是否完整皆可生成相应的酶并能够作用于显色剂,其原因是利用了受体菌中 lac 操纵子与它的互补作用。这样当运用有 lac 操纵子组建的质粒或其他载体时都需注意同时应用适合的受体菌,否则将不能达到目的。现以 pUR222 与 pUC 系列质粒为例来说明 lac 启动子质粒的应用。此二质粒的结构见图 8.5。

pUR222 与 pUC 系列质粒来自 pBR322 质粒,有 pBR322 的 Ap^r 基因与复制起

点部分,都是多拷贝的。所插入的 lac 启动子与 lacZ 部分基因是来自 M13mp 类噬菌体衍生物的 DNA。pUR222 是利用 M13mp2 PstI 片段中的 lacP-lacO-lacZ, 以后又加入了 M13mp7 的多聚连接子。pUC 系列质粒则直接利用 M13mp 系列噬菌体衍生物 DNA。因此所得到的 pUC 系列质粒的特点与一些 M13mp 系列 DNA 结构有相应关系, 如 pUC-18 的 lacZ 部分序列就与 M13mp 18 完全一致。这两种质粒都保持了原来 M13mp 系列中多聚连接子的特点, 故适用于不同酶切下来的基因插在 lacZ 基因的产物的 N 末端。如果插入的基因小于 50 个核苷酸, 而在连接上又符合原来 lacZ 基因的框架关系, 则 pUR222 或 pUC 系列质粒都能保持原 β 半乳糖苷酶产生后作用于显色剂的颜色反应[如用 IPTG 诱导加 X-Gal 显色则为蓝色, 如用麦康凯 (MacConkey) 培养基, 因其中有中性红与结晶紫故产生粉红色]。其中的多聚连接子已经计算过编码框架关系。只是当加入 100 以上核苷酸时因片段过长不论框架是否正确, 都会影响此酶的显色反应, 如生成大量 β 半乳糖苷酶, 用定量分析法还可测出此酶的产量。

利用 pUR 222 与 pUC 系列质粒中的 lac 操纵子等部分序列一般可完成以下三个方面工作, 一是用于筛选插入外源基因。二是可用来测定插入外源基因的核苷酸序列。在分析中既可以用末端终止法又可用化学测序法。三是用于外源基因的表达。lacP 的启动速率不很高, 但可经诱导来加强, 又可以与 lacZ 基因连接在一起, 形成外源蛋白与 β

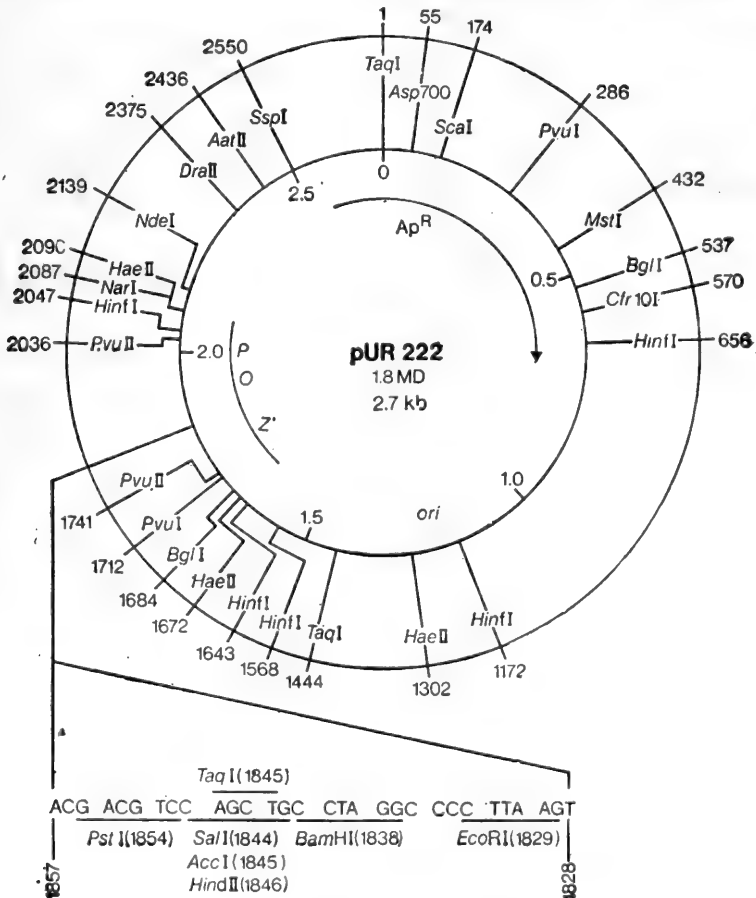


图 8.5a pUR222 结构图。

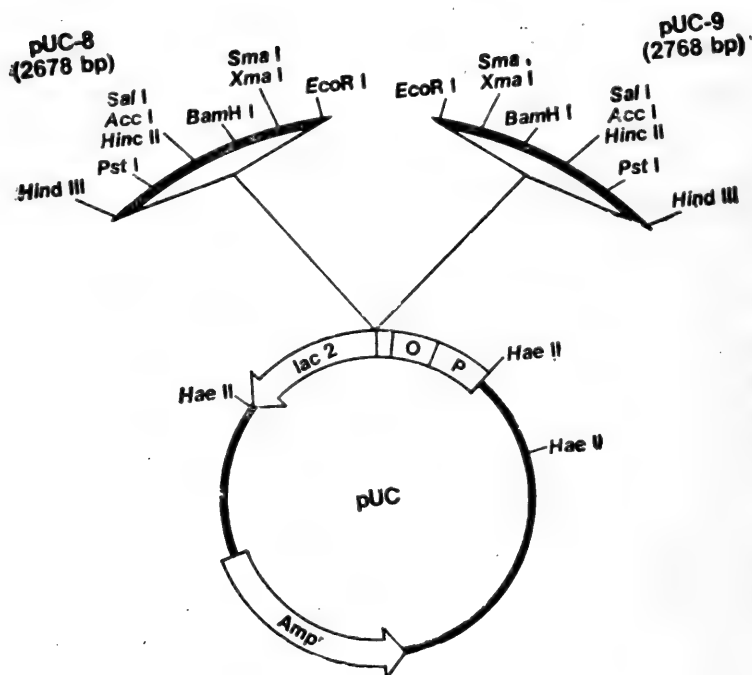


图 8.5b pUC8 与 9 的结构图。

半乳糖苷酶的融合蛋白,以保持获得的外源蛋白的稳定。

二、trp 启动子与人工合成 tac 启动子的应用

trp 启动子是大肠杆菌体系中的一个强启动子。tac 启动子则是人工合成的,由核苷酸序列中 trp 启动子的 -35 区部分与 lac 启动子的 -10 区的部分组建而成,它的作用比 trp 或 lacUV5 都强。目前有商品出售的相应质粒有 pDR720 (带有 trp 启动子)及 pDR540 (带有 tac 启动子),它们都是利用 pK0-1 重新组建而成,见图 8.6。

pK0-1 也是由 pBR322 而来,它带有 pBR322 的复制起点和氨基苄青霉素抗性基因,没有四环素抗性基因。在相应部位插有一个半乳糖激酶基因 (GalK), 此基因的表达产物可用简单的酶法测定。当有外源基因表达时,可用此酶的生成量多少来推算表达水平;同时,在此基因前加上不同的启动子,也可用此酶的生成量推算出加上的启动的强弱。pDR700 与 pDR 540 中的启动子虽有不同,但在 GalK 基因前却有一 Sma-I 限制性内切酶位点,从此处可连接上不同启动子。在组建好的 pDR700 和 pDR540 中, trp 或 tac 启动子后各有一多酶切点和一个单一酶切点 BamH1。从此处可插入外源基因,只要连接时框架正确,便可使外源基因表达。

trp 与 tac 都是强启动子,在组建时应注意其负调节控制。一个强启动子如平时没有负调节因子来控制其作用,则组建好的重组质粒往往不稳定,影响质粒的复制作用,甚至逐渐消失。前面已讲过,在 trp 操纵子上并没有直接相连的 trpR 基因,在组建和应用

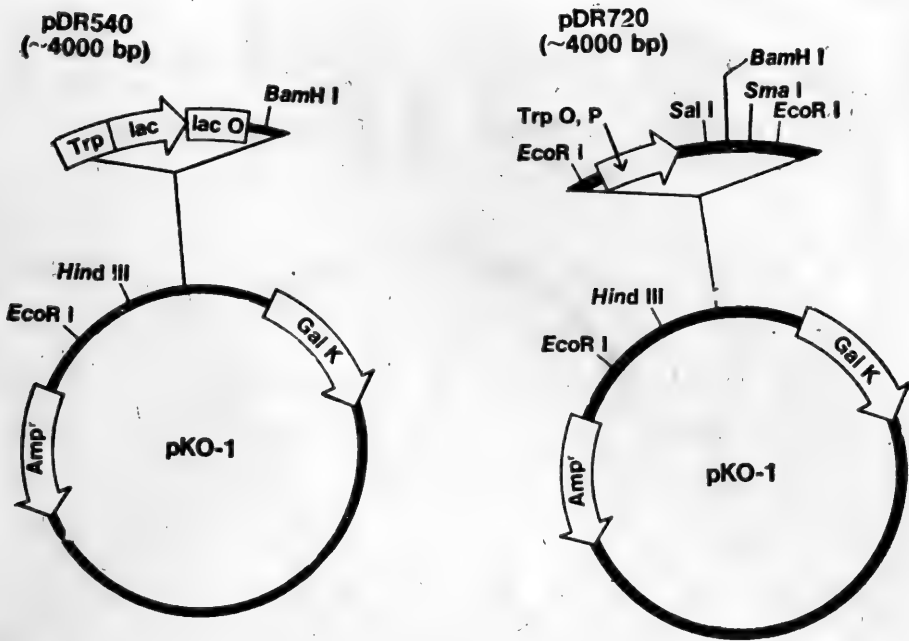


图 8.6

pDR700 时,是利用宿主受体菌 C600, galK⁻ trpR⁺ 大肠杆菌。为了去阻遏,可加入 3-β-吲哚丙烯酸 (IAA) 作为诱导剂,来代替色氨酸。tac 启动子是人工合成的,其序列中既有 trp 启动子,又有 lac 启动子,其负调节仍是用 lac 操纵子的 lacI 系统,所以由 lac 组建的质粒都应用有 lac 启动子质粒所需的宿主菌,去抑制的诱导物则用 IPTG。

pDR540 中的 tac 启动子虽是人工合成的,但并不是直接用核苷酸酶促人工合成,而是从某些质粒上剪切下来,又重组形成的。下图说明 tac 启动子的特点(图 8.7)。

启动子	-35 与 -10 区	-55	-50	-40	-30	-20	-10	-1+1
lac UV5	18	CTCATTAGGCACCC	AGGCTTTACACTTTATGCTTC	CGGCTCGTATAAATG	TGTGTGG ^{**} AATTGTGTA			
trp	17	CTTACTCCCCATCCCCCTG	TTGACAATTAATCAT	CGAAGTACTTAACTAGTACGCAGGGGCA				
trp-lac(tac)	16	CTTACTCCCCATCCCCCTG	TTGACAATTAATCAT	CGGCTCGTATAAATG	TGTGTGG ^{**} AATTGTGTA			

图 8.7 lacUV5 trp 及 tac 启动子部分序列。

图上的 trp 启动子片段是从质粒 pDR33 中剪切来的,在 trp 启动子的 -35 区与 -10 区之间有一个 TaqI 酶切点。lacUV5 是野生型大肠杆菌用紫外线诱变得来,其中 lac 启动子核苷酸发生了突变, -10 区中的同感序列由 TATGTTG 变为 TATAATG,后者正是此区同感序列中最好的和出现最多的一个。在 lac 启动子 -35 区与 -10 区

之间有一个 Pvu-II 切点。Pvu-II 切口为 -CCGG-,它与 Taq-1 切口特点 -TCGA- 有相似的一CG-序列,故可相互粘合。tac 启动子的形成正是取 trp 启动子中有 -35 区的片段和 lac 启动子中有一 10 区的片段,在以上两酶切点处连接而成。根据图上三个启动子序列的比较可以看出, lacUV5 启动子 -35 区与 -10 区之间有 18 个核苷酸,

trp 启动子中则为 17 个核苷酸,而 tac 启动子中最短,为 16 个核苷酸。经过用同一基因表达比较,已知 tac 启动子比 trp 启动子强 3.0 倍,比 lacUV5 启动子强 11 倍。但应注意,并非所有合成的人工启动子都可以增强基因表达的速率,如 tac 启动子原有两种,一为 tacl,即上述的 tac 启动子,有增强作用;而另一种 tacII 则远不如 tacl,甚至有些合成启动子,不但不能提高反而降低基因表达水平。

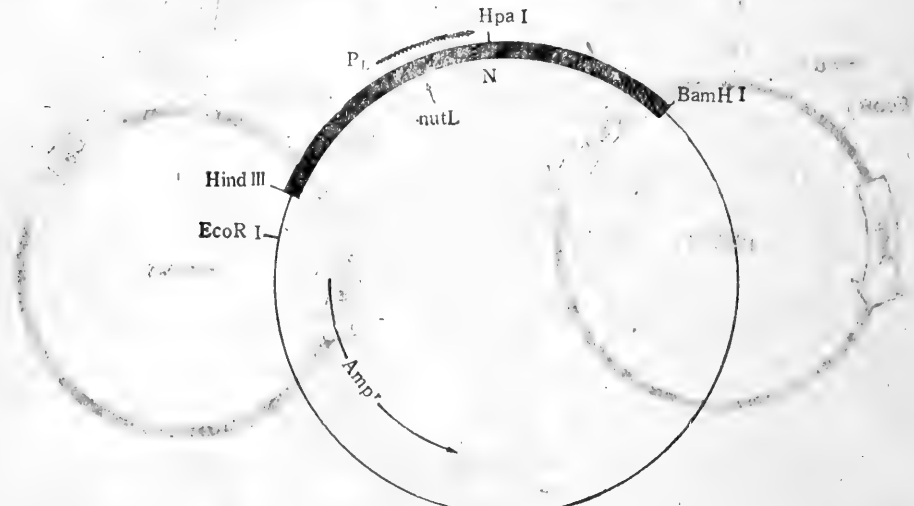


图 8.8a pKC30 结构图。

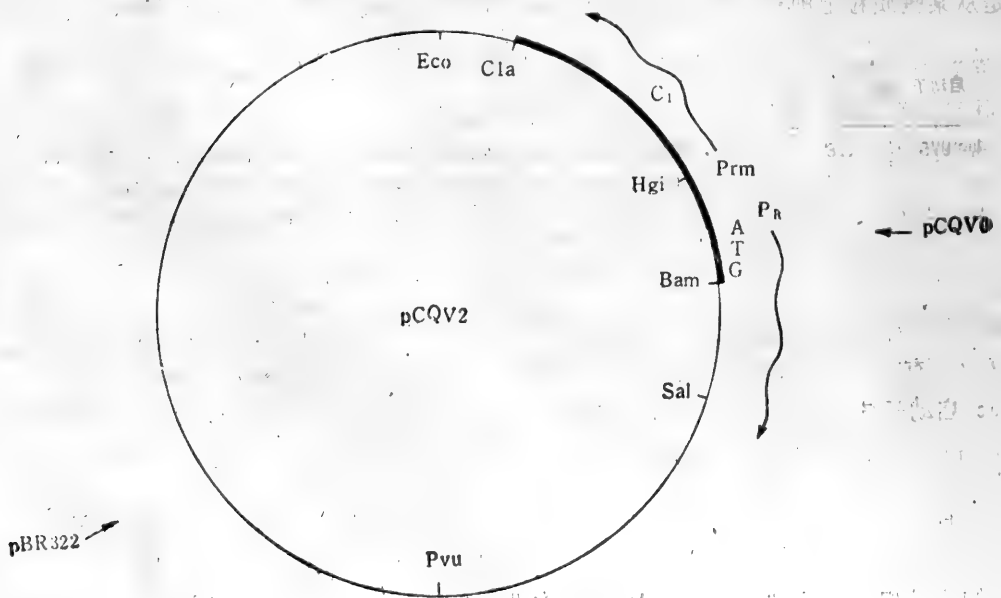


图 8.8b pCQV2 结构图。

三、 λ 噬菌体中 P_L 与 P_R 两个启动子的应用

P_L 和 P_R 是 λ 噬菌体中两个很强的启动子,可用在大肠杆菌体系的基因高效表达。已组建好的表达质粒有 pKC30 和 pCQV2,见图 8.8。

这两个质粒皆来源于 pBR322,都是把 λ 噬菌体 DNA 的有关片段组建在四环素抗性基因上,破坏了原有的四环素抗性基因。 P_L 和 P_R 启动子都是 λ 噬菌体中进行初期转录用的强启动子,在调节控制上又都受 CI 蛋白的抑制。由 P_L 启动子开始转录的第一个蛋白质基因为 N 基因。在组建 pKC30 时剪切下来的 λ DNA 片段上除有 P_L 启动子外,还有 N 蛋白基因部分。这样就保留了全部基因表达所需的元件。把外源基因连接在 N 蛋白基因的 N 末端编码的 HpaI 酶切口处,可利用 N 蛋白基因的起始密码翻译出相应的蛋白质。pCQV2 质粒上带的是 P_R 启动子,此质粒与 pKC30 不同,在组建时不仅重组上启动子,而且还重组了 CI 蛋白基因。这样应用 pCQV2 就比 pKC30 简单,不必采用有 CI 蛋白基因表达的受体菌。pCQV2 的另一个特点是,在组建时要注意剪切下 Cro 蛋白结构基因的全部编码序列,仅保留了起始密码子,而且起始密码子又正好与 Bam HI 酶切口重合,这样只要从 Bam HI 酶切口处连接上符合翻译框架的外源基因,就可以利用 P_R 启动子转录翻译出不带任何 Cro 蛋白部分的外源蛋白。与翻译有关的 SD 序列还是利用原来 Cro 蛋白结构基因前的一个。以 pKC30 与 pCQV2 相比,前者得到的外源蛋白往往带有少量由 N 蛋白来的其他短肽,而 pCQV2 得到的是单纯的外源蛋白。

应用 λ 噬菌体中的 P_L 或 P_R 启动子时需要注意 CI 蛋白的特点。由于 P_L 及 P_R 都是很强的启动子,在未进行基因表达时,必须用 CI 蛋白来控制。CI 蛋白类似 lacI 蛋白的作用,通常容易抑制 P_L 或 P_R 的启动作用,但它的去阻遏方法却不同,CI 蛋白是一种温度敏感蛋白,在 42℃ 时 CI 蛋白即失活,但此温度并不影响其他酶蛋白的活性。故利用 P_L 或 P_R 启动子做基因高效表达时,总是先用 37℃ 培养,达到扩增的目的后,提高温度至 42℃,破坏 CI 蛋白,此时 P_L 或 P_R 启动子开始工作,在很短时间内即可完成外源基因的高效表达。

总结以上几个组建的基因表达质粒可以看出,研究基因表达的元件和应用都是很重要的。生物体有复杂的调控基因表达机制,目前能应用于组建表达质粒的还只是在基因水平上的调控基因和因子等。在生物细胞进行基因表达时则常比已知的理论还要复杂,如现在组建的表达质粒只注意转录翻译的起始元件,实际上转录终止部分也很重要。另外,有些外源基因产物是表达用细胞不很相容的蛋白,尽管理论上应得到高效表达并获得高产量,但当遇到以上问题时,有时就不能得到满意的结果,这说明有关基因表达的理论还有待深入,才能解决那些尚存在的悬而未解的问题。

第三节 翻译框架与融合蛋白

基因表达研究中发现不同生物体中应用的基因密码子不同,其中也包括起始密码不同,它与外源基因表达与否或表达高低有关,但利用时只有以人工合成适当的密码子才能解决。对此本节不多讨论,只对有关框架与融合蛋白两个问题作简要介绍。

一、基因表达与翻译框架

实现外源基因的表达, 必须注意在拼接时保持密码子框架的正确。所以做基因表达工作进行基因重组, 比一般分子克隆要复杂得多。一般分子克隆不需要考虑密码子框架问题, 在连接方向上, 顺向或反向也无关紧要, 但对研究基因表达则必须框架正确。因为几乎不可能在一个结构编码基因前找到既有合适酶切位点, 又符合密码子框架的切口, 因此必须运用特殊剪切、削接或补接等方法(见第四章)。在有些基因表达质粒的酶切图上, 为此还标出原有基因密码子框架, 以便连接外源基因用, 如图 8.5 中的 pUR222。

当连接外源基因时, 如不是一般拼接, 而是削接或补接, 常会出现核苷酸数量增加或减少。在连接后, 即使框架正确表达翻译的结果中也可能减少几个氨基酸或增加几个氨基酸。只要这些氨基酸的改变不影响外源蛋白的活性, 仍可应用。如采用不定点的削接或补接, 从理论上则只有三分之一的机率得到正确框架, 因为密码子只有三种编排的可能, 即 1、2、3, 4、5、6; 或 2、3、4, 5、6、7; 或 3、4、5, 6、7、8 三种框架。框架不正确当然得不到有活性的外源蛋白。

二、融合蛋白

外源基因表达中生成融合蛋白是常见的。有些是插入外源基因的部位在原有一些蛋白编码基因之中, 基因表达后, 自然生成这种杂合蛋白, 既包括原有蛋白部分, 又有外源蛋白部分。最常见的是利用 lacZ β 半乳糖苷酶基因, 此酶基因编码甚多, 有 1,021 个氨基酸密码。此 lacZ 基因大多不是应用全编码序列, 只有几十个氨基酸或几百个氨基酸编码, 其他情况类似。也有只增加几个氨基酸的, 但不属于融合蛋白。融合蛋白是指该杂合蛋白既有原蛋白功能, 又有外源蛋白功能的结合体。

(一) 融合蛋白对外源蛋白的保护作用

融合蛋白能否被利用于基因工程生产, 需视产品的用途, 如果不是药物, 有其他蛋白功能在内不影响产物的应用, 也适用于生产, 因为融合蛋白能保护外源蛋白在生成后不被表达细胞消化掉。外源基因表达, 特别是高效表达, 往往会在表达细胞中堆积大量外源蛋白, 将不利于该细胞的生长, 因此表达细胞必然要排斥这种蛋白, 消化这种蛋白。事实上许多真核生物基因就难以在原核生物细胞中获得稳定的高产蛋白。融合蛋白可以帮助外源蛋白稳定, 留在表达细胞内, 少被表达细胞破坏。lacZ β 半乳糖苷酶就很适于应用, 保护外源蛋白, 从保留下来的 β 半乳糖苷酶的部分分子的大小来看, 分子保留愈大, 愈有保护力, 当然对外源蛋白活性的影响也大。至于形成的融合蛋白是否需要在抽提后再裂解分开, 则视外源蛋白产品的需要而定。

(二) 酵母基因表达体系中的融合蛋白

酵母是一种简单的真核生物, 用它的细胞进行基因表达研究也有融合蛋白问题。但在酵母中有不少表达的外源蛋白或多肽可分泌出细胞, 在分泌过程, 细胞内有某些特异酶

能切下相连的前导序列肽。这种功能在原核细胞中也有发现。可见这类融合蛋白的生成并不妨碍获得纯外源蛋白。

第四节 几种基因表达体系的比较

目前应用最多的基因表达体系是大肠杆菌、酵母和哺乳动物细胞三类，它们在应用上各有优缺点。现将三者的差别列表如下：

表 8.1 三类基因工程表达体系的比较

表达体系	产物	产生部位	培养方式	提纯	产物活性	潜在危险性
大肠杆菌	多肽, 蛋白或融合蛋白	菌体内	容易 部分可获高产	一般	对原核者好 真核者稍差	不大
酵母	多肽, 蛋白 糖化蛋白	菌体内或分泌 出细胞	容易 可高产	菌体内稍复杂	真核的接近 天然	不大
哺乳动物细胞	完整糖化蛋白	分泌出细胞	较难 成本高 可高产	简单	几乎为天然 产物	需注意有致 癌因素

参 考 文 献

- [1] Glass, R. E., Gene Function — *E. coli* and its Heritable Elements Croom Helm, London, 1982.
- [2] Pribnow, D., Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter, *Proc. Nat. Acad. Sci.*, **72**, 784, 1975.
- [3] Shine, J. and Dalgarno L., Determinant of cistron specificity in bacterial ribosome, *Nature*, **254**, 34, 1975.
- [4] Baner, C. E., Attenuation in Bacterial Operons Gene Function in Prokaryotes, Cold Spr. Har. Lab., 1983.
- [5] Yanofsky, C., Prokaryotic mechanisms in eukaryotes? *Nature*, **302**, 751, 1983.
- [6] Kgoury, G. and Gruss, P., Enhancer elements, *Cell*, **33**, 313, 1983.
- [7] Guarent, L., Yeast promoters-positive and negative elements, *Cell*, **36**, 799, 1984.
- [8] Rutter, V., Construction and properties of a new cloning vehicle allow direct screening for recombinant plasmids, *Molec. gen. Genet.*, **178**, 475, 1980.
- [9] Koenen, R. U. *et al.*, pUR222, a vector that uses phage for cloning and rapid chemical sequencing of DNA, *Nucl. Acid. Res.*, **9**, 4087, 1981.
- [10] Vieira, J. and Messing, J., The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers, *Gene*, **19**, 259, 1982.
- [11] Russell, D. R. and Bennet, G. N., Construction and analysis of in vivo activity of *E. coli* promoter hybrids and promoter mutants that alter the -35 to -10 spacing, *Gene*, **20**, 231, 1982.
- [12] de Boer, H. *et al.*, The tac promoter: A functional hybrid derived from the trp and lac promoters, *Proc. Nat. Acad. Sci.*, **80**, 21, 1983.
- [13] Shimatake, H. and Rosenberg, M., Purified regulatory protein CII positively activates promoters for lysogenic development, *Nature*, **292**, 128, 1981.
- [14] Queen, C., A vector that uses phage signals for efficient synthesis of proteins in *E. coli*, *J. Molec. Appl. Genet.*, **2**, 1-10, 1983.

第九章 大肠杆菌体系中的基因表达

侯云德

大肠杆菌体系是表达外源基因的一个重要宿主细胞系统。在大肠杆菌中有关基因表达的调控机制,包括基因转录、翻译以及翻译后修饰,目前已研究得比较清楚。大肠杆菌体系中的基因表达研究,不仅在理论上可以阐明原核基因的表达原理,更重要的是在实用上可以大量生产自然界难于提取的多肽,用于医药或其他目的。就这方面讲,真核基因在大肠杆菌中的表达研究,尤其显得重要。

自1974年以来,一系列的实践已经证明,许多外源原核基因以及一些真核基因,如胰岛素、生长激素、干扰素、白细胞间质素2等,都能在大肠杆菌中获得有效表达,并大量生产。

要使真核基因在大肠杆菌中获得有效表达,必须首先了解真核基因与原核基因在结构以及表达调控上的差异,真核基因在大肠杆菌体系中表达的必要条件以及影响其表达的因素,真核基因在大肠杆菌中表达的方式等。

第一节 真核基因与原核基因在结构和表达调控上的差异

如表9.1所示,真核基因与原核基因的结构与表达调节在许多方面是不相同的,其中最主要的是,真核基因的不连续性以及其调控信号不能为原核细胞所正确识别。大多数原始真核基因不能在大肠杆菌体系中有效表达的主要原因有如下几点:

(1) 真核基因有内含子,将结构基因分割成几部分。由此,产生了两个问题:第一,真核基因平均来说要比原核基因大得多,例如小鼠二氢叶酸还原酶基因超过42kb;第二,大肠杆菌不能拼接由内含子所分隔的原始转录物,使之产生可翻译的mRNA。当然,如果先由mRNA反转录成cDNA,再进行克隆和表达,就可以解决这一问题,例如卵白蛋白基因,人 γ 干扰素基因等。另外,少数真核基因不含内含子,例如组蛋白基因、人 α 、 β 干扰素基因等。

(2) 缺少能为大肠杆菌RNA多聚酶所识别的DNA序列(启动子),因而不能被有效转录;即使有小量转录物,又由于缺少与大肠杆菌16S核糖体RNA 3'端相匹配的序列(SD序列),也不能被有效地翻译。

(3) 即使克服了上述障碍,翻译的真核蛋白还有可能被大肠杆菌本身的蛋白酶类所破坏。

据此,一个真核基因要在大肠杆菌中获得有效表达,必须具备如下三项基本要求:

(1) 真核基因的编码区没有被内含子所分隔。

(2) 真核基因必须置于大肠杆菌启动子的下游,后者能有效地被大肠杆菌RNA多聚酶所识别。

表 9.1 真核基因与原核基因在结构和表达调控上的差异

差 异	真核基因	原核基因
基因结构	大多数是不连续的,有内含子	绝大多数是连续的,无内含子
转录酶	I、II、III 三类	有各自的 RNA 聚合酶
转录后加工	转录形成前 RNA, 拼接后为成熟 RNA	转录产生成熟 RNA
启动基因调控序列	-12--32 位 (TATA 序列) -70--100 位保持区 最上游转录增强序列	-10 位 (Pribnow 序列) -35 位保守区
翻译起始信号	有的需要 5' 端 7 甲基鸟苷帽子结构	在 ATG 前有 3-12bp 的 SD 序列
转录终止信号	3' 端终止区有 AAUAAA 序列	有依赖 rho 蛋白和不依赖 rho 蛋白两种形式
翻译后加工	能糖基化 (Asn-X-Ser/Thr)、 切割信号多肽; 某些 RNA 病毒的长翻译物 可被切割成几个短的多肽	甲硫氨酸 NH ₂ 端乙酰基可被水解; 1 个或几个 NH ₂ 端残基可被去除; 可切割信号多肽;不能糖基化

(3) 真核基因 mRNA 必须相对稳定和有效地翻译, 另外翻译的外源性蛋白不易被大肠杆菌的蛋白酶类所降解。

可见, 只要掌握原核基因表达调控的基本原理, 许多真核基因不难在大肠杆菌体系中获得有效表达。

第二节 影响真核基因在大肠杆菌体系中表达的因素

基因表达的目的往往是希望获得足够量的产物, 所以, 要使基因高效表达, 除了了解真核基因在原核细胞中表达的必要条件之外, 还要了解影响表达的因素, 这样才能设计组建高效表达载体, 用于制备疫苗或生产活性多肽。

一、启动子

启动子是指能指导 RNA 多聚酶结合到 DNA 上, 并起始 RNA 合成的 DNA 区段。不同的启动子具有不同的效率: 强的启动子起始 mRNA 的合成具有较高的效率; 弱的启动子则只有较低的效率。启动子的强弱是决定基因表达水平的主要因素。

关于启动子的工作原理, 对 lac 启动子研究得比较详细。lac 启动子的核苷酸序列中包含有两种蛋白与之特异性相作用的必要信息: 一是 RNA 聚合酶, 另外是一种正性激活蛋白, 称为分解代谢产物基因激活蛋白(简称 CAP 或 CRP 或 CGA)。转录起始的一般过程是, 先由 CAP 与 cAMP 形成复合物与启动子的特异性位点处相结合, 后者又刺激 RNA 多聚酶与启动子的另一处特异性位点处相作用; 如果这时阻遏蛋白没有结合在操纵基因上, 而 4 种核苷酸又均存在的话, lac mRNA 的转录即可开始。CAP-cAMP 复合物的识别和结合位点是 -55 和 -68 位之间的对称性序列 (图 9.1)。RNA 多聚酶可能在 +19 和 -24 位之间形成开放性复合物。图 9.2 及图 9.3 列出了 29 种启动子的 +1 到 -50 位的序列, 包括大肠杆菌、噬菌体的启动子。由此可见, 其中有两处高度保持区段, 一是在转录起始点上游大约 10bp 处, 有一段相似的序列 TATRATR 或 TATAAT, 称为 Pribnow 序列; 另一处位于转录起始点上游大约 35bp 处。这两个区段对决定启动子

的强度是重要的,因为凡转录效率降低的变种,其序列与上述两个保守区的同源性也差。然而,启动子中中等度保守的区段对其强弱也有关系,并且分隔上述保守区的核苷酸数目对启动子的效率也有重要影响。阻遏蛋白可以识别并结合到与 RNA 多酶酶识别位点不同的序列上,以阻止 RNA 多聚酶-lacP 开放性复合物的形成,从而抑制转录。转录终止也受 DNA 序列中的信号所控制,这一序列的特点是在转录终止位点前有一GC 丰富区,并有一折叠对称,紧接着是 AT 丰富区。但是,也有通过蛋白来控制转录终止的,如 rho 因子,λ 噬菌体的 N 基因产物等。

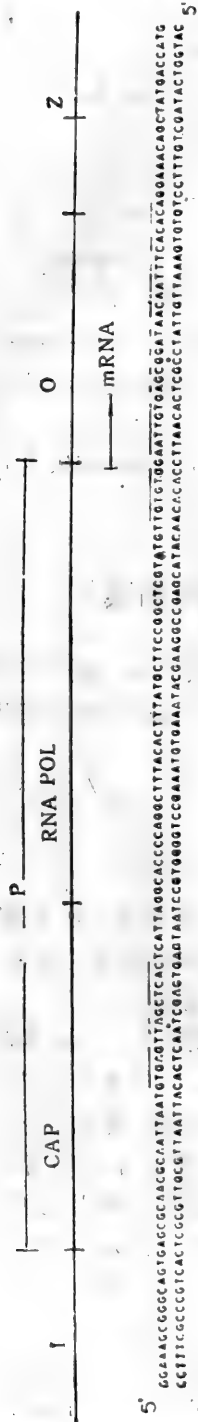


图 9.1. lac P, O 全序列, lacI C 端及 lacZ N 端部分序列。
1. lacI C 端。 2. lacZ N 端。 P. lac 操纵基因。 O. lac 操纵子。 CAP. CAP 位点。 RNA POL. RNA 聚合酶作用位点。序列上方横线为对称性序列。

判断某一启动子的强弱,应当测定其起始相应 mRNA 合成的效率。但是,这种测定方法很难做到。通常是根据其相应蛋白产物的表达水平间接地推算出某一启动子的效率。这种方法不能完全正确地比较不同启动子的效率。这是因为不同 mRNAs 含有不同的 5' 端非编码的领先序列,后者能影响 mRNA 的翻译效率。因此,蛋白产物的表达水平既反映了启动子的强度,又反映了 5' 端非编码的领先序列的组成与长度对翻译影响的程度,或者反映了两者的联合效果。

要使真核基因在大肠杆菌中获得高效表达,必须将欲表达的基因置于强的启动子的下游。常用的启动子有 lac、lpp、trp、tac、λPL、ompF、bla 等。不同启动子的调节机制是不相同的。

lac 启动子: 如前所述, lac 操纵子受两种机制调节。在缺乏乳糖(或其它诱导物)时,由于 lac i 基因产物(lac 阻遏物)与操纵基因相结合,转录受到抑制。另一方面,通过 CAP 发挥正调节作用。在葡萄糖缺乏时,CAP 与 cAMP 形成复合物,与启动子的上游区段相结合,从而刺激转录。所以 lac 启动子在乳糖存在时去抑制,或者在细菌培养基中加异丙基-β-D-硫代半乳糖苷(IPTG, Isopropyl-β-D-thiogalactoside)也可以去抑制,后者可与 lac 阻遏物相结合而去除之。当 lac 启动子以多拷贝存在时, lac 阻遏物往往不能抑制其转录。但是,在过量产生 lac 阻遏物的细菌中(1⁹), lac 启动子的作用可以被抑制。lac 启动子是中等强度的启动子。

trp 启动子: trp 启动子受 trp 阻遏物的调节,在细菌培养基中加 3' 吲哚乙酸(IAA)或 3-β-吲哚丙烯酸可以去阻遏。与 lac 系统不同的是,单拷贝 trp 阻遏物基因产生的蛋白足以使 trp 操纵子完全受抑制,所以,当 trp 阻遏物到达一定浓度时,即使是多拷贝的 trp 启动子也被抑制。trp 启动子是较强的启动子。

tac 启动子: tac 启动子是 trp 启动子 -35 区段与 lac 启动子 -10 区段和 lac 操纵基因融合而成。所以它受 lac 阻遏物的调节,由 tac 启动子组建的质粒,在 I⁺ 宿主菌中加 IPTG 可以去阻遏。tac 启动子也是常用的较强的启动子。

λ P_L 启动子: λ P_L 启动子是常用的强启动子,受 cI 阻遏物的调节。在 cIts857 溶源性细菌中,当 28—32℃ 时, cI 基因可以正常表达,抑制 λ P_L 的作用,当转移到 42℃ 时, cI 基因不能表达, λ P_L 也就解除抑制,在短时间内,可使其控制的基因进行表达。

lpp 启动子: 它是大肠杆菌脂蛋白的基因调控序列,也是常用的高效表达启动子。如连接上 lacO, 可受 lac 阻遏物的调控。

启动子名称	-50	-40	-30	-20	-10	+1
lac	TAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA					
lac I	GACACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCG					
gal	ATTTATTCCATGTACACTTTTCGCATCTTTGTTATGCTATGGTTATTTCA					
araBAD	AGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATA					
araC	T TTCTGCCGTGATTATAGACACTTTTTGTTACGCGTTTTTGTGCATGGCTTTG					
trp	TCTGAAATGAGCTGTTGACAATTAATCATCGAACTAGTTAACTAGTACGCA					
tRNA ^{tyr} ₁₈₂	TTCTCAACGTAACACTTTTACAGCGGCGCGTCATTTGATATGATGCGCCCCG					
λ P _L	TTATCTCTGGCGGTGTTGACATAAAATACCCTGGCGGTGATACTGAGCACA					
λ P _R	CTAACACCGTGCGTGTTGACTATTTTACCTCTGGCGGTGATAATGGTTGCA					
λ P _{RM}	CAACACGCACGGTGTTAGATATTTATCCCTTGCGGTGATAGATTTAACGTA					
λ P _O	TACCTCTGCCGAAGTTGAGTATTTTTGCTGTATTTGTCATAATGACTCCTG					
λ P _{R'}	TTAACGGCATGATATTGACTTATTGAATAAAAATTGGGTAAATTTGACTCAA					
λ c17	TGGTGTATGCATTTATTTGCATACATTC AATCAATTGTTATAAATTGTTATC					
i ³⁵ P _R	CAAGAAAACTGTATTTGACAAAACAAGATACATTGTATGAAATACAAGAAA					
i ³⁵ P _{RM}	TTTGTCAAATACAGTTTTTCTTGTGAAGATTGGGGTAAATAACAGAGGTG					
fd X	CTCTTAATCTTTTTGATGCAATTGCGTTTTGCTTCTGACTATAATAGACAGG					
fd II	ACAAAACATTAACGFTTACAATTTAAATATTTGCTTATAACAATCATCCTG					
fd IIb	GTTTGAATCTTTGCCTACTCATTACACCGGATG TACTTAAAAATATATG					
fd VIII	TGATACAAATCTCCGTTGTACTTTGTTTTCGCGCTTGGTATAATCGCTGGG					
ϕ X 174 A	AAATAACCGTCAGGATTGACACCCTCCCAATTGTATGTTTTTCATGCCTCCA					
ϕ X 174 B	GCCAGTTAAATAGCTTGCAAAATACGTGGCCTTATGTTACAGTATGCCCA					
ϕ X 174 D	GAGATTCTTTGTTGACATTTTAAAAGAGCGTGGATTACTATCTGAGTCCG					
SV40	GAATGCAATTGTTGTTGTTAACTTGTTTATTGGAGCTTATAATGGTTACAA					

图 9.2 启动子序列(1)。+1 为 mRNA 起始点;横线为相似区。

β 内酰胺酶基因启动子: 它本身不能通过生长条件的改变来调节,是一种较弱的启动子。

此外,强的启动子,组成性高水平的转录可以影响质粒 DNA 的复制,从而引起质粒的不稳定性。在强的启动子的下游安置强的转录终止信号序列,就可以避免这个问题。

二、基因剂量

除少数例外,一般来说细菌内基因拷贝数增加,基因产物也可增加。增加基因拷贝数的方法一般有下列几种:

启动子名称	-50	-40	-30	-20	-10	+1
spc	AGCCGTTTATTTTCTACCCATATCCTTGAAGCGGTGTTATAATGCCGCG					
str	GATCGTTGTATATTTCTTGACAACCTTTTCGGCATCGCCCTAAAATTCGGCG					
T7A1	AAAAGAGTATTGACTTAAAGTCTAACCTATAGGATACTTACAGCCA					
T7A2	AAACAGGTATTGACAACATGAAGTAAACATGCAGTAAGATACAAATCG					
T7A3	ACAAAACGGTTGACAACATGAAGTAAACACGGTACGATGTACCACA					
biop 98	GTTGTTAATTCGGTGTAGACTTATAAACCTAAATCTTTTAAATTTGGTTTA					

图 9.3 启动子序列(2)。

(1) 采用松弛控制复制型质粒: 这类质粒的复制与蛋白的继续合成无关, 但依赖 DNA 聚合酶 I 的活性; 在对数生长期, 每个染色体可以有 10—200 个质粒。例如 ColE1、pMB1 质粒, 或含有 ColE1、pMB1 复制子的其他重组质粒。采用这类复制子的重组质粒, 在大肠杆菌的拷贝数一般可达 20—30 个。严紧控制复制型质粒则相反, 这类质粒的复制紧紧地与宿主细胞的繁殖联系在一起。这样, 每个细胞只允许 1 个或少数几个质粒存在, 例如 pSC101。

(2) 采用 pAT153 系统代替 pBR322 系统: 一般表达载体常用 pBR322 的衍生物, 如果改用 pAT153, 则可使质粒拷贝数增加 1.5—3.0 倍, 这是因为 pAT153 移去了 pBR322 的 Hae II B、H 区段, 而在该区段中包含有控制质粒拷贝数的序列。作者在研究人 α D 型干扰素基因的表达时, 曾将 pBR322 系统(pBV867)改成 pAT153 系统 (pBV181), 结果质粒拷贝数增加约 3 倍, 干扰素表达量增加约 5 倍。

(3) 采用失控复制型载体 (runaway plasmid): 有一种质粒 pKN 402, 它是 R1 drd19 的衍生物, 在 30°C 时每个细菌可容纳 20—50 个拷贝。但是在 35°C 以上繁殖时, 质粒的复制失去控制, 在 3 小时末, 在质粒上编码的 β 内酰胺酶的表达量增加了 400 倍。曹旭在研究 T4 DNA 连接酶的表达时, 也获得类似的结果。为便于外源基因的插入, 将 pKN402 进行了改造, 其中之一是 pAS2。如果把一个带有 lac 或 tac 启动子的能表达的基因克隆到 pAS2 上。带有这一杂交质粒的 lac^I 细菌在 42°C 培养几小时后, 质粒的拷贝都可增加到每个细菌 1,000 个以上, 然后加入 IPTG 以灭活 lac 阻遏物, 这样, 基因产物就可以大大增加。这类质粒称为温度诱生性失控复制型载体。

(4) 采用温度诱生性 λ 噬菌体-Col E1 杂交质粒 pKC16: 这一质粒含有 Amp^r 和 Kan^r 基因, 一段 λ DNA (从 λ N 通过 λ P, 包括其复制子和 λ clts857 基因) 以及包含有复制区和免疫区的 Col E1 片段。当培养细菌的温度上升至 42°C 时, λ 转录开始, λ 复制系统被启用, 质粒的拷贝数增加至大约每细胞 250 个拷贝。把外切酶 III 基因克隆到 pKC16 中后, 含有这一重组质粒的细菌先在 32°C 培养, 然后升温至 42°C, 能使外切酶 III 的产量大大增加。

三、核糖体结合位点

大肠杆菌核糖体结合位点对真核基因在该细菌中高效表达是十分重要的。大肠杆菌核糖体结合位点包括起始密码子 AUG 和位于 AUG 上游 3—11 个核苷酸处一段 3—9 个核苷酸序列, 称为 Shine-Dalgarno 序列, 简称 SD 序列, 它与大肠杆菌 16SrRNA 的 3'

端是互补的。一种 mRNA 在大肠杆菌中转译的效率受以下三种因素的左右:

(1) mRNA 中 SD 序列与大肠杆菌 16 SrRNA 3' 端的互补程度。

(2) AUG 与 SD 序列之间的距离以及其序列对翻译效率有明显影响。

(3) AUG 之前 20 个到之后 13 个核苷酸有其适宜的序列, 特别是在 AUG 之后的 1 个核苷酸可以影响与核糖体的结合。领先序列对 mRNA 的翻译也有明显影响。

四、基因产物的稳定性

真核基因的产物在原核细胞中往往不稳定, 易被大肠杆菌的酶类所降解。所以, 即使原始的表达量很高, 由于很快在菌体内被降解, 而实际的产量却很低。有几种办法可以提高真核基因产物的稳定性。

(1) 组建融合基因, 产生融合蛋白: 也就是说, 表达的多肽, 其 N 端是大肠杆菌的多肽序列, C 端是真核多肽。许多融合蛋白与自然的真核蛋白相比较, 在细菌体内比较稳定, 不易被细菌酶类所降解。在某些情况下, 融合蛋白经过化学处理可以释放出具有生物学活性的非融合的真核多肽。

由于融合蛋白中含有一段原核多肽序列, 这可能会影响真核蛋白的免疫原性, 所以一般不能作为人体使用的注射用药, 但可作为抗原用。

(2) 利用大肠杆菌的信号多肽, 把真核基因产物搬运到胞浆周质的空隙中。大肠杆菌的外膜脂蛋白, N 端有 20 个氨基酸的信号多肽, 后者可使表达的脂蛋白前体分子搬运到胞浆膜上, 信号多肽随即被集中在胞浆膜上的酶所识别并切割, 使脂蛋白可以集中在胞浆周质的空隙中, 以组装细菌外膜。一般认为在胞浆周质的空隙中, 外源蛋白不易被细菌酶类所降解, 因而能增加外源多肽的表达量。据此, 如果把一个外源基因与大肠杆菌的信号多肽组成融合多肽, 借大肠杆菌信号多肽的功能, 可把外源基因运送到胞浆周质的空隙中。Masui 组建的 p1N-1B 系列载体即属于这一类。

(3) 利用某些真核多肽中自身的信号多肽, 把真核多肽搬运到胞浆周质的空隙中。某些真核信号多肽如大白鼠胰岛素原的信号多肽可被细菌所识别和切割。作者也发现人 α D 干扰素的信号多肽可以被大肠杆菌所识别并切割, 但切割点不很正确。

(4) 采用位点特异性突变的办法, 改变真核多肽二硫键的位置, 从而增加了多肽的稳定性。例如, 人 β 型干扰素多肽在菌体内不稳定, 采用位点特异性突变的方法, 把多肽键中编码 Cys 的 TGT 中的第一个 T 改为 A, 成为 AGT, 即改变成 Ser, 这种 IFN- β_{175c} 多肽的稳定性就大为提高。

(5) 采用酶缺陷型大肠杆菌。采用缺陷蛋白酶的一种或多种成分的大肠杆菌, 作为表达外源基因的宿主菌。如 Lon⁻ 菌, 可以减少真核蛋白在菌体内的降解。

五、编码多肽密码子的组成情况

一种氨基酸可由几种密码子来编码, 例如, UCU、UCC、UCA、UCG、AGU、AGC 等 6 种密码子均编码丝氨酸; CUU、CUC、CUA、CUG、UUA、UUG 等均编码亮氨酸 (Leu) 等等。而真核基因与原核基因对编码同一种氨基酸所喜用的密码子却不尽相同,

这可能与不同的宿主系统中不同种类 tRNA 的浓度有关。所以,对真核系统中喜用的密码子,在原核细胞中的翻译效率就有可能下降。为了提高表达水平,在人工合成真核基因时,应当考虑使用大肠杆菌所喜用的密码子。

六、表达产物的构象

某些在真核系统中是一种可溶性蛋白,但是在大肠杆菌中表达的却是不溶性蛋白。这很可能是蛋白的构象发生了改变,或由于不合适的二硫键结构,或不合适的折叠所引起的。

七、表达产物的分子量大小

总结迄今真核基因在大肠杆菌中表达的情况可以看到,较小的蛋白,例如分子量为 20,000 左右的蛋白,常可获得高水平的表达,而更大的蛋白,似乎细菌不能耐受,表达水平很低。

八、原核增强序列

近年来相继发现,某些 DNA 序列可以增加某些哺乳动物基因及某些病毒基因组的转录,称为增强子。增强子的作用是顺式的 (cis-acting),无严格方向性,与启动子可以相隔一段距离,并有组织特异性。其作用机制尚不很清楚,核心序列为 GTGGXXXG (X=A 或 T)。一般认为增强子序列是一种 RNA 多聚酶分子进入 DNA 的入口处,进入后能向两个方向移动,所以它对两个方向都可作用,关于是否有原核增强子,目前尚不能肯定。

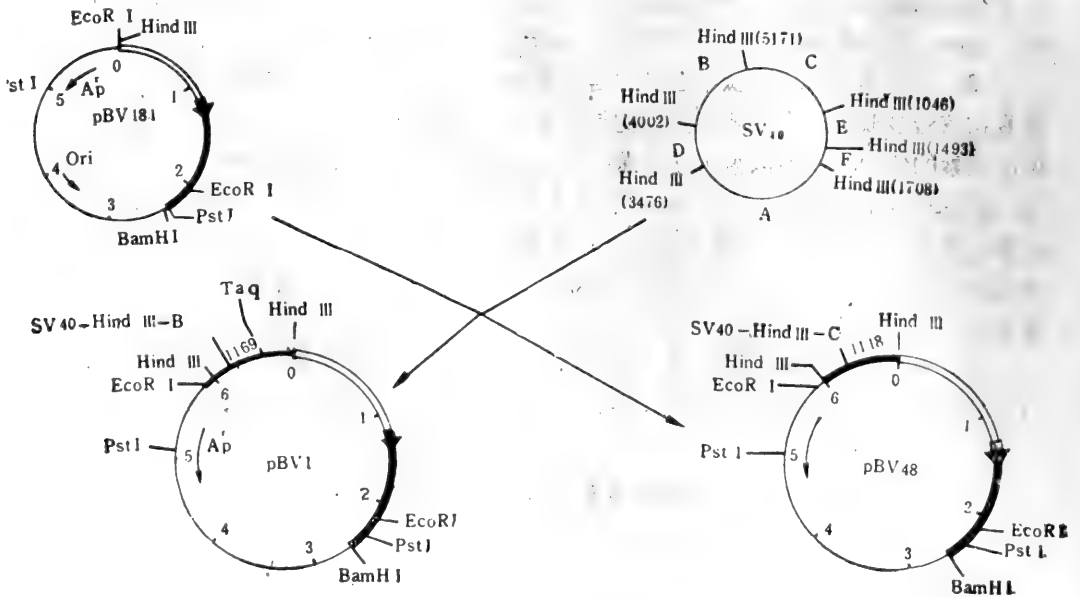


图 9.4 pBV1 及 pBV48 构建图。

1984年作者发现,如将 SV40 DNA Hind III B 片段(1,169 bp)置于 pBV181 中 P_L 启动子的上游(图 9.4),对其下游的干扰素基因表达有明显的增强作用,表达水平平均提高 7 倍。分析 SV40 DNA Hind III B 片段序列发现此片段中 (A + T)% 为 63.5%, G(A 或 T)_xG (x = 5—18) 有 20 处。

表 9.2 SV40 DNA Hind III B 片段对人 αD 型干扰素基因在原核细胞中表达的影响

在 HB101 中干扰素表达水平 (单位/升 $\times 10^{-6}$)				在 BMH71-18 中干扰素表达水平 (单位/升 $\times 10^{-6}$)			
实验次数	pBV1	pBV48	滴度比	实验次数	pBV181	pBV1	滴度比
1 (n = 4)	14.1	3.05	4.62	1 (n = 4)	31.9	10.7	2.98
2 (n = 2)	8.5	0.8	10.56	2 (n = 3)	23.9	5.85	4.09
3 (n = 4)	11.1	2.4	4.64	3 (n = 2)	65.4	3.75	17.44
4 (n = 4)	5.5	1.1	5.19	4 (n = 3)	32.9	3.7	8.89
				5 (n = 6)	6.8	3.2	2.13
平均滴度比			6.3	平均滴度比			7.1

第三节 真核基因在大肠杆菌中表达的方式

如图 9.5 所示,真核基因在大肠杆菌中的表达方式概括起来有下述三种。

一、融合基因-融合蛋白

这种融合蛋白的氨基端是原核序列,羧基端是真核序列。表达融合蛋白的优点是基因操作简便,在菌体内比较稳定;缺点是只能作抗原用。

表达融合蛋白的必要条件是,真核基因的读码框架必须与原核基因的读码框架相符合。使原核基因与真核基因的框架相匹配,有如下数种方法:

(1) 如果外源真核基因序列已经清楚,表达载体中要与之相连接的原核基因的序列也已经清楚时,要设计选择一定的酶切位点,使它们相连接后框架符合。

(2) 利用末端转移酶通过相互匹配的同种寡核苷酸的延伸进行连接 (dA-dT 或 dG-dC)。由于核苷酸延伸的数目长短不一,这样自然可以形成三种读码框架的重组体,经表达选择就可以选出框架正确、有表达的重组体。

(3) 建成一套三种所谓位相载体,插入位点分别形成三种读码框架。这样,在插入外源基因后,三种重组质粒中保证有一种质粒的读码框架是正确的。

(4) 如果真核基因的两端有较长的侧翼时,可用 Bal 31 外切酶,随机消化,然后修平,装上连接子,与载体连接。由于是随机消化,可以自然形成三种读码框架的重组体。

(5) 加上专门设计的人工合成的 DNA 接头或定位突变方法,使真核基因的读码框架与载体上原核基因的读码框架相吻合。

二、非融合基因-非融合蛋白

所谓非融合蛋白是指在细胞内合成的蛋白以要表达的真核蛋白 mRNA 的 AUG 为起始,在其氨基端不含任何细菌多肽序列(图 9.5)。为此,表达非融合蛋白的操纵子必须构建:原核启动子—原核 SD 序列—真核基因起始密码子、结构基因及终止密码子。为此目的,可以改建一种表达载体,含有原核启动子和细菌 SD 序列;在启动子的上游和 SD 的下游 3—10 个核苷酸处,有单一性限制性内切酶切点,使用时可以将其切下,分离出来,这称为携带式启动子。要表达的真核基因则根据其基因结构的不同情况,采用多种修饰酶、限制性内切酶,切去真核基因 ATG 前的非编码序列,然后把携带式启动子直接置于真核基因的 ATG 之前。如果要表达的真核基因编码是一种分泌性蛋白,而成熟蛋白的第一个密码子又不是 ATG,那么需要人工合成一个带有限制性内切酶位点和 ATG 的接头,把真核结构基因与 SD 序列连接起来,连接时不能改变原来真核基因的读码框架。此外,也可以改建一种表达载体,使之带有原核启动子、细菌 SD 序列和 ATG 密码子。其中 SD 序列和 ATG 在启动子的下游,SD 序列与 ATG 之间有适当的距离,而 ATG 之后有单一的限制性内切酶切点,经消化后突出部分用 S1 核酸酶去除,使在 ATG 之后可以直接连接外源结构基因。在此,外源结构基因先经限制性内切酶和修饰酶加以改造,必要时加一段人工接头,在与 ATG 连接之后不致改变原来真核基因的核苷酸序列和读码框架。此外,还可以采用定位突变技术,来表达非融合蛋白。

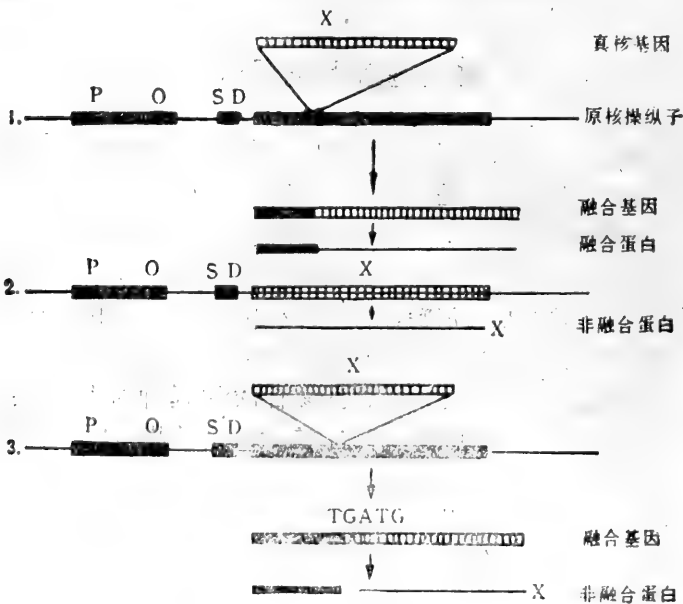


图 9.5 真核基因在大肠杆菌体系中表达方式示意图。

P. 原核启动子。O. 原核操纵基因。SD. 核糖体结合位点。X. 真核基因或多肽。
TGATG. 终止密码与起始密码子的重叠序列。粗黑线为原核操纵子部分。

三、融合基因-非融合蛋白

这是作者在研究干扰素基因在大肠杆菌体系中表达调控时建立的一种方式。如图 9.5 所示,这一表达模式是:原核启动子—SD 序列—原核 ATG—原核结构基因—TGATG—真核结构基因。这一表达模式是利用 TGATG 序列,一方面 TGA 终止第一种(原核)读码框架的翻译,另一方面 ATG 起始第二种(真核)读码框架的翻译,其中 A 是公用的。作者证明,第二种读码框架的翻译效率比第一种读码框架的翻译效率高约 6 倍。

第四节 真核基因在大肠杆菌体系中的表达设计

根据上述真核基因在大肠杆菌中表达的必要条件、影响表达的因素、表达的方式以及各个特定真核基因结构的特点,不同基因表达的设计是多种多样的,但是,表达设计的核心是如何利用好原核启动子。所以,以下分别介绍不同启动子表达真核基因的设计,包括表达载体的组建与不同真核基因的表达方案。

一、lac 启动子

lac 启动子是原核细胞中研究得较详尽的基因表达调控系统,也是在原核基因工程中最早使用的启动子。如前所述, lac 操纵子受 lac 阻遏物的负调节及 CAP 的正调节,在组建 lac 启动子的表达载体时,常保留 lac P、O 部分,去除 CAP 的识别区段。1976 年 Polisky 等将含有部分 lac 操纵子的转导型噬菌体 (λ p lac5) DNA 的 Eco R1 片段插入到 Col E1 衍生质粒中,组建成一种含有 lac P、lac O 和大部分半乳糖苷酶基因的杂交质粒。在此基础上又组建成含有较小的“携带式” lac 启动子片段的质粒,即将含有 lacP、O 和 β 半乳糖苷酶头 8 个密码子的 203bp 的 Hae III 片段与用 Eco R1 切开补平的 pBR 322 DNA 平端连接所组成。由于这种连接使连接处恢复了 Eco R1 切点,用 EcoR1 消化又可将这段 lac 启动子取下,所以称为携带式。含有这种质粒的细菌可以组成性地合成 β 半乳糖苷酶,在含 X-gal 的培养基上可以形成蓝色菌落。这是由于 lacO 的多拷贝消耗了 lac 阻遏物,导致染色体上 β 半乳糖苷酶的去抑制而引起的。lac 启动子上由于 CAP 位点突变,已对 CAP 的作用不敏感。为了表达非融合蛋白又进一步改建成含有 95bp 的 Alu I 片段的表达载体,后者含有 lac UV5,去除了 CAP 结合位点、lacO 以及大部分 SD 序列,刚好去除了 β 半乳糖苷酶的 ATG。以后,不同实验室组建成类似结构的 lac 启动子表达载体,常用的有 pUR222 (Rüther, 1981)、pPC ϕ (1, 2, 3) (Charnay 等, 1978)、pOP203-13 (Fuller, 1981)、p β -gal 13C (Goeddel 等, 1979)、pOMPO (Charnay 等, 1979)、pGH6 (Goeddel 等, 1979)、pWR 系列(郭礼和等, 1984)等。

pUR222: 2.7kb, amp^r, 带有 lac P、O 以及 Z 基因。在 Z 基因中有多连接子切点: Pst I、Taq I、Acc I、Hinc II、Sal I、BamHI 及 Eco R1, 在此位点中插入外源基因,如果框架正确,可以表达融合蛋白(图 9.6)。

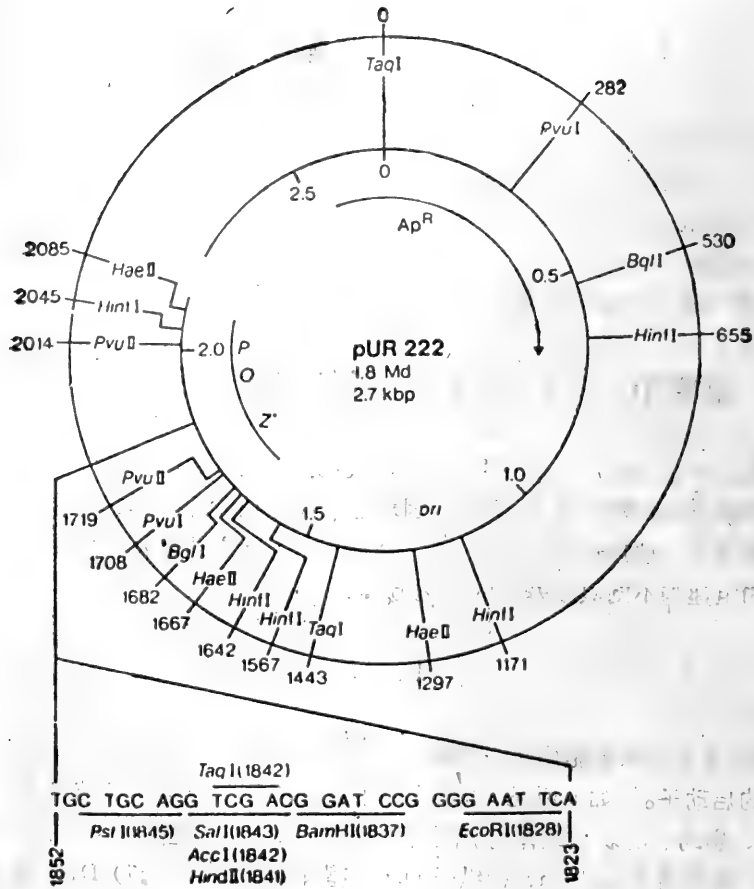


图 9.6 pUR222

pPC ϕ 1, 2, 3: 4.4kb, Amp^r, 由 pBR322 衍生而来。在单一切点 Eco R1 处插入外源基因可以表达带有 β 半乳糖苷酶头 8 个氨基酸的融合蛋白。共有三种读码框架。

pOP203-13: ~4.5 kb, 来源于 pBR322, 带有 lac UV5 启动子和 Eco R1 单一切点。在此处插入外源基因, 如果框架正确, 可以表达为融合蛋白, 后者含有 β 半乳糖苷酶头 7 个氨基酸和 Eco R1 连接子编码的几个氨基酸(图 9.7)。

p β -gal 13C: ~7kb, 由 pBR322 衍生而来, 带有 lac 启动子以及 lac Z 基因大片段, 后者编码 β 半乳糖苷酶 N 端头 1,005 个氨基酸, 在单一切点 Eco R1 处插入外源基因, 如果框架正确, 可以表达大的融合蛋白, 在细菌内比较稳定(图 9.8)。

一些真核基因采用 lac 表达载体在大肠杆菌中进行了有效的表达, 分别介绍如下。

1. 生长激素释放抑制因子 (somatostatin)

生长激素释放抑制因子是只有 14 个氨基酸的短肽。由于它的氨基酸序列早已清楚, 又有敏感的放射免疫检测方法, 所以它成为真核基因在大肠杆菌体系中表达成功的第一个例子。1977 年 Itakura 等按照生长激素释放抑制因子的氨基酸序列, 人工合成了它的基因, 两端具有 Eco R1 和 BamHI 粘末端, 可与 lac 启动子的表达载体连接(图 9.9)。在其基因 Eco R1 位点前有一段较长的 β 半乳糖苷酶片段, 可组合 β 半乳糖苷酶-生长激

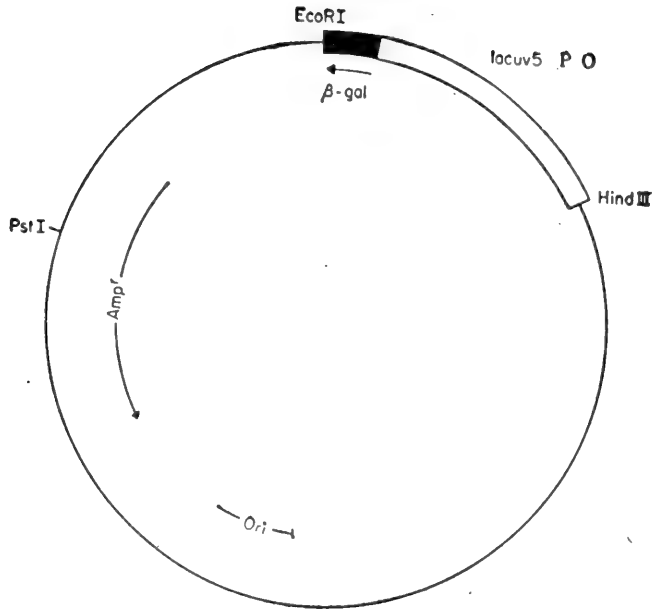


图 9.7 pOP203-13

素释放抑制因子的融合蛋白,然后在试管内用溴化氰在甲硫氨酸处切割,成为有活性的生长激素释放抑制因子。在外源基因与表达载体连接时,前者的读码框架与 β 半乳糖苷酶的框架是一致的。采用溴化氰在甲硫氨酸处切割以去除融合的细菌多肽的办法,已广泛地用于其他多肽的表达。

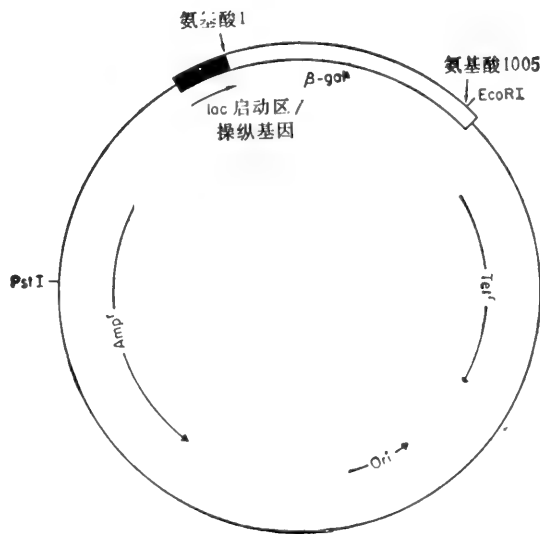


图 9.8 p β -gal 13C

似,只有两处不同,即在 27 位小鼠是 His,人是 Tyr;在 31 位,小鼠是 Gln,人是 Gly。Shine 等(1980)采用 lacP、O 和 β 半乳糖苷酶在 N 端融合的办法成功地使小鼠 β 内啡肽在大肠杆菌体系中获得表达。他们以小鼠促肾上腺皮质激素 (ACTH) 和促脂肪激素 (LPH) 的前体肽的 cDNA 克隆作为 β 内啡肽 cDNA 的来源,这一 cDNA 片段含有 β 内啡肽的全部编码序列,但 C 端缺 Gln。组建时 C 端要补上 Gln 以及终止密码子 TGA。然后与 β 半乳糖苷酶基因连接,并使读码框架吻合。由于 β 内啡肽的第 5 位氨基酸是 Met,所以不能用 CNBr 切割。可以用另一种方法切割以去除融合部分。在 β 内啡肽基因的前面有 β 促黑素细胞激素的基因,其中有一个对胰酶切割敏感的 Arg 位点,而在 β 内啡肽内部则无此切点。在 β 内啡肽中的 Lys,则可设法避免胰酶的作用。这样经过胰酶消化可以获得具有活性的 β 内啡肽。

4. 卵白蛋白

1978 年 Mercereau-Pujalon 等采用具有 EcoR1 单一切点(在 β 半乳糖苷酶的第 8 个密码子后)的 lac 启动子表达载体,利用平末端连接卵白蛋白 cDNA 克隆的 Hha1 片段。框架吻合,形成 β 半乳糖苷酶-卵白蛋白融合蛋白。在此蛋白中,原卵白蛋白 N 端的 5 个氨基酸已被去除,换上 β 半乳糖苷酶 N 端的 8 个氨基酸。这是分子量大于 40,000 的真核基因在大肠杆菌中表达的首次报告。但是表达水平不高,这可能是因转录或翻译效率不高,或者被菌体内的蛋白酶类所降解,也可能是密码子的使用不是大肠杆菌所常用的。

5. 人生长激素 (HGH)

1979 年 Goeddel 等将 HGH 基因置于 lac 启动子(双拷贝)(pGH6)下游,获得了有效表达。他们首先从 HGHcDNA 中克隆出 551bp 的 HaeIII 片段,后者编码 HGH 的 24—191 位氨基酸;另外人工合成一个基因片段,包括 ATG 以及编码 1—24 位氨基酸的序列。两个片段分别克隆到 pBR322 上,cDNA 片段经 dC-dG 连接插入 pBR322 的 Pst-1 位点处;人工合成片段有 EcoR1 及 Hind III 两末端,插入 pBR322 的相应位点处;从后者取出 EcoR1-HaeIII 的 77bp 片段,从前者取出 HaeIII-Xma1 的 512bp 片段,两者连接,用 EcoR1 和 Sma I 消化,使之一端为 EcoR1 粘末端,另一端为平末端,然后插入 pGH6 表达载体(lac 启动子为双拷贝),从而获得表达。它与胰岛素和 β 内啡肽融合多肽不同,HGH 是可溶性的。在 lac SD 与 HGH 的 ATG 之间的距离为 11 bp 时,表达较好。大肠杆菌产生的 HGH 与自然的一样具有相同的活性,氨基酸的序列与二硫键的位置也与自然的一样,只是在 N 端多一个 Met。

由上可见,不同的真核基因应采取不同的表达方案。迄今为止,采用 lac 表达载体来表达的真核蛋白还有多种,例如大白鼠胰岛素原(分子量 12,000)、胸腺素 $\alpha 1$ (3,100)、亮内啡肽(600)、 α 新内啡肽(1,200)、 β 内啡肽(3,200)、Rous 肉瘤病毒(RSV)蛋白激酶(src)(60,000)、单纯疱疹病毒(HSV)胸苷激酶(TK)(43,000)、家兔 β 珠蛋白(19,000)、人 β 干扰素(20,000)、人 α_1 和 α_2 干扰素(20,000)、流感病毒血凝素(61,000)、乙型肝炎病毒核心抗原(HBcAg)(19,000 或 22,000)、乙型肝炎病毒表面抗原(HBsAg)(22,600)、多瘤病毒 t 抗原(20,000)、SV40t 抗原(19,000)等。

当使用 lac 表达载体时,为了寻找 SD 与 ATG 之间适合距离,而表达的基因产物又缺少检测方法时,可以采用如下方案: 将外源基因的 3' 端与 β 半乳糖苷酶基因相连接,然后把携带式 lac 启动子置于这一融合基因的前面,在连接前用 Bal31 消化,使

SD 与外源基因 ATG 之间有不同距离, 然后在 X-gal 培养基上选出蓝色较深的菌落, 即选择到 SD 与 ATG 之间有适合距离, 表达水平较高的菌落, 最后再切去 β 半乳糖苷酶部分。

二、trp 启动子

大肠杆菌 trp 操纵子调控一连串 5 个基因, 其基因产物可催化合成色氨酸。图 9.10 是 trp 操纵子的示意图。

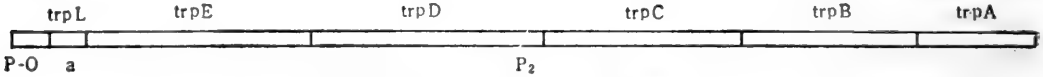


图 9.10 大肠杆菌 trp 操纵子示意图。

a. trp 弱化基因; P₂ 低水平内部启动子。

如将大肠杆菌 trp 操纵子组建在一个质粒上, 5 个基因的表达水平很高, 在用 IAA 诱导之后, 表达产物可以占菌体总蛋白的 20—25%, 可见 trp 是一个很强的启动子。自 1980 年以来, 多种带有 trp 启动子的表达载体相继组建成功。这种载体有两类, 一类是用于表达融合蛋白的, 另一类是用于表达非融合蛋白。用于表达融合蛋白的载体含有 trp 启动子、操纵基因 trpO、领先序列和弱化基因 (attenuator)、全部 trpE 和部分 trpD 基因, 或者再加上多连接子位点。由于 trpE、D 基因长度不同又可组建成多个衍生物, 将外源基因插在多连接子位点处, 框架要吻合, 在大肠杆菌 trpR⁺ 宿主细胞中 (例如 C 600trpR⁺), 用 IAA 诱导, 可以获得高度表达。用于表达非融合蛋白的载体含有 trpP、O 以及转录起始后一段序列, 或包括 S.D 序列、多连接子位点。将外源基因插在多连接子位点处, SD 与 ATG 之间距离合适, 在 trpR⁺ 宿主细胞中, 在 IAA 的诱导下, 可以高效表达。

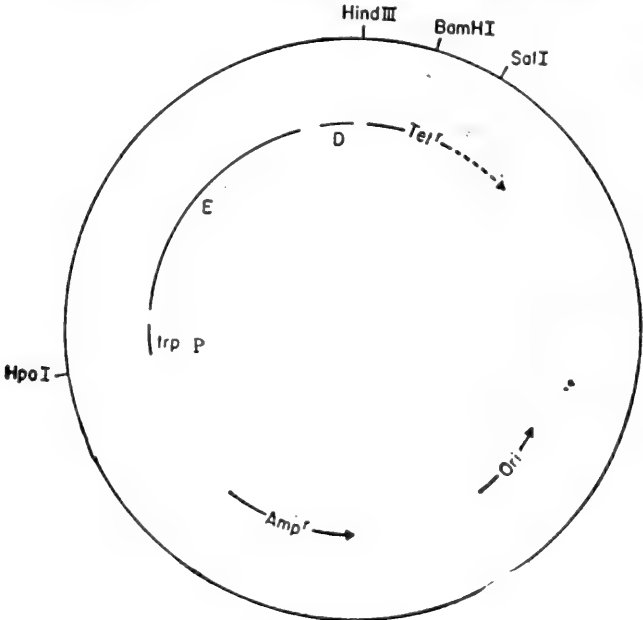


图 9.11 p trp ED5-1

为便于基因操作,现有商品 *trp* 启动子片段出售,这种 *trp* 启动子片段的结构为:单一限制性内切酶切点—*trpP*、*O*—领先序列—*SD*—多连接子位点。

常用的表达融合蛋白的 *trp* 载体如下。

ptrpED5-1: 由 Hallewell 等(1980)组建, 6.7kb, 含有 *trp* 启动子, 在 *trpD* 中有 *Hind III* 单一切点, 在此位点插入外源基因可以表达融合蛋白, 后者包括 *trpD* 蛋白 N 端 75 个氨基酸以及外源多肽(图 9.11)。Tacon 等(1980)将此质粒加以改造, 外源基因可以有三种框架插入 *Hind III* 单一切点, 融合蛋白中包括 *trpE* 的头 7 个氨基酸。经过改造的三种框架质粒称为 *pWT111*, *pWT121*, *pWT131*, 复制子与 *ptrpED5-1* 是相同的, 为 *pMB1*, 4.8kb, 带有 *Amp^r* 基因。

pNCV: 由 Goeddel 等 (1980) 组建, 4.5kb, 复制子是 *pMB1*, 带有 *trp* 启动子, *trpΔLE1413*, 并把 *trpLE* 的终止密码去除, 接上 *EcoRI*-*Pst*.1-*EcoRI* 的人工接头, 保留原来 *pBR322* 的 *tet^r* 基因。 *trp* 启动子缺少弱化基因区段; 甚至含有过量的色氨酸时, 也能部分诱生(图 9.12)。

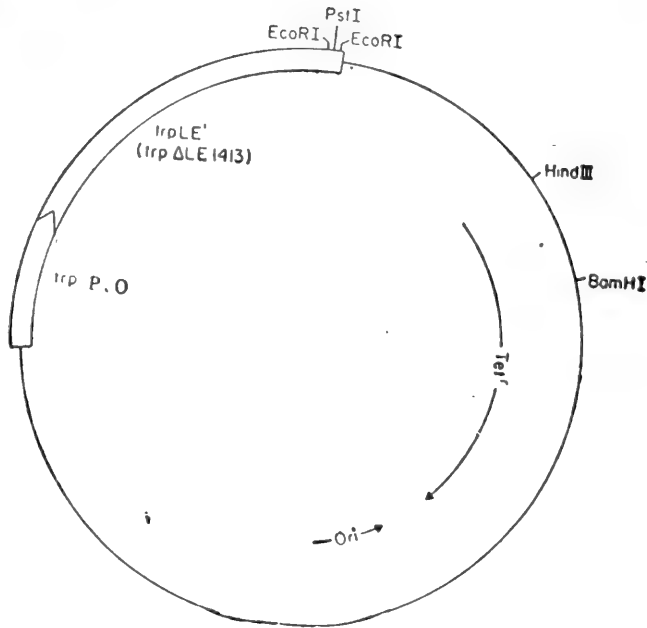


图 9.12 *pNCV*

常用的表达非融合蛋白的 *trp* 载体如下。

pDR720: 由 Bennett 实验室组建(1982), 约 4.0kb, 含有 *trpP*、*O* 以及 *SalI*、*BamHI*、*SmaI* 3 个单一切点, 分别位于 +14, +20, +27bp 处(图 9.13)。

ptrpL1: 由 Edman 等(1981)组建, 约 4.6kb, 带有 *trp* 启动子和 *trp LSD* 序列。在 *trp LSD* 序列下游 3 bp 处有单一 *ClaI* 切点。在启动子上游有 *HindIII* 单一切点。用 *ClaI* 切割后, 造成的突出部分用大肠杆菌 DNA 多聚酶 I Klenow 片段修补, 然后再用 *Hind III* 消化可以分离出带有 *trp* 启动子的片段。这一携带式启动子可以放在要表达的真核基因前面来表达非融合蛋白。另一方面, 在此质粒的 *Cla I* 切点处也可以直

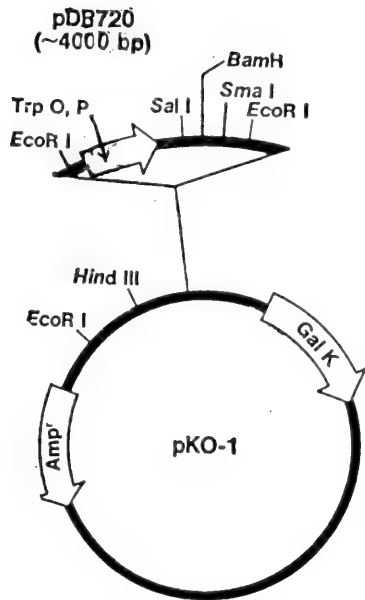


图 9.13 pDR720

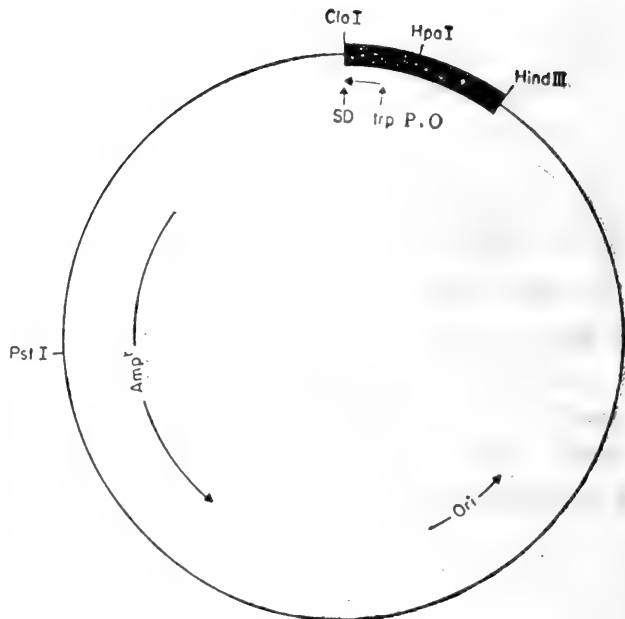


图 9.14 ptrpL1

接插入外源基因,使外源基因的 ATG 直接置于 SD 序列的下游(图 9.14)。

一些真核基因在 *trp* 表达载体中获得了有效表达,现就几个主要的介绍如下。

1. 鸡瘟流感病毒血凝素 (Emtage 等, 1980; Heiland 等, 1981)

将鸡瘟流感病毒血凝素 (HA) 基因的 cDNA 克隆,插入 pWT121 的 Hind III 位点处,所表达融合蛋白的组成应是: *trpE* 的头 7 个氨基酸,来自连接子的 6 个氨基酸,来自 HA cDNA (dT)₁₉ 尾的 6 个苯丙氨酸,来自 HA cDNA 5' 端非编码区的 7 个氨基酸,来自 HA 信号序列和结构基因的 558 个氨基酸,来自 C 端 Hind III 接头的 5 个氨基酸,一共是 589 个氨基酸。但是表达的多肽,根据 PAGE 的结果,却比推算的要小,说明表达的多肽在菌体内发生了某种加工。其表达量很低,这可能是真核基因中信号多肽未去除的缘故。如果把 HA 中的信号多肽去除,可以大大提高其表达量。

2. 口蹄疫病毒 (FMDV) VP1 (Kleid 等, 1981)

采用上述 pNCV 一类质粒,质粒上有 *trp* 启动子, *trp* ΔLE1413 与 FMDV 的 VP1 基因相融合。其特点是,有 190 个氨基酸的 *trp* LE 片段是不溶性的,可以抵抗蛋白水解;其次,由于去掉了弱化基因位点,以致使任何 mRNA 的二级结构导致减弱的作用可以避免 (Yanofsky, 1981)。表达的融合蛋白 *trp* LE-VP1 为 44,000, 大约占全细菌蛋白的 17%。

3. 人 αA 干扰素 (Goeddel 等, 1980)

为了去除人 αA 干扰素基因中的信号序列,利用在 αA 干扰素基因的第一和第二个密码子中有 Sau3a 的特点,首先分离出 34bp 的 Sau3a-AvaII 片段,加上一个人工接头含有 EcoR1 粘末端和 ATG 起始密码子,然后与其他的片段相连接,组成完整的人 αA 干扰素基因(无信号序列) EcoR1-Pst·1 片段。将此片段插入 *trp* 表达载体,后者含有 *trp* P、O 和 SD 序列以及 EcoR 1 位点。结果获得高水平表达 (2.5×10^8 u/L) (图

9.15)。

4. 人 β 干扰素 (Goeddel 等, 1980)

由于人 β 干扰素基因中缺少合适的酶切点, 所以要表达成熟人 β 干扰素基因(去信号序列)必须采取另一种设计方案。如图 9.16 所示, 首先从含有人 β 干扰素 cDNA 全序列的质粒上切出 HhaI 片段, 变性后与人工合成的一段引物进行退火, 后者含有 ATG 和编码成熟人 β 干扰素头 4 个氨基酸的序列, 然后用 DNA 多聚酶 Klenow 片段进行修复合成; 与此同时, 酶的 3'-5' 外切酶活性可以去除 3' 端的突出部分使成平末端, 然后用 PstI 分离出 141bp 片段, 再与基因的其余部分连接成为完整的成熟人 β 干扰素基因。将此基因插入 trp 表达载体, 后者带有 trp P、O 和 SD, 从而获得高效表达。进一步研究表明, SD 与 ATG 之间的距离为 8bp 时, 表达量最高。

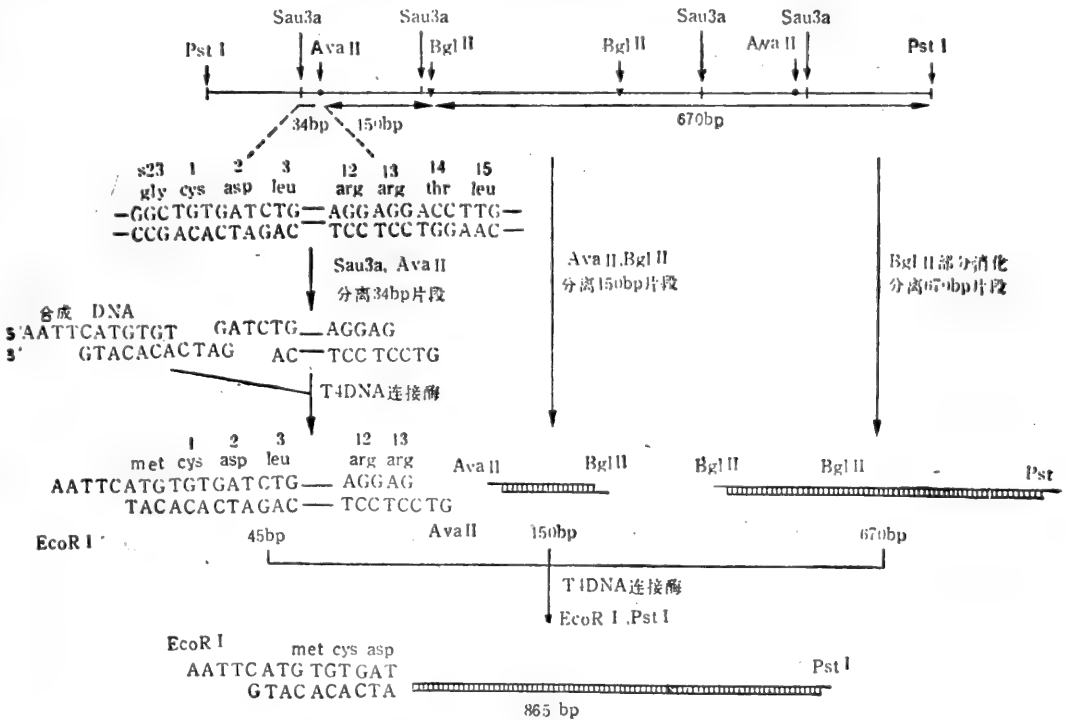


图 9.15 人 α A 干扰素基因的构建(引自 Goeddel 等, 1980)。

采用 trp 启动子来表达的真核蛋白还有: 人生长激素 (HGH) (分子量为 24,000)、Rous 肉瘤蛋白激酶 (src) (60,000)、滤泡性口腔是病毒 G 蛋白 (62,000)、乙型肝炎病毒表面抗原 (HBsAg) (22,600)、乙型肝炎病毒核心抗原 (HBcAg) (22,000)、人 α 、 β 干扰素 (20,000)、人 γ 干扰素 (17,000)、人血清白蛋白 (67,000)、凝乳酶 (35,000) 等。

三、tac 启动子

tac 启动子是由 try 启动子 -35 区、lac UV-5 与启动子 -10 区段以及 lac O 所组成, 受 lac 阻遏物调节, 在 lac I^r 宿主菌中, 启动子受抑制, 加 IPTG 后可以抑制。

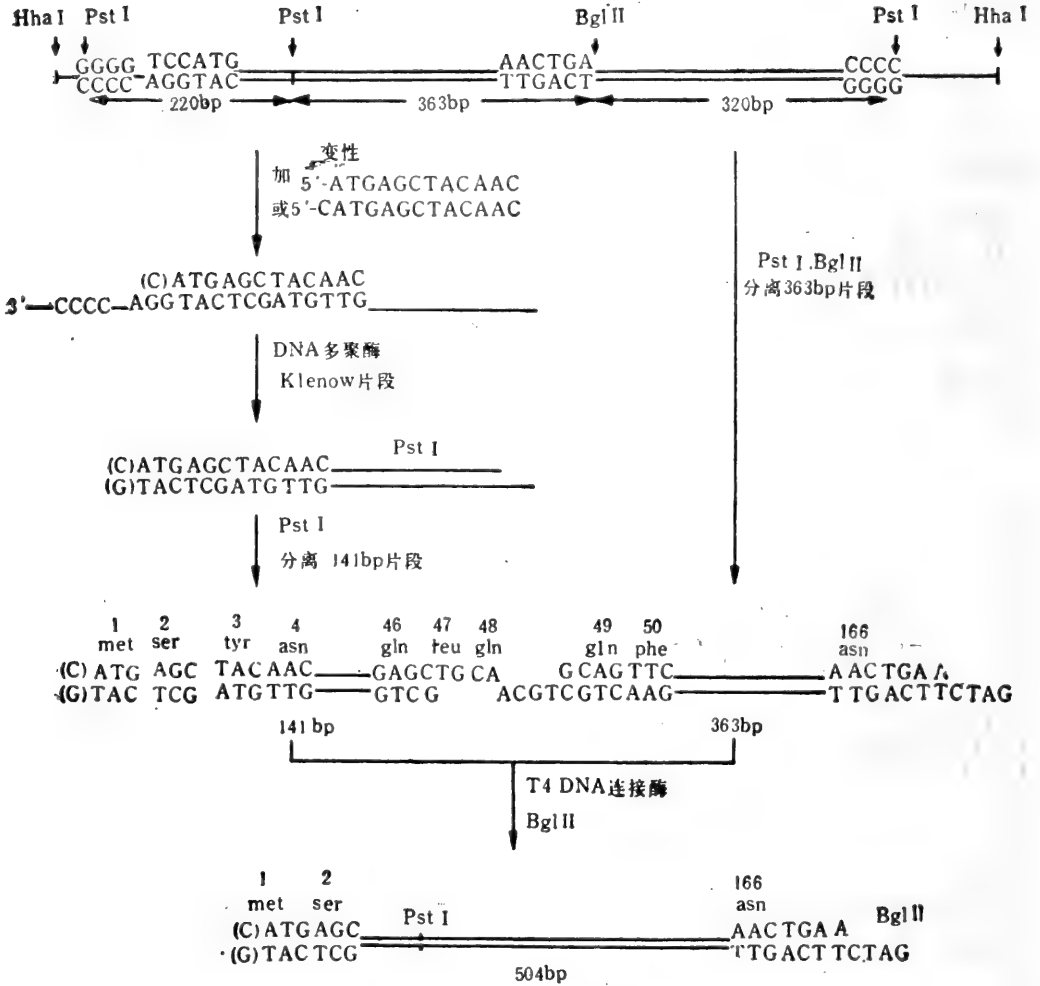


图 9.16 人 β 干扰素基因构建图。

常用的 tac 表达载体如下。

pDR540: 带有 tac 启动子, 在 +35 位, lac SD 下游 5bp 处有单一 BamHI 切点, Amp^r 以及 Gal K 编码区(图 9.17) (Russel 等, 1982)。

ptac12: 约 2.6kb, 含 tac 启动子, 用 EcoRI 和 pVU II 消化后可以分离出含有启动子的约 260 bp 片段。这段 DNA 有 tac 启动子和 lac 领先序列。pVU II 位点在 lac Z 基因 SD 序列下游 5bp 处。这种携带式启动子可以插入待表达基因的起始密码子的前面, 来表达非融合蛋白(图 9.18) (Amann 等, 1982)。

安淑芳等 (1985) 采用 tac 启动表达了 HBcAg; 吴淑华等 (1985) 采用 tac 启动子表达了人 β 干扰素。

四、P_L 启动子

在正常噬菌体感染周期中, P_L 启动子控制 DNA 的早期左向转录, 通过 N 基因到

int。P_L 是一种强启动子，受 cI 基因产物(阻遏物)的调节控制，在随后的感染中又受 cro 蛋白的调节。当 P_L 启动子克隆到质粒时，常使质粒不稳定，这可能是因为高水平的转录对 DNA 复制产生有害的影响，克服的办法是将带有 P_L 启动子的质粒转化到能够抑制 P_L 转录的宿主菌中。在实际工作中，常采用带有 cIts 变异的溶源性细菌 (cIts857)。在 28—32℃ 培养时，产生 cI，抑制 P_L 转录；培养温度上升至 42℃ 时，cI 基因不能正常产生 cI，P_L 解除抑制，在短期内可以大量转录。一般来说 P_L 启动子比 lac 启动子要强 8—10 倍。

组建的 P_L 表达载体有两种，一种是表达融合蛋白的，另一种是表达非融合蛋白的。

pKC30：约 6.4 kb，带有 λP_L 启动子，在 +321 位点处有 HpaI 单一切点。该质粒是将 2.4 kb 的 λDNA Hind III-BamH1 片段插入 pBR322 的 Hind III 和 BamH1 位点处构成，插入区段含有 P_L 启动子序列、N 基因产物识别序列 (nutL)、N 基因本身以及强的依赖 r_{zho} 的转录终止信号 t_o。HpaI 位点位于

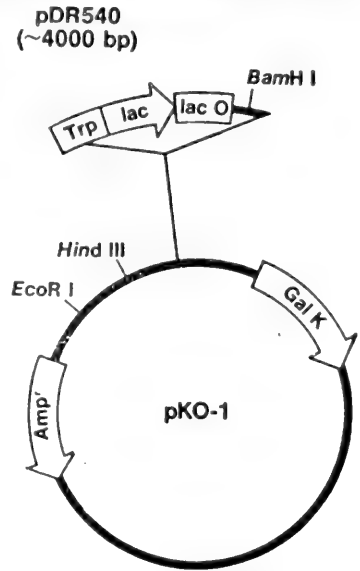


图 9.17 pDR540

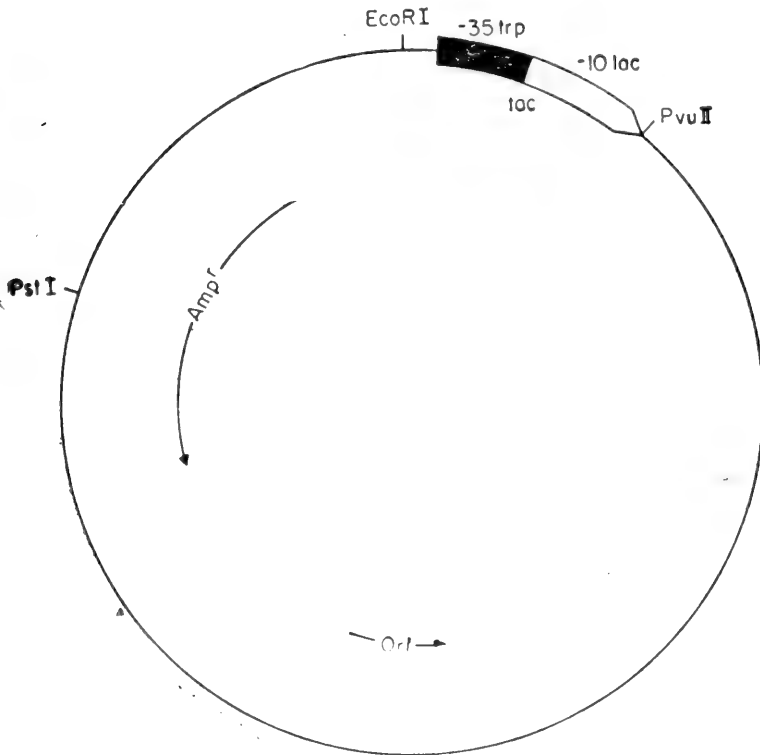


图 9.18 ptacl2

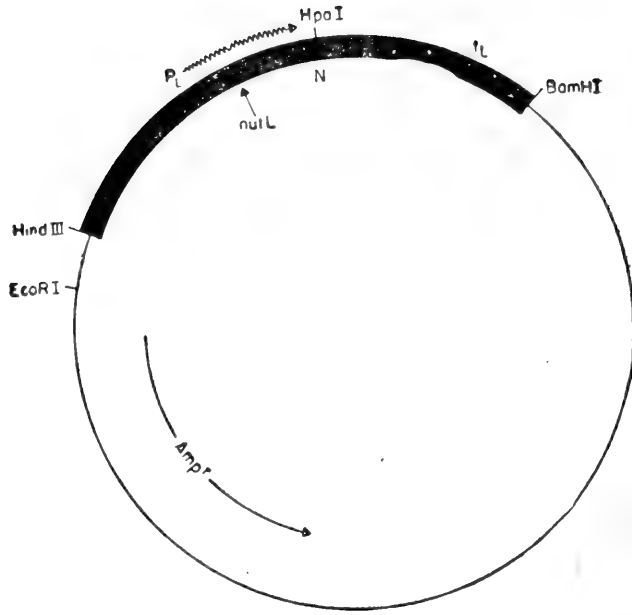


图 9.19 pKC30

N 基因编码区内(图 9.19) (Shimatake 等, 1981)。

pLC24: 3.3kb, 含有 λ P_L 启动子和 MS2 噬菌体多聚酶基因及核糖体结合位点 (SD)

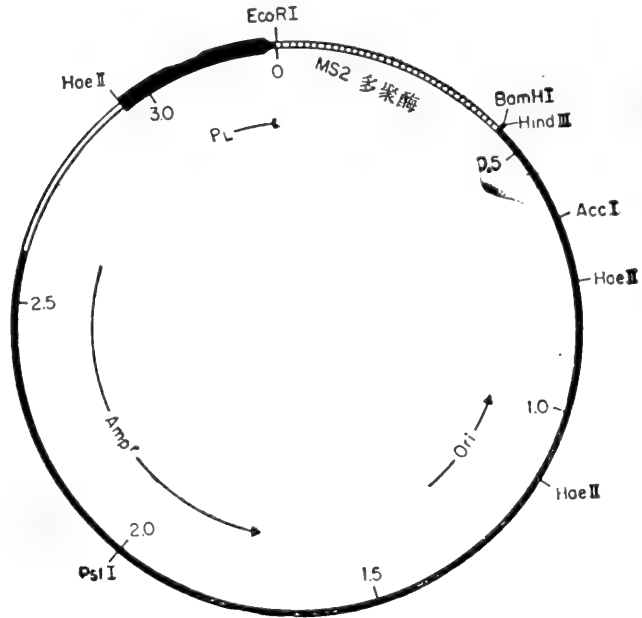


图 9.20 pLC24

序列)。在 BamH1 或 Hind III 位点处插入外源基因,如果框架正确,可以表达融合蛋白,后者含有 MS2 多聚酶的98个氨基酸(图 9.20) (Derynck 等, 1980)。

pCP 系列: pCP3 (7.05kb)、pCP39 (~5.2kb) 和 pCP40 (~5.2kb) 均带有 P_L 启动子、Amp^r 基因以及失控复制质粒 pK402 序列。在 P_L 启动子下游有多人工接头,可以插入外源基因。但是在 P_L 启动子的下游没有 SD 序列,插入的外源基因必须带有 SD 序列以及真核基因的 ATG 才能表达。由于有失控复制质粒 pKN402 序列,在提高培养温度至 42°C 时,质粒的拷贝数可以明显增加,因而基因产物也可明显增加 (图 9.21) (Remaut 等, 1981)。

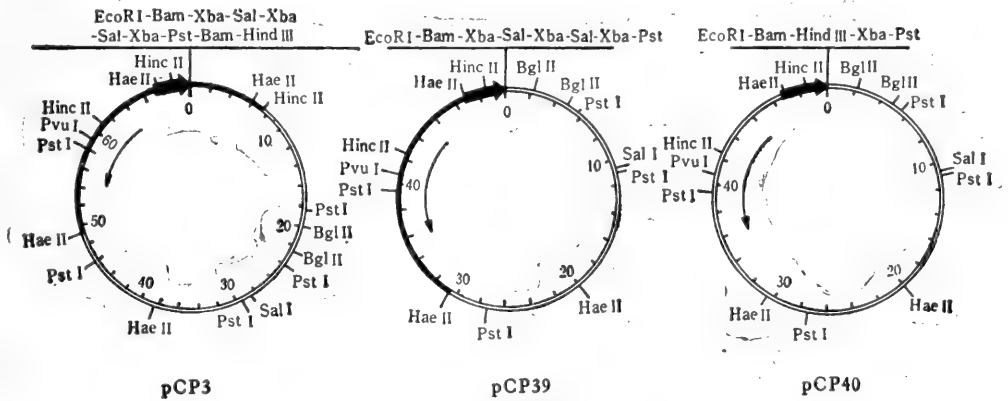


图 9.21 pCP 系列

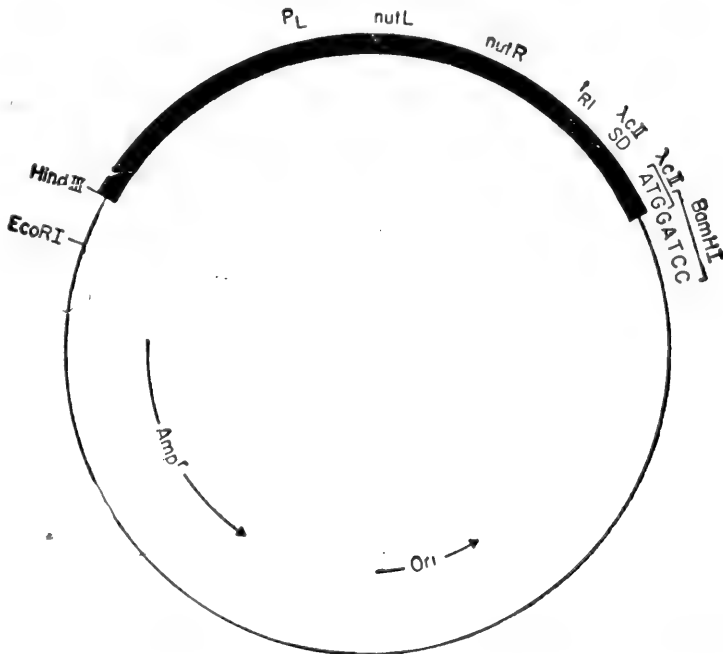
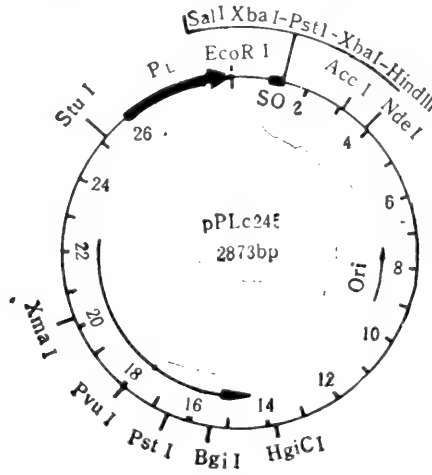


图 9.22 pAS1

pAS1: 约 5.0kb. 它带有 λP_L 启动子、λcII、SD 序列以及 λcII 基因的 ATG。这一质粒来源于 pKC30, 即在 pKC 30 的 Hpa I 位点处插入了 λcII 基因, 再用外切酶

消化直至保留 λ c II 基因的 ATG, 而 ATG 中的 G 则是 Bam H I 位点 (G^1 GATCC) 的第一个核苷酸。使用时用 Bam H I 切开, 用 S I 核酸酶去除突出部分。要表达的基因使改造成缺少第 1 个密码子的平末端。两者平末端连接, 可以表达非融合蛋白 (图 9.22) (Shatzman 等, 1982)。

pPLc245: 2873bp, 它含有 λ P_L 启动子、MS2 噬菌体复制酶基因的 SD 序列和 ATG。其中 ATG 位于 SD 序列下游第 6 个核苷酸处, 而 ATG 中的 G 是 Sal I 位点 (G^1 TCGAC) 的第 1 个核苷酸, 在 Sal I 之后接着还有 Xba I、Pst I、Xba I 和 Hind III 位点。使用时, 先用 Sal I 消化, 再用 S I 核酸酶切去突出部分, 使之成为 ATG 平末端, 外源基因与之平末端连接, 可以表达非融合蛋白 (图 9.23) (Remaut 等, 1983)。



在 ATG 起始密码子附近的序列:

AGGATTACCC ATG TCGACCCTCTAGAGCCTGCAGCGTCTAGAGCCAAGCTT

SD Sal I Xba I Pst I Xba I Hind III

图 9.23 pPLc245

下述真核基因在 P_L 表达载体中能获得有效表达。

1. SV40t 抗原

SV40t 基因的 5' 端没有合适的酶切点, 用 Bal 31 外切酶随机消化, 使 t 基因第 2 个密码子的第一个核苷酸 G (ATG GAT...) 与补齐的 Ava I 位点融合 (...CCCGA), 或 ...CCCGAGAT... 用 Ava I 消化, Mung bean 核酸酶补平, 与 pAS1 平末端连接, 表达非融合的 t 抗原, 产量约为菌体蛋白的 10% 左右。

2. 小鼠金属硫蛋白 (metallothioneine) (MT)

小鼠 MT 基因第二个密码子中有一个 Ava II 酶切位点 (5'...ATG¹GAC CCC...3'), 该基因用 Ava II 切开, DNA 多聚酶补平; pAS1 用 BamH I 消化, Mung bean 核酸酶修平, 两者连接可以表达非融合的小鼠 MT。

3. 人 α D 型干扰素

作者等采用 pKC30 表达载体以及 TGATG 序列, 使人 α D 型干扰素在大肠杆菌中获得高效表达。

在 pKC30 中 P_L 启动子下游有一单一 Hpa I 切点。用 Hpa I 切开 pKC30, 与

人 α D 干扰素 cDNA 的 Hae III 片段连接, 由于 Hpa I 刚好在编码 N 蛋白第 59 个氨基酸的核苷酸后切开, 而干扰素基因的 Hae III 切点却位于编码干扰素信号多肽的第二个氨基酸的第一个核苷酸后, 这样的连接使人 α D 型干扰素基因由原来的正确的第一种读码框架改为第二种读码框架。因此, 在第 8 个信号多肽处就遇到 TGATG, 重新恢复了第一种正确读码框架, 去除了 N 蛋白融合部分 (pBV867) (图 9.24)。

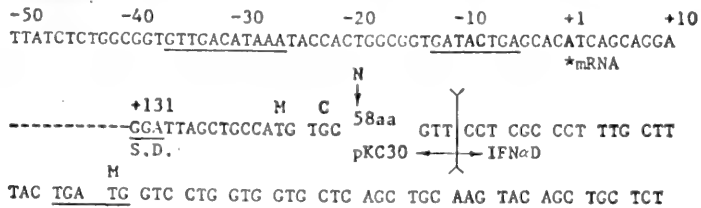


图 9.24 pBV867 部分序列。

采用 P_L 表达载体表达的真核蛋白还有许多, 例如人 β 干扰素(20, 000)、FMDV-VP1 (24,000) 等。

五、 β 内酰胺酶启动子

β 内酰胺酶基因是 pBR322 上的 Amp^r 基因, 全序列已经清楚, 在第 181 和第 182 密码子中有一 Pst 1 切点。 β 内酰胺酶启动子是一个较弱的启动子。利用 pBR 322 或其衍生物如 pAT 153 中 Amp^r 操纵子的 β 内酰胺酶启动子也可作为真核基因的表达载体, 在 Pst 1 位点处插入外源基因可以表达为 181 个氨基酸的 β 内酰胺酶和外源多肽的融合蛋白。这种表达载体有如下特点。

第一, pBR322 中 Pst1 位点常常是真核 cDNA 经 dG-dC 尾连接的克隆位点, 由于加尾的长度不同, 在插入外源基因后的框架与 β 内酰胺酶的框架有 1/3 的机率吻合, 因而表达为融合蛋白, 所以这种载体可以说是克隆和表达的两用载体。表达的水平虽然不高, 但常可通过检测外源基因是否表达来挑选出插入全长度外源基因的重组质粒。作者在克隆人 α D 型干扰素基因时就是采用这个办法。

第二, 可在 Pst 1 位点组成三种框架的表达载体。如 pKT 系列 (287, 288, 289) 即是 pBR322 的衍生物。在 Pst1 切开后, 用 Bal31 消化, 再加上 Pst 1 连接子而组成三种框架的表达载体。

第三, 由于 β 内酰胺酶是一种分泌性蛋白, 大肠杆菌可以识别其分泌信号多肽序列而切割之, 表达的融合蛋白可以被搬运到细菌周质中, 不易被细菌酶类所降解。

采用这一表达载体表达的真核蛋白有大鼠前胰岛素原 (12,000)、人前胰岛素原 (12,000)、大鼠前激素原 (22,000)、大鼠催乳激素原 (22,000)、牛生长激素 (21,000)、猪肠促胰液肽 (3,000)、人 α 干扰素 (20,000)、尿激酶 (35,000)、小鼠二氢脱氧还原酶 (22,000)、SV40t 抗原 (19,000)、HBcAg (22,000) 和 HBsAg (22,600) 等。

六、lpp 启动子

lpp 是大肠杆菌脂蛋白的启动子。由 lpp 组建的表达载体有如下共同特点。

+21 +40 +60 +80 +100
 ----ATCTAGAGGGTATTAAATGAAGACTAACTAACTGGCCGGGTAACTCTGGGTTCTACTCTGCTGGCAGGTTGCTCCAGCAAGGCT----
 MetLysAlaThrLysLeuValLeuGlyAlaValIleLeuGlySerThrLeuLeuAlaGlyCysSerSerAsnAla
 10 20

(A)

EcoRI HindIII BamHI
 A1 ---ATCTAGAGGGTATTGATGAAGGGAAATCCCAAGCTTGGATCCGGCTGAGCAACGACGCTG---
 MetLysGlyAsnSerLysLeuGlySer---

EcoRI HindIII BamHI
 A2 ---ATCTAGAGGGTATTAAATGAAGGGAAATCCCAAGCTTGGATCCGGCTGAGCAACGACGCTG---
 MetLysGlyLysGluPheClnAlaIrrIleArg---

EcoRI HindIII BamHI
 A3 ---ATCTAGAGGGTATTAAATGAAGGGAAATCCCAAGCTTGGATCCGGCTGAGCAACGACGCTG---
 MetLysGlyCylProSerLeuAspPro---

+100 +120
 ---GCTGGCTGCTAGCAAGCT--- (C) ---GCAGGTTGCTCCAGCAACGGCTAAATCGATCGAATCCCAAGCTGGATCCGGCTGAGC---
 AlaGlyCysSerSerSerAsnAla
 20 30

(B)

EcoRI HindIII BamHI
 B1 ---GCTGGAAATCCCAAGCTTGGATCCGGCTGAGC---
 AlaGlyAsnSerLysLeuGlySerGly

EcoRI HindIII BamHI
 B2 ---GCTGGAAAGAAATCCCAAGCTTGGATCCGGCTGAGC---
 AlaGlyLysGluPheClnAlaIrrIleArg

EcoRI HindIII BamHI
 B3 ---GCTGGGGAAATCCCAAGCTTGGATCCGGCTGAGC---
 AlaGlyCylIleProSerLeuAspPro

EcoRI HindIII BamHI
 C1 ---GCAGGTTGCTCCAGCAACGGCTAAATCGAATCCCAAGCTGGATCCGGCTGAGC---
 AlaGlyCysSerSerSerAsnAlaLysIleAspArgAsnSerLysLeuGlySerGly

EcoRI HindIII BamHI
 C2 ---GCAGGTTGCTCCAGCAACGGCTAAATCGAATCCCAAGCTTGGATCCGGCTGAGC---
 / AlaGlyCysSerSerSerAsnAlaLysIleAspArgLysGluPheClnAlaIrrIleArg

EcoRI HindIII BamHI
 C3 ---GCAGGTTGCTCCAGCAACGGCTAAATCGAATCCCAAGCTTGGATCCGGCTGAGC---
 AlaGlyCysSerSerSerAsnAlaLysIleAspArgGlyIleProSerLeuAspPro

图 9.25 pIN-1 系列体系。A、B、C 代表多连接于不同的插入部位。1、2、3 代表 3 种不同框架。

第一,脂蛋白是大肠杆菌最丰富的蛋白,因此, *lpp* 是大肠杆菌最强的启动子之一。

第二,脂蛋白是菌体的外膜蛋白,原始合成物是脂蛋白原,N端有 20 个氨基酸序列的信号肽,信号肽可被细菌识别并切割,脂蛋白可被搬运到细菌周质。所以,如果外源基因连接在信号区之后,外源蛋白可以籍此而被运送到细胞周质中。

Masui 等(1983)组建了一系列不同用途的 *lpp* 表达载体。

1. pIN-1 系列载体

大肠杆菌 *lpp* 基因编码由 78 个氨基酸组成的脂蛋白原,N端有 20 个氨基酸组成的信号多肽。其 mRNA 3' 端的 2/3 具有丰富的稳定的干环结构,3' 端有不依赖 *rho* 的转录终止信号。*lpp* 基因的表达是组成性的,其启动子区段(45bp) A-T 极为丰富,在 ~-160bp 处, A-T 也很丰富。大肠杆菌 *lpp* 基因的这些结构上的特点在其他细菌 *lpp* 基因中是高度保守的,因此,这些特点可能与其有效的转录、翻译和 *lpp* mRNA 的高度稳定有关。

pIN-1 系列载体的组建是为了高水平地、组成性地表达外源基因。根据外源基因插入部位的不同,又可分为 A、B、C 三类,每类又有三种不同框架的插入点,以 1、2、3 表示之。在插入处有三个单一切点: *Eco*R1、*Hind* III 和 *Bam* HI。这类系列载体的部分序列和质粒的物理图详见图 9.25 和图 9.26。

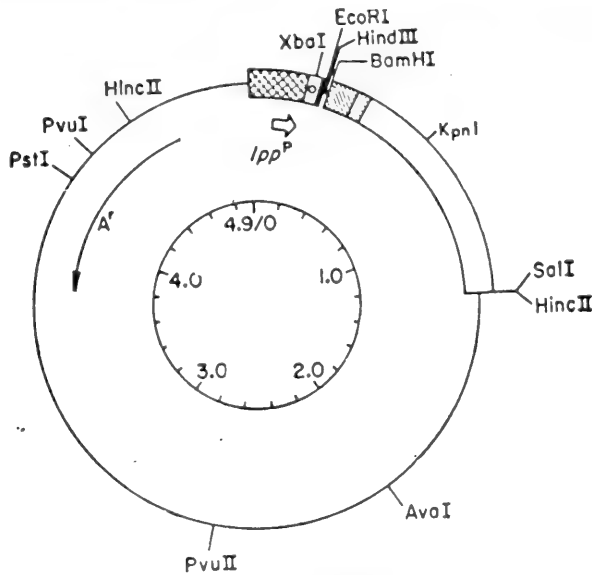


图 9.26 pIN-1 酶切图。

A 位点: 包含有 *Eco*R1、*Hind* III 及 *Bam* HI 的 22 个 bp 插在紧接着 ATG 之后,并组成三种框架,在单一切点之后均有终止密码子。如果插入的外源基因没有终止密码子的话,也不致于读过插入的基因序列。此外,正常脂蛋白的不依赖于 *rho* 的转录终止位点依然保留,以有效地终止外源基因的转录。表达载体导源于 pBR322,保留 *Amp*^r 基因。

B 位点: 单一酶切位点的序列与 A 一样,但是插入在紧接着脂蛋白信号多肽编码区之后。外源基因产物是一种与脂蛋白信号多肽的融合物,由于信号多肽的存在,外源基

因产物可以被搬运到细菌周质，同时信号多肽被切割。使外源蛋白局限在细菌周质的优点是，该处不易被蛋白酶类所降解，并且又易于提纯。

C 位点：所不同的是三个单一酶切位点的序列（22bp）插入成熟脂蛋白的编码区。不论那一种框架，表达的基因产物为信号多肽+脂蛋白第 8 个氨基酸+外源多肽。翻译后信号多肽被切割后，杂交蛋白可以插入菌体外膜。

由图 9.26 可见，在 p1N 系列载体中 lpp 基因的 SD 序列上游，有单一 Xba 切点。如果外源基因带有自身的 SD 序列的话，也可以在此切点插入。外源基因的 N 端最好有 EcoRI、Hind III 或 Bam HI 酶切点，其中 BclI、Bgl II 或 Sau 3 A 的粘末端可与 BamHI 的粘末端相匹配。由于载体中三个单一酶切点的序列是 EcoRI-Hind III-Bam HI。因此，EcoRI-RcoRI、EcoRI-Hind III、EcoRI-Bam HI、Hind III-Hind III、Hind III-Bam HI 或 Bam HI-BamHI 6 种片段均可插入载体。两端有不同酶切粘末端外源基因还可确保插入的方向。如果要插入的外源基因没有适当的酶切位点，则可用 DNA 连接子或平末端连接。如果外源基因的长度过大又没有合适的酶切位点，则可用 Bal 31 随机消化，选择适合长度，修平或补齐，加连接子或平末端连接。平末端连接时，三种单一酶切点均可利用，其中以 Hind III 比较有利，因为 Hind III 切开，补齐或修平后，平末端连接外源基因，用 EcoRI 和 BamHI 切割可以取下插入的外源基因。如果要插入的外源基因的框架尚不清楚时，可以用三种框架的混合质粒进行克隆，其中必有一个的框架是正确的。

必须指出的是，采用这类载体表达的外源多肽，在其 N 端，由于连接子序列的参加，常有少数几个额外的氨基酸。据此，要去掉 N 端的多余氨基酸残基，才能获得与自然产物一样的多肽。其办法是去除 XbaI 和克隆的单一酶切点之间的片段，而代之以一段人工合成的 DNA，后者含有表达所必须的核苷酸序列。上述 XbaI 的位点位于 ATG 上游 14 个 bp 处。

采用 p1N-1 载体表达人生长激素，产量可达菌体蛋白的 10—15%。大白鼠肉瘤病毒癌基因 (V-Ha-ras) 插入 p1N-1 的 Hind III 位点，可以成功地表达 p21。

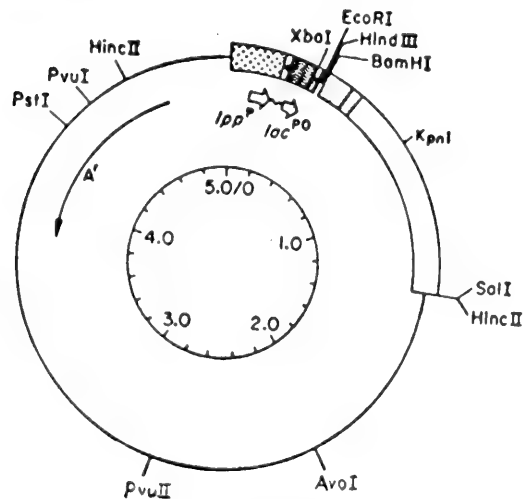


图 9.27 p1N-II 酶切图。

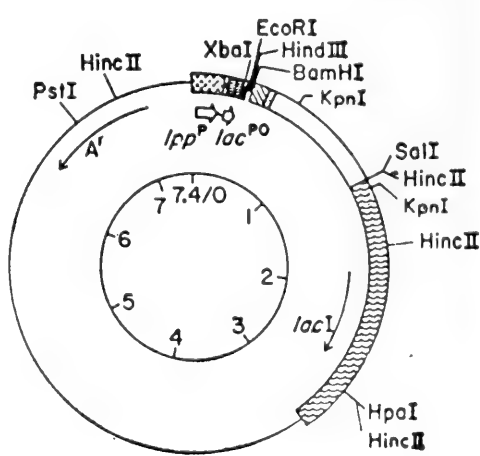


图 9.28 p1N-III 酶切图。

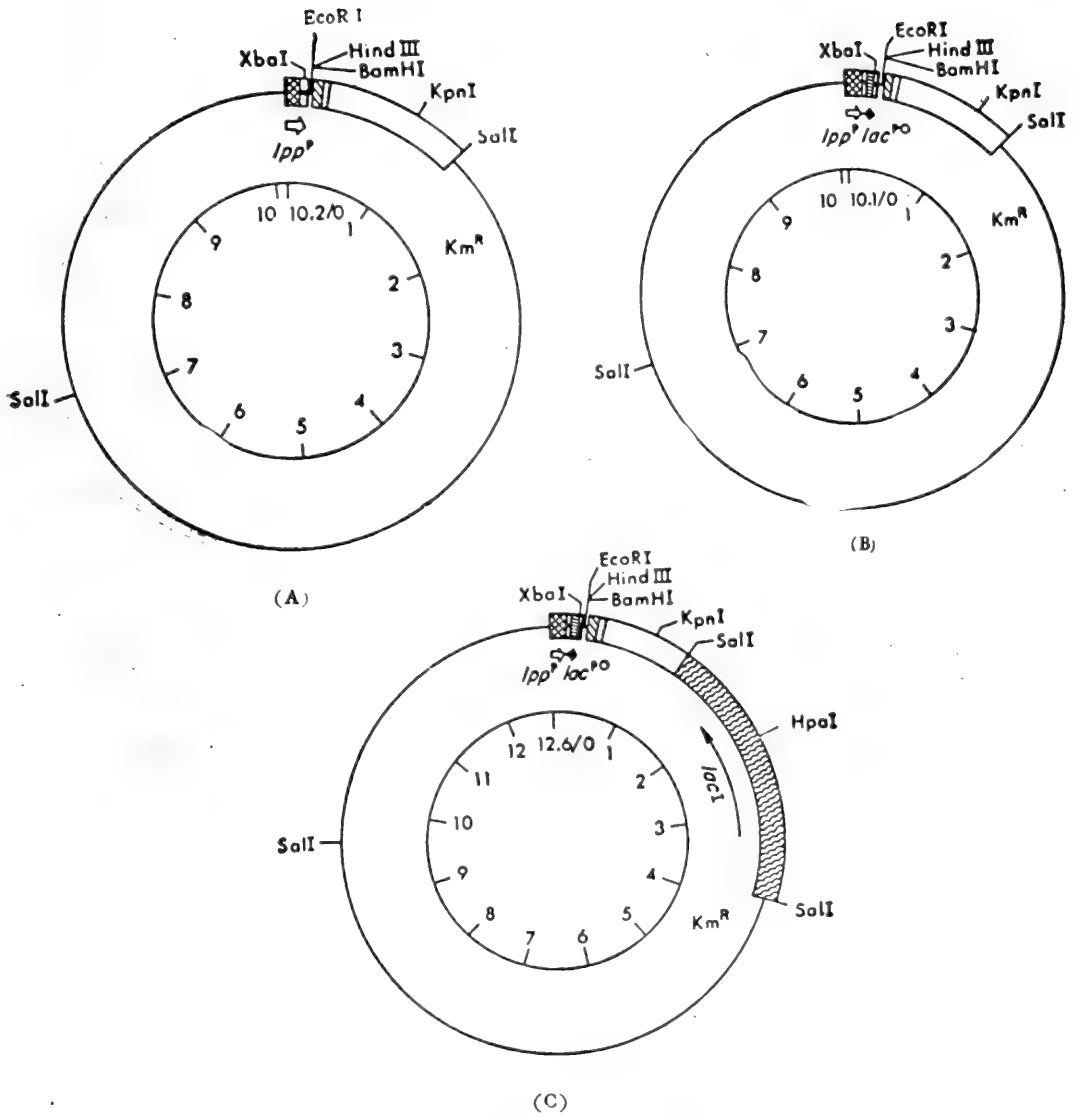


图 9.29 p1M 系列载体。
A. p1M-I; B. p1M-II; C. p1M-III。

2. p1N-II 系列载体

p1N-I 系列载体表达外源基因是组成性的。如果要表达的外源蛋白对细菌本身有害,则不能用 p1N-I 载体。因此, p1NII 的组建解决了上述问题。在此,外源基因的表达只在诱导剂存在下才能进行。

p1N-II 载体导源于 p1N-I,即在 p1N-I 的 XbaI 处插入了一段 95bp 的 AluI DNA 片段,后者含有 lacUV5 P、O 区段。详见图 9.27。插入区段刚刚在脂蛋白 SD 序列的上游。这种载体只有在 IPTG 存在的条件下才能表达,而 p1N-II 的宿主菌必须能过量产生 lac 阻遏物,即 lac I 基因产物。宿主菌可用 JA221 (hsdR, ΔtrpE5, leu B6, lacY, recA, thi),含有 F' (lac I⁺, lacZ⁺, lacY⁺, lacA⁺, proA⁺, proB⁺),因为它带

有 $lac I^o$ 基因,可以过量产生 lac 阻遏物。

如果在 $XbaI$ 处插入 113bp 的 $BstNI$ - $Hae III$ 的 $lac P$ 、 O 片段, 组成 $pIN-II$ (lac^{PO} , 113bp) 载体, 其表达外源基因的水平高于 $pIN-II$ ($lac^{P-O95bp}$) 载体, 原因尚不清楚。

3. $pIN III$ 系列载体

$pIN II$ 的缺点是必须用具有 $lac I^o$ 基因的宿主菌。 $pIN III$ 是在 $pIN II$ 上插入了一段 $lac I$ 基因, 因此, 它可在任何细菌中应用(图 9.28)。

采用 $pIN III$ 载体, 成功地高效表达人 α 干扰素。

4. pIM 系列载体

pIN 系列载体导源于 $pBR 322$, 其复制子来自 $pMB 1$, 在一般情况下, 拷贝数为 ~ 20 拷贝/细胞。为了获得高效表达, 用温度复制失控型质粒替代 $pBR322$ 。用 $pIN-I$ 、 II 、 III 分别组成 $pIM-I$ 、 II 、 III (图 9.29)。

这类载体的特点是, Km^r 对温度敏感, 一般操作在 $30^\circ C$ 进行。在 $30^\circ C$ 时, 质粒拷贝数为 ~ 15 拷贝/细胞, 升温至 $35^\circ C$ 时, 可以增加至 2,000 拷贝/细胞。

5. pIC 载体

当外源基因的序列清楚时, 可用 pIN 和 pIM 系列载体。但是, 当要克隆表达的基因序列不清楚时, 上述质粒就不适用。 pIC 载体可以弥补这一不足。 $pICIII$ 是将 $lacZ$ 基因插入 $pIN IIIA3$ 的 $BamHI$ - $SalI$ 位点处, 然后将 $Hind III$ 位点改为 $Sma I$ 位点。此处插入外源基因, 与 β 半乳糖苷酶基因 $5'$ 端融合, 如果外源基因有一个开放的读码框

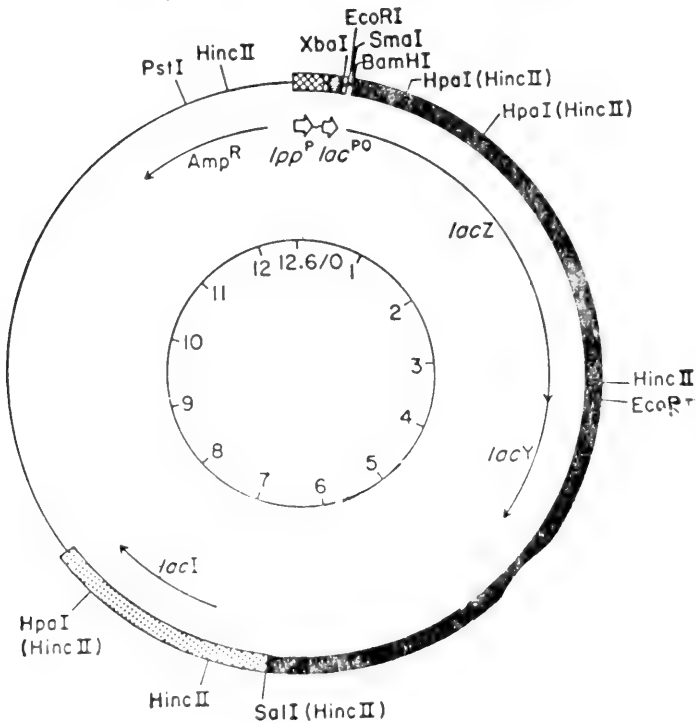


图 9.30 pIC 酶切图。

架与起始密码子相吻合,并且长度为 $3n + 2$ 的话,载体的表型由 lacZ^- 变为 lacZ^+ 。融合蛋白的表达受 lac 阻遏物的调节,大的融合蛋白比较稳定。如果有表达,可以用 Xba I 或 EcoRI 和 BamHI 双酶解,切下外源基因,再将它插入 p1N 或 p1M 系列载体(图 9.30)。

参 考 文 献

- [1] 侯云德,病毒基因工程的原理及方法,人民卫生出版社,1985。
- [2] Hamer, D. H. and Rosenberg, M. J., Gene Expression UCLA Symposia on Molecular and Cellular Biology, New Series, Vol. 8, Alan R. Liss, Inc., New York, 1983.
- [3] Harris, T. J. R., Expression of Eukaryotic Genes in *E. coli*, in Genetic Engineering, 4, ed. by R. Williamson, 1983.
- [4] Brammar, W. J., Vectors Based on Bacteriophage Lambda, in Genetic Engineering, 3, ed. by R. Williamson, 1982.
- [5] Lawn, R. M., The Cloning and Expression of Human Interferon Genes, in Genetic Engineering, Principles and Methods, Ed. by Setlow, J. K. *et al.*, 1982.
- [6] Shatzman, A. *et al.*, Use of Phage λ Regulatory Signals to Obtain Efficient Expression of Genes in *E. coli*, in Experimental Manipulation of Gene Expression, 1—14. Acad. Press, 1983.
- [7] Masui, Y. *et al.* Multipurpose Expression Cloning Vehicles in *E. coli*. in Experimental Manipulation of Gene Expression, 15—32, Acad. Press, 1983.
- [8] 侯云德等,利用 TGATG 序列在大肠杆菌中直接表达非融合的 $\text{L}\alpha\text{D}$ 型干扰素,病毒学报 1,295,1985。
- [9] Oppenheim, D. S. *et al.*, Translational coupling during expression of the tryptophan operon of *E. coli*, *Genetics*, **95**, 785, 1980.
- [10] Remaut, E. *et al.*, Improved plasmid vectors with a thermoinducible expression and temperature regulated runaway replication, *Gene*, **22**, 103, 1983.
- [11] Pestka, S., The human interferons-from protein purification and Sequence to cloning and expression in bacteria: before, between and beyond, *Arch. Biochem. Biophys.*, **221**, 1983.
- [12] Davis, B. D. *et al.*, The mechanism of protein secretion across membranes, *Nature*, **283**, 433, 1980.
- [13] Boer, H. A. *et al.*, Construction of a tandem trp-lac promoter for efficient and controlled expression of the human growth hormone gene in *E. coli*, in Promoter Structure and Function, ed. by Rodriguez, R. L. *et al.*, Praeger Pub. N. Y. 1982.
- [14] Miller, J. H. *et al.*, The Operon, CSH, 1978.
- [15] Goeddel, D. V. *et al.* Expression in *E. coli* of chemically synthesised genes for human insulin, *PNAS*, **76**, 106, 1979.
- [16] Grantham, R. *et al.*, Codon catalog usage is a genome strategy modulated for gene expressivity, *N. A. R.*, **9**: 443, 1981.
- [17] Iserentant, D. *et al.*, Secondary structure of mRNA and efficiency of translation initiation, *Gene*, **9**, 1, 1980.
- [18] Itakura, K. *et al.*, Expression in *E. coli*, of a chemically synthesized gene for the hormone somatostatin, *Science*, **193**, 1056, 1977.
- [19] Talmadge, R. *et al.*, Cellular location affects protein stability in *E. coli*, *PNAS*, **79**, 1830, 1982.
- [20] Yanofsky, C., Attenuation in the control of expression of bacterial operons, *Nature*, **289**, 751, 1981.

第十章 酵母体系中的基因表达

申 同 健

酵母是单细胞真核生物,也是重要的工业微生物。酵母繁殖迅速,可以廉价地大规模培养,而且没有毒性,很适合作为基因工程菌。因此,研究酵母体系中的基因表达,既有一定的学术价值,又有重要的实用意义。

近年来,随着各种酵母质粒的构建和酵母转化技术的建立,酵母体系的基因表达研究有了较快的进展。已经有多种高等生物的基因在酵母中实现了表达,特别是某些在细菌系统中表达不良的真核基因,如乙型肝炎表面抗原基因,在酵母中表达良好,已可用于大规模生产。还在酵母中成功地建立了几种有分泌功能的表达系统,可以把所表达的产物直接分泌出酵母细胞以外,从而大大简化了产物的分离提纯工艺。酵母系统的研究历史虽然还不长,有不少问题尚未完全解决,但迄今取得的成果已经表明,酵母系统有着巨大的发展潜力,经深入研究开发,将有光明的前景。

在各种酵母中,以啤酒酵母 (*Saccharomyces cerevisiae*) 的应用历史最为悠久,研究资料也最丰富。本文介绍的内容,均以啤酒酵母为对象。

第一节 酵母表达系统概述

和其他表达系统一样,酵母表达系统也包括载体、启动子等控制序列和寄主细胞三个主要部分。

一、载 体

酵母载体是可以携带外源基因在酵母细胞内保持和复制,并随同酵母分裂传递到子代细胞的 DNA 或 RNA 单位。载体中最主要的成分自身能在酵母中进行复制的序列,以及某种带有选择性标记功能的序列。此外,为了应用上的方便,还常在酵母载体中加上大肠杆菌质粒的成分。

(一) 载体的复制序列

酵母载体的复制序列主要来自酵母质粒和酵母染色质。依照复制序列的性质,可将酵母载体分为以下四类。

1. YE_p 类(其中 Y 代表酵母, E 代表游离基因)

此类载体的复制序列为酵母 2 μ 质粒成分。2 μ 质粒是多数啤酒酵母中天然存在的质粒,长 6.3kb。由于该质粒中有二段 0.6kb 的反向重复序列引起的重组,2 μ 质粒通常的两种异构体的形式存在,它们都可同样地独立于酵母染色体进行复制。

将 2 μ 质粒中的复制起点区 (replication origin) 及 REP3 基因区的序列引入载体,

就足以维持在大多数酵母株系中的复制能力。但对于某些无内源 2μ 质粒的酵母株,需要更长的 2μ 序列引入载体,才能使它顺利地复制和扩增。

含有 2μ 成分的载体相对比较稳定,在细胞中的拷贝数约为 20—50,某些高拷贝的载体可达 100—200,目前应用最为广泛。

2. YRp 类(其中 R 代表复制)

此类载体的复制序列是非 2μ 来源的自主复制序列 (autonomous replication sequence, 简写 ARS),可来自酵母染色体,也可来自其他生物,甚至人的染色质。此类质粒稳定性稍差,拷贝数亦少,一般在 5—10 之间。

3. YCp 类(其中 C 代表中心粒)

此类载体的复制序列亦属 ARS,并且含有酵母染色体中心粒成分,有的还引入酵母染色体端粒 (telomere) 成分。它们作为一种类似染色体的单位参与细胞有丝分裂和减数分裂,在子代细胞中精确地分配,因此显示出很好的稳定性。不过,它们的拷贝数只有 1 个。

4. YIp 类(其中 I 代表整合)

此类载体含有可与酵母染色体重组的序列,可以完全整合到酵母染色体中。因此,它依靠酵母染色体进行复制,并有很好的稳定性。

酵母转座子 Ty 系列可以整合到酵母染色体上的许多部位。它本身的结构和基因组均与反转病毒相似。由于可能获得多种整合位点,酵母转座子是一种有前途的载体。

除 DNA 外,某些酵母还含有一种天然存在的、可自行复制的双链 RNA,称为 Killer 因子,它们编码一些蛋白质,可杀死敏感的酵母株。这种双链 RNA 亦可能发展成载体。

(二) 选择性标记

YRp 类载体稳定性较差。YE_p 类虽然比较稳定,但并非绝对稳定,当酵母高速增长或多次传代后,即会观察到质粒丢失现象。为了防止发生丢失,最简单的方法是引入适当的选择性标记,同时使酵母在一定的选择压力下传代。

最常使用的选择性标记是某些氨基酸或核苷酸合成酶系基因。这类基因突变产生的营养缺陷型菌株是很易通过诱变筛选得到的。例如:亮氨酸合成酶系中的 *leu2* 基因突变产生的 *leu2⁻* 营养缺陷型菌株,色氨酸合成酶系中的 *trp1* 基因突变产生的 *trp1⁻* 营养缺陷型菌株,尿嘧啶合成酶系中的 *ura3* 基因突变产生的 *ura3⁻* 营养缺陷型菌株,等等。

当在不含亮氨酸的培养基 (*leu⁻* 培养基)上培养 *leu2⁻* 营养缺陷型菌株时,因其 *leu2* 基因失效不能自己合成亮氨酸,故不能生长。如引入带有野生型 *leu2* 基因的载体,使之恢复合成亮氨酸的能力,就可以在 *leu⁻* 培养基上生长,一旦它们失去带有 *leu2* 基因的载体,则又不能生长。这里,载体上的 *leu2* 基因就是选择性标记,它与 *leu⁻* 培养基条件的选择压力共同作用,迫使 *leu2⁻* 酵母菌株不能丢失胞内的载体。

对于 *trp1*, *ura3* 等选择标记,其作用方式也是一样的。

利用对抑制酵母生长的药物具有抵抗力的抗性基因亦可构成选择标记。其性质类似大肠杆菌的抗氨基青霉素、抗四环素等抗性标记。如将氨基糖苷转磷酸酶的基因引入酵母载体,就构成抗 G418 或抗新霉素、卡那霉素等选择标记。又如将二氢叶酸还原酶 cDNA 引入载体,就构成抗氨基蝶呤的选择性标记。

YIp 类载体带有的野生型酶基因序列,如野生型 *ura3* 序列,可以通过重组整合到 *ura3⁻* 菌株染色体上对应的突变 *ura3* 基因位置上,与此同时载体上 *ura3* 附近的其他基因也被带到酵母染色体原 *ura3* 邻近的位置上。对于某些基因的表达,整合载体是最适用的载形式。

(三) 大肠杆菌质粒序列

从大肠杆菌中制备质粒远较从酵母中容易。因此,酵母质粒的加工和制备大部分是通过大肠杆菌进行的,只有在最后阶段才转入酵母。这就要求酵母载体也同时具有在大肠杆菌中复制和增殖的能力。为此,常向酵母载体中引入 pBR322 的 *ori* 部分和 *amp^r* 或 *tet^r* 部分,这样构成的载体,既可在酵母中复制,又可在大肠杆菌中繁殖,因此称为酵母-大肠杆菌穿梭载体或穿梭质粒。

使酵母载体能够在大肠杆菌中繁殖,还有利于从酵母转化子中回收质粒进行鉴定。尤其对于整合型的质粒更为重要。

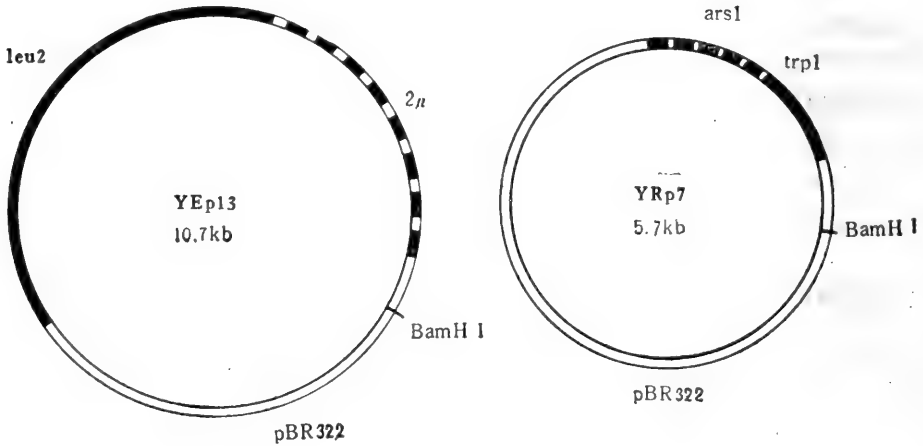


图 10.1 两种酵母-大肠杆菌穿梭质粒。

质粒名称	YEp13	YRp7
酵母复制单位	2μ	arsI
酵母选择性标记	leu2	trp1
大肠杆菌质粒序列	pBR322	pBR322
	amp ^r , tet ^r	amp ^r , tet ^r

二、启动子及其他控制序列

启动子是基因表达的关键构件。外源基因在酵母中的表达,绝大多数需要配以酵母自身的启动子,此外,还包括启动子 5' 上游区必要的调节控制序列,以及一部分编码序列(如分泌型启动子)。基因 3' 端的终止序列亦有一定的重要性。

(一) 启动子序列

在酵母基因表达研究中应用的启动子种类甚多,兹就有代表性的数种列入表 10.1。表

表 10.1 常见的酵母启动子

启动子类型	所属的基因(组)	基因符号
组成型	磷酸甘油酸激酶	PGK
	甘油醛-3-磷酸脱氢酶	GAP
	醇脱氢酶	ADH 1,2
	烯醇酶	ENOL
	磷酸三碳糖异构酶	TRI
诱导型	细胞色素 c	CYC 1
	金属螯合硫蛋白	CUP 1
	酸性磷酸酶系	PHO 5,3
	乳糖酶系	GAL 1,10
分泌型	转化酶	SUC 2
	α 因子	MF α 1
	酸性磷酸酶	PHO 5
复合型		TRI-SUC2
		GAL1-SUC2
		GAL1-PHO5

中的大部分启动子及有关的基因序列均已报道。只要在启动子后面直接插入外源基因，原则上就能获得表达。

天然状态各酵母基因表达效率差别甚大。某些糖代谢基因如 PGK 和 GAP 的表达产物可达总蛋白的 1—5%。当它们具有多个拷贝时，表达可高达总蛋白的 30% 以上。另一些基因如 CYC1 的产物则不超过细胞的 0.1%，因此，常称 PGK 一类启动子为高效启动子。

当在高效启动子后面连接外源基因后，其表达率常常剧烈下降，其中原因尚不完全清楚。有证据表明，外源基因的转录活力较低，生成的 mRNA 稳定性亦差。估计这些高效启动子自己的基因编码中某些序列可能影响其转录效率，其 mRNA 与表达产物的特异结合也可能提高稳定性而增加转译数量。

组成型启动子（包括上述所谓高效启动子）原则上在酵母生长各个时期都能发挥作用，但是，当外源基因产物对酵母细胞有不利影响时，过早表达常使酵母培养物生长不良以致达不到所需的细胞密度，从而影响了总表达产量。通过采用某种代谢控制方法，如更换碳源、改变温度等，可以部分地缓解这类问题。

诱导型启动子的表达受诱导物或诱导条件的特异性影响，可以在较大范围内改变表达效率，如 PHO5 在低磷酸盐浓度时及 CUP1 在加有铜离子时，表达效率可以改变数十倍以至百倍以上。寻找高效的诱导启动子是提高酵母表达效率的重要途径。

分泌型启动子也属于组成型，其中起作用的部分还包括该基因 5' 端一段编码序列，即信号肽部分以及前导肽部分的编码序列。这部分肽的结构特点，帮助了后面的表达产物分泌出酵母细胞，并在适当的部位由胞内蛋白酶加工切掉表达产物与前导肽之间的肽键，产生正确的表达产物分子。将表达产物分泌出细胞以外，既可以避免它们对酵母细胞可能产生的不利影响，也避免了提纯产物时破碎细胞、分离蛋白的困难，有重要的实用意义。但应注意分泌型启动子后面外源基因的连接部位，不是与启动子序列紧接，而是接在前导肽编码的后面。

复合型分泌启动子是由各种高效启动子后面连接各种有效的信号肽序列或前导肽序列的编码组成,后者可以来自酵母,也可以来自其它生物。复合型分泌启动子的表达和分泌效率常较单独分泌型启动子更优越。

(二) 终止序列

对于真核基因的表达,基因编码区后接的终止序列有重要作用。终止序列保证了转录产物(mRNA)在适当部位终止和加上 polyA 尾部,这样形成的 mRNA 可能比较稳定并被有效地转译。在酵母中表达的人工合成基因一般不含有终止序列,必须借用外加的终止序列或载体上现成的终止序列(如 2μ 质粒上的终止序列)。常用的外加终止序列有 ADH1、CYC1、MFA1 和 PGK。要注意对终止序列不适当的删除会导致终止功能的剧烈下降,因此,一般不对终止序列作特殊的剪切。

三、寄主酵母菌株

寄主株系对外源基因表达的质量关系甚大。不论建立新的表达系统,或是对某特定基因进行表达研究,均应谨慎选择所用的酵母菌株。

对于表达用的酵母株通常应考虑以下几点:

(1) 生长力强。使用非分泌型启动子,表达产量直接与酵母菌密度有关,故应选择生长力强的菌株,即使在一定的选择压力下,亦应易于达到 $OD_{600} \geq 10$ 的菌体密度。

(2) 内源蛋白酶较弱。外源基因表达产物易于受到酵母胞内的蛋白酶分解。内源蛋白酶活力低的菌株造成的破坏较小。

(3) 菌株性状稳定。常用的寄主菌株几乎都是突变株,其有用性状的回复突变显然对表达不利,故应避免回复突变率高的不稳定株系。如对所用株系性状稳定情况不了解,应注意定期对菌种作分离纯化以保证菌株纯度。最好使用二倍体或多倍体菌株。

(4) 分泌力强。不同菌株分泌能力亦有差别。如使用分泌型启动子,最好选用分泌力强的株系,文献报道的超分泌突变株(SSC)分泌产量可提高十倍左右。

此外,使用某种载体或启动子时对菌株有特殊要求,如某些选择性标记要求菌株为相应的营养缺陷型,使用 α 因子或 a 因子启动子要求菌株为 α 型或 a 型等等。

第二节 酵母表达系统的构建

一、酵母启动子的提取

如果没有现成的启动子供应用,则需从酵母基因文库中提取,一般首先提取启动子所在的基因,然后从中分离启动子。

若有合适的突变株和适当的选择手段,用遗传互补方法筛选目的基因是最简捷的。此外,若能得到其他生物类似的基因片段,利用物种间同种基因结构上一定的相似性,亦可以前者作为探针从基因文库中筛选到目的基因。

然而,对于任何基因或启动子的筛选都适用的是利用化学合成的寡核苷酸探针的方

法,只要知道目的基因或启动子的一部分序列,就可以使用这个方法。以下介绍该方法的一般步骤。

(一) 酵母基因文库的建立

酵母 DNA 的复杂性较小,以质粒为基础构建基因文库即可达到要求。按常规方法提取酵母总 DNA (可以包括线粒体 DNA 在内,也可不包括),用 *Sau3* A1 部分地酶解,蔗糖梯度超离心或凝胶电泳分离 2—7kb 部分,组入 pBR 322 或 YEpl3 的 *Bam*H1 位点,转化到新制备的大肠杆菌 HB101 或其他合适受体菌的感受态细胞中即得。每微克 DNA 应得到至少 10^4 个菌落,背景则不应超过 10—20%;如背景过高,可试用碱性磷酸酯酶将 pBR322 或 YEpl3 的 *Bam*H1 切口末端磷酸除去,然后再做连接。

(二) 启动子基因探针的合成

近年来,化学合成脱氧寡核苷酸的技术已经成熟。有条件的实验室,可自行合成所需的片段,一般的实验室亦可委托有核酸合成的单位或厂商代为合成。脱保护的产物应予纯化,并有序列检查的数据。

如对目的基因的序列数据掌握较多,则可选择最适合作探针的片段,选择的的原则是:

(1) 最好是基因内的重复序列,这样可以增加杂交强度。

(2) 片段中最好有较高的 GC 比例,及较长的 GC 连续区,以增加杂交强度;但须避免均聚体区域,以选择密码区片段较好。

(3) 如有计算机序列库,则可挑选那些与已知酵母其他基因及基因文库载体、大肠菌 DNA 中相似性最小的序列,以减少假结果的数量。

(4) 如条件允许,则可在目的基因序列中选择二段以上作为探针,对二或三个不同探针均呈阳性的菌落有极大的含有目的基因的可能性。

(三) 菌落杂交

杂交过程请参照本书第五章中菌落杂交部分。以下只着重指出使用寡核苷酸探针的特殊的地方,假设探针链长为 20nt 左右,G + C 占 50%。

为保证筛选成功,一次筛选的菌落数可相当或略高于 2×10^4 。用 8—10 个 $\phi 9$ 氨苄青霉素-YT 平皿即够。硝酸纤维素膜 (0.45 μ m 孔径, $\phi 8.5$) 转移时间不少于 3 分钟,碱变性和中和时间勿短于 5 分钟。烘干后的膜应在清水中浸透,将干燥菌渣尽量漂净后阴干。

有时,菌落中质粒 DNA 经二次扩增后再做杂交效果较好。为此,待硝酸纤维素滤膜吸附菌落后,将菌落朝上放在新的氨苄青霉素-YT 平皿上 37°C 保温 4—8 小时,再转到新的氯霉素-YT 平皿上 37°C 保温 16—24 小时,再将滤膜做变性、烘干处理。

预杂交在 45°C 进行 2—10 小时,杂交在 45°C 进行 16—24 小时,在 50°C 漂洗三次,每次 10 分钟,使用 $5 \times$ SSC;然后在 30°C 做三次,每次 15 分钟,使用 $2 \times$ SSC。

预杂交溶液组成与长链 DNA 探针所用稍有不同,组成如下:

tris-HCl (pH7.5)	10m mol/L
EDTA (pH7.5)	6m mol/L

NaCl	0.8mol/L
ATP	0.1mmol/L
焦磷酸钠	1mmol/L
Denhardt 溶液	1 ×
SDS	1%
变性 <i>E. coli</i> DNA	0.1mg/ml

杂交的滤膜放在 -70°C 对 X 光底片曝光 1—3 天(加增感屏), 由得到的杂交斑点可确定阳性菌落的位置。

(四) 探针标记

1. 5' 端标记

通常化学合成的寡核苷酸 5' 端无磷酸, 用 γ - $[^{32}\text{P}]$ -ATP 和激酶标记 5' 端最为方便。反应条件如下:

寡核苷酸探针 (10—25nt)	0.05—0.1 μg
10 × T4 激酶反应液	2 μl
γ - $[^{32}\text{P}]$ -ATP ($\geq 3000\text{Ci}/\text{mmol}$)	100 μCi
T4 多核苷酸激酶	10—20 u
H ₂ O	补足 20 μl

37 $^{\circ}\text{C}$ 保温 30 分钟, 85 $^{\circ}\text{C}$ 加热 3 分钟。

制得的探针通过一细长的葡聚糖 G-25 凝胶柱纯化后估算其比放射强度。如转化率较高, 探针亦可不经纯化直接使用。

2. 3' 端标记

如 5' 端标记的放射强度不足, 可考虑进行 3' 端标记。用 α - $[^{32}\text{P}]$ dATP 的加尾产物可以提供高达 $5-10 \times 10^9 \text{cpm}/\mu\text{g}$ 的比放射性, 是 5' 端标记的 10—50 倍。

反应条件如下:

寡核苷酸探针	0.01 μg
甲磺酸钾 (1mol/L, pH7.0)	1 μl
CoCl ₂ (10mmol/L)	1 μl
二巯基乙醇 (10mmol/L)	1 μl
α - $[^{32}\text{P}]$ -dATP ($\geq 3000\text{Ci}/\text{mmol}$)	1—200 μCi
末端转移酶	20u
水	补足 10 μl

37 $^{\circ}\text{C}$ 保温 1 小时, 注意防止蒸发。

产物很易用 oligo dT-纤维素纯化。加入 100 μl 0.2% SDS 于反应产物中, 100 $^{\circ}\text{C}$ 加热 3—5 分钟, 用 0.5ml 0.5mol/L NaCl 稀释后, 加到一个 $\phi 0.7 \times 1$ 的 oligo dT-纤维素柱上[预先用 0.5mol/L NaCl-10mmol/L Tris-HCl (pH7.4) 平衡], 用同样溶液洗柱至无放射性流出后, 换 0.2% SDS 洗脱, 柱温可升至 55—60 $^{\circ}\text{C}$, 收集主要的放射性峰, 离心除去可能混杂的纤维素细渣后, 4 $^{\circ}\text{C}$ 保存, 亦可直接使用。

若用葡聚糖 G-25 柱纯化产物, 则须先用酚处理除去蛋白杂质, 以防较长的单链

标记物在柱上严重吸附。

(五) 第二次菌落杂交

第一次菌落杂交选出的阳性菌落不可能判断十分准确,每个菌落与其他菌落混杂的可能性很大。为此,将每个杂交点对应的菌落区切下 $\phi 0.5$ 的一小块琼脂,放在 1ml YT 培养液中振荡均匀。吸取 0.1、1、10 μ l 各稀释到 100ml,分别涂布在新的平板上。选择长有 40—400 个菌落的平板作第二次菌落杂交。

第二次杂交与第一次基本相同,探针也可用第一次杂交后回收者,此时自显影时间应适当延长。第二次杂交将得到一批单独的阳性菌落,将它们转移扩大培养,即可用于提取质粒 DNA。

(六) 目的基因的筛选

由于受寡核苷酸探针长度的影响,杂交的特异性是有限的,因此,很可能得到多个阳性菌落,其中只少数几个含有所要的目的基因。为得到它们还需要进一步筛选。

根据目的基因的结构特点,筛选可以用不同的方式进行,通常首先进行区带杂交和酶解图谱的检查。

1. 区带杂交

根据已知的结构数据,用适当的限制性内切酶水解得到的各个质粒 DNA,在凝胶电泳后做 Southern 吸印,仍用前述探针杂交。在正确链长范围有强杂交带者,有可能是所要的基因,将它们制备一定数量。

2. 酶解图谱

根据已知的结构数据,用适当的限制性内切酶水解可能是所要的基因的区带,观察酶解图谱。如果酶解图谱与已知数据相符,即是所要的基因;如果只有部分相符,则须进一步鉴定。

(七) 序列鉴定

序列测定是最重要的鉴定。即使酶解图谱正确,也需要对关键部位的序列做检查,以了解可能发生的点突变等细节变化。为此,将接近启动子或前导序列部分做亚克隆(subcloning),用化学法或 M13 方法测定其序列。序列鉴定正确方可认为得到了所要求的基因或启动子。

某些情况下,合成的寡核苷酸探针可以当作序列测定的引物,作为 M13 引物的补充大幅度增加可读序列的范围。因此,在设计探针时不妨考虑这一用途。

二、表达载体的构建

表达载体至少应含有酵母启动子及有关的表达控制序列,并有适于外源基因插入的位点。现以几种 α 因子表达载体为例说明表达载体的构建过程。

(一) 普通表达载体的构建

此类载体只要求能方便地引入外源基因并进行表达,对表达产物的组成,尤其其 N 末

-170
AGTG

-160 -150 -140 -130 -120 -110 -100 -90
CAAGAAAACCAAAAAGCCACACAGGTTTTGGATAACTACATATATAAGAGGGCCTTTTGTCCCATCAAAATGTTACTGTT

-80 -70 -60 -50 -40 -30 -20 -10
CTTACGATTCATTTACGATTCAGAAATAGTTCAAACAAGAAGATTACAACACTATCAATTTTCATACACAATATRAACGACCCAAA

1 10 20 PstI 30 40 50 60
AGA ATG AGA TTT CCT TCA ATT TTT ACT GCA GTT TTA TTC GCA GCA TCC TCC GCA TTA GGT GGT
met arg phe pro ser ile phe thr ala val leu phe ala ala ser ser ala leu ala ala
1 10 20

70 80 90 100 110 120
CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC
pro val ASN THR THR thr glu asp glu thr ala gln ile pro ala glu ala val ile gly tyr
30 40

130 140 150 160 170 180
TCA GAT TTA GAA GGG GAT TTC GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG
ser asp leu glu gly asp phe asp val ala val leu pro phe ser ASN SER THR asn asn gly
50 60

190 200 210 220 230 240
TTA TTG TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT TTG GAT
leu leu phe ile ASN THR THR ile ala ser ile ala ala lys glu glu gly val ser leu asp
70 80

250 260 HindIII 270 280 290 300
AAA AGA GAG GCT GAA GCT TGG CAT TGG TTG CAA CTA AAA CCT GGC CAA CCA ATG TAC
lys arg glu ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
90 102

310 320 HindIII 340 350 360
AAG AGA GAA GCC GAA GCT GAA GCT TGG CAT TGG CTG CAA CTA AAG CCT GGC CAA CCA ATG TAC
lys arg glu ala glu ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
111 123

370 380 HindIII 400 410 420 430
AAA AGA GAA GCC GAC GCT GAA GCT TGG CAT TGG CTG CAA CTA AAG CCT GGC CAA CCA ATG TAC
lys arg glu ala asp ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
132 144

440 450 HindIII 460 470 480 490
AAA AGA GAA GCC GAC GCT GAA GCT TGG CAT TGG TTG CAG TTA AAA CCC GGC CAA CCA ATG TAC
lys arg glu ala asp ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
153 165

500 510 520 530 Sali 540 550 560 570
TAA GCCCGACTGATAACACACTGTAGATGTAACAAGTCGACTTTGTTCCCACTGTACTTTTAGCTCGTACAAAATACAAT
stop

580 590 600 610 620 630 640 650 660
ATACTTTTCATTTCTCCGTAACAACCTGTTTTCCCATGTAATATCCTTTTCTATTTTCGTTTCGTTACCAACITTTACACAT

670
ACTTTATATAGCTAT

图 10.2 α 因子基因序列

氨基酸是否增减并不严格要求。

对于 α 因子来说,它是分泌型启动子,外源基因应在其前导肽编码后部插入。恰好天

然 α 因子基因在此部位有一 Hind III 切点,可用于插入外源基因。因此,只需选出上述含有 α 因子基因的亚克隆质粒即可作为这类表达载体。 α 因子基因结构见图 10.2,它的一种表达载体 pYA-1 见图 10.3。

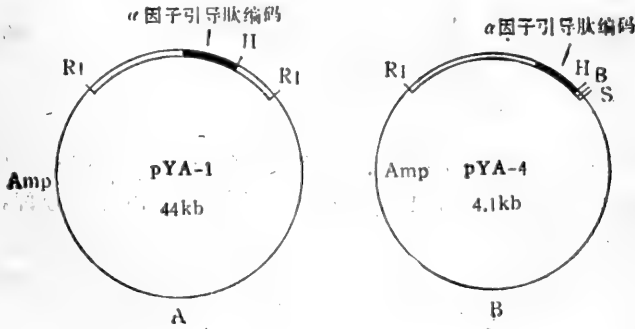
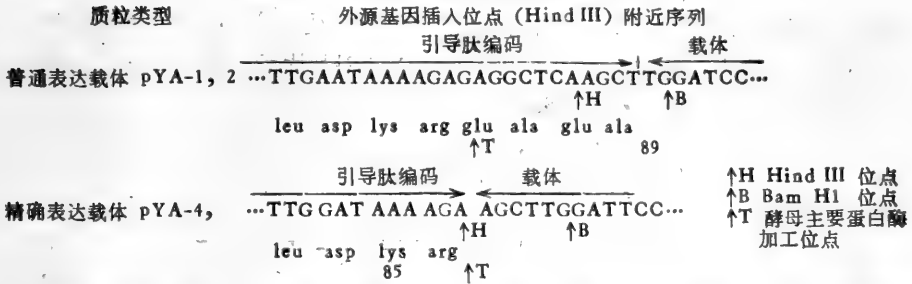


图 10.3 α 因子分泌型表达载体的构建。



(二) 表达具有正确 N 末端氨基酸序列的外源基因表达产物的载体——精确表达载体的构建

此类载体要求在启动子或引导肽序列编码的适当部位有内切酶位点,以利于接入外源基因,并使它在表达和加工后 N 端氨基酸序列与天然产物相同,既无多余的氨基酸,也无缺失的氨基酸。

例如,在构建 α 因子精确表达载体时,应在其 Arg85 密码子后面引入一内切酶位点,因为 α 因子前导肽最主要的断裂位点在 Arg 85,如在该处引入外源基因,表达产物加工断裂后可以给出正确的 N 端氨基酸。

为了在 α 因子 Arg85 密码子后面引入内切酶位点,首先应在该处将 α 因子基因切断,然后装上适当的内切酶位点序列。这里将使用一种技术,可以不依赖限制性内切酶,在任何已知序列的指定部位将基因精确切断。这一技术与后面叙述的表达质粒构建例 2 一样,是由寡核苷酸引导的点突变技术发展而来。

1. 合成与欲切断部位序列相应的寡核苷酸

按照预定剪切处上游一侧的 5' 末端序列合成 10—20 此长度的寡核苷酸。在此例中,合成的是 Arg85-Ser81 编码相应的负链 15 核苷酸,即 dTCTTTTATCCAAAGA (见图 10.2)。

2. 单链 α 因子基因序列的准备

用 α 因子基因中 1.5kb EcoRI-Hind III 片段,以分子克隆方法接入 M13mp10 RF 相应位点中并制备单链噬菌体核酸。

3. 引物杂交

在 $100\mu\text{l}$ 的 0.1mol/L NaCl 、 $0.01\text{mol/L Tris-HCl}$ (pH7.5) 中,加入上述 M13 DNA $30\mu\text{g}$, 合成的 1.5 核苷酸引物 90 pmol , 65°C 保温 20 分钟,然后放入冰浴中冷却。

4. 引物延长

上述溶液中再补加入四种 dNTP 各 10 nmol , 100mmol/L DTT $10\mu\text{l}$, DNA 聚合酶 Klenow 片段 $5-10\text{u}$, 20°C 保温 10 分钟。

5. 单链部分的删除

加入 $10\times\text{S1}$ 反应液(0.5mol/L NaCl , 0.5mol/L NaAc pH4.8, 0.05mol/L ZnCl_2) $15\mu\text{l}$, 0.01mol/L MgCl_2 $15\mu\text{l}$, 补水至 $150\mu\text{l}$, 放置冰浴中至少冷却 5 分钟,然后加入 S1 酶 10u , 0°C 反应 10 分钟,立即加入 1mol/L Tris 中和至 pH7.5,加入等体积酚(0.1 mol/L tris 饱和)抽提,上清液用氯仿处理,乙醇沉淀,沉淀用 75% 和 95% 乙醇各洗一次,溶于 $15\mu\text{l TE}$ 缓冲液。

6. α 因子基因回收

向 TE 缓冲液中加入 $2\mu\text{l}$ 限制性内切酶用 $10\times$ 缓冲液, EcoRI 5u , 37°C 保温 1 小时。产物用低熔点琼脂糖凝胶电泳分离,切取 1.2kb 片段, 65°C 熔化后,用酚抽提 2 次,氯仿抽提一次,乙醇沉淀,产物溶于 $5\mu\text{l TE}$ 缓冲液。

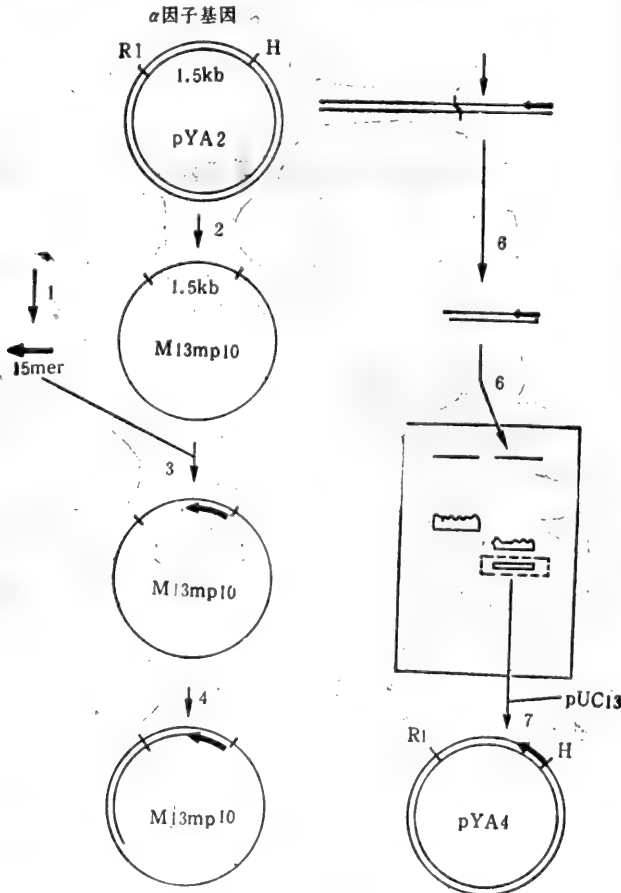


图 10.4 表达载体 pYA42 的构建。

7. 组建成表达载体

15 μ l 反应混合物中,含有 10 \times 限制性内切酶缓冲液 1.5 μ l, pUC13DNA 1 μ g。Hind III 酶 5u, 37 $^{\circ}$ C 保温 1 小时,加入 4 种 dNTP 各 4nmol, DNA 聚合酶 Klenow 片段 2 单位, 20 $^{\circ}$ C 保温 10 分钟, 80 $^{\circ}$ 加热 5 分钟, 再补加 EcoRI 5u, 37 $^{\circ}$ C 保温 1 小时, 加入 7.5mol/L NH₄Ac 9 μ l, 95% 乙醇 55 μ l, -20 $^{\circ}$ C 冷冻 1 小时, 离心收集沉淀, 乙醇洗涤, 干燥 1 分钟, 溶于 10 μ l TE 缓冲液。

取此 TE 溶液及步骤 6 TE 溶液各 3 μ l, 加入 1 μ l 10 \times 连接酶反应液, 2 μ l 水及 1 μ l T4 连接酶(10u), 14 $^{\circ}$ C 过夜。转化 *E. coli* HB 101, 用菌落杂交法选出重组质粒, 即相当图 10.3B 的 pYA4 为精确表达载体, 在 α 因子 Arg 85 密码子后面紧接 Hind III 位点。

上述实验步骤总结为图 10.4。

三、外源基因的引入和表达质粒的构建

将外源基因正确地引入表达载体即成为表达质粒。表达质粒可以先在大肠杆菌质粒中装配, 然后引入酵母质粒, 亦可直接在穿梭质粒中构建, 在穿梭质粒基础上构成的表达质粒应在酵母细胞中直接进行表达。

表达质粒的构建有两种情况, 分别举例说明如下。

(一) 对产物 N 端序列无严格要求

当表达产物 N 端氨基酸可以增、减若干不影响表达目的时(如某些酶的生产), 引入外源基因较为简单, 只需在普通表达载体的启动子后面插入外源基因序列即可。

例如用醇脱氢酶启动子 (ADH1) 表达载体, pAAR 6 表达固氮酶基因 *nifH*。在 pAAR 6 的 ADH1 启动子后面的单一 Hind III 位点处用该酶切断, 用绿豆核酸酶修齐切口, 直接接入修齐的固氮酶基因 H 的 Hinf I 酶解片段, 得到的含有 *nifH* 片段的质粒经选择即得到固氮酶表达质粒 pJB 81 (见图 10.5)。由于平端连接的插入片段有两种取向, 必须对重组产物进行选择以得到取向正确的一种。

(二) 要求生成具有正确 N 端序列的表达产物

此时要求外源基因的起始密码正好位于启动序列的相应部位, 或第一个氨基酸密码子正好与前导肽加工位点的编码相连, 不能出现误差。

如引入的基因是化学合成者, 则在合成时应考虑与精确表达载体的引入部位配合。

在一般情况下, 插入基因时须做一定处理才可能得到符合要求的连接, 现将几种处理方法分别举例说明如下。

例 1. 只对基因加工: 当引入精密表达载体时, 只须加工基因的 5' 端。如向 α 因子表达载体 pYA41 引入人胰岛素原基因(图 10.6), 处理步骤如下:

(1) 合成人胰岛素原基因 phe 1-gln 4 编码的正链 12 核苷酸 dTTTGTC AATCAG 作引物。

(2) 将人胰岛素原基因用 EcoRI 和 BamHI 从质粒 pBCA 中切出, 再经分子克隆转入 M13mp II RF, 制备相应的单链 DNA。

(3) 单链 DNA 与 12 核苷酸引物杂交。

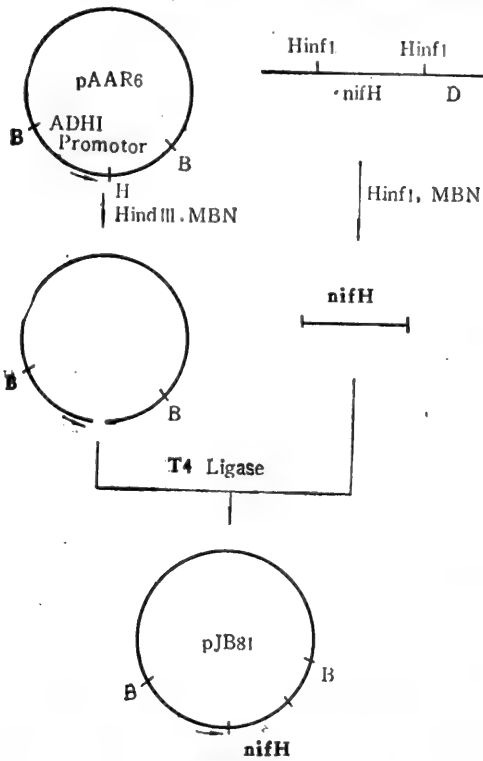


图 10.5 将固氮酶H基因引入表达载体。

(4) 引物延长。

(5) 单链部分删除。

(6) 处理后的人胰岛素原基因回收。

以上 (2)–(6) 各步条件均可参照本章二(二)中的第 2–6 各步。

(7) 将处理后的人胰岛素原基因组入表达质粒 pYA41:

pYA41 用 Hind III, S1 酶在温和条件下水解, 酚-氯仿处理后, 再用 BamHI 处理, 失活后, 乙醇沉淀, 溶于 10 μ l TE 缓冲液。

取上述处理后的 pYA41 (平末端—BamHI 粒性末端线状载体) 0.1 μ g, 与步骤 (6) 得到的 5' 端删除的人胰岛素原基因 (平末端—BamHI 粒性末端片段) 0.2 μ g 混合, 用 T4 连接酶连接。

连接产物转化 *E. coli* JM101, 随机选取 3–4 个重组菌落制备质粒 DNA。

(8) 序列鉴定。重点检查 pYA41 与胰岛素原基因连接处的序列, 应该得到 α 因子

Arg85 密码子与胰岛素 phe1 密码子紧密连接的接点, 约有 50% 以上的质粒是正确的。

例 2. 同时对表达载体和外源基因进行加工: 将启动子或前导肽编码后面和外源基因 5' 端前面的多余序列, 同时删去以得到正确的连接。通常采用的是寡核苷酸互补删除法, 如向 α 因子表达载体 p57 引入人 α 干扰素基因(见图 10.7)。主要步骤如下:

(1) 将人 α 干扰素基因插在 α 因子表达质粒 p57 中: 用 XbaI-DNA 聚合酶 Klenow 大片段与 PstI 先后作用, 从质粒 pLeIFDtrp11 中切出 0.8 kb 的人 α 干扰素基因片段, 用间接方法接在 α 因子表达载体 p57 的引导肽编码后面。此时干扰素基因取向是正确的, 但前导肽 Arg85 密码子与 α 干扰素 cys1 密码子之间有 24 个多余核苷酸。

(2) 将上述连接产物引入 M13mp8, 得到单链 DNA。

(3) 合成所要求的正确接点序列的寡核苷酸引物。正确连接即前导肽 Arg85 与 α 干扰素 cys1 连接, 其前后氨基酸顺序为:

82 83 84 85 1 2 3 4
 ...leu asp lys arg cys asp leu.pro...

相应的编码是:

...TTG GAT AAA AGA TGT GAT CTC CCT...

合成寡核苷酸为它的互补链, 接点两端各 12 个核苷酸共 24 核苷酸:

5'-AGG GAG ATC ACA TCT TTT ATC CAA.

使用前将 5' 端磷酸化。

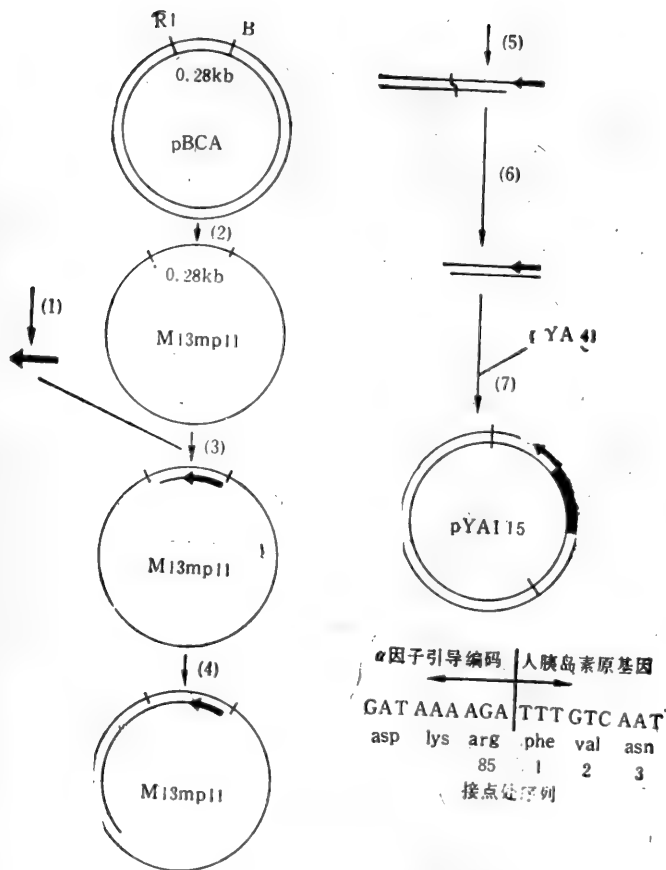


图 10.6 人胰岛素原基因引入表达载体 pYA115。

(4) 单链 M13 DNA 与 24 核苷酸杂交。此时形成一较大的单链环形凸出，相应于应删除的部分。

(5) 引物延伸：(4)、(5)二步可参看本节二(二)。注意加足够的 DNA 聚合酶 Klenow 片段使之合成全部负链，然后加入 T4 连接酶使之首尾封闭。

(6) 蔗糖梯度超离心或 S1 酶作用。使未形成闭环双链的 M13 DNA 水解，以降低背景。如筛选技术熟练，本步亦可省略。

(7) 转化 *E. coli* JM101 及筛选。纯化的闭环 DNA 转化 *E. coli* JM101，用合成的正确接点处 15 核苷酸序列 5'-AGATCACATCTTTTA 作探针，在 30—45°C 不同温度下作噬斑杂交。完全正确的序列应有最高的杂交温度，而不完全删除或未删除的序列因错配而大大降低杂交温度。

(8) 连接产物的序列鉴定。按 Messing 方法直接在 M13 上测定接点序列。从二个最高杂交温度的噬斑中，有一个给出预期的接点序列。

(9) 构建表达质粒。将上述 α 因子- α 干扰素片段从 M13RF 切出，组入适当载体。

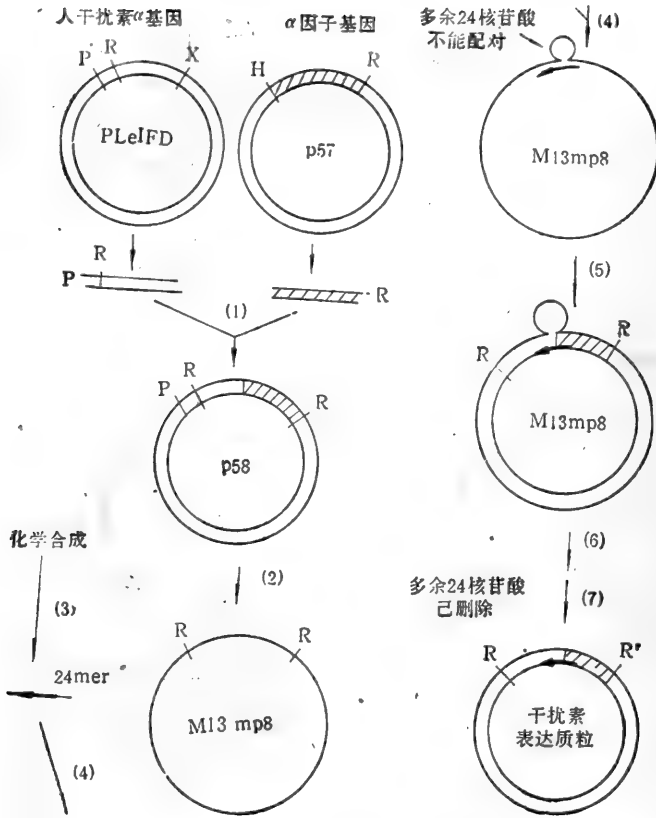


图 10.7 将人干扰素 α 基因引入表达载体。

例 3. 与例 2 类似,亦属同时对表达载体和外源基因进行加工,但在质粒上直接进行:以 α 因子-人胰岛素原基因连接为例(见图 10.8)加以说明。

(1) 将质粒 pBCA 中的人胰岛素原基因插在质粒 pYA-2 的 α 因子前导肽编码后面,中间保留一个 Hind III 位点。注意在 α 因子 Arg85 密码子与人胰岛素原 Phe1 密码子之间,有 24 个多余的核苷酸。

(2) 在 Hind III 处将两个基因断开。

(3) λ 外切酶作用。

λ 外切酶从 DNA 的 5'-P 末端逐个降解,形成含有 3' 单链末端的双链 DNA 中间产物。此 3' 单链区可用特定引物杂交(见下文)并背向 3' 端点做延伸修复,最后残留的未修复部分仅仅是引物 5' 端外侧的 3' 单链区,可以很方便地删除(见图 10.8(5))。

(4) 合成寡核苷酸引物。共需二个引物,各相当接点两侧 5' 端应保留序列。在此例中,合成了相当前导肽 Arg 85—Ser 81 编码负链的 15 核苷酸 dTCTTTTATCCAAAGA 和人胰岛素原 Phe1—Gln4 编码正链的 12 核苷酸 dTTTGTCAATCAG。用前将其 5' 端磷酸化。

(5) 引物与 λ 外切酶作用后的 3' 单链区质粒 DNA 杂交,并用 DNA 多聚酶 Kle-

now 片段延长修复。

(6) 用 S1 酶温和地作用, 将未修复部分删除。

(7) 删除区二侧平端连接, 形成环形质粒。

(8) 转化 *E. coli* JM101, 并筛选。连接产物中有相当大一部分是所要求的产物, 制备 10 个左右的质粒 DNA 用 Hind III 水解实验, 凡不水解者即选出做接点检查。

(9) 接点检查。上述选出的质粒, 按例 1 的原理检查接点。正确的接点应是 α 因子基因 arg85 密码子与人胰岛素原基因 Phe1 密码子的直接连接者, 即上述二个引物 5' 端直接对接者(双链)。如(5)~(7)条件掌握合适, 可望 30~70% 选出的质粒有正确的接点。

本法中当 S1 酶作用过度时, 会引起修复末端缺口处降解, 可用绿豆核酸酶代替 S1 酶。

本法与例 2 比较, 方便和效率方面是其特点, 如在启动序列和接入基因的 5' 端之间有较长的片段需要删除时, 本法更为适用。

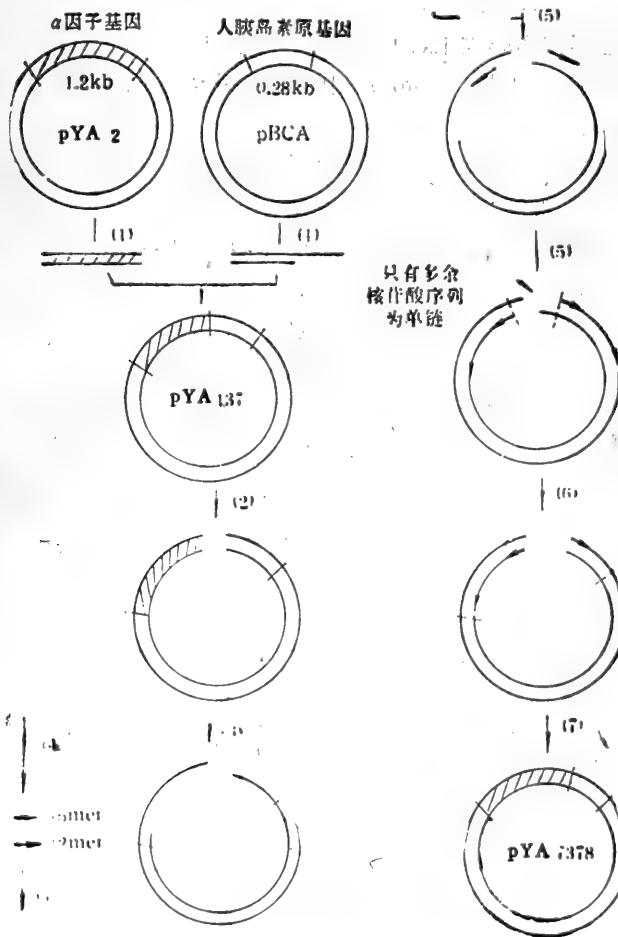


图 10.8 将人胰岛素原基因引入 α 因子表达载体。

第三节 酵母体系中的基因表达

一、酵母菌的转化

与细菌和动物细胞相似,酵母细胞也能够接受外部 DNA 并发生转化。由于酵母有较厚的细胞壁,一般先用特殊的酶类将壁溶解,变成原生质体再进行转化,这样可以有较高的转化效率,对多种质粒都适用,但需要较长的再生细胞壁时间。近年来,还发展了借助各种盐类如 LiAc、LiSCN 进行转化的方法,操作简便,节约时间,但效率较低,且并非每种质粒都可以得到转化子。现将两种方法介绍如下,质粒为 YEp13 系列,标记为 leu2, 寄主菌采用 5573-1B, 系 (α , leu2, trp1, his4) 多缺陷型。

(一) 酶法转化

(1) 寄主菌斜面活化 1 次,接种一环菌苔至 50ml YEPD 培养基, 30℃ 振荡培养至 $OD_{600} = 1-2$, 约有 2×10^7 细胞/ml。

(2) 室温下将培养物离心收集,用 1mol/L 山梨醇 20ml 洗涤 1-2 次。

(3) 重新将酵母细胞悬浮在 1mol/L 山梨醇 5ml 中,加入 5 μ l 2-巯基乙醇, 0.1ml 蜗牛溶壁酶 (glusulase), 30℃ 保温 1 小时。

(4) 室温下低速离心 (2,000 r/min) 5 分钟,收集酵母细胞,用 1mol/L 山梨醇 10ml 洗一次,再用山梨醇-Ca 溶液 10ml 洗一次,悬浮在 1ml 山梨醇-Ca 溶液中,动作应极轻。

(5) 在一个小试管中加入 1-5 μ g 酵母质粒 DNA (总体积勿超过 25 μ l), 加入 0.1ml 上述细胞悬液,室温放置 5 分钟。

(6) 加入 1ml PEG-Ca 溶液,室温放置 10 分钟。

(7) 低速离心收集细胞,悬浮在 1ml 山梨醇-Ca 中,转入 15ml 试管。

(8) 向试管中加入 8ml 45℃ 的再生琼脂,摇动后,倒在加有底层琼脂的平板上。

(9) 30℃ 保温 4-5 天。

(10) 选择深层生长的菌落转接至同样培养基的斜面上, 37℃ 保温 2 天。

(11) 用存活的转化子鉴定质粒,并作保存。

(二) 整体细胞转化

(1) 寄主菌活化 1 次,接种一环菌苔至 100ml YEPD 培养基中。 30℃ 振荡培养至 $OD_{600} = 0.8-1$ (菌浓度为 $0.5-1 \times 10^7$ 细胞/ml)。

(2) 室温下将酵母培养物离心,菌体用水洗涤 1 次。

(3) 将酵母悬浮在 0.5ml LiAc 溶液中, 30℃ 保温 1 小时,轻轻振荡。

(4) 在一小试管中将酵母质粒 DNA 5-10 μ g 与小牛胸腺 DNA (超声处理,但不能用 RNA 代替) 40-60 μ g 混合,乙醇沉淀,干燥后立即加入 20 μ l 水润湿。

(5) 加入上述 LiAc 处理的酵母悬液 0.1ml, 30℃ 保温 30 分钟,轻轻振摇。

(6) 加入 0.7ml PEG-Li 溶液, 30℃ 保温 30 分钟轻轻振摇。

(7) 42℃ 处理 5 分钟。

(8) 短暂离心。沉淀的酵母细胞用 TE 缓冲液洗涤 2—3 次。

(9) 将细胞悬浮于 0.1ml TE 缓冲液,铺在底层琼脂平板上, 30℃ 培养 2 天。

以上转化方法均应做对照。

对照 1. 不加质粒 DNA (其他一切相同)。应无任何菌落出现,如果在培养期间出现较多的菌落,又非污染的话,说明培养基成分有问题,或酵母株有混杂,或酵母已出现了回复突变,应重新分离菌株。

对照 2. 酶法转化,还应在酶处理后直接把 1/20 酵母原生质体溶液放在 YEPE 培养基上观察。如菌数 $> 10^3$, 说明酶处理不够,应改换酶制剂或处理较长时间。

(三) 转化实验用培养基及溶液

1. YEPE

蛋白胨	10g
酵母提取物	5g
加水至	450ml
20% 葡萄糖	50ml

分别灭菌,然后混合。

2. 山梨醇-Ca 溶液

山梨醇 (2.5 mol/L)	40ml
Tris-HCl (pH7.5, 0.1mol/L)	
-CaCl ₂ (0.1mol/L)	10ml
H ₂ O	50ml

分别灭菌,然后混合。

3. PEG-Ca 溶液

Tris-HCl (pH7.5, 0.1mol/L)	
-CaCl ₂ (0.1mol/L)	1ml
PEG 4000 (45% 水溶液)	9ml

分别灭菌,用前临时混合。

4. 再生琼脂

琼脂 (7.5%)	33ml
山梨醇 (2.5mol/L)	33ml
10 × YNB-氨基酸-硫酸铵	10ml
葡萄糖 (50%)	3.3ml
YEPE 液体	1.7ml

分别灭菌,然后混合。(避免多次灭菌)

5. 底层琼脂

琼脂	20g
水	800ml
10 × YNB-氨基酸-硫酸铵	2.5g
水	100ml

葡萄糖 20% 100ml
 分别灭菌,然后混合。

6. $10 \times$ YNB-氨基酸-硫酸铵

YNB-氨基酸	2.5g
硫酸铵	5g
水	100ml

灭菌后冰箱保存。

7. YNB-氨基酸混合物(用于 leu 标记质粒)

YNB (酵母含氮碱)	50g
腺苷硫酸盐	0.6g
尿嘧啶	0.6g
L-trp, L-his, L-met, L-arg	各 0.6g
L-tyr, L-ile, L-lys	各 0.9g
L-phe	1.5g
L-glu, L-asp, L-val, L-thr L-ser	各 3.0g

混合碾细备用。

8. LiAc 溶液

LiAc	0.1mol/L
Tris-HCl(pH7.5)	0.01mol/L
EDTA(pH7.5)	1mmol/L

灭菌备用。

9. PEG-Li 溶液

PEG-4,000	40%(w/v)
LiAc	0.1mol/L
Tris-HCl (pH7.5)	0.01mmol/L
EDTA (pH7.5)	1mmol/L

灭菌备用。

二、酵母质粒的检测

未转化的寄主菌在选择培养基上不能生长,而转化后可以生长,这一事实说明转化酵母菌可能含有互补于选择标记的质粒。如果该质粒上携有外源基因,则可直接测定表达产物。在某些场合,如表达产物甚少,产物不能分泌,测定反应不甚灵敏时,逐一测定表达产物较为费时费力,则以先确定转化酵母细胞确实含有所要求的质粒较为合适。

(一) 酵母质粒的微量法提取

(1) 在选择培养基中培养转化的酵母菌,一般需 $\phi 9$ 平板的 1/4 表面。

(2) 从培养基上刮下酵母菌苔,悬浮在 1ml 1mol/L 山梨醇-0.01mol/L EDTA(pH8.0) 中,离心 2 分钟。

(3) 将沉淀的细胞再悬浮于 0.9mol/L 山梨醇-50mmol/L 磷酸钾缓冲液 (pH7.5)-1.4mmol/L 2-巯乙醇中,加入新配制的 0.1% Zymolase 60,000 水溶液 25 μ l,37 $^{\circ}$ C 保温 30 分钟,离心 1 分钟。

(4) 沉淀悬浮于 0.15mol/L NaCl-0.01mol/L EDTA (pH8.0) 中,加入 20% SDS 25 μ l,立即混匀,加入 1% RNaseA 5 μ l¹⁾,室温放置 20 分钟。

(5) 加入 1 μ l DEPC 或 50 μ l 10 mmol/L PMSF,60 $^{\circ}$ C 保温 10 分钟,置冰浴中冷却。

(5) 加入 50 μ l 5mol/L KAc,0 $^{\circ}$ C 放置 30 分钟,高速离心去掉残渣。

(7) 上清液加入 1ml 无水乙醇,10 $^{\circ}$ C 放置 10 分钟,离心收集沉淀。

(8) 沉淀干燥,并溶于 50 μ l TE 缓冲液,在 -20 $^{\circ}$ C 保存。

(二) 质粒 DNA 的鉴定

以上提取的质粒 DNA 样品相当不纯,一般直接电泳得到的 DNA 区带比较模糊。为得到较明确的结果,可采取以下方法:

1. Southern 吸印及杂交

将上述提取物取 1/5 进行电泳,然后用常规方法作 Southern 吸印,以 pBR322 序列或插入的外源基因序列作探针,进行杂交鉴定。此法可避免酵母内源杂质干扰,结果常很清晰。

2. 酵母质粒的回收

因为所用的质粒多属穿梭质粒,故可在 *E. coli* 中回收。方法是上述提取物 1/5 与 *E. coli* HB101 感受态细胞 500 μ l (制备方法见有关章节),0 $^{\circ}$ C 保温 1 小时,42 $^{\circ}$ C 处理 2 分钟,加 2ml LB 培养基,37 $^{\circ}$ C 振荡 1 小时,铺氨苄青霉素-LB 平板,37 $^{\circ}$ 过夜,应有大量转化子。对照用未转化的酵母 DNA (YEPD 平板上生长)同时操作,应得不到转化子。

回收的转化子可用微量法提取 DNA,作酶解鉴定,应得到预期的电泳区带。

(三) 酵母菌的培养和表达产物的测定

酵母菌培养比较容易。在完全培养基上酵母生长迅速。30 $^{\circ}$ C 1—2 天即得到最大密度。用选择培养基则酵母生长较缓慢,常不能获得最大密度,但质粒的拷贝数有时却较高,按菌体计算的表达产率也较高。

酵母表达产物的测定包括转录产物、转译产物和分泌产物测定。

1. 转录产物测定

酵母 mRNA 一般比较稳定,测定转录产物不十分困难。当需要了解转录调节情况,或蛋白产物测定困难时,转录产物测定就更有必要了。转录产物测定步骤如下。

(1) 酵母总 RNA 提取。转化的酵母菌在 1L 选择培养基上长至 OD₆₀₀ = 1。离心收集菌体,用冷水洗 1 次。-70 $^{\circ}$ C 冰冻酵母细胞 1 小时,加入 0.15mol/L NaCl、50mmol/L Tris-HCl (pH7.4)、5mmol/L EDTA、5% SDS 10ml,混合均匀,加入酚-氯仿-异戊醇 (1:1:0.02, V/V) 10ml,振荡至少 10 分钟,室温放置 10 分钟。

在下相中加入 40 mesh 玻璃珠,强力振荡 3 分钟。低速离心,吸出水层,放入硅烷化

1) 当需要 rRNA 区带时,不必加 RNaseA。

试管中。加入 1/10V 3mol/L NaAc(pH7.0)、2V95% 乙醇, -70℃ 过夜, 高速离心 30 分钟收集沉淀, 乙醇洗涤 2 次, 干燥。

(2) 总 mRNA 提取。用 oligo dT- 纤维素方法(见本书上册 mRNA 提取一章)。

(3) mRNA 电泳分离。甲基汞-琼脂糖凝胶电泳可得到最好的分离效果。但甲基汞毒性太大, 可考虑使用甲醛-琼脂糖凝胶电泳。后者对多数较短的 mRNA 分离效果亦较好(电泳方法见本书上册 RNA 电泳一章)。

(4) Northern 吸印及杂交鉴定。RNA 吸印法与 DNA 基本相同, 只是不必经过碱变性, 且最好不经溴乙锭染色。

用外源基因作为探针, 测定表达产物的区带有很好的特异性, 几乎没有背景干扰。在多数情况下, 杂交区带都足够清晰。区带显影的密度可以很方便地用光密度计测定。如给予适当的内标准, 可以做到半定量, 足以反映外源基因转录的情况。

2. 外源基因转译产物测定

转译产物测定一般指胞内该蛋白质的定性, 定量测定。酵母细胞壁较厚, 首先必须将细胞破碎才能进行胞内成分测定

(1) 酵母细胞的破碎。使用 French 压力室可以破碎任何年龄的细胞, 使用酶消化的方法则需用年轻的细胞。

酵母培养至 $OD_{600} = 0.5$, 离心收集菌体, 悬浮在 1/10V 的 0.9 mol/L 山梨醇和 25 mmol/L Tris-HCl (pH7.5)-1.4mmol/L 2-巯乙醇溶液中, 加入 Zymolase 60,000 至 50 $\mu\text{g}/\text{ml}$, 25℃ 保温 1 小时, 离心, 菌体用 0.1% Triton X-100-25mmol/L Tris-HCl (pH 7.5)-1mmol/L PMSF 处理 10 分钟, 高速离心澄清, 上清液即是酵母细胞提取物。

(2) 表达产物抗原活力测定。对于细胞粗提物之类的样品, 最好用免疫方法测定表达产物如采用放射免疫或酶标免疫方法。由于免疫方法灵敏度高, 特异性强, 即使很粗的样品也不必预先纯化或高倍浓缩(具体方法请参照本册第 5 章)。

在条件许可时, 用 ^{35}S 标记表达产物而用非标记的免疫血清亦可以获得明确的结果, 省去了自制标记抗体的麻烦。最简单的标记方法是: 转化酵母并培养到 $OD_{600} = 0.5$ 以后, 离心收集酵母细胞, 将其悬浮在酵母含氮碱-混合氨基酸培养液中(成分参照本节一(一), 但除去 L-甲硫氨酸), 补加葡萄糖至 2%, 加入 ^{35}S -L-甲硫氨酸 10—40 μCi , 30℃ 振荡培养 1 小时, 离心收集菌体, 按上述方法破碎细胞, 得到抽提物, 按常规作免疫沉淀。注意沉淀必须充分洗涤, 并作空白对照。本法的主要缺点是定量困难。

(3) 表达产物的电泳测定。胞内成分十分复杂。直接做 SDS-PAGE 常得到太多的区带。上样过多则区带分辨力下降, 上样过少则产物不易检出, 只有产物的含量较高(总蛋白 2% 以上)时, 才较易于找到产物区带。用 ^{35}S -L-甲硫氨酸(或 ^{35}S -L-半胱氨酸)标记后, 灵敏度大大增加, 但亦限于检出那些分子量与胞内其他蛋白相差较多的产物。若产物分子量恰与胞内某大量蛋白相近, 则 ^{35}S 标记背景常常过强以致难于检出产物区带。

较好的方法是把凝胶中的蛋白区带用电泳转印到硝酸纤维素薄膜上, 再用放射性或酶标记的抗体处理, 或依次用非标记抗体和放射性或酶标记的第二抗体处理, 即所谓 Western 吸印——免疫测定方法(可参考本册第五章)。此法不仅可给出清晰的产物区带, 而且还可判断产物分子量是否正确, 以及表达产物是否遭受了不正确的加工或降解。

(4) 表达产物的生物活力测定。如产物有较易测试的生物活力(如酶活力、免疫活力

等等),应尽可能予以测定,可利用粗提液,或经过简单纯化后测定。具有预期的生物活力是表达产物最实际的判断指标。

(5) 表达产物的结构分析。免疫活力和生物活力并不能说明全部结构细节。在必要情况下,应对表达产物的结构作进一步分析。

首先必须将产物高度纯化,在 HPLC 的两种分析柱上(如反相柱和离子交换柱,或反相柱和疏水分析柱)显示单一的峰,或达到其他类似分辨力的纯度指标。

然后,视要求或检测一段氨基酸序列(Edman 法,或利用蛋白质序列分析仪)或检测蛋白酶解图谱或溴化氰肽谱。

如产物中有不止一对二硫键,应考虑二硫键错配的可能。

如产物中有糖化、酰化等修饰,应考虑酵母的加工系统与高等生物细胞有所不同。

3. 分泌产物的测定

分泌产物可用培养液直接测定,为减少背景,常采用无蛋白培养基(参照本节一(三)),分泌产物测定与胞内产物各项测定基本相同。

同时测定胞内和培养液中表达产物抗原活力或生物活性,即可计算出分泌效率。

4. 周质部分的表达产物测定

有些表达系统分泌产物大部分在周质,该处产率可通过测定总细胞含量和原生质体含量的差额算出。测定原生质体含量所用的原生质体,可按本节 2(1) 制备,至 Triton X-100 处理前一步。

参 考 文 献

- [1] Maniatis, T. *et al.*, Molecular Cloning, 1982.
- [2] Struhl, K., The new yeast genetics, *Nature*, **305**, 391, 1983.
- [3] Harashima, S., Yeast Host-Vector System and Its Use for the Study of Yeast Plasmid-Stability and Gene Cloning, *Genetic Engineering and Biotechnology*, **441**, 1983.
- [4] Hinen, A. *et al.*, Transformation of yeast, *Proc. Natl. Acad. Sci. USA*, **75**, 1929, 1978.
- [5] Kurjan, J. and Herskowitz, I., Structure of a yeast pheromone gene (MF α), *Cell*, **30**, 933, 1982.
- [6] Singh, A. *et al.*, synthesis, secretion and processing of α -factor-Interferon fusion proteins in yeast, *Nucl. Acids. Res.*, **12**, 8927, 1984.
- [7] Shen, T. J. *et al.*, Recent Methods for Sequencing Synthetic Oligonucleotides and Expression, *Proc. of 1984 Symp. on Gen. Engineer. and Biotech.*, **18**, 1984.
- [8] 申同健等,酵母 α 因子表达系统的构建,生物化学杂志,待出版。
- [9] 申同健等,人胰岛素原基因在酵母中的分泌表达,同上。
- [10] Clarke, L. and Carbon, J., Isolation of a yeast centromere and construction of functional small circular chromosomes, *Nature*, **275**, 504, 1980.
- [11] Beggs, J. D., Transformation of yeast by replicating hybrid plasmid, *Nature*, **275**, 104, 1978.
- [12] Coff, C. G., Expression of calf prothymosin in *saccharomyces cerevisiae*, *Gene*, **27**, 35, 1984.
- [13] Miyanchara, A. *et al.*, Expression of hepatitis B surface antigen gene in yeast, *Proc. Natl. Acad. Sci. USA*, **80**, 1, 1983.
- [14] Berman, J. *et al.*, Expression of a nitrogen-fixation gene encoding a nitrogenase subunit in yeast, *Gene*, **35**, 1, 1985.
- [15] Ammerer, G., Expression of genes in yeast using the ADC I promoter, *Methods in Enzymology*, **101**, 192, 1983.
- [16] Mellor, J. *et al.*, Factors affecting heterologous gene expression in *saccharomyces cerevisiae*, *Gene*, **33**, 215, 1985.
- [17] Smith, R. A. *et al.*, Heterologous protein secretion from yeast, *Science*, **229**, 1219, 1985.
- [18] Bitter, G. A. and Egan, K. M., Expression of heterologous genes in *saccharomyces cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. *Gene*.

32, 263, 1984.

- [19] Henderson, R. C. A. *et al.*, The Transformation of bresing yeast with a plasmid containing the gene for copper Resistance, *Currens Genetics*, 9, 133, 1985.
- [20] Collins, M. L. and Hunsalcer, W. R., Improved hybridization assays employing tailed oligonucleotide probes, *Anal. Biochem.*, 151, 211, 1985.

第十一章 哺乳动物细胞受体系统中的基因表达

任 贵 方

使用重组 DNA 技术表达外源基因所用的受体细胞有两大系统：其一是原核细胞，诸如大肠杆菌和枯草杆菌等(见第二章)，其二是真核细胞，包括酵母(见第三章)及哺乳动物细胞。本章叙述哺乳动物细胞表达外源基因的原理和技术。

哺乳动物细胞自 50 年代以来已广泛用于病毒学的研究，诸如病毒分离、培养、鉴定、诊断以及多方面的病毒学基础研究，在病毒学的发展史上起了巨大的作用。在实用方面，特别是在制备病毒疫苗方面产生了巨大的经济和社会效益，脊髓灰质炎活疫苗、麻疹活疫苗、乙型脑炎灭活疫苗以及狂犬病、腮腺炎等灭活疫苗的问世和推广，控制了这些传染病的流行，保护了人民的健康，这些成就的取得都和动物细胞的应用密切不可分。随着重组 DNA 技术的发展，对一些在动物细胞中不能分离、难以培养或由于其他原因难以用传统的方法制备疫苗的病毒，诸如乙型肝炎病毒、EB 病毒、甲型肝炎病毒、小儿及成人腹泻病毒等等，有可能利用重组 DNA 技术达到制备疫苗的目的。近几年乙型肝炎病毒基因工程细胞疫苗的发展最为显著，故本章将以乙肝病毒重组 DNA 在哺乳动物细胞中的表达为例进行叙述。近年来利用动物细胞研究基因工程产品除乙肝病毒外，已有人绒毛膜促性腺激素(HCG)、黄体酮化激素(LH)、滤泡刺激激素(FSH)、人生长激素(HGH)、第八因子(Factor VIII)、组织纤维蛋白溶酶原激活物(TPA)等均获得不同程度的表达。由此可见，动物细胞和重组 DNA 技术相结合，具有更广泛的造福于人类的前景。

使用动物细胞作受体表达外源基因具有下列诸种优势：由于外源基因的表达产物可自重组 DNA 转化的细胞分泌到细胞培养液中，细胞培养液成分完全由人控制，从而使产物的纯化变得容易。动物细胞分泌的基因产物是糖基化的，接近或类似于天然产物。细胞在一次接种后可长期培养(15—120 天)，基因产物源源不断地自细胞分泌出来，从而可定期连续收获，大大降低了劳务和成本。这些优势使动物细胞用于外源基因的表达的研究越来越得到人们的重视。

目前用于表达外源基因的细胞均为传代细胞，一般认为传代细胞均是恶性化的细胞，因而对使用这类细胞生产重组 DNA 的产品是否有致癌的问题，自然而然地产生了疑虑。在此似有必要追述一下哺乳动物细胞用于生产生物制品的简史及这类基因产物质量控制的要点，以解除人们的疑虑，为广泛应用动物细胞于重组 DNA 产品的研究开辟途径。

1954 年世界公认只有来源于正常动物的组织细胞才可用于疫苗的生产，即使如此，在已有的活疫苗生产中，例如脊髓灰质炎病毒活疫苗因使用原代恒河猴肾细胞常常带有 SV40 病毒的污染，麻疹活疫苗因使用鸡胚细胞也常常混有鸡白血病病毒的污染。60 年代另一类可以有限传代而染色体正常的二倍体细胞问世了，二倍体细胞能否用于疫苗生产，事实上是经过了数年的争论才被认可。绿猴传代细胞(如 CV-1 或 Vero 细胞系)在 70 年代后期开始用于脊髓灰质炎及狂犬灭活疫苗的生产，其安全性控制标准是每十代做细胞检查。方法是在裸鼠或去胸腺小鼠或注射抗胸腺血清的小鼠，接种 10^7 细胞后观察三

周不致瘤。由于病毒在细胞内繁殖,细胞的致瘤基因有可能和病毒整合,这种病毒若用于制备疫苗,特别是活疫苗是应十分慎重的。因此,如果是使用病毒载体制备基因工程活疫苗,上述质量控制的标准还是应严格遵守。80年代以来,随着单克隆抗体技术及重组DNA产物的相继出现,更广泛地使用传代细胞制备产品又成了人们争论的问题,争论的焦点是致癌。尽管目前用于生产重组DNA产品的细胞均是传代细胞,也可能在动物体内均是致癌的,传递肿瘤的物质也公认是核酸,然而动物细胞所提供的基因产物均是蛋白质,问题是提纯后的产品中是否混有以及混有多少可能有潜在致瘤性的细胞DNA。以乙肝基因工程动物细胞疫苗而论,美国基因技术公司和综合基因公司分别用中华地鼠卵巢和小鼠C-127细胞制备的乙肝疫苗,在其纯制品中含细胞DNA量均小于1pg,前者已经美国食品药品管理署(FDA)同意进行临床试验,多数肿瘤专家认为疫苗制品中含细胞DNA量小于10pg,产品对人是安全的。世界卫生组织1985年12月已拟定草案(Petriciani, 1986),草案规定动物细胞生产的乙肝疫苗含细胞DNA量每剂量中应小于10pg(1986年放宽至100pg),至于动物细胞本身在动物试验中是否致瘤的问题,似无必要考虑。

第一节 哺乳动物细胞表达外源基因概述

使用多种技术,如钙沉淀技术、显微注射、脂质体等技术可将任何一种外源DNA导入动物细胞。DNA进入细胞后可能有多种归宿,一是和细胞DNA发生整合,这一过程称为转化,转化的机率约有 10^{-5} — 10^{-6} (Graham等, 1973)。转化的机制尚不十分清楚,可能与细胞的生理状态有关。另一种情况是外源DNA借助与其相连接的特殊病毒DNA,如牛乳头瘤病毒(bovine papilloma virus),可以不与细胞DNA整合或大部分不整合而存在于细胞染色体外,借助于BPV DNA在特定的细胞内可以复制而不形成完整病毒的特性,连接于其DNA前后的外源DNA也可复制并表达(Wang等, 1983)。再一种情况是外源DNA导入细胞后不发生前述的情况,最终被细胞清除。以转化细胞而论,在转化细胞中外源基因随即成为细胞DNA组成的一部分,在特定条件下,诸如外源DNA的结构,包括外源DNA所编码的基因是否完整,启动子、扩增基因、增强子的有无和强弱以及结构的合理性等等,受体细胞的营养缺陷性状、选择标记及与之相匹配的选择培养基的使用,正确的转化和细胞克隆技术,最终还需视外源DNA和细胞染色体DNA整合的状态以及整合于细胞DNA的外源基因拷贝数等等。如果这些条件均能满足而且合理,外源DNA即可随细胞繁殖代谢过程而获得表达。外源基因所编码的多肽蛋白质,随着细胞繁殖代谢而合成,但又不为细胞所需,从而可以源源不断地自转化细胞向外分泌,不需裂解细胞释放产物。

细胞基因工程产品的要求主要在两方面,其一是高效表达,否则制备有效的特异性产品不经济;其二,产品必须高度纯化以保证使用安全。

第二节 哺乳动物细胞表达外源基因的必需条件

如上概述,外源基因在哺乳动物细胞中获得表达,需要一些必不可少的或特定的条

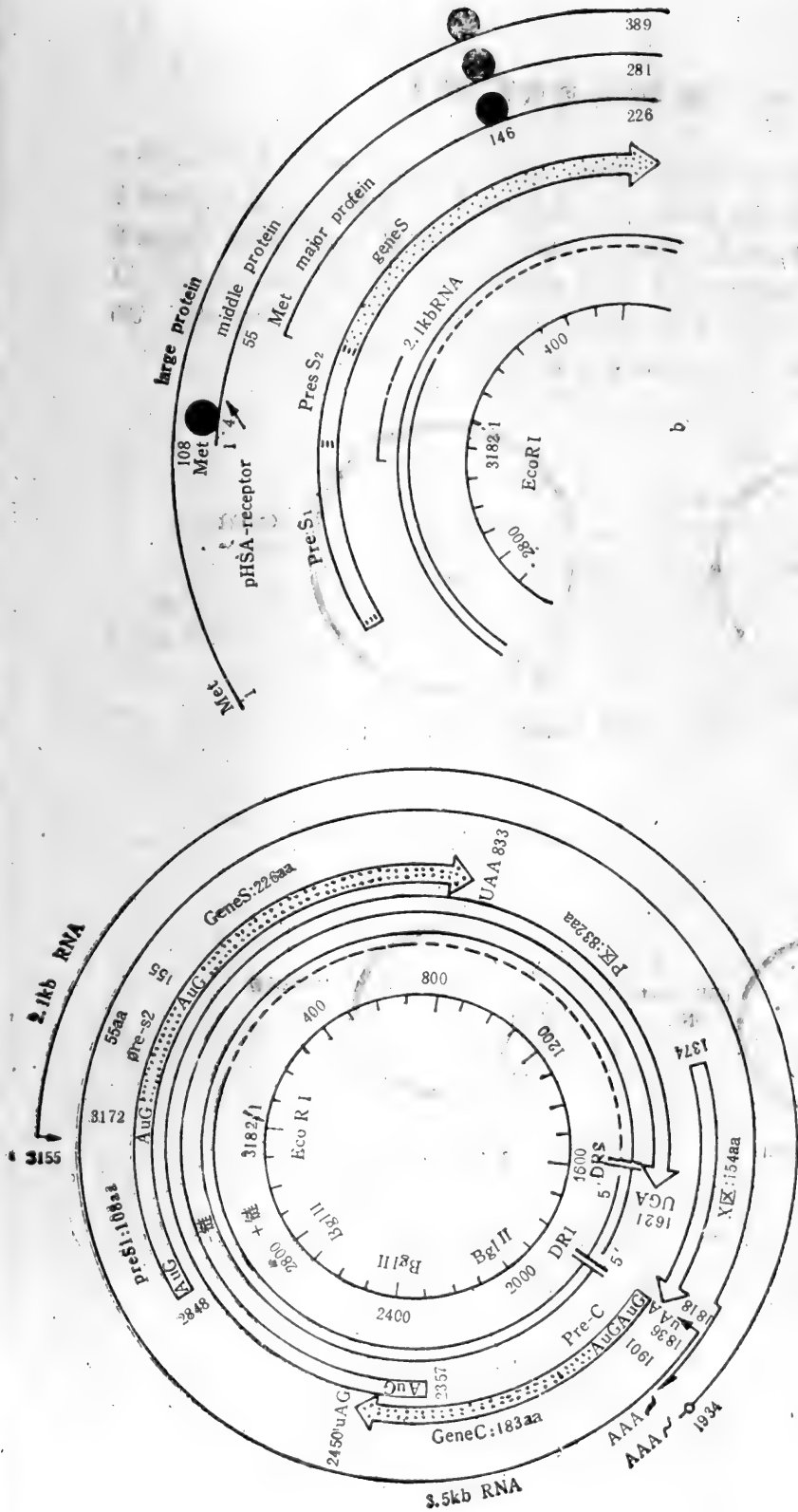


图 11.1 乙型肝炎基因组的结构与功能示意图。

a. 粗黑圆示乙型肝炎 DNA 的阴性链 (-Strand), 其内的粗黑圆示阳性链 (+Strand), 与其相连的点线示 DNA 缺失部分。阴性链密码各种结构蛋白: 前 S1 基因 (Pre-S1) 密码 108 个氨基酸 (a.a.), 前 S2 基因 (Pre-S2) 密码 55aa, 表面抗原基因 (Gene S) 密码 226aa, 均是构成病毒外膜并具有免疫原性的抗原。核心抗原基因 (Gene C) 密码 183aa, X 区基因 (Region X) 密码 154aa, P 区基因 (Region P) 密码 832aa, DNA 阴性链转录的 2.1 kb RNA 是外膜主蛋白的信息 RNA, 3.5kb 的 RNA 可能编码 DNA 多聚酶。

b. 大蛋白 (Large protein) 由 Pre-S1、Pre-S2 和 Gene-S 密码, 含 42,000 的糖化多肽 (GP42) 和 39,000 的非糖化多肽 (P39)。中蛋白 (Middle Protein) 由 Pre-S2 和 Gene S 密码, 含 33,000 和 36,000 两个糖化多肽 (GP33, GP36)。其中 Pre-S2 密码的糖化多肽具有人聚血清白蛋白受体 (PHSA-receptor) 功能。主蛋白 (Major Protein) 由 Gene-S 密码, 含 27,000 糖化多肽 (GP27) 及 24,000 非糖化多肽 (P24), 是构成表面抗原 (HBsAg) 的主要成分 (原图见 Tiotlais, P, et al., *Nature*, 371, 489, 1985)。

件,现分述于下。

一、外源 DNA 或供体 DNA

导入细胞的外源 DNA 称之为供体 DNA,它可以通过提取、人工合成或从 mRNA 反转录获得。经过限制性内切酶处理插入某一细菌质粒载体转化受体菌,获得克隆或称无性繁殖系。使用这种克隆之前要求对其结构与功能尽可能了解,诸如全基因组的碱基排列顺序。限制性内切酶酶切图谱以及这一全基因组的各个特异性基因编码区段等等。以乙型肝炎病毒 DNA 为例,HBV-DNA 的结构与功能已相当清楚 (Tiollais, 1985) (图 11.1)。现将三个亚型 ayw、adw 及 adr 的初级克隆图示如下(图 11.2)。

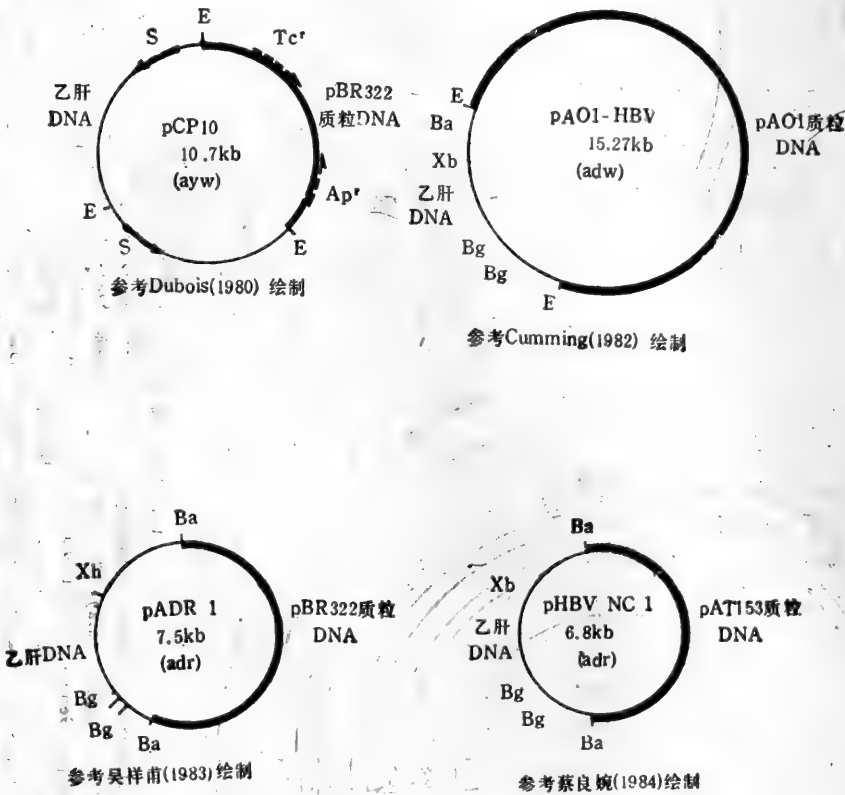


图 11.2 三个亚型乙肝 DNA 初级克隆结构图。

E = EcoR-I。 Ba = BamHI。 Bg = BgIII。 Xb = XbaI。 Xh = XhoI。 S = 乙型肝炎表面抗原基因。
Ap^r = 氨苄青霉素抗性基因。 Tc^r = 四环素抗性基因。

已有的结果表明,以上各个克隆系直接用于转化哺乳动物细胞表达均不理想。有的认为不表达 (Dubois 等, 1980), 有的虽有表达但量很小 (阮力等, 1983), 使用首尾相连的双拷贝 HBV-DNA 的克隆质粒如 pCP10 或 pTHBV-1 表达量也达不到高效 (Dubois 等, 1980; Christman 等, 1982)。为此, 欲获高效表达的 HBV-DNA 克隆质粒必须进

行进一步的改造,后面将述及。为了与改造后的 HBV-DNA 克隆质粒相区别,上述克隆系均称之为初级克隆。

二、载 体

用于转移外源 DNA 的运载体称为载体。用于转化动物细胞的载体大体上分为两类,下面以乙型肝炎病毒 DNA 的克隆为例加以说明。

1. 初级克隆用载体

常用的有 pBR322、pBR325 以及 pAT153 等。

2. 表达用载体

Berg(1981) 和 Mulligan (1980, 1981) 等人构建了一系列的病毒 DNA 表达用载体。这类载体含有大肠杆菌质粒 pBR 322 的 2.3 kb DNA, 包括其复制起点和 Ap^r 基因。SV40 病毒 315 bp 的片段, 包括 SV40 的复制起点、早期启动子及 72bp 的重复序列, 后者具有增强子的作用。SV40 的 629 bp 片段, 含剪接位点以及 SV40 早期转录物的终止和多聚 A 化的 966 bp 片段(图 11.3A)。作者在 SV 40 复制起点片段下游的 Hind 粘末端插入大肠杆菌的黄嘌呤鸟嘌呤磷酸核糖转移酶 (XGPR) 的基因 (Eco-gpt), 或氨基糖苷磷酸转移酶 (AGT) 的基因 (Neo)。Subramani 等 (1981) 将小鼠的二氢叶酸还原酶 (DHFR) 的基因 (dhfr) 插入同一位点, 构造成功带有不同选择标记的表达用载体。利用这些载体将外源基因插入其适当部位即可达到高效表达的目的(图 11.3 B 示三种表达用载体)。其原理及实例详述于下节。

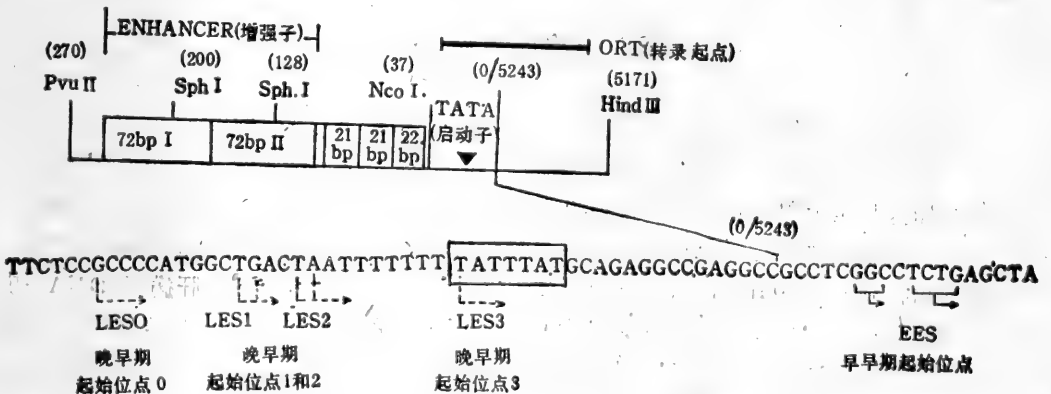


图 11.3A 图示 SV40 Pvu II-Hind III 片段的早期启动子, 转录起点和增强子(参考 Baty(1984) 文献绘制)。

另一种载体是用人或鼠的摄金蛋白 (metallothionein) 基因插入大肠杆菌质粒中, 例如 pMT 质粒是将小鼠的摄金蛋白基因插入 pMT 质粒, 构成内含摄金蛋白的结构基因及其强启动子。将一个外源基因插入摄金蛋白启动子后, 可望获得高效表达。在 pMT 质粒 MT 的下游插入牛乳头瘤病毒 (BPV) 的全长或 69% 的 DNA 片段, 利用 BPV 可以在细胞核内复制, 但又基本上不与细胞 DNA 整合, 也不装配成完整病毒, 可扩增至百余个拷贝的特性, 带动其旁侧的区段 MT 及外源基因同时扩增。利用上述原理 Hsiung

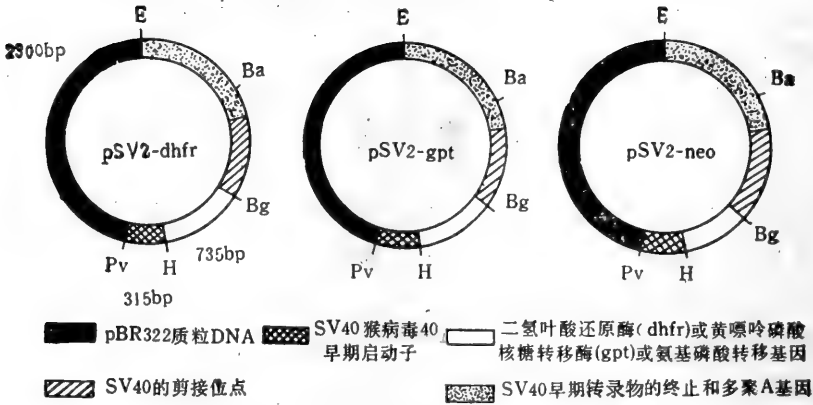


图 11.3 B 三种表达用载体结构图。

(1984)等人将 HBV-DNA 的 1.35 kb 的 BamHI 片段, 含完整的乙肝表面抗原基因插入鼠 MT 基因启动子后的 Bg III 切口, 转化小鼠 C-127 细胞, 获得 1—2mg/L 的 HBsAg 高产细胞系。谢彦博(1986)等人利用人的摄金蛋白(HuMT-II)启动子, 经多层次的剪切连接, 最终构造成 pdMTsAg-5 重组质粒, 转化 C-127 细胞, 也获得 HBsAg 的高产细胞系。并利用 MT 基因可被重金属离子诱导而提高表达量的原理, 对表达阳性细胞加入 $4\mu\text{mol/L}$ 的 ZnCl_2 和 CdCl_2 使乙肝表面抗原的产量达到 0.7mg/L 的水平, 较初表达量提高 1.81 倍。

三、重组质粒的改建

供体基因的初级克隆质粒往往由于缺乏启动子, 或虽有启动子而不够强, 或者由于读码框架不合理等均会造成表达不高或不能表达, 因而需要进行必要的改建。为了达到高效表达的目的, 改建所用的载体 DNA 的选择也十分重要。最常用而行之有效的载体 DNA, 以乙型肝炎表面抗原基因为例, 有 pSV2-dhfr、pMT-BPV、以及罗氏肉瘤病毒的长末端序列构造的 pAV-2(Standring, 1984)。改建所用供体 DNA 是取自乙肝病毒 DNA 初级克隆质粒的某一片段, 将此片段用其粘末端或平末端插入载体质粒中, 经转化大肠杆菌获得改建后的重组质粒, 再进行酶切电泳鉴定。当证实其转录方向与设计相同时, 此新改建的重组质粒即可用于转化动物细胞。下面以 HBV-DNA 为例介绍几种使用 pSV2-dhfr 载体的改建方案及方法。

1. 粘末端连接法改建质粒

pSV2-dhfr (见图 11.4A) 为利用 SV40 启动子后的 Hind III 酶切位点, dhfr 后的 Bgl II 酶切位点, 或 SV40 多聚 A 片段前部的 BamHI 酶切位点, 均可考虑为插入供体 DNA 的部位。供体 DNA 来自质粒 pTHBV-1, 用 Bgl II 内切酶不完全水解, 用琼脂糖凝胶电泳分离 2.7kb 片段。pTHBV-1 是 ayw 亚型的以 EcoRI 内切酶为起点的首尾相连的双拷贝 HBV-DNA (见图 11.4B)。每一个拷贝上有 3 个 Bgl II 酶切点, 完全水解时可以产生 A、B、C 3 个片段(图 11.4C)。在不完全水解时获得的 2.7 kb 的片段中

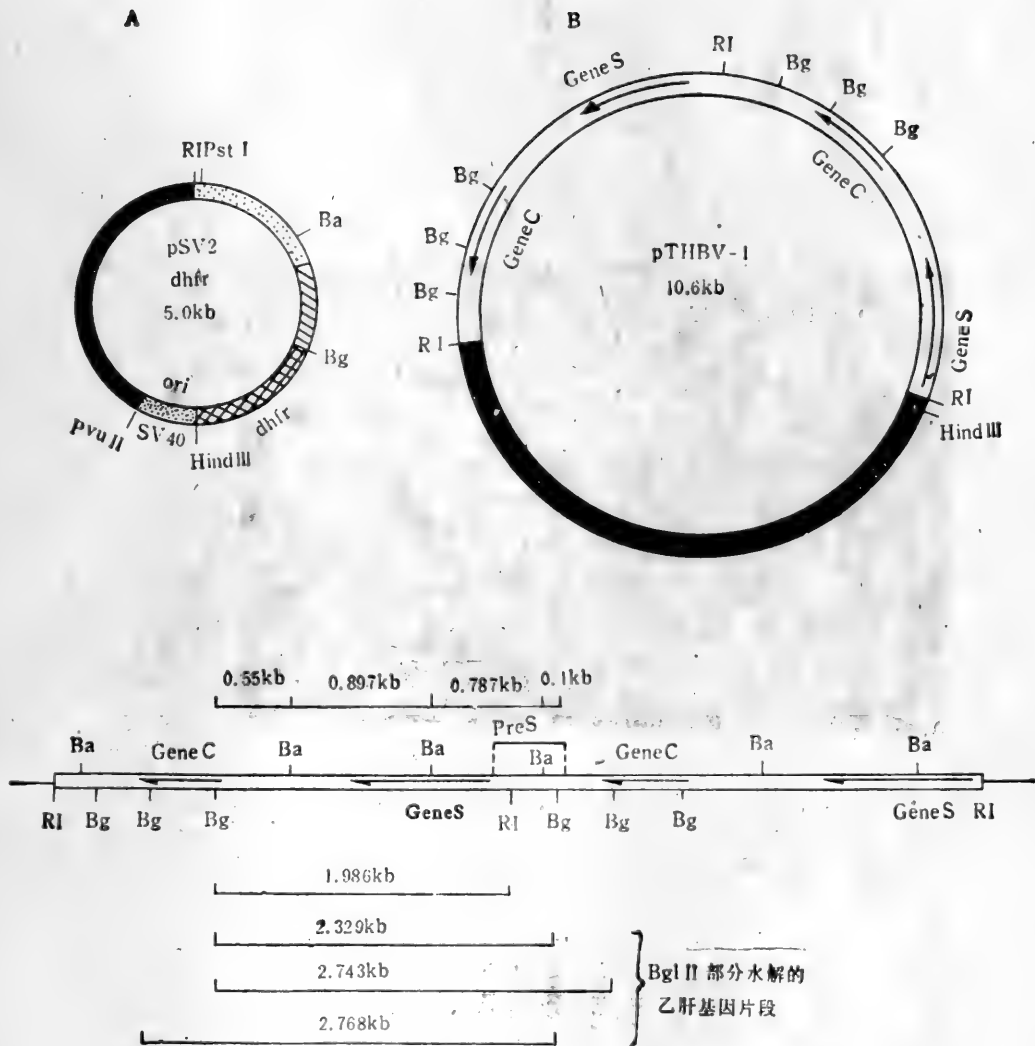


图 1.4 pSV2-dhfr 和 pTHBV-1 的结构示意图。

▨ SV40 (复制起点), 0.32kb。 ▩ 二氢叶酸还原酶基因, 0.73kb。 ▧ SV40 的剪接位点, 0.63kb。 ▤ SV40 的多聚A位点, 1.0kb。 ▣ pBR322 (包括复制起点), 2.3kb, Gene C: 乙肝核心抗原基因。 Gene S: 乙肝表面抗原基因。 RI: EcoRI。 Bg: Bgl II。 Ba: Bam HI。 参考Subramani(1981), Berg (1981), Christman (1982) 绘制。

实际包含有两种大小不同而又相近的片段：一种是 2.743kb 片段，它是 Bgl II C、A 两片段的未完全水解产物，含有完整的前 S (包括 Pre-S1 及 Pre-S2) 及完整的 S 基因；另一种是 2.768kb 片段，它是 Bgl II A、B 两片段的未完全水解产物，含有不完整的前 S (Pre-S1 不完整，Pre-S2 完整)，完整的 S 基因及不完整的 C 基因。将回收的所有 2.7kb 片段利用同一内切酶的粘末端插入载体 DNA pSV2-dhfr 的 Bgl II 切口处，方法是将 pSV2-dhfr 用 Bgl II 完全水解，再用碱性磷酸酶处理防止自身环化。载体 DNA 与供体 DNA 的用量比例为 1:3，二者混合加入 T4 DNA 连接酶，10℃ 反应 20 小时，转化大肠杆菌 HB101，在含有氨苄青霉素 (Amp) 的琼脂平皿上任意挑出 100 个转化菌落，用快速提

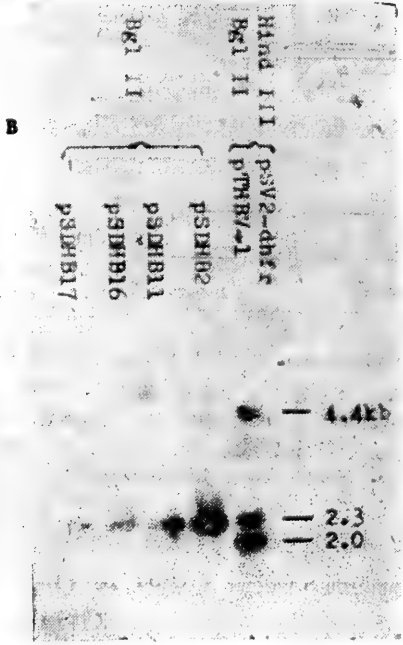
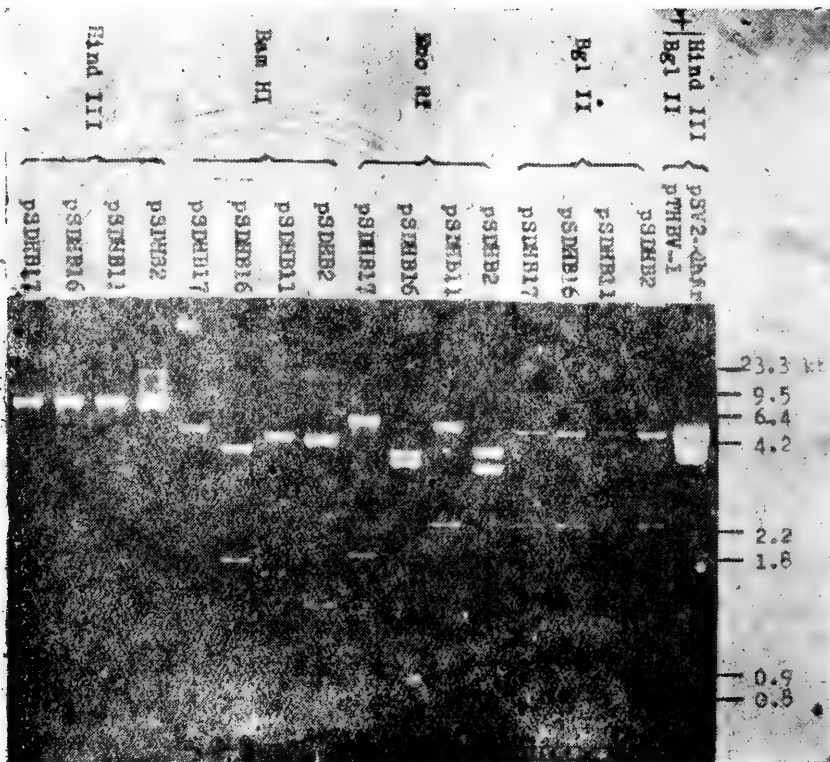


图 11 5 A. 四种乙型肝炎 DNA 重组质粒的酶切鉴定。 B. 同上重组质粒的核酸杂交鉴定。

取法提取所有转化菌的质粒 DNA，然后与载体质粒 DNA 在电泳上作比较，选出比载体质粒分子量明显大的重组质粒。用 Bgl II 完全水解重组质粒，证明含有 2.3 kb 的 HBV-DNA 片段，说明重组成功。再用 EcoRI 水解重组质粒，筛选出含有 2.743kb 和

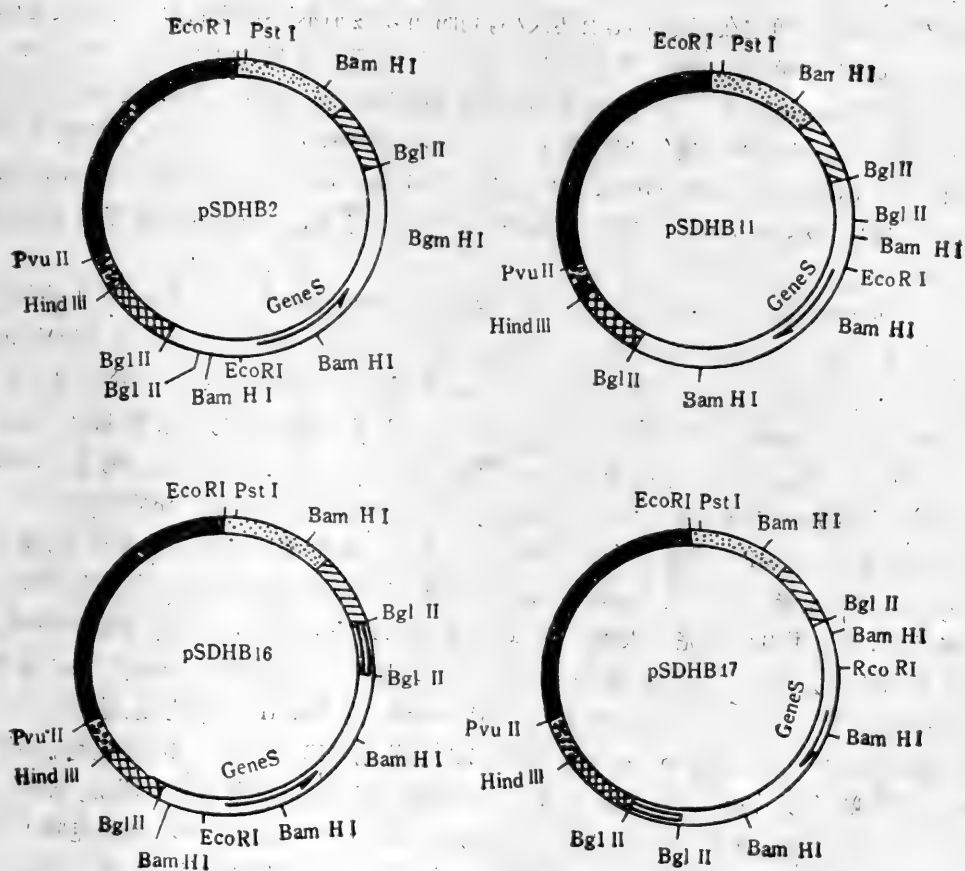


图 11.6 四种新重组质粒的结构示意图。

▨ SV40 (复制起点)。▤ 二氢叶酸还原酶基因。▧ SV40 的剪接位点。▩ SV40 的多聚A位点。■ pBR322 (复制起点)。pSDHB2 及 pSDHB11 含 2.743 kb 的 HBV DNA 片段。pSDHB16 及 pSDHB17 含 2.768 kb 的 HBV DNA 片段。所有这些片段均由 HBV DNA 用 Bgl II 不完全消化而来。同一片段的转录方向在不同的质粒中不同。

2.768kb 片段的且插入的拷贝数和连接方向不同的重组质粒，挑出其中具有代表性的 4 种 (pSDHB2、pSDHB11、pSDHB16 和 pSDHB17) 重组质粒进行进一步鉴定。

重组质粒的鉴定一般使用限制性内切酶水解重组质粒即可作出，必要时使用核酸转膜杂交法鉴定。现略述如下：

(1) 酶切电泳鉴定。上述 4 种质粒是用 2.743 和 2.768kb 片段的混合 HBV-DNA 与 pSV2-dhfr 重组产生。在 HBV-DNA 上无 Hind III 内切酶切点，在 pSV2-dhfr 上有 1 个切点。用 Hind III 水解 4 种质粒都只产生约 7.7kb 的一条带。Bgl II 内切酶水解除产生 pSV2-dhfr 的 5.0kb 的一条带外，尚产生 2.3kb 和 0.4kb 的两条带。此结果说明四种质粒均由一个拷贝 HBV-DNA 的 2.743kb 或 2.768kb 片段与 pSV2-dhfr 所构成。EcoRI 内切酶在 pSV2-dhfr 上只有一个切点，在 HBV-DNA 的 2.7 kb 的两种片段

上也有一个切口,但切口两侧的 HBV-DNA 片段的大小不同,所以 EcoRI 水解此四种质粒都产生两条带,通过在电泳上比较其大小,同时可以鉴别不同质粒的插入方向及是哪种片段。EcoRI 水解 pSDHB2 产生 4.1kb 和 3.6kb 片段,水解 pSDHB16 产生 4.0kb 和 3.7kb 片段,表明前者含有 2.743kb,后者含有 2.768 kb 片段;二者 S 基因的转录方向均和 dhfr 基因一致。EcoRI 水解 pSDHB11 产生 5.3kb 和 2.4kb 片段,水解 pSDHB17 产生 5.7kb 和 2.0kb 片段,表明 pSDHB 11 含有 2.743 kb HBV-DNA 片段,pSDHB17 含有 2.768kb HBV-DNA 片段,二者 S 基因的转录方向均与 dhfr 基因的转录方向相反(图 11.5A)。

(2) 核酸转移杂交鉴定。将 Bgl II 完全水解的四种质粒 DNA 和同一内切酶完全水解的 pTHBV-1 对照均转移到硝基纤维素薄膜上,与 [α^{32} P]dCTP 标记的 HBV-DNA 探针杂交。结果表明,这四种质粒产生的 2.3kb 片段均系特异性的 HBV-DNA Bgl II 完全水解片段(图 11.5B)。综合上述鉴定结果可给出四种重组质粒的示意图(图 11.6)。

使用 adw 或 adr 亚型的 HBV-DNA 初级克隆进行重组质粒改造(张一鸣等, 1985), 由于 adw 的初级克隆 pAol-HBV 是由 EcoRI 处剪切重组的全长 HBV-DNA 一个拷贝的细菌质粒, EcoRI 切断了 HBV-DNA 的前 S 基因,如需要完整的前 S 基因,就要用 EcoRI 从初级克隆切下全长 3.2kb 的 HBV-DNA 基因组,经电泳分离,再用 T4DNA 连接酶环化,然后用 Bgl II 完全水解。因 adw 亚型 HBV-DNA 只有 2 个 Bgl II 切点,故经水解后即可去掉 C 基因的主要部分,而留下 2.8kb 的带有完整的 Pre-S 和 S 基因的 HBV-DNA 片段。同上法将此片段插入 pSV2-dhfr 的 Bgl II 切口处,经过酶切鉴定即可获得改造后的重组质粒 pSDHBw (图 11.7B)。使用 HBV-DNA adr 亚型时,由于原插入的 3.2kb 全长片段是 BamHI 片段,已将 HBV-DNA 的多聚 A 基因切断,如既要前置 S 及 S 基因完整,又要多聚 A 完整,则需使用 BamHI 从 pHBV-NCI 初级克隆上剪切下 3.2kb 片段,加以环化,再用 Bgl II 切开,同上法插入 pSV2-dhfr 质粒,即可获得改造后的 adr 亚型的重组质粒 pSDHBr-1 (图 11.7A)。

pSDHBw 和 pSDHBr-1 的酶切鉴定:原则上和前述方法类似,即先用 Bgl II 完全水解,经电泳证明有 2.8kb 的插入片段,再选用在 pBR322 2.3kb 片段上和插入的 HBV 序列上有单一切口的内切酶,计算好插入片段的正反面不同在酶切后会出现大小不同的片段的情况。鉴定时 pSDHBw 用 EcoRI 切, pSDHBr-1 用 BamHI 切,从各自的两条大小不同的电泳带判断这两种重组质粒均属正向。

上述重组质粒改建的方法所用供体 DNA 均是 HBV-DNA 的 Bgl II 大片段,除含有 Pre-S 和 S 基因外尚含有 HBV 的启动子,且置于 dhfr 基因后,可能使 SV40 的强启动子未充分发挥作用。为了纠正上述改建中可能存在的缺点,可将 HBV-DNA 的 S 基因直接插入 SV40 启动子后,由于供体 DNA 和受体 DNA 的粘末端不同,需使用平末端连接方法(见下述)。

2. 平末端连接法改建重组质粒

使用上述 pSDHBr-1 作供体 DNA,用 Hind III 和 Xho I 完全水解, S1 酶削平粘末端,再用 T4DNA 连接酶环化,pSDHBr-1 去掉 Hind III-Xho I 小片段后成为 pSVHBr,其 SV40 启动子直接和乙肝 S 基因 5' 端相连接。在 S 基因的起始信号 ATG 前只有 23 bp 的旁侧序列,这种结构可能有利于发挥 SV40 早期启动子的作用。为利用 dhfr 基

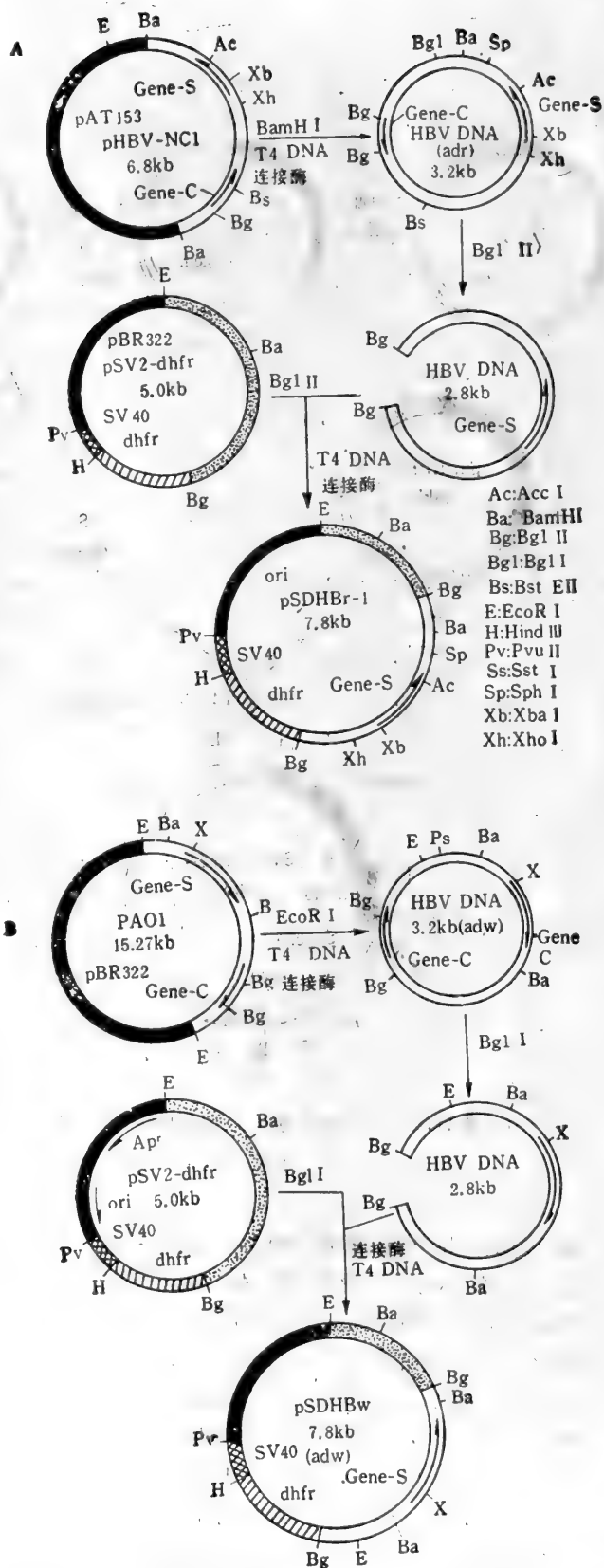


图 11.7 粘末端连接构建 adr, adw 亚型乙肝重组质粒。

A. 重新构建 pSDHBr(adr)重组质粒。 B. 重新构建 pSDHBw(adw) 重组质粒。

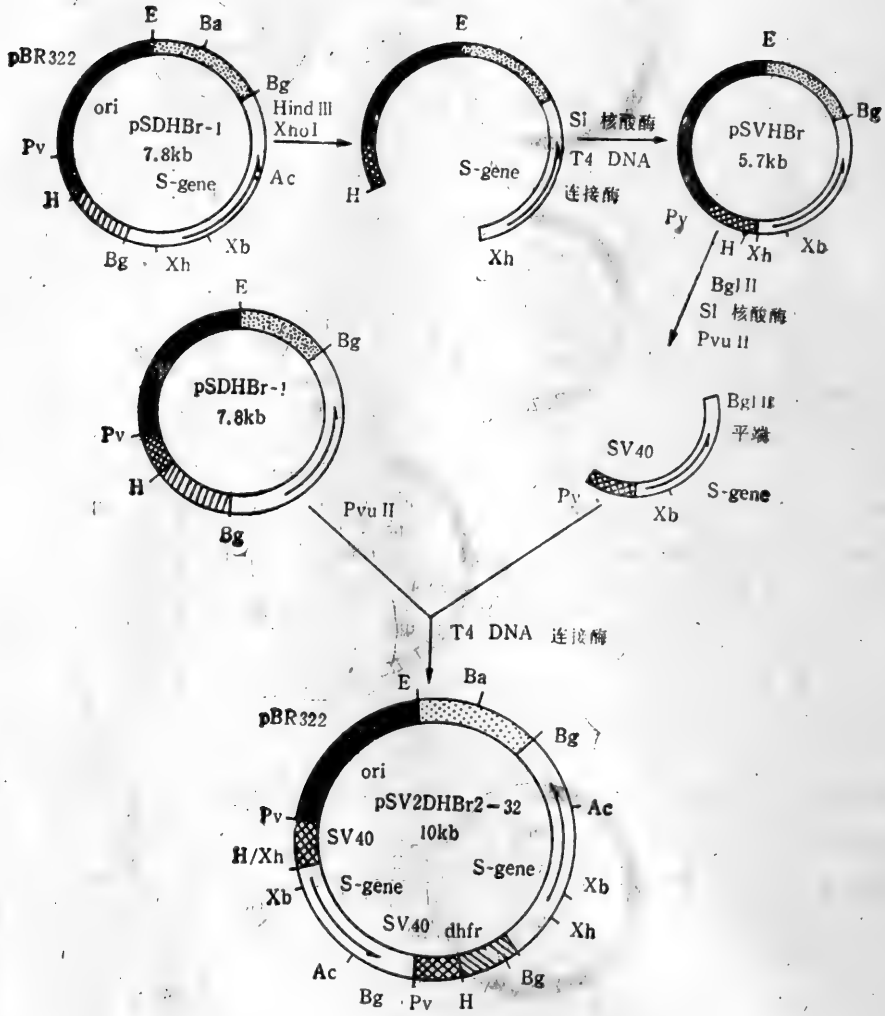


图 11.8 平末端连接构建 pSV2DHBr2-32 质粒。

因，再以 pSVHBr 为供体用 Pvu II 和 Bgl II 完全水解，S1 酶削平粘末端，电泳回收 PvuII-Bgl II 约 2.1 kb 的小片段，含有 SV 40 启动子及乙肝的 S 基因，将此片段插入 pSDHBr-1 的 PvuII 切口处，构建成一个双拷贝 S 基因的质粒，从 PvuII 起向 3' 方向的排列顺序是 SV40 → 无 Pre-S 的 S 基因 → SV40 → dhfr → 带完全 Pre-S 的 S 基因，命名为 pSV2DHBr-2-32 (图 11.8)，其大小为 10kb，同前法鉴定结果为正向。

将 pSV2 DHBr2-32 用 BglIII 完全水解，使其缺失掉 3' 端的带 Pre-S 的 S 基因，即构成 pSV2 DHBr1-32，此质粒只带有一个无 Pre-S 的 S 基因。

使用上述设计组建的乙肝重组质粒可分类如下：

(1) 利用 HBV-DNA Pre-S 和 S 基因及其启动子和 dhfr 扩增基因组建的质粒有

pSDHB2、pSDHB16 (ayw)、pSDHBw(adw) 及 pSDHBr(adr)。

(2) 利用 SV40 启动子直接和 S 基因连接及 SV 40 启动的 dhfr 基因组建的质粒为 pSV2DHB1-32。

(3) 既利用 SV40 启动子直接和无 Pre-S 的 S 基因连接,也利用完整的 Pre-S 与 S 基因及其自身启动子及 dhfr 扩增基因构造的质粒为 pSV2DHB2-32。

以上这些改建后的乙肝重组质粒均可获得 $1\mu\text{g/ml}$ 左右的 HBsAg 的表达,其中 pSV2·DHB1-32 表达量最高。

四、受体细胞

接受外源基因的重组 DNA 从而表达外源基因的动物细胞谓之受体细胞。受体细胞目前只能用传代细胞而不能用原代细胞,原因是后者生活代数短,不能适应基因工程技术导入 DNA、选择、克隆、建立细胞系以及连续传代收集产物等之用。传代细胞分为两大类:一类是营养缺陷型 (auxotrophic) 细胞,例如 Ltk^- 、 CHO-dhfr^- 细胞。另一类是非营养缺陷型细胞,这类细胞很多,诸如来自绿猴肾的传代细胞 CV-1、Vero、小鼠的 3T3、C127-1 等等。下面将分别介绍两类细胞接受外源基因,并提供选择的原理和使用条件。

1. 营养缺陷型细胞

细胞的核酸代谢途径有两条:其一是正常代谢途径,细胞利用葡萄糖经磷酸戊糖途径合成 5-磷酸核糖 (R-5-P),再经磷酸核糖焦磷酸激酶催化与 ATP 作用合成 1-焦磷酸-5-磷酸核糖 (PRPP),然后经过 10 步代谢反应而合成次黄嘌呤核苷酸 (IMP)。如果在上述 10 步代谢反应中由于细胞某一基因缺失或由于核酸代谢药物的阻断,常使所需要的酶缺失,细胞尚可利用第二条途径,即应急途径完成核酸代谢以保证细胞的存活和繁殖。在应急途径中已知的酶有次黄嘌呤鸟嘌呤磷酸核糖转移酶 (HGPRT)、腺嘌呤磷酸核糖转移酶 (APRT) 以及胸腺嘧啶核苷激酶 (TK) 等。HGPRT 和 APRT 可以利用外源的次黄嘌呤和腺嘌呤合成次黄嘌呤核苷酸 (IMP)。TK 可以利用外源的胸腺嘧啶脱氧核苷合成胸腺嘧啶脱氧核苷酸 (dTMP) 从而完成嘌呤嘧啶的代谢过程,保证细胞生长。营养缺陷型细胞就是指那些缺失了 HGPRT 或 APRT 或 TK 以及二氢叶酸还原酶 (DHFR) 基因的传代细胞。HGPRT、APRT、TK 基因缺失的细胞的应急途径不通,而 DHFR 缺失的细胞则是影响到细胞核酸代谢的正常途径中断。前者在正常培养基中能够正常生长和繁殖,而后者则需利用其未缺失的 HGPRT、APRT 以及 TK 的存在,在培养基中加入次黄嘌呤或腺嘌呤及胸苷才能维持这类细胞的正常生长和繁殖。利用这类营养缺陷型细胞经导入相应的基因弥补它们的缺陷,使用相应的核酸代谢阻断剂形成各自的选择系统,可从百、十万个细胞中能选出一个转化细胞,其妙用自不待言。

下面就两种不同的营养缺陷型细胞具体说明其原理及使用条件。

(1) 小鼠胸腺嘧啶激酶基因缺失细胞 (Ltk^-)。这是一种双缺失细胞,即完全缺失了细胞 DNA 双链上的 tk 基因 (tk^-)。它的正常核酸代谢途径无障碍,因而在正常培养基中可以生长良好。当使用氨基蝶呤或氨甲蝶呤等核酸代谢阻断剂使其正常的核酸代谢途径中

断,即使在培养基中加入胸苷和次黄嘌呤,细胞由于缺失 tk 基因及胸苷激酶终至死亡。如果利用基因工程技术将克隆的 tk 基因[通常用疱疹病毒 tk 基因克隆质粒 pX1(Enquist, 1979),它是由 pBR322 和 HSV-tk 重组而成]导入细胞,当 HSV-tk 和细胞 DNA 发生整合后, Ltk⁻细胞即转化成 Ltk⁺ 细胞,其转化率虽只有 10⁻⁵—10⁻⁶,即十万—百万分之一,在加入 HAT (hypoxanthine, aminopterin, thymidine) 培养基后,未转化的 tk⁻细胞很快死亡,而 tk⁺ 细胞得以存活并长成细胞集落。如果在 HSV-tk 导入细胞的同时将另一外源基因的重组 DNA 也导入细胞(此法称为共转化),或将外源 DNA 事先和 tk 基因重组入同一质粒中,将此重组 DNA 导入细胞(单转化),在使用上述选择培养基条件下,可以借 tk 基因标记获得 tk⁺ 并能表达外源基因的克隆细胞系。

(2) 中华地鼠卵巢二氢叶酸还原酶基因缺失细胞 (CHO-dhfr⁻) (Chasin, L. 赠)。这也是一种完全缺失了细胞 DNA 双链上的二氢叶酸还原酶基因 (dhfr⁻) 的双缺失细胞。在动物细胞正常的核酸代谢途径中,甲酰四氢叶酸参与十步反应中的两步,提供甲酰基完成 1-甘氨酸胺-5'-磷酸核苷 (GAR) 到 1-甲酰甘氨酸-5'-磷酸核糖 (FGAR),及 5-氨基-4-氨基甲酰咪唑-5'-磷酸核苷 (AICAR) 到次黄嘌呤核苷酸 (IMP) 的代谢反应。在胸苷酸 (dTMP) 的代谢中也需甲酰四氢叶酸提供甲基而使脱氧三磷酸尿苷 (dUTP) 甲基化而完成。二氢叶酸还原酶 (DHFR) 是催化二氢叶酸还原成四氢叶酸所必需的酶,它的缺失使细胞失去了四氢叶酸的供应,细胞的正常核酸代谢途径必然中断,此种细胞在未转化前在正常培养基中也不能生长,必须加入次黄嘌呤(或腺嘌呤)和胸苷,借助其应急途径才能生长。如果将克隆化的小鼠 dhfr 基因导入细胞,dhfr 基因一旦和细胞 DNA 整合, dhfr⁻ 细胞则转化为不缺失的 (dhfr⁺) 细胞。在导入 dhfr 基因后从培养基中撤去次黄嘌呤(H)和胸苷(T),转化的 dhfr⁺ 细胞即可生长,而未转化的依然是 dhfr⁻ 的细胞则迅速死亡。利用这一原理,如果把一个外源基因与 dhfr 基因重组入同一质粒,用其重组 DNA 转化 CHO-dhfr⁻ 细胞,撤除原培养基中的 H 和 T,即可获得 dhfr⁺ 并能表达外源基因的克隆细胞系。

2. 非营养缺陷型(代谢途径正常)细胞

所有的动物传代细胞除已证明是营养缺陷型细胞之外均属此类,因其种类多未必均

表 11.1 从三系猴肾细胞选择非营养缺陷型细胞

选择试剂 ($\mu\text{g/ml}$)	营养缺陷性状	猴肾细胞		
		LLC-MK2	Vero	CV-1
5'-溴脱氧尿嘧啶 30	TK ⁻ +	(P4) +*	(P4) +*	(P3) +*
8-杂氮鸟嘌呤	HGPRT ⁻ +	-	-	-
二氨基嘌呤 50	APRT ⁻ +	-	-	-
霉酚酸 25	XGPRT ⁻ +	-	-	-
HAT	TK ⁺ 和 HGPRT ⁺ +	-	+	+
MXAAT	XGPRT ⁺ +	-	±	-

HAT: 次黄嘌呤 15 $\mu\text{g/ml}$, 氨基蝶呤 2 $\mu\text{g/ml}$, 胸苷 10 $\mu\text{g/ml}$ 。

MXAAT: 霉酚酸 25 $\mu\text{g/ml}$, 黄嘌呤 250 $\mu\text{g/ml}$, 腺嘌呤 25 $\mu\text{g/ml}$, 氨基蝶呤 2 $\mu\text{g/ml}$, 胸苷 10 $\mu\text{g/ml}$ 。

+ 示细胞生长, - 示细胞死亡, ± 示细胞不完全死亡。P_x 示细胞传代代数。

经过检查,是否有缺失以及缺失何种酶应事先作检查,在证实其确无缺失时方可使用,以免误事。表 11.1 示出几种细胞的检查结果。

从表 11.1 可见三种猴肾传代细胞的营养缺陷性状(阮薇琴等,1984,1985)。正常的动物细胞具有 TK、HGPRT、APRT 等酶,但不存在黄嘌呤鸟嘌呤磷酸核糖转移酶(XGPRT),因之正常细胞分别在含有 5-溴脱氧尿苷(BUDR)、8-杂氮鸟嘌呤,2-氨基嘌呤(2-Ap)及霉酚酸的培养基中不能生长,其中 BUDR 效应慢,需连续传代加药观察。表 11.1 中三种细胞以 CV-1 细胞较为理想,CV-1 对 BudR、8-Aza 及 2Ap 均敏感,在 HAT 培养基中能生长,说明 CV-1 细胞无 TK、HGPRT 及 APRT 的缺陷,确是无营养缺陷的细胞。CV-1 细胞对 MPA 敏感,在 MXAAT 选择培养基中也不生长,说明其不存在 XGPRT,适宜用 Eco-gpt 作标记进行转化和用 MX (mycophenolic acid 及 xanthine) 选择培养基选择(图 11.9)。LLC-MK2 细胞对 BudR 有抵抗,在 HAT 培养基中不生长,怀疑可能有 TK 或其他酶的缺失。Vero 细胞对 BudR 抵抗,对 MXAAT 有一定抗性,二者均不是理想的无营养缺陷的正常细胞。

使用正常细胞进行转化表达外源 DNA 时,需使用显性标记(dominant marker)作选择标记,详见下节。

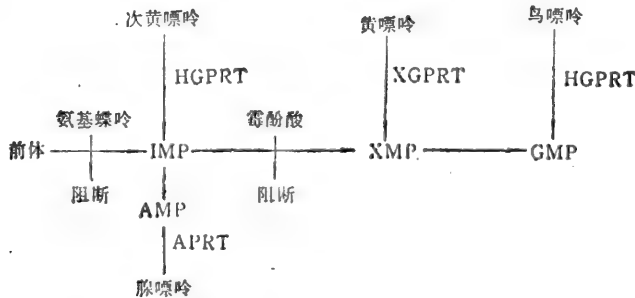


图 11.9 gpt 选择系统的嘌呤代谢[参考 Mulligan (1981) 文献绘制]。

五、选择标记和选择培养基

人工将外源 DNA 重组质粒导入动物细胞希望获得基因表达,也即产生基因编码的蛋白质,这是基因工程的目的。前已述及 DNA 转化细胞的机率很低。要想从几十万细胞中选出一至数个转化并有基因表达而且产量较高的细胞,如果没有选择标记几乎是不可能的。因此选择标记无论是用什么受体细胞(如大肠杆菌、酵母或动物细胞)都是必需的。对动物细胞而言,由于细胞的种类多、性质不同,使用的选择标记也各异。选择培养基一般是用细胞核酸代谢的阻断剂,同时加入欲利用的细胞代谢途径(如正常或应急途径)所需的底物,并使选择标记和选择培养基相匹配,每种细胞和选择标记、选择培养基相匹配构成各自的选择系统。常用者列举如下。

1. Ltk⁻ 细胞-HSV-tk 标记基因-HAT 选择培养基

Wiglar 等人使用这一标记系统十分有效 (Wigler 等, 1979a, 1979b)。在这一系统中 tk 基因导入并整合于细胞 DNA, HAT 选择培养基中的 aminopterin 阻断细胞正常代谢途径, 迫使细胞动用其 HGPRT 及新获得的 TK。利用 HAT 中的 hypoxanthine 及 thymidine 通过应急通路使 tk⁺ 细胞存活, tk⁻ 细胞死亡从而达到选择的目的。所用 HAT 的最终浓度 H15 μg/ml, A 2.2 μg/ml, T10 μg/ml, 事前配成 100 倍母液, 用时按 1% 比例加入生长液。此系统的优点是细胞易于培养、克隆, 缺点是 tk 基因只有选择作用并无扩增功能, 但可与其他系统联合使用而达到高效表达的目的。另外当把 A 的浓度加大 10 倍, T 的浓度减少 10 倍, 可迫使细胞表达稳定 (阮力等, 1984)。

2. CHO-dhfr⁻ 细胞-dhfr 标记基因-MTX 选择培养基

这一系统中 dhfr 基因不但可作为选择标记, 同时也可作为扩增基因。CHO-dhfr⁻ 细胞被 dhfr 基因转化后, 从细胞培养基中撤去 H (或 Adenine) 和 T, 转化后 dhfr⁺ 细胞即可在正常培养液中生长。当加入 MTX 后, 随着浓度加大 (一般从 10⁻⁷ mol/L 开始, 依次增加到 3—5 × 10⁻⁷ mol/L、10⁻⁶ mol/L), dhfr 基因的拷贝数即扩增, 产物 DHFR 也大量增加, 这种基因扩增涉及 dhfr 基因两侧的区段, 如果在其前或后插入外源基因, 就有可能同时使外源基因得到扩增。我们在乙肝 DNA 表达试验中利用这一原理, 可使 HBsAg 的产量扩增数倍到数十倍。

3. Vero 细胞-dhfr 标记基因-MTX 选择培养基

绿猴传代细胞 Vero 细胞系并不缺失 dhfr 基因, 在正常细胞培养液中可以生长良好。但试验证明它对氨甲蝶呤 (MTX, 0.02 μg/ml) 敏感, 其原因不明, 可能和细胞原有的 dhfr 基因不能快速适应 MTX 的加入量有关。因此重新导入 dhfr 基因使其和细胞 DNA 整合即可试用 MTX 选择出对 MTX 有抗性的细胞 (阮薇琴等, 1985)。用 pSDHB2 和 pSDHB 16 两种质粒转化 Vero 细胞, MTX 的终浓度从 0.01 μg/ml 渐增至 0.2 μg/ml, 结果选出 4 系能表达 HBsAg 的细胞系。

4. CV-1 细胞-Eco-gpt 选择标记-MX 选择培养基

5. Vero 细胞-Neo 选择标记-G418 选择培养基

这两个系统中所用的细胞均是无营养缺陷型细胞, 来自大肠杆菌 XGPRT 的基因 (Eco-gpt) 及氨基糖苷转移酶 (AGT) 的基因 (Neo)。在正常动物细胞中并没有这些酶或基因, 使用这些基因作标记称为阳性标记。当把这些基因导入并与细胞 DNA 整合, 即可对相应的选择培养基产生抗性而被选择出来, 那些未整合有这些基因的细胞同时死亡。和 Eco-gpt 相匹配的选择培养基是霉酚酸 (M) 和黄嘌呤 (X), 其中 M (终浓度 25 μg/ml) 阻断 IMP → XMP 的细胞代谢途径 (图 11.9), 只有整合有 Eco-gpt 基因的细胞才可借其编码的 XGPRT 利用 X (终浓 250 μg/ml) 合成 XMP 而使细胞生长, 用这一系统使 HBsAg 在 CV-1 细胞中获得表达 (阮薇琴等, 1984)。和 Neo 相匹配的选择培养基是 G418 (Colbere-Garapin, 1981), G418 是一种蛋白合成抑制剂, 借其 2-去氧链霉胺的作用抑制真核细胞蛋白的合成。导入并整合了 AGT 的基因的细胞可以抵抗 G-418 的作用而存活, 反之在 G418 作用下全部死亡, 在此系统中 G418 的终浓度为 400—600 μg/ml。

以上两个系统的共同缺点, 一是选择细胞克隆较难, 二是因没有扩增基因, 外源基因表达量不高。前者和细胞的生长有关, 后者则为 Eco-gpt 及 Neo 基因本身的性质所限。

六、重组 DNA 转染细胞

1. 重组质粒的制备

无论是外源基因与选择标记重组的质粒 DNA 或是只含选择标记基因的质粒 DNA, 转化细胞后获得质粒菌种均冻干保存, 用前活化质粒菌种。

(1) LBTMG 培养基的配制。

Bacto trypton	10g
粉剂酵母浸出液	5g
NaCl	5g
1mol/L Tris-HCl	10ml
1mol/L MgCl ₂	1ml
加水至	1,000ml

调整 pH 至 7.5, 高压灭菌, 冷却后再加入 40% 的葡萄糖溶液 5ml 及注射用维生素 B₁ 10mg。

(2) 活化带质粒菌种。将菌种接种于 5ml 的 LB TMG 培养基中, 加相应抗生素 (100 μ g/ml 氨苄青霉素或 100 μ g/ml 氨苄青霉素加 20 μ g/ml 四环素), 37 $^{\circ}$ C 静置培养过夜。

(3) 重组 DNA 的扩增。

- 5ml 活化菌液转种于 1,000ml LBTMG 培养基中(含相应抗生素, 量同上)。
- 37 $^{\circ}$ C 振荡培养 6—12 小时, 细菌浓度达到 OD₅₅₀ = 0.9—1 之间。
- 加氯霉素 (250 μ g/ml)。
- 37 $^{\circ}$ C 振荡培养 12—24 小时。

(4) DNA 的提取与纯化。

a. 上述菌液在 4 $^{\circ}$ C 5,000 r/min 离心 10 分钟, 弃上清液, 用预冷的蔗糖缓冲液洗 1 次, 细菌重悬于 15ml 预冷的蔗糖缓冲液中 (50mmol/L Tris-HCl, pH8.5, 50mmol/L Na₂EDTA, 15% 蔗糖), 置冰浴中。

b. 加 2ml (15—20mg) 新鲜配制的溶菌酶水溶液, 迅速混匀, 冰浴中 (0 $^{\circ}$ C) 放置 5 分钟, 至菌体稍变粘稠。

c. 加 30 ml Triton X-100 裂解液 (50 mmol/L Tris-HCl, 50 mmol/L EDTA, pH8.5, 0.1% Triton X-100), 0 $^{\circ}$ C 下放置 10—20 分钟, 使菌液变粘稠。

d. 4 $^{\circ}$ C 下 17,000r/min 离心 1 小时, 收集清亮上清液。

e. 加等体积饱和酚抽提 1 次, 回收水相, 再加氯仿-异戊醇 (24:1) 抽提取 1 次, 仍回收水相。

f. 加 1ml (1—2mg) 加热处理过的 RNA 酶 (100 $^{\circ}$ C, 10 分钟), 37 $^{\circ}$ C 作用 1 小时。

g. 加 1/4 体积的 5 mol/L NaCl 及 1/3 体积的 30% 的聚乙二醇 (Mr = 6000), -10 $^{\circ}$ C 过夜。

h. 4 $^{\circ}$ C, 10,000 r/min 离心 1 小时, 分离上清液。

i. 沉淀物重悬于 2ml 的 50mmol/L Tris HCl (pH8.5)、250mmol/L NaCl、5mmol/L

L EDTA 缓冲液中,用饱和酚及氯仿-异戊醇(24:1)各提取 3 次。

j. 加 2.5 体积无水乙醇沉淀质粒 DNA (液氮 5 分钟,或 -70°C 15 分钟,或 -20°C 过夜), 4°C , 15,000 r/min 离心 15 分钟。

k 沉淀 DNA 用 70% 乙醇洗 3 次,抽干后溶于 TE 液中备用。

2. DNA 转染用液体的配制

(1) 2 倍 HeBS: 50mmol/L Hepes, 280 mmol/L NaCl, 1.5mmol/L Na_2HPO_4 , 准确调 pH 至 7.1 ± 0.05 。

(2) 2mol/L CaCl_2 。

以上液体可在滤过灭菌后分装于 -30°C 保存备用。

3. 选择培养基的配制

(1) HAT 选择培养基:

次黄嘌呤 0.11mmol/L

氨基蝶呤 0.5mmol/L

胸腺嘧啶脱氧核苷 40mmol/L

分别用 1mol/L NaOH 溶解,加无菌双蒸水配成 100 倍浓缩母液,用时按 1/100 加入细胞培养液。

(2) MTX 选择培养基。

氨甲蝶呤 10^{-4} mol/L 用 1mol/L NaOH 溶解,加水配制而成,用时按需要稀释加入细胞生长液。

(3) MX 选择培养基。

霉酚酸 25 $\mu\text{g}/\text{ml}$

黄嘌呤 250 $\mu\text{g}/\text{ml}$

均用 1mol/L NaOH 溶解,加水配成 100 倍浓缩母液,用时按 1/100 加入细胞培养液。

(4) G418 (Gibco 产品) 400 $\mu\text{g}/\text{ml}$, 用无菌蒸馏水配制成 100 倍浓缩母液,用时按 1/100 或 1.5/100 加入细胞生长液。

以上除已注明者外均为 Sigma 产品,配成的液体用前均在 -20 — -30°C 冻存。

4. 转染用细胞的准备

(1) 冻存的细胞经复苏后传代, 37°C 培养。

(2) 生长液均用 DME 加 10% 小牛血清及青链霉素各 100 单位/ml 生长液。

(3) Ltk^- 细胞需在生长液中加入 30 $\mu\text{g}/\text{ml}$ 的 BudR, 连续传 3 代以去除自发回变的 Ltk^+ 细胞。

CHO-dhfr⁻ 细胞需在生长液中加入:

次黄嘌呤 0.03mmol/L

胸腺嘧啶脱氧核苷 0.03mmol/L

脯氨酸 10^{-4} mol/L

甘氨酸 0.1mmol/L

各自先配成 100 倍浓缩母液
用时按 1/100 加入生长液

(4) 传代细胞用前连续在正常培养液中快速传 3 代(2 天传代 1 次,细胞浓度可适当加大)。

(5) 在实验前一日每瓶 (100ml 容量方瓶,底面积 36cm^2) 接种 $1-2 \times 10^6$ 细胞,使

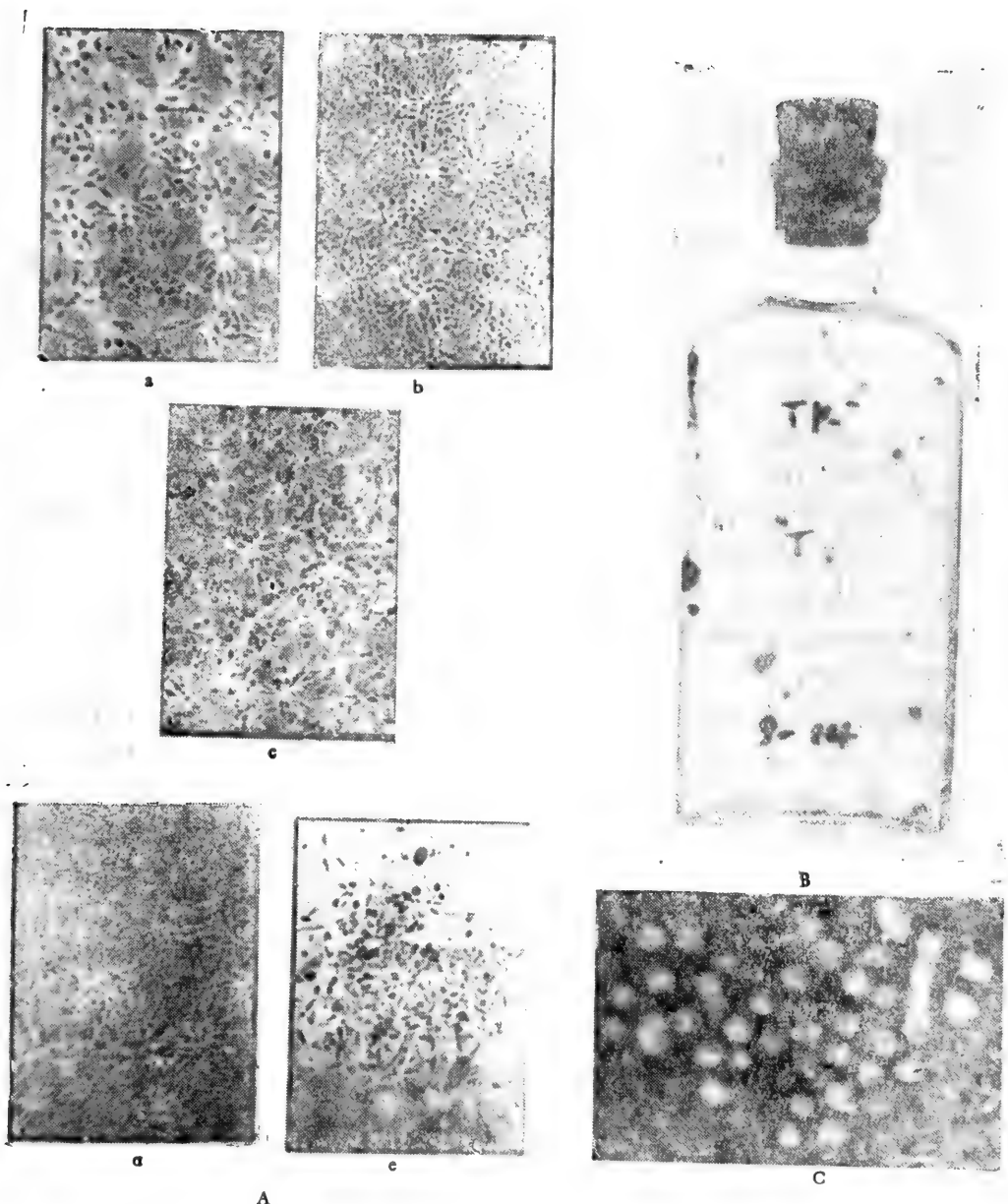


图 11.10 A. 转化过程中的细胞状态。a. 吸附 DNA 时。b. 加 HAT 时。c. 加 HAT 3 天后。d. 未加 DNA 及 HAT 的细胞对照。e. 二周后长出的 TK⁺ 细胞灶。B. 转化成功的 TK⁺ 细胞灶结晶紫染色。C. 哺乳动物细胞分泌的 HBsAg, 1EM 200,000X。

在实验当天长成分布均匀但不完全成片的单层细胞。

5. DNA 转染细胞步骤

(1) 在实验前 4 小时细胞换液, 仍在 37°C 培养。

(2) 将 1 μ g 的标记基因 DNA, 10 μ g 的外源基因重组 DNA 及 10 μ g 的携带 DNA (常用鲑鱼精子 DNA) 溶于 10mmol/L Tris-HCl(pH7.1) (共转化)。若标记基因已和外源基因组建在同一质粒, 则只加后两种 DNA。补加同样液体及 0.125ml 的 2 mol/L

CaCl₂ 液,使总量为 1ml。将此液缓缓加入 2 倍的等量的 HeBS 液,随加随轻轻混匀,室温静置 15—30 分钟,随时作镜下观察,待出现肉眼可见的轻度混浊,镜下可见 DNA 磷酸钙细小颗粒时立即使用。

(3) 每瓶细胞中加入 1ml DNA 磷酸钙悬液,和瓶内生长液轻轻混匀,37℃ 培养 8—12 小时,倒净瓶中液体,加入新生长液 10ml,继续在 37℃ 静置培养。

(4) 经 24—48 小时再次换入新生长液,同时加入选择培养基。这要视细胞生长速度而定,细胞生长快需提前在 20—24 小时即换液,否则可延至 48 小时再换液。或者在 48 小时消化分种多瓶细胞,使用 CHO 细胞时在 DNA 转染后第 1 次换液即可撤除次黄嘌呤和胸苷。但还需加入脯氨酸和甘氨酸,因后者为 CHO 细胞所必需。

6. 转化细胞克隆

(1) 细胞经 DNA 转化,加入选择培养基后约 2—3 周,未转化的细胞在头 1—2 周陆续死亡脱落,转化细胞长成肉眼可见的细胞集落(见图 11.10A (a、b)),镜下可见数十至数百个细胞成向外生长的圆型集落,在集落细胞未老化前视机进行克隆。

(2) 细胞克隆方法。文献报告均用一特制的金属小环置于细胞集落周围,底部密封,环内加入少量细胞消化液,消化剥离的细胞用吸管吸出进行培养。国内因暂时限于条件,可用下述方法克隆。

事先拉制弯头毛细吸管,头部口细而平,吸取小量消化液(0.25% 胰酶 1 份,EDTA 4 份混合,若细胞不易自瓶壁剥离可加大胰酶比例)对准事前已用笔标好的细胞集落,边刮边吸将细胞尽量吸入吸管中。然后移至 96 孔平底细胞培养板中,每孔加液量不大于 0.15 ml。置 CO₂ 孵箱中培养,生长液的 pH 要调至 7.1 左右,CO₂ 的浓度要求 5%。开始时每 5—7 天换液,其后视细胞生长情况改为 2—3 天换液 1 次,待细胞长成单层经过消化转种至 24 孔细胞培养板,仍在 CO₂ 孵箱中培养,待长成单层后再消化转种至小瓶中培养。为使克隆细胞快速成活生长,在克隆后可将生长液中的选择培养基减量或暂时撤消。

7. 细胞生长液中外源基因表达产物的检测

以乙型肝炎病毒表面抗原(HBsAg)为例。受乙肝病毒重组 DNA 转化的细胞,如果条件合适即可分泌 HBsAg,但不同细胞克隆系的分泌量有所不同。为在多个克隆系中及时找到高效表达 HBsAg 的细胞系,必须及时检测,以便淘汰不表达或表达低的克隆系。

(1) RPHA 或 Elisa 检测 HBsAg: 当克隆细胞在 24 孔培养板中长成单层时,及时换入 1ml 生长液,24 或 48 小时后收液待测。细胞传代,收集的液体用 RPHA 或 Elisa 检测 HBsAg,方法见下述常规。RPHA 法以最终稀释度++为阳性,一般达到 1:200 左右仍为阳性视作高产细胞系,Elisa 法则以最终稀释度的 P/N 值 ≥ 2.1 者为阳性。RPHA 法检测快速简便,可为粗筛使用,但需注意阳性及阴性对照每次试验都不可缺少,且需多次检测。

(2) RIA 法: 使用 Abbott 产 AUSRIA-II 125 试剂盒最佳。在初筛获得高产细胞系后,摘其高者用此法检测。细胞生长液定时收获,细胞消化传代同时计数,收集的液体作 1:50、1:100、1:500、1:1000 稀释,进行检测。所获结果与该试剂阳性对照(含 HBsAg 20ng/ml \pm 5ng) 相比计算 HBsAg 表达产量。

(3) 免疫电镜观察。按常规制作标本进行电镜观察(见图 11.10c)。

8. 高效表达 HBsAg 的转化细胞系的建库

经多次检测不同细胞系不同代数的细胞培养液, 确定 1—2 个确属高产细胞系后, 立即传代冻存 30—50 支作为种子细胞库。每支冻存细胞约含 10^6 细胞, 冻存液用生长液 77%、小牛血清 15%、二甲基亚砜 8%。每支液量 1ml, 安瓶密封冻存于液氮液相中, 若作为生产用细胞可用种子细胞库细胞传代增殖, 视生产需要建立工作细胞库。

七、图示几种转化表达 HBsAg 的系统

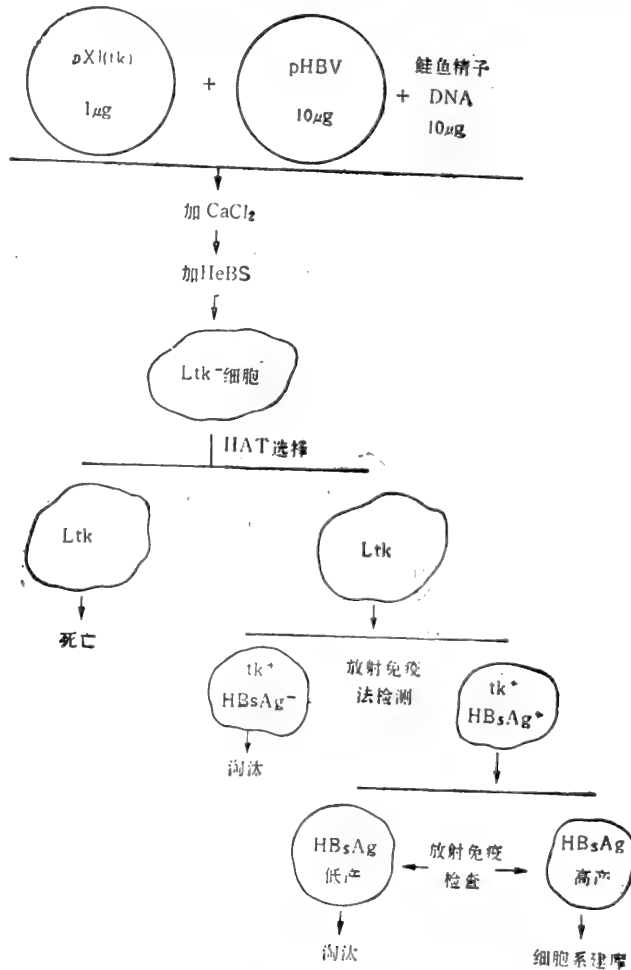


图 11.11 Ltk⁻ 细胞-tk 标记基因-HAT 选择培养基系统。

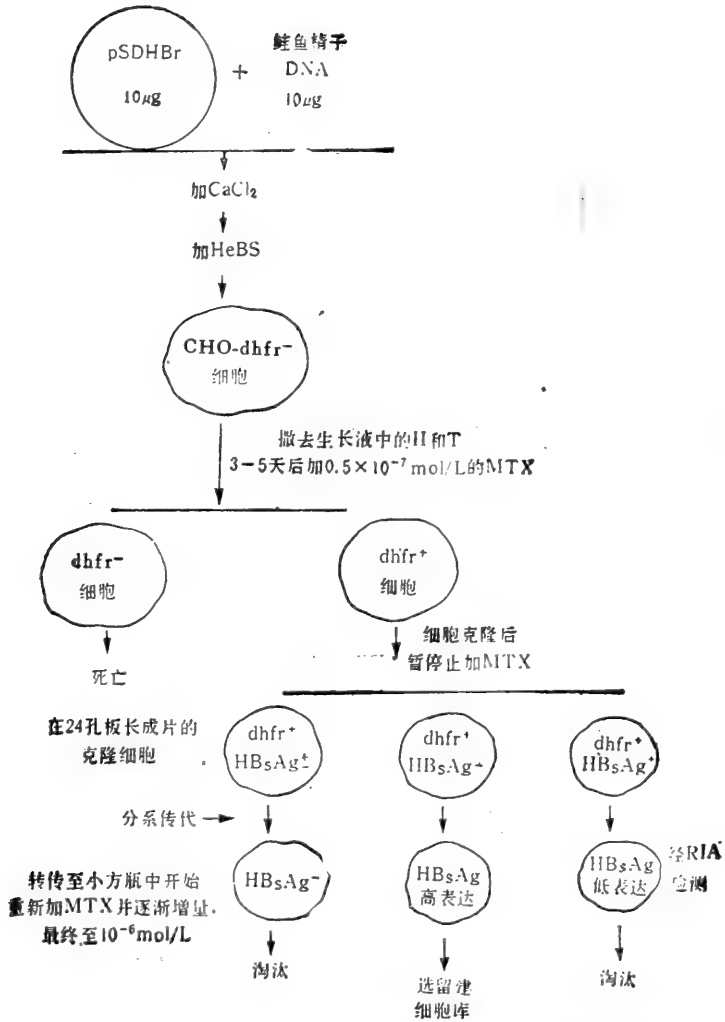


图 11.12 CHO-dhfr⁻ 细胞-dhfr 选择标记-MTX 选择培养基系统。

oleic Acids Res., **12**, 915, 1984.

- [12] Berg, P., Dissections and reconstructions of genes and chromosomes, *Science*, **213**, 296, 1981.
- [13] Christman, J. K. *et al.*, Amplification of expression of hepatitis B surface antigen in 3T3 cotransfected with a dominant-acting gene and cloned viral DNA, *PNAS USA.*, **79**, 1815, 1982.
- [14] Colbere-Garapin, F. *et al.*, A new dominant hybrid selective marker for higher eukaryotic cells, *J. Mol. Biol.*, **100**, 1, 1981.
- [15] Cumming, I.W. *et al.*, Isolation, characterization and comparison of recombinant DNAs derived from genomes of human hepatitis B virus and woodchuck hepatitis virus, *PNAS USA.*, **77**, 1842, 1980.
- [16] Dubois, M.F. *et al.*, Excretion of hepatitis B surface antigen particle from mouse cells transformed with cloned viral DNA, *PNAS*, **77**, 4549, 1980.
- [17] Enquist, L. W. *et al.* Construction and characterization of a recombinant plasmid encoding the gene for the thymidine kinase of herpes simplex type 1 virus, *Gene*, **7**, 335, 1979.
- [18] Graham, F. L. *et al.*, A new technique for the assay of infectivity of human adenovirus 5 DNA, *Viral*, **52**, 456, 1973.
- [19] Hirshman, S.Z. *et al.*, Expression of cloned hepatitis B virus DNA in human cell culture, *PNAS USA.*, **77**, 5507, 1980.
- [20] Mulligan, R. C. *et al.*, Selection for animal cells that express the *E. coli* gene coding for xanthine-guanine phosphoribosyltransferase, *PNAS USA.*, **78**, 2072, 1981.
- [21] Mulligan, R. C. *et al.*, Expression of a bacterial gene in mammalian cells, *Science*, **209**, 1422, 1980.
- [22] Hsiung, N. *et al.*, Efficient production of hepatitis B surface antigen using a bovine papilloma virus metallothionein vector, *J. Mol. Appl. Genetics*, **2**, 297, 1984.
- [23] Petriccioni, J. C. *et al.*, The safety of products derived from continuous, cell lines. DRAFT. WHO., 1986.
- [24] Rutter, W. J. *et al.*, The nucleotide sequence of hepatitis B viral genome and the identification of the major viral genes, *Animal Virus Genetics* (Eds, Fields, B. *et al.*) Academic Press, New York, 57-70, 1980.
- [25] Standing, D. N. *et al.* Transcription of hepatitis B surface antigen gene in cultured murine cells initiates within the presurface region, *J. Viral.*, **50**, 563, 1984.
- [26] Subramani, S. *et al.*, Expression of the mouse dihydrofolate reductase complementary deoxyribonucleic acid in simian virus 40 vectors, *Mol. Cellular Biol.*, **1**, 854, 1981.
- [27] Tiollais, P. *et al.*, The hepatitis B virus, *Nature*, **317**, 489, 1985.
- [28] Wang, Y. *et al.*, Enhanced production of hepatitis B surface antigen in NIH 3T3 mouse fibroblasts by using extrachromosomally replication of bovine papillomavirus vector, *Mol. Cellular Biol.*, **3**, 1032, 1983.
- [29] Wigler, M. *et al.*, DNA-mediated transfer of the adenine phosphoribosyltransferase (*aprt*) locus in to mammalian cells, *PNAS USA.*, **76**, 1373, 1979a.
- [30] Wiglar, M. *et al.*, Transformation of mammalian cells with genes from procaryotes and eucaryotes, *Cell*, **16**, 777, 1979b.

附录一 大肠杆菌的常用培养基

1. M9 培养基(每升)

Na ₂ HPO ₄	6g
KH ₂ PO ₄	3g
NaCl	0.5g
NH ₄ Cl	1g

用 NaOH 调 pH 至 7.4, 121℃ 灭菌 20 分钟, 冷后加入:

1mol/L MgSO ₄	2ml
20% 葡萄糖	10ml
1mol/L CaCl ₂	0.1ml

以上三种溶液分别过滤灭菌或消毒。

在配制琼脂培养基时, 应分别配制 1/2 体积的 2.4% 的琼脂, 1/2 体积的 2 倍强度的 M9 盐溶液, 分别灭菌, 冷后混合并加入上述三种溶液。

2. M63 培养基(每升)

KH ₂ PO ₄	3.0g
K ₂ HPO ₄	7.0g
(NH ₄) ₂ SO ₄	2g
FeSO ₄ ·7H ₂ O	5mg (1mg/ml 溶液 5ml)

用 KOH 调 pH 至 7.0, 121℃ 灭菌 20 分钟, 冷后加入:

1mol/L MgSO ₄ ·7H ₂ O	1ml
维生素 B ₁ (50mg/ml)	2ml
20% 葡萄糖	10ml

以上三溶液单独过滤或蒸汽灭菌。

配制固体培养基方法同 M9 培养基。

以上两种为合成培养基。为满足营养缺陷型生长的需要加入所需维生素至终浓度 1 μg/ml, 氨基酸至 40 μg/ml (L 型) 或 80 μg/ml (D、L 型)。

3. LB 培养基(每升)

细菌蛋白胨	10g
细菌酵母膏粉	5g
NaCl	10g

用 NaOH 调 pH 至 7.2。固体培养基加琼脂 15g/L。

4. YT 培养基

细菌蛋白胨	8g
细菌酵母膏粉	5g
NaCl	5g

用 NaOH 调 pH 至 7.2。

5. 麦康凯 (McConkey) 培养基(每升)

蛋白胨	20g
乳糖	10g
胆酸钠 (No.3)	1.5g
NaCl	5g
琼脂	13.5g
中性红	0.03g
结晶紫	0.001g

该培养基已被广泛用于发现和分离肠道杆菌的常规培养基之一，国内外均有商品出售，一般不必自己配制。

6. 穿刺琼脂培养基(每升)

牛肉膏粉	3g
蛋白胨	5g
琼脂	6g

加热溶化后分装小管，121℃ 灭菌 15 分钟。

7. 远藤氏 (Endo) 培养基(每升)

蛋白胨	10g
乳糖	10g
Na_2HPO_4	3.5g
Na_2SO_3	2.5g
碱性复红	0.4g
琼脂	15g

附录二 大肠杆菌的常用菌株

菌株	基 因 型
LE292	HfrH arg Eam rpoB galT:: ($\lambda\Delta$ [int-FII])
LE392	F ⁻ supE supF hsdR galK trpR metB lacY
MC 1000	F ⁻ araD139 Δ (aroABC-leu) 7679 galU K Δ (lac) \times 174 rpsL thi
MC 4100	F ⁻ araD139 Δ (araE-lac) u169 rps L150 rec A1 flb B 5301 deoC1 pts F25 rbs R
Cla	野生型
C600	F ⁻ sup E44 tonA21 thr ⁻¹ leuB6 lacY1
Q358	F ⁻ hsd Rk ⁻ hsdMk ⁺ supE ϕ 80 ^r
Q359	F ⁻ hsd Rk ⁻ hsdMk ⁺ supE ϕ 80 ^r (p2)
JM103	Δ (lac-pro) thi strA supE endA sbcB hsdR-F('traD36 proAB lac I ^g Z Δ M15)
K802	hsd R ⁻ hsd M ⁺ gal met supE
HB101	F ⁻ hsd S20 ($r_{\bar{b}}$ $m_{\bar{b}}$) recA13 ara-14 proA2 lac Y1 galK2 rpsL20 xy 1-5 mtl-1 supE44
BHB 2688	F ⁻ recA λ ^r (λ Eam4b2 red3 imm 434 Clts Sam7)
BHB 2690	F ⁻ recA λ ^r (λ Dam15b2 red3 imm 434 Clts Sam7)
X1776	F ⁻ tonA53 dapD8 minA1 glnV44 (sup E42) Δ (gal-uvrB)40 λ ⁻ minB2 rfb-2 gyrA25 thyA142 oms-2 metC65 oms-1 (tte-1) Δ (bioH-asd) 29 cycB2 cycA-1 hsd R2
CSH45	F ⁻ Δ (lac) thi trpR(λ CI857 sam7)
1046	F ⁻ recA supE supF hsd ^{s-} met
RR1	F ⁻ hsdR hsdM galK2 lacY1 ara14 proA2 ara14 strA

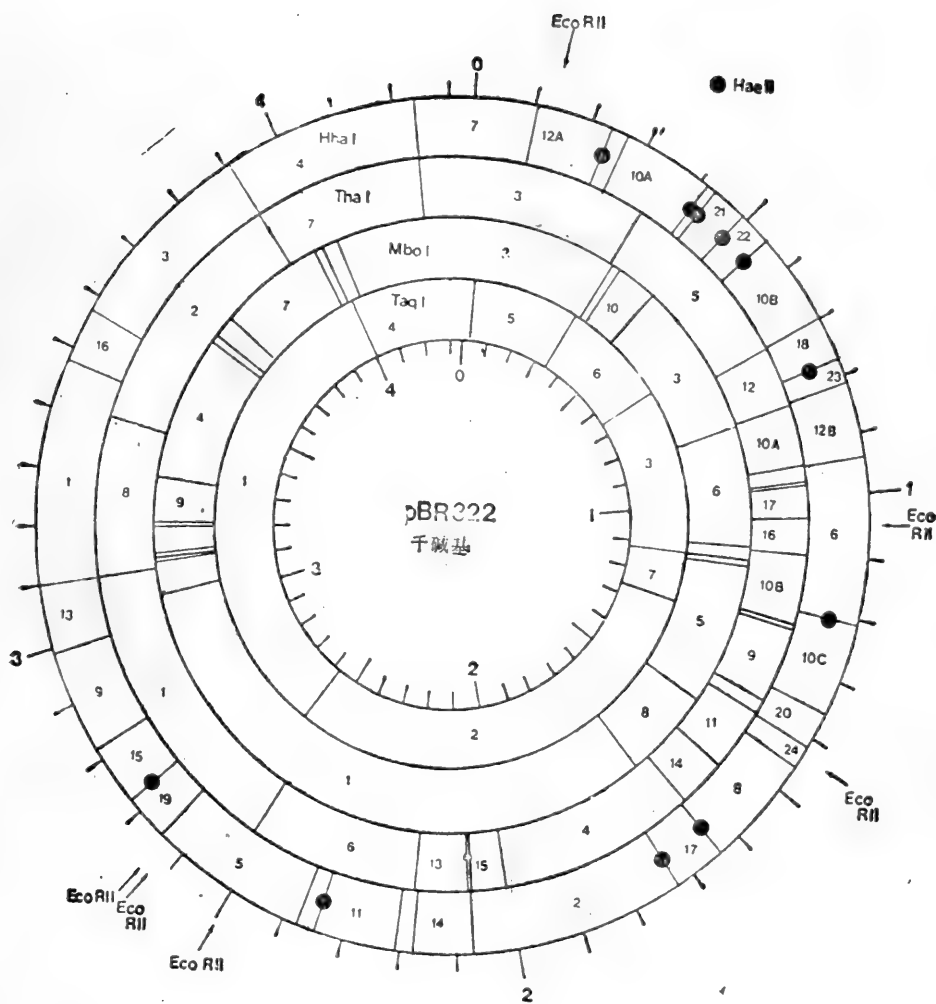
附录三 大肠杆菌遗传学中常见的基因型

基因型	表 现 型	基因型	表 现 型
ace	乙酸盐利用	min	微细胞
ade	需要腺嘌呤	mlt	甘露醇利用
ala	需要丙氨酸	nal	萘啶酮酸敏感性
arg	需要精氨酸	nir	硝酸盐还原酶活性
aro	需要芳香族氨基酸	pan	需要泛酸
asn	需要天冬酰胺	phe	需要苯丙氨酸
asp	需要天冬氨酸	pho	碱性磷酸酶活性
azi	叠氮化物抗性	phs	硫化氢产生
bio	需要生物素	pls	磷脂生物合成
cys	需要半胱氨酸	pol	DNA 聚合酶活性
cyt	需要胞嘧啶	pro	需要脯氨酸
div	细胞分裂	pts	磷酸转移酶活性
end	核酸内切酶	pur	嘌呤生物合成
fla	鞭毛生物合成	put	脯氨酸利用
gal	半乳糖利用	pyr	嘧啶生物合成
glt	需要谷氨酸	rec	遗传重组能力
gln	需要谷酰胺	rha	鼠李糖利用
glp	甘油磷酸酯利用	rif	利福平抗性
gly	需要甘氨酸	rpo	RNA 聚合酶活性
gua	需要鸟嘌呤	rpsE	链霉素抗性
his	需要组氨酸	rpsL	链霉素抗性
hlf	高频溶源化	ser	需要丝氨酸
hsd R	寄生特异性(限制性内切酶)	spc	同 rpsE
hsd M	寄生特异性(修饰酶)	str	同 rpsL
hut	组氨酸利用	thi	需要硫胺素 (B1)
ile	需要异亮氨酸	thr	需要苏氨酸
ilv	需要异亮-缬氨酸	thy	需要胸腺嘧啶
lac	乳糖利用	ton	噬菌体 T1 抗性
leu	需要亮氨酸	trp	需要色氨酸
lon	依赖于 ATP 的蛋白酶	tsx	噬菌体 T6 抗性
lys	需要赖氨酸	tyr	需要酪氨酸
mal	麦芽糖利用	ura	需要尿嘧啶
man	甘露糖利用	uvr	紫外线敏感
mel	蜜二糖利用	val	需要缬氨酸
met	需要蛋氨酸	xyl	木糖利用

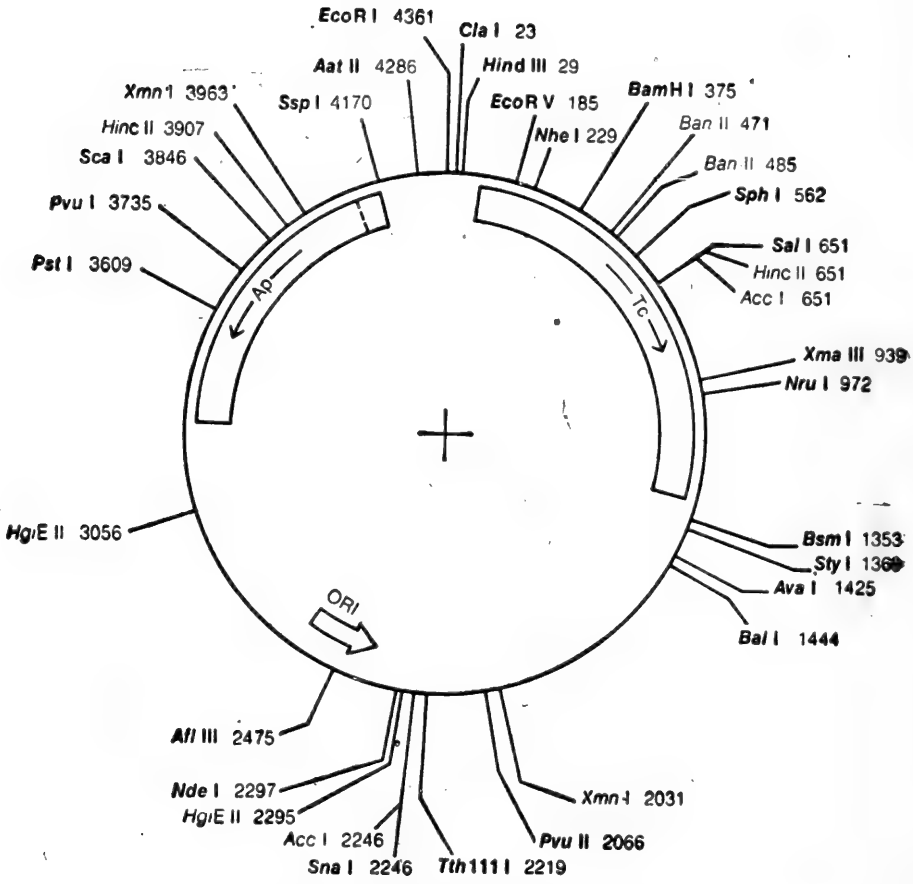
附 图

	AATT	ACGT	AGCT	ATAT	CATG	CCGG	CGCG	CTAG	GATC	GCGC	GGCC	GTAC	TATA	TCGA	TGCA	TTAA
▼□□□□									Mbo I*							
□▼□□□		Mae II				Hpa II*		Mae I		HinPI*				Taq I		
□□▼□□			Alu I				FnuD II		Dpn I		Hae III	Rsa I				
□□□▼□										Hha I						
□□□□▼					Nla III											
A▼□□□T			Hind III		Afl III		Mlu I Afl III	Spe I	Bgl II Xho II							
A□▼□□T														Cla I		
A□□▼□T				Ssp I							Stu I Hae I	Sca I				
A□□□▼T					HspCI					Hae II				Hsi I		
C▼□□□G					Nco I	Xma I Ava I		Avr II			Xma III Cfr I Gdi II			Xho I Ava I		Afl II
C□▼□□G				Nde I												
C□□▼□G			Pvu II NspB II			Sma I	NspB II									
C□□□▼G							Sac II		Pvu I							
C□□□□▼G																Pst I
G▼□□□C	EcoRI						BssH II	Nhe I	BamHI Xho II	Ban I		Ban I		Saf I		
G□▼□□C		Aha II								Nar I Aha II			Acc I	Acc I		
G□□▼□C				EcoRV		Nae I								Hinc II		Hpa I Hinc II
G□□□▼C		Aat II	Sac I Ban II HqiA I		Sph I NspCI					Bde I Hae II	Apa I Ban II	Kpn I			HglAI*	
T▼□□□A								Xba I	Bcl I		Cfr I					
T□▼□□A														Asu II		
T□□▼□A		SnaB I					Nru I			Fsp I	Bal I Hae I					Dra I
T□□□▼A												Gof II				

附图 1 识别 4 个和 6 个核苷酸序列的常用限制性内切酶。



附图3 pBR322 的限制性酶切点图(之二)。



附圖 3 pBR322 的限制性酶切点图(之三)。

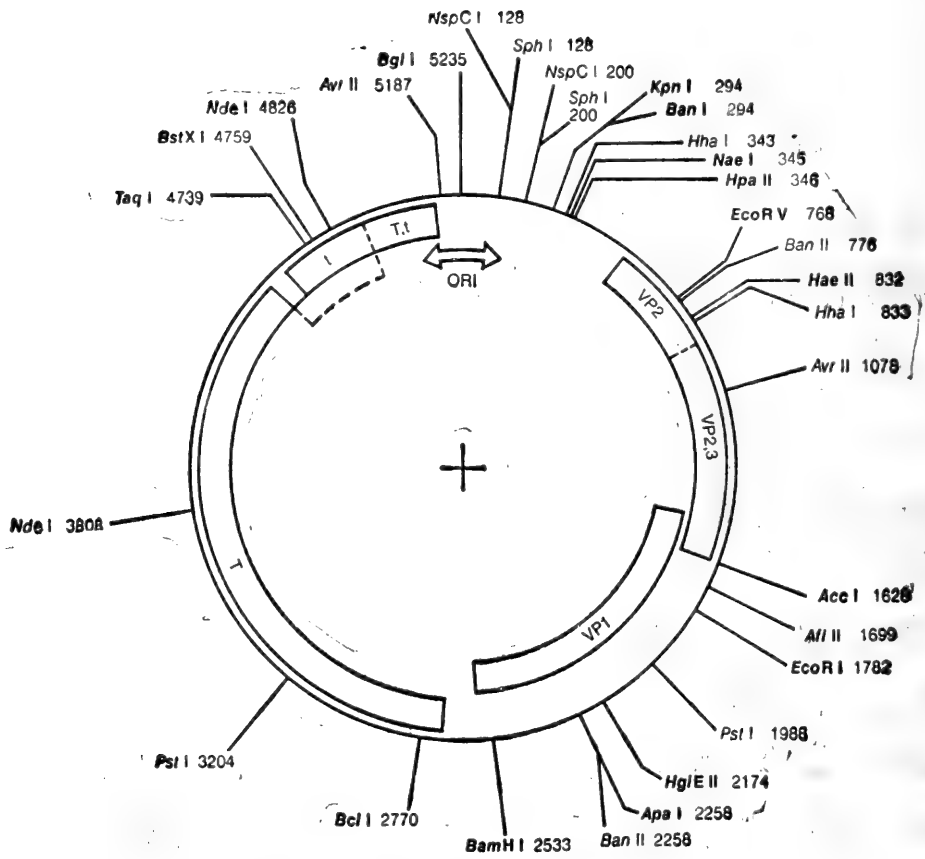
酶	切点数	定 位				酶	切点数	定 位			
Aat II	1	4286									
Afl III	1	2475									
Ava I	1	1425									
Bal I	1	1444									
BanH I	1	375									
Ban I	1	1353									
Cla I	1	23									
EcoR I	1	4361									
EcoR V	1	185									
Hind III	1	29									
Nde I	1	2297									
Nhe I	1	229									
Nru I	1	972									
Pst I	1	3609									
Pvu I	1	3735									
Pvu II	1	2066									
Sal I	1	651									
Sca I	1	3846									
Sna I	1	2246									
Sph I	1	522									
Ssp I	1	4170									
Sty I	1	1369									
TthII I	1	2239									
Xba III	1	939									
Acc I	2	651	2246								
Ban II	2	471	485								
HaeII II	2	2295	3056								
Hinc II	2	651	3907								
Xan I	2	2031	3963								
Bgl I	3	329	1163	3402							
Dra I	3	3232	3251	3943							
Rsa I	3	164	2282	3847							
Dra II	4	523	1430	1480	4343						
Eco47 III	4	232	494	775	1727						
Fsp I	4	260	1356	1454	3508						
Gau I	4	811	1983	1983	3453						
Hae I	4	401	769	929	1283						
Hae II	4	413	434	548	1205						
Har I	4	562	1816	2110	2475						
Nsp7524 I	4										
Gdi II	5	295	399	531	939	3756					
Hae I	5	230	1489	2970	3223	3558					
TthIII II	5	7	1922	3049	3082	3088					
Aha II	6	413	434	548	1205	3904					
BstNI I	6	130	1059	1442	2502	2623					
Cfr I	6	2636									
NspB II	6	295	399	831	939	1444					
		3756									
		1139	2066	2185	2815	3060					
		4001									
Cfr10 I	7	160	401	410	769	929					
		1283	3448								
Hae I	7	918	990	1047	1444	2408					
		2439	2951								
Taq I	7	24	339	652	1127	1268					
		2575	4019								
Ava II	8	799	887	1136	1439	1481					
		1760	3506	3728							
Dde I	8	1581	1743	2285	2750	3159					
		3325	3865	4291							
Hya I	8	276	587	1174	1465	2291					
		2789	3950	4035							
Xho II	8	375	1667	3116	3127	3213					
		3225	3993	4010							
Ban I	9	76	119	413	434	548					
		766	1205	1289	3316						
Bsp1288	10	276	471	485	587	1174					
		1855	2291	2789	3950	4035					
HinI I	10	532	852	1006	1304	1525					
		2031	2375	2450	2946	3383					
Hae II	10	901	957	1546	1570	1800					
		2228	3178	3594	3967	4287					
Hci I	10	170	534	1258	1484	1812					
		2120	2155	2854	3550	3901					
Hae II	11	232	413	434	494	548					
		775	1205	1644	1727	2349					
Hae I	11	2719									
		390	649	944	976	1240					
		1390	2004	2181	2577	3155					
		3905									
Hbo II	11	464	738	1809	1601	2354					
		3125	3216	3971	4849	4158					
		4354									
Sin I	12	375	376	1097	1667	3042					
		3116	3120	3213	3226	3690					
		3993	4811								
Fob I	12	112	133	987	1032	1681					
		1770	1848	2009	2150	3348					
		3529	3816								
Mph I	12	136	408	453	1307	1528					
		2085	2094	3219	3446	3842					
		4068	4083								
Sau96 I	15	172	524	799	887	1136					
		1260	1439	1481	1760	1949					
		3410	3489	3506	3728	4344					
Alu I	16	15	30	686	1089	1999					
		2056	2067	2116	2135	2416					
		2642	2778	3035	3556	3656					
		3719									
ScrF I	16	130	170	534	1059	1258					
		1442	1484	1812	2120	2155					
		2502	2623	2636	2854	3550					
		3901									
Mae III	17	125	213	881	1148	1808					
		1831	1917	2130	2225	2832					
		2895	3011	3294	3625	3683					
		3836	4024								
Bbv I	21	226	615	773	1406	1430					
		1559	1562	1685	2065	2068					
		2114	2211	2380	2398	2817					
		2882	2885	3091	3419	3608					
		3785									
Hae III	22	173	296	400	524	532					
		596	830	919	940	991					
		1048	1261	1445	1949	2489					
		2500	2518	2952	3410	3690					
		3757	4344								
Hbo I	22	349	376	467	826	1098					
		1129	1144	1461	1668	3042					
		3117	3128	3136	3214	3226					
		3331	3672	3690	3736	3994					
		4011	4047								
SfaNI I	22	134	204	247	393	405					
		658	1026	1033	1421	1673					
		1682	1769	1847	1910	2151					
		2267	2322	2343	2563	3615					
		3825	4055								
FnuDI II	23	346	782	817	946	973					
		1078	1039	1105	1234	1244					
		1389	1415	1537	1634	2006					
		2075	2077	2180	2521	3102					
		3432	3925	4257							
Nla IV	24	76	119	329	378	413					
		434	524	548	766	887					
		1205	1254	1289	1324	1438					
		1481	1760	2505	2544	3316					
		3410	3451	3662	4252						
Hpa II	26	161	170	387	402	411					
		534	694	770	930	1020					
		1258	1284	1485	1665	1812					
		2121	2155	2682	2829	2855					
		3045	3449	3483	3550	3660					
		3902									
Hnl I	26	115	175	379	598	797					
		865	981	1157	1228	1266					
		1293	1479	1793	1851	1987					
		2072	2102	2364	2590	2647					
		2914	3314	3395	3525	3731					
		4342									
Nla III	26	5	352	490	563	709					
		748	934	1051	1179	1194					
		1251	1458	1592	1817	1881					
		1946	2111	2216	2476	3196					
		3687	3697	3775	3811	4204					
		4309									
Hba I	31	101	233	261	414	435					
		495	549	791	776	816					
		947	1206	1357	1419	1455					
		1645	1728	2076	2179	2209					
		2350	2383	2653	2720	2820					
		2994	3103	3496	3589	3926					
		4258									
FnuHI I	42	226	297	300	578	581					
		615	722	773	938	1023					
		1106	1163	1208	1287	1406					
		1409	1416	1430	1559	1562					
		1685	1766	2065	2068	2114					
		2211	2264	2380	2398	2401					
		2519	2674	281							

TTCTCATGTTTACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTATCACAGTTAAATGCTAACGCAGTCAGGCACCCGTGTATGAAATCTAACAAT
 100
 AAGAGTACAAACTGTCGAATAGTAGCTATTGCGAAATTACGCCATCAAAATAGTGTCAATTTAACGATTGGCTCAGTCCGTGGCACATACTTTAGATTGTTA
 200
 GCGCTCATCGTCATCCTCGGCACCGTCACCCGTGATGCTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCTCTTCGGGGATATCGTCCATTCCG
 CCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGCATCCGTATCCGAAGCAATACGGCCATGACGGCCCGGAGAACCCCTATAGCAGGTAAGGC
 300
 ACAGCATCGCAGTCACTATGGCGTGTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCCGACCCGTTCTCGGAGCACTGTCCGACCCGTTTGGCCG
 TGTGCTAGCGGTGAGTATACCGCAGCAGCATCGCGATATACGTAACACGTTAAAGATACGGGTGCCGAAGAGCCGCTGACAGGCTGGCGAAACCGGC
 400
 CCGCCAGTCTGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGGCATCATGGCGACCACCCGCTCTGTGGATCTCTACGCCGGACGCATCGTG
 GCGGGTACGAGCAGCGAAGCGATTGAACACCGGTGATAGCTGATGCGCTAGTACCCTGTTGGCGAGGACACTAGGAGATGCGGCTCGGTAGCAC
 500
 GCGCGCATCACCGGCCACAGGTGCGGTTGCTGGCGCTATATCGCGCATACCCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATAGCGCTT
 CCGCGTAGTGGCCCGGTTCCACGCCAACGACCGCGGATATAGCGGCTGATGGCTACCCTTACGCCGAGCGGTGAACCCGAGTACTCGCGAA
 600
 GTTTCGGCGTGGGTATGGTGGCAGGCCGTTGGCGGGGACTGTTGGCGCCATCTCCTTGCATGCACATTCTTGGCGGGCGGTGCTCAACGGCCCTC
 CAAAGCCGACCCATACCACCTCCGGGACCGGCCCTGACAACCCGCGTAGAGGAACGTACGTGTAAGGAACGCCCGCCACGAGTTGCCGGAG
 700
 AACCTACTACTGGGCTGCTTCTAATGCAGGAGTGCATAAAGGAGAGCGTGCACCGATGCCCTTGAGAGCCTCAACCCAGTACGCTCTTCCGGTGGG
 TTGGATGATGACCCGAGGAAGGATTACGCTCCTCAGCTATTCCCTCTCGCAGCTGGTACGGAACTCTCGGAAGTTGGGTAGTCAGGGAAGGCCACC
 800
 CCGGGGATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCAITTTCCGGCAGGA
 GCGCCCGTACTGATAGCAGCGCGTAACTAGCAGAGAAGATAGTACGTTGAGCATCCTGTCCACGGCCGTGCGGAGACCCAGTAAAAACCGGCTCT
 900
 CCGCTTTCGCTGGAGCGCAGCATGATCGGCTGTGCTTGCCTATTCCGAATTTGACGCCCTCGCTCAAGCCTTCGTCAGTGTCCGCCACCAAA
 GCGAAAGCGACCTCGCGCTGCTACTAGCCGACAGCGAACCCATAAGCCTTAGAACGTGCGGGAGCGAGTTCGGAAGCAGTACACAGGGCGGTGGTTT
 1000
 CGTTTCGGCGAGAAGCAGGCCATTATCGCCGCATGGCGGCGACGCGCTGGGCTACGTTGCTGGCGTTCCGACGCGGAGGCTGGATGGCTTCCCA
 GCAAAGCCGCTTTCGTCGGTAATAGCGCCGATACCBCGGCTGCGCGACCCGATGCAGAACGACCGCAAGCGCTCGCTCCGACCTACCGGAAGGGGT
 1100
 TTATGATCTTTCGCTTCCGGCGGCATCGGGATGCCCGCTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGACACCTTCAAGGATC
 AATACTAAGAAGAGCGAAGCCGCGTAGCCCTACGGGGCAACGCTCGGTACGACAGGTCCGCTCATCTACTGCTGGTAGTCCCTGTGGAAGTCTCTAG
 1200
 GCTCGCGGCTTACCAGCCTAAGTTCGATCACTGGACCGGTGATCGTACGGCGATTATGCCGCTCGGGAGCACATGGAACGGGTGGCATGGATT
 CGAGCCCGGAGAAATGGTGGATTGAAGCTAGTGACCTGGCGACTAGCAGTCCGCTAAATACGGCGGAGCCGCTCGTGACCTTCCCAACCGTACCTAA
 1300
 GTAGGCGCGCCCTATACCTTGTCTGCCCTCCCGCGTGTGCTGCGCGTGCATGGAGCGGGCCACCTCGACCTGAATGGAAGCCGGGCGACCTCGCTAA
 CATCCGCGCGGGATATGGAACAGAGCGAGGGGCGCAACGACGCGCCACGTACTCGGCCGGTGGAGCTGGACTTACCTCGGCCGCGTGGAGCGATA
 1400
 CCGATTACCACTCCAAGAATTGGAGCAATCAATTTGCGGAGAACTGTGAATGCGCAAAACCAACCTTGGCAGAACAATCCATCGCGTCCGCGCATC
 GCCTAAGTGGTAGGTTCTTAACTCGGTAGTTAAGAAGCCCTCTTGACACTTACCGGTTGGTGGGAACCGTCTTGTATAGTAGCGCAGGCGGTAG
 1500
 TCCAGCAGCCGACGCGCGCATCTCGGGCAGCGTGGGCTTGGCCACGGGTGCGCATGATGCTGCTCTGCTGTTGAGGACCCGCTAGGCTGGCGGG
 AAGTGTGCGCGTGGCCCGGTAGAGCCGTCGAACCCAGGACCGGTGCCACGCGTACTAGCACGAGGACAGCAACTCTGGGCGGATCCGACCGCCC
 1600
 GTTCCCTTACTGGTATGAGAATGAATACCGATACCGAGGCAACGTGAAGCGACTGCTGCTGCAAAACGCTCTGCGACTGAGCAACAACATGAATGGT
 CAACGGAATGACCAATCGTCTACTTAGTGGCTATGCGCTGCTTGCACCTGCTGACGACGACGTTTTGACAGCGTGGACTCGTGTGTTACTTACCA
 1700
 CTTCCGTTCCCGTGTTCGTAAGCTCGAAACCGGAAAGTCAGCGCCCTGACCAATTATGTTCCGGATCGCATCGCAGGATGCTGCTGGCTACCCTGT
 GAAGCCAAAGGCACAAAGCATTTCAGACCTTTGCGCTTCACTGCGGGACGTGGATATAAAGGCCAGACGTAGCGTCCCTACGACGACCGATGGGACA
 1800
 GGAACACCTACATCTGTATTAACGAAGCGTGGCATGACCCCTGAGTATTTTTCTGCTGGTCCGCGCATCCATACCBCAGTGTGTTACCCTCACAA
 CCTTGTGGATGAGACATAATTGCTTCGCGACCGTAACCTGGGACTACTAAAAGAGACAGGCGCGGTAGGATGCGGCTCAACAAATGGGAGTGTG
 1900
 GTTCCAGTAACCGGCATGTTATCATCAGTAACCCGATCGTGAGCATCTCTCTCGTTTATCGGATATCATTACCCCATGAACAGAAATCCCCCTT
 CAAGGTCAATGGCCCGTACAAGTAGTAGTATTGGGATAGCACTCTAGGAGAGAGCAAAAGTAGCCATAGTAATGGGGTACTTGTCTTAAAGGGGAA
 2000
 ACACGGAGGATCAAGTGACCAAAACAGGAAAAACGCCCTTAACTGGCCGCTTATCAGAAGCCAGACATTAACGCTTCTGGAGAACTCAACGAGC
 TGTGCCCTCGTAGTCTACTGGTTGCTCTTTTGGCGGAAATTGACGGGCGAAATAGTCTTCGGTCTGTAATGCGAAGCACTCTTGGATGCTCG
 2100
 TGGACCGGGATGAACAGGACAGACATCTGTGAATCGCTTACGACCCAGCTGATGAGCTTACCAGCTGCTCGCGCTTTCGGTGTGACGGTGA
 ACCTGCGCTACTTGTCCGCTGTAGACACTTAGCGAAGTGTGTCGCACTACTCGAAATGGCGTGCAGGAGCGCGCAAGCCACTACTGCCATTT
 2200
 CCTGTGACACATGACAGTCCCGGAGACGGTCAAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGGTCAGCGGGTGTGGCGGG
 GGAGACTGTGACGTGAGGGCCCTGCCAGTGTGCAACAGACATTGCGCTACGGCCCTCTGTTTCGGGAGTCCCGCAGTCCGCCACACCGCC
 2300
 TGTGGGGCCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTACTGGCTTAACTATGGGATCAGAGCAGATTGACTGAGAGTGCACCATAT
 CAGCCCGCGCTCGTACTGGTCACTGCTATCGCTACCACTGACCGAATTGATACCGCGTAGTCTGCTAACATGACTCTCAGCTGGTATA
 2400

附图 5 pBR322 的核苷酸序列。

GCGGTGTAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTCGGCTCGGTCTCGGCTG
 CGCCACACTTTATGGCGTGTCTACCGATTCTCTTTTATGGCGTAGTCCGGGAGAAGGGCAAGGAGCGAGTACTGAGGCGACGGAGCCACCAAGCCGAC
 2500
CGGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACCGAGAAAGAATGAGGAGCAAAAGGCCAGCAAAAGG
 CGCGCTCGCCATAGTCGAGTGAGTTCCGCGATTATGCCAATAGGTGCTTATAGTCCCTATTGCGTCTTCTTGACTACTCGTTTTCCGGTCTGTTTTCC
 2600
CCAGGAACCGTAAAAAGCGCGGTTGCTGCGGTTTTTCCATAGGCTCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
 GGTCTTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGACTGCTCGTAGTGTTTTAGCTGCGAGTTCACTCTCCACCGCTTT
 2700
CCCGACGAGACTATAAAGATACCAGGCGTTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACC GGATACCTGTCCGCTTT
 GGGCTGCTCGATATTTCTATGGTCCGCAAAAGGGGACCTTCGAGGGAGCACCGGAGAGGCAAGGCTGGGACGGCGAATGGCCTATGGACAGGGGAAA
 2800
CTCCCTTCGGGAGAGCTGGCGTTTTCTCAATGCTCAGCGTGTAGGTATCTCAGITCGGTGAGGTGCTTCCGCTCAAGCTGGGCTGTGTGCGACGAAACCC
 GAGGAAAGCCCTTCGACCCGCAAAAGAGTTACGAGTGGACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCCAGCCGACACAGGTGCTGGGG
 2900
CGGTTCAGCCGACCGCGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCGGTAAAGACAGCACTTATCGCCACTGGCAGCAGCCACTGTAACAG
 GCAAGTCCGGCTGGCGACCGGAATAGGCCATTGATAGCAGAAGTCAAGTTGGGCCATTCTGTGTGAATAGCGGTGACCGTCTCGGTGACCATGTCT
 3000
CATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTCGGCTCTG
 GTAATCGTCTCGCTCCATACATCCGCCAGGATGTCTCAAGAACTTACCACCGGATGATGCGGATGTGATCTTCTGTCTATAAACCATAGACCGGAGAC
 3100
CTGAAAGCGATTACCTTCGAAAAAGAGTGTGAGCTTGTATCCGGCAAAACAAACCACCGCTGTAAGCGGTGTTTTTTTGTGTCAGCAGCAGATTA
 GACTTCGGTCAATGGAAGCCTTTTCTCAACCATCGAGAAGTAGGCCGTTTGTGTGGGAGCATTGCCACCAAAAACAAACGTTCTGCTCTAAT
 3200
CGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGAGCTCAGTGAACGAAAACTCAGTTAAAGGATTTTGGTATGAG
 CGCGCTTTTTTCTAGAGTCTTCTAGGAACTAGAAAAGATGCCCGAGCTGCGAGTCACTTGTCTTTGAGTGAATTCCTTAAACCAGTACT
 3300
ATTATAAAAAGGATCTTCACTAGATCCTTTTAAATAAAAATGAAGTTTAAATCAATCAAAGTATATAGAGTAACTTGGTCTGACAGTTACCAA
 TAATAGTTTTTCTAGAACTGGACTAGGAAAATTAATTTTACTTCAAATTTAGTTAGATTTCATATATATCATTTGAAACAGACTGTCAATGGTT
 3400
TGCTTAATCAGTAGGACACCTATCTCAGGATCTGCTATTTCTGTTTATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGGAGGGCT
 ACGAATAGTCACTCCGTGGATAGAGTCGTAGACAGATAAAGCAAGTGGTATCAACGGACTGAGGGGACGACAGACTTATGATGCTATGCCCTCCCGA
 3500
TACCATCTGGCCCAAGTGTGCAATGATACCGCGAGACCCAGCTCACC GGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGAG
 ATGGTAGCCGGGTCAGCAGCTTACTATGGCGCTCTGGTGGCAGTGGCCGAGGTTCAAATAGTCTTATTTGGTGGTCCGCTTCCCGCTCGGCTC
 3600
AAGTGGTCTGCACTTTATCCGCJCCATCCAGTCTTAATTTGTTGCCGGGAAGCTAGAGTAAAGTGTGCGCAGTTAATAGTTTGGCGAACGTTGTT
 TTCACGAGGAGCTTGAATAGGCGGAGGTAGGTCAGATAATTAACAACCGCCCTTCGATCTCATCAAGCGGTCAATTATCAAACGGTTGCAACAA
 3700
GCCATTGCTGACGGCATCGTGGTGTACGCTCGTTCGTTGGTATGGCTTATTAGCTCCGGTTCACACGATCAAGGGAGTTACATGATCCCCATGT
 CGGTAACGAGCTCCGTAGCACCACTGCGAGCAGCAAAACATACGGAAGTAACTCGAGGCCAAAGGGTTGCTAGTCCGCTCAATGACTAGGGGTACA
 3800
TGTGCAAAAAAGCGGTTAGCTCCTTCGCTCCGATCGTTGTGAGAAGTAAAGTGGCCGAGTGTATCACATCGTTATGGCAGCACTGCATAATTC
 ACAGCTTTTTTCGCAATCGAGGAAGCCAGGAGGCTAGCAACAGCTTCAATCAACCGCGTCAACAATAGTGAATACCAATACCGCTCGTACGATTAAG
 3900
TCTACTGTATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCACTGAGAAATAGTGTATGCGGGCAGCCAGTTGCTCTTGC
 AGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCCTCATGAGTTGGTTCAGTAAGACTCTATAGTGATACCGCGCTGGCTCAACGAGAAGC
 4000
CCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTAAAGTGGCTCATCATTGAAAACGTTCTTCGGGGCAAAACTCTCAAGGATCTTAC
 GCGCGAGTTGTCGCTATTATGGCGGGTGTATCGTCTGAAATTTTACAGGATGTAACCTTTTGAAGAAGCCCGCTTTTGAGAGTTCTAGAAATG
 4100
CGCTGTGAGATCCAGTTCGATGAACCCACTCGTGCACCAACTGATCTTCAGCATCTTTACTTCCAGCGGTTTCTGGGTGAGCAAAAACAGGAAAG
 GCGACAACCTTAGGTCAAGCTACATGGGTGAGCAGTGGGTGACTAGAAGTCGTAGAAAATGAAAGTGGTCCAAAGACCCACTCGTTTTGCTCTTC
 4200
GCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCACTCTTCCTTTTCAATATTAAAGCATTATCAGGGTATTGTT
 CGTTTTACGGCGTTTTTCCCTTATCCGCTGTGCTTTACAACCTATGAGTATGAGAAGGAAAAGTTATAATAACTTCGTAATAGTCCCAATAACA
 4300
CTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGCTCAAGAAACCA
 GAGTACTCGCTATGTATAAATACATAAATCTTTTATTTGTTATCCCAAGGCGGTGTAAGGGGCTTTTCCAGGTGGACTCGAGATCTTTGGT
 4360
TTATTATCATGACATTACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTTTCAAG AA
ATAATAGTACTGTAATGGATATTTTATCCGCATAGTGTCTCCGGAAAAGCAGAAGTTC TT4362

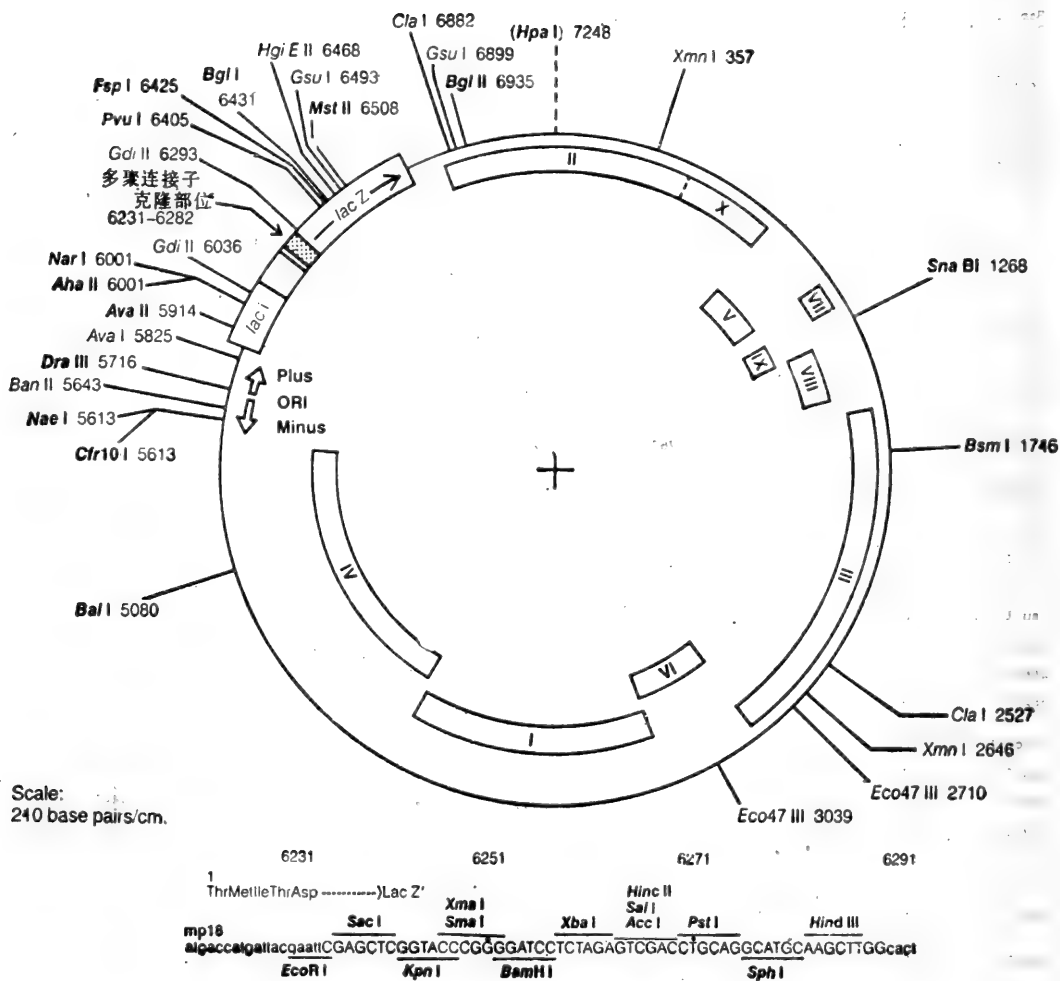
附图 5 (续图)



附图 6 SV40 DNA 的限制性酶切点图。

酶	切点数	定 位				酶	切点数	定 位					
Acc I	1	1628				Mbo II	16	461	1106	1501	2851	2861	
Afl II	1	1699					2875	2905	3661	4036	4417		
Apa I	1	2258					4430	4495	4904	4973	5000		
BamH I	1	2533					5015						
Ban I	1	294				Nla IV	16	156	228	294	387	777	
Bcl I	1	2770					1017	1404	1507	1561	1975		
Bgl I	1	5235					2131	2171	2258	2533	2797		
BstX I	1	4759					4696						
Cfr10 I	1	345				BstW I	17	160	177	232	358	910	
EcoR I	1	1782					964	1408	1535	2528	2897		
EcoR V	1	768					3146	3247	3335	4008	4069		
Eco47 II	2	832					4892	5092					
Esp I	1	1710				Nla III	17	38	129	201	334	384	
Hae II	1	832					561	1913	2542	2774	2840		
HgiE II	1	2174					2856	3234	3582	3866	4052		
Hpa II	1	346					4373	4797					
Kpn I	1	294				SerF I	17	160	177	232	358	910	
Nae I	1	345					964	1408	1535	2528	2897		
Sfi I	1	5234					3146	3247	3335	4008	4069		
Taq I	1	4739					4892	5092					
Avr II	2	1078	5187			Hae III	19	6	331	361	660	709	
Ban II	2	776	2258				742	756	935	1235	1462		
Hha I	2	343	833				1507	2259	2799	3172	3201		
Nde I	2	3808	4826				4862	5191	5234	5243			
Nsp7524 I	2	128	200			Dde I	20	287	2722	1337	1436	1641	
Pst I	2	1988	3204				1711	1857	2375	2793	2846		
Sph I	2	128	200				2929	3057	3290	4387	4499		
							4638	4808	4858	4918	5228		
Dra II	3	587	2258	2797		Bbv I	22	340	442	568	613	616	
Nco I	3	37	333	560			619	649	664	667	733		
Nsi I	3	126	198	3583			784	802	1541	1715	1987		
Pvu II	3	270	1716	3506			2461	2652	3505	3508	3844		
Xho II	3	2533	4099	4769			4312	5127					
Ssm I	4	1891	2583	2682	4528	Fnu4H I	24	340	442	568	613	616	
Bsp1286	4	776	1288	2258	3385		619	649	661	664	667		
Hpa I	4	499	519	2666	3733		733	784	802	1541	1715		
Hph I	4	1085	1245	1381	3237		1987	2461	2652	3505	3508		
Nsp6 II	4	270	662	1716	3506		3844	4312	5127	5241			
Jva II	6	557	588	1018	2013	3538	Alu I	36	271	406	435	612	639
		5118						792	804	1047	1097	1350	
Bin I	6	873	2138	2533	2534	4099		1494	1543	1581	1634	1709	
Gau I	6	653	965	1623	1743	3380		1717	1840	1986	2272	2279	
		3647						2523	2651	3426	3436	3477	
Hind III	6	1046	1493	1708	3476	4002		3507	3788	4003	4160	4310	
		5171						4643	5126	5172	5226		
SfaN I	6	125	197	688	2621	4881	Hnl I	51	2	8	206	310	319
		4925						363	658	719	724	744	
Ssp I	6	1303	1823	3256	4127	4280		1088	1150	1341	1402	1435	
		5083						1596	1643	1664	1805	1859	
Hinc II	7	470	499	519	2057	2297		2198	2440	2449	2481	2492	
		2666	3733					2495	2510	2519	2711	2720	
Stu I	7	360	708	741	755	1234		2743	2788	2795	3379	3388	
		1461	5190					4028	4082	4442	4472	4517	
Mbo I	8	874	2138	2534	2771	3716		4807	4917	4937	5024	5079	
		4100	4710	4770				5139	5193	5203	5208	5232	
Sty I	8	37	333	560	1078	2812		3238					
		3455	4409	5187									
Hinf I	10	1739	2824	2848	3373	3610							
		4376	4459	4568	4592	5135							
Fok I	11	80	1442	2455	3311	3395							
		3654	3702	3945	4804	4926							
Hae I	11	5036											
		330	360	708	741	755							
		934	1234	1461	3200	4861							
Sau96 I	11	5190											
		557	908	1018	1507	2013							
		2258	2259	2798	3171	3538							
		5118											
Tth111 II	11	423	3497	3627	3640	3689							
		3728	3950	4345	4582	4822							
		4868											
Dra I	12	1607	1798	2363	2727	3157							
		3472	3543	3589	3907	4646							
		5057	5147										
Hae I	12	1079	1459	1610	1779	2160							
		2580	3434	3748	4105	4483							
		5112	5188										
Rsa I	12	295	970	1467	3072	3225							
		3576	3870	3927	3942	4168							
		4876	4987										
Huv III	14	352	820	866	1333	1477							
		1844	1871	2036	2424	2639							
		3238	3997	4273	4678								

附图7 SV40 DNA 限制性酶切点的定位。



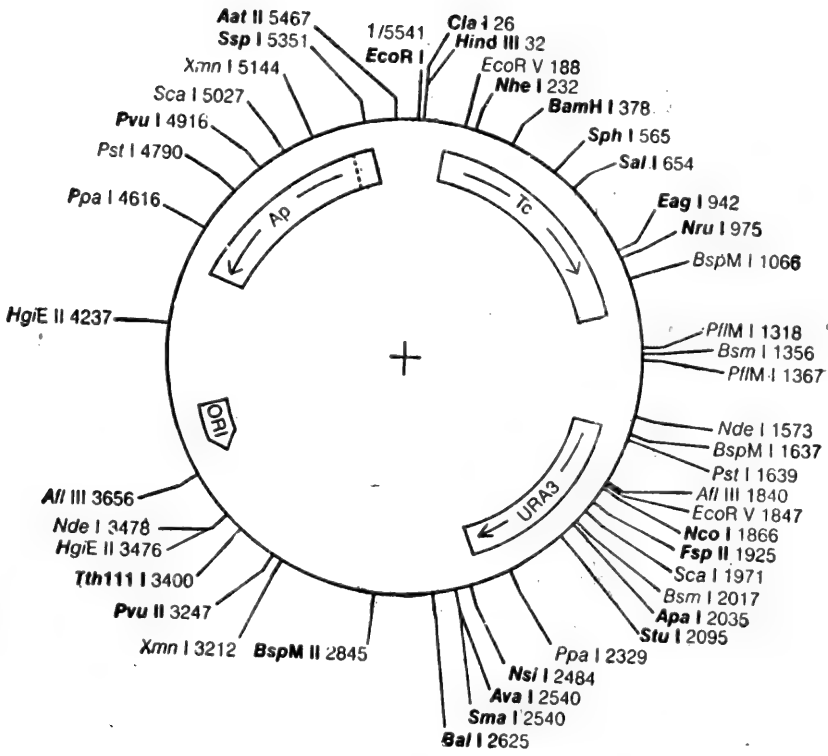
附图 8 M13mp18 DNA 限制性酶切点图。

酶	切点数	定 位				酶	切点数	定 位				
Acc I	1	6264										
Aha II	1	6001										
Ava II	1	5914			Nla III	16	149	196	1107	1300	1799	
Bal I	1	5080					2034	2157	2855	3532	3619	
BamH I	7	6252					3718	5179	6222	6277	6857	
Bcl I	1	6431					6965					
Bcl II	1	6935			Fnu4H I	17	931	1367	1394	2285	2280	
Bsu I	1	1746					2312	2357	2521	3132	4871	
Cfr10 I	1	5613					5500	5514	5536	5923	6052	
Dra III	1	5716					6355	6428				
EcoR I	1	6231			FnuD II	18	43	347	1119	1176	2466	
Fsp I	1	6425					2355	3409	3599	3952	4313	
Hae II	1	6468					4994	5489	5513	5533	6029	
Hinc II	1	6264					6031	6689	6752			
Hind III	1	6282			Hpa II	18	314	965	1095	1924	2378	
Kpn I	1	6243					2396	2552	3370	3842	4018	
Mst II	1	6508					5614	6159	6248	6463	6481	
Nae I	1	5613					6838	6961	7021			
Nar I	1	6001			Hph I	18	1376	1503	1774	1909	2398	
Pat I	1	6270					2542	2581	2620	2626	2635	
Pvu I	1	6405					4847	4923	5117	5706	5947	
Sac I	1	6237					5980	7005	7031			
Sal I	1	6204			Nla IV	18	1062	1249	1541	1551	1803	
Sma I	1	6249					2051	2374	2392	5644	5656	
SnaB I	1	1268					5677	5867	6001	6131	6243	
Sph I	1	6276					6252	6465	6477			
Xba I	1	6258			Rea I	19	173	280	1022	1165	1769	
							1796	1889	1905	1970	2133	
Ava I	2	5825	6247				3467	3668	4190	4380	5384	
Ban II	2	5643	6237				5486	6244	6843	7165		
Cla I	2	2527	6882									
Eco47 III	2	2710	3039		Mae II	22	770	789	942	1269	1577	
Gcl I	2	6036	6293				1688	2546	2818	3151	4404	
Gsu I	2	6493	6899				4447	4635	5038	5456	5463	
Xan I	2	357	2646				5607	5717	5760	5772	6309	
							6578	6774				
Afl III	3	195	3618	3717								
Cfr I	3	5080	6036	6293	Mae III	24	145	939	1103	1377	1718	
HqIA I	3	4743	5465	6237			1775	1790	2095	2543	2621	
Nde I	3	2723	3803	6846			2744	2910	3211	4322	4385	
Pvu II	3	5960	6053	6375			4610	5111	5529	5541	6314	
Xho II	3	2220	6252	6935			6334	6554	6580	6638		
					Hba I	26	44	1011	1085	1177	1470	
Bln I	4	1382	2221	6252	6253		2195	2467	2711	3040	3096	
Fok I	4	239	3547	6361	7244		3408	3598	4312	4995	5490	
Hae I	4	2244	5080	5344	6681		5503	5512	5534	5560	5568	
Hcl I	4	1924	6247	6248	6838		6002	6030	6095	6426	6447	
HspB II	4	1631	5960	6053	6375		6562					
Sau96 I	4	5724	5914	5938	6396	Hinf I	27	136	216	490	511	723
							2011	2497	2845	3258	3418	
Bsp1286	5	2088	4743	5465	5643	6237	3742	3838	4072	4117	4349	
Dra I	5	189	472	4622	6784	7074	5120	5329	5375	5438	5766	
Mae I	5	3827	5565	6259	6861	6979	5788	6041	6262	6624	6885	
							6904	7041				
Hae II	6	2710	3039	5559	5567	6001						
		6446										
Hsp7526 I	6	195	3531	3618	3717	6276						
		6856										
Vsp I	6	499	2660	5023	5213	6767						
		6788										
Aan I	7	1242	5677	6001	6131	6243						
		6465	6477									
BotM I	7	1014	1966	5941	5998	6137						
		6328	6455									
Hya I	7	526	2164	2479	3237	4083						
		5158	6687									
Mbo I	7	1382	1714	2221	6253	6406						
		6502	6936									
EfaM I	7	25	388	1354	3979	4850						
		6546	6559									
Tth111 II	7	3395	3446	4257	4506	4671						
		5878	6893									
Gbv I	10	931	1367	2521	3132	4871						
		5536	5923	6052	6355	6428						
ScrF I	11	1014	1924	1966	5941	5998						
		6137	6247	6248	6328	6455						
		6838										
Mbo II	12	781	977	2218	3912	4075						
		4271	4937	5255	5587	6391						
		6504	6805									
Taq I	13	336	975	1127	1508	1949						
		2528	3455	3694	4665	5683						
		6235	6265	6883								
Mae III	15	1396	2245	2554	5081	5239						
		5345	5414	5725	5939	6037						
		6294	6396	6513	6682	7023						

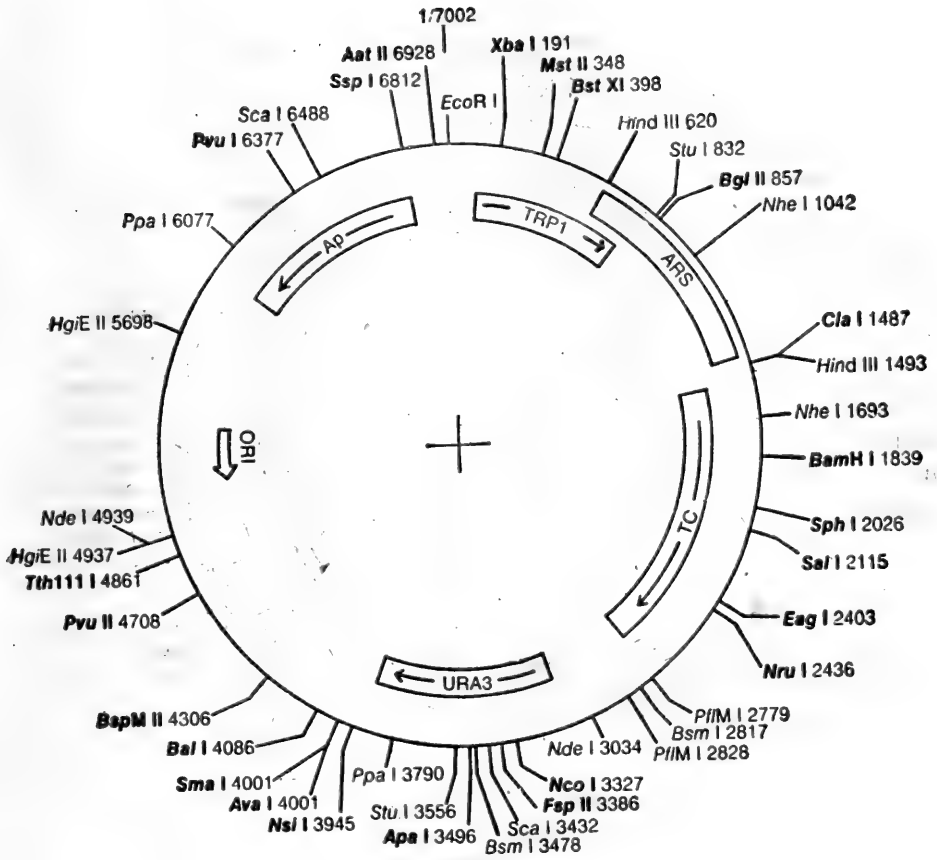
附图9 M13mp 18 DNA 限制性酶切点的定位。

酶	切点数	定 位			酶	切点数	定 位						
Apa I	1	10086			Fsp I	15	463	2503	4270	5155	6979		
Nae I	1	20040					11563	11690	13355	16046	21805		
Nar I	1	45679			Pvu II	15	21826	27949	32683	34821	42380		
Nhe I	1	34679					209	1917	2385	2526	3058		
SnaB I	1	12189					3637	7831	12099	12162	16078		
Xba I	1	24508					19716	20059	20695	22991	27412		
Xho I	1	33498			Nal I	18	1326	2206	3260	4193	6496		
Avr II	2	24322	24396				6877	7584	7978	8056	8859		
Eco47 III	2	20995	37057				10609	10777	13934	14903	21260		
Kpn I	2	17053	18556				26623	28618	36040				
Mst II	2	26717	34318		Afl III	20	458	628	5548	11281	15372		
Sac I	2	24772	25877				17791	18284	19996	20952	22220		
Sal I	2	32745	33244				24133	24168	26528	32764	39395		
Tth III	2	11202	36120				42086	42363	43762	44501	46982		
Xba III	2	19944	36654		Sep I	20	8469	24075	24746	25257	25624		
Afl II	3	6540	12618	42630			26313	29154	29364	29514	33570		
Dra II	3	2815	28797	48473			35001	38974	41905	42825	43263		
Pvu I	3	11933	26254	35787	EcoR V	21	650	2084	6681	8084	8822		
Sma I	3	19397	31617	39888			13435	14023	17767	18385	21269		
Sna I	3	15260	18834	19473			22948	26821	28198	28211	33587		
Nco I	4	19329	23901	27868	44248		39352	41273	41541	41576	42231		
Sac II	4	20320	28530	21606	40386	Gdi II	21	2739	5601	6008	8366	10580	
BamH I	5	5505	22346	27972	34499	41732	13481	14575	16416	18547	19284		
EcoR I	5	21226	26104	31747	39168	44972	19332	19944	20239	20323	20928		
Nru I	5	4590	28050	31703	32407	41808	20988	22025	35465	36654	39458		
Rsr II	5	3000	6041	13983	19288	22242	45214						
Sca I	5	16421	18684	25685	27263	32802	415	1606	2531	5505	6422		
Bgl II	6	415	22425	35711	38103	38754	22346	22425	24511	27027	27972		
BesH II	6	38814		4126	5627	14815	29593	30426	34499	35711	38103		
		3522					38664	38754	38814	39576	41732		
		28008					47773						
Esp I	6	10297	10682	11661	16518	20744							
Sph I	6	39450											
Spe I	6	2212	12002	23942	24371	27374	Xmi I	24	33	1151	2319	8490	10111
Stu I	6	39418											
		12434	31478	32997	39992	40596							
		40614											
Asu II	7	18048	25884	27980	29150	30396	Ban I	25	1180	1365	2331	5407	5663
EcoR II	7	34331	42637										
Hind III	7	581	10086	19763	21570	24772							
		25877	39453										
		23130	25157	27479	36895	37459							
		37584	44141										
Nlu I	7	458	5548	15372	17791	19996	Gsu I	25	419	1531	5021	5872	6810
Mde I	7	20952	22220										
		27630	29883	33679	36112	36668							
		38357	40131										
Avs I	8	4720	19397	20999	27887	31617	HgiA I	28	5619	6002	9485	10295	11950
Bcl I	8	33498	38214	39888									
		8844	9361	13820	32729	37352							
		43682	46366	47942									
Aac I	9	2190	15260	18834	19473	31301							
		32745	33244	40201	42921								
Ant II	10	5105	9394	11243	14974	29036							
Dra III	10	40806	41113	42247	45563	45592							
		2954	5613	6635	8999	14477							
		30365	31909	41479	47312	48434							
Sty I	10	19329	21211	23901	24322	24396							
		27868	28793	35016	36505	44248							
BstE II	13	5687	7058	8322	9024	13348							
		13572	13689	16012	17941	25183							
		30005	36374	40049									
BstX I	13	2855	6706	8413	8850	10915							
		13263	14338	18029	19741	21622							
		34596	38292	46434									
		90	8460	16294									
Dra I	13	25436	26132	26665	32703	36302							
		36530	38833	47429									
Hae I	13	23947	24323	24397	24589	24541							
		25119	26124	34680	35037	37255							
		44654	47335	47706									
HaeE II	14	1785	2250	5903	6555	12513							
		13954	15877	17433	20244	26435							
		35595	35639	37999	42048								
Hpa I	14	732	5267	5708	7948	8199							
		11583	14991	21902	27316	31807							
		32217	35259	39606	39834								
Nsi I	14	10325	27206	27372	28432	30342							
		30989	32967	33682	34208	35868							
		36665	36671	37769	38307								
Cla I	15	4198	15583	16120	26616	30289							
		11990	32962	33584	34696	35050							
		36969	41363	42020	43824	46438							

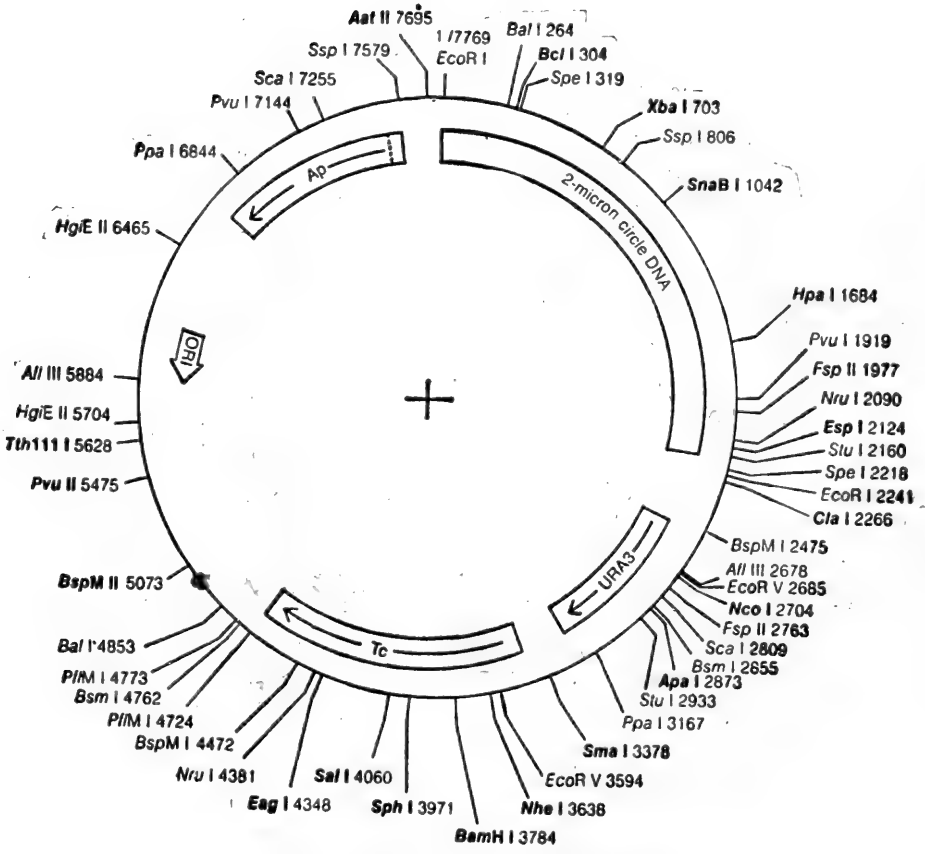
附图 11 λ DNA 限制性酶切点的定位。



附图 12 YIP5 DNA 的限制性酶切点图。



附图 13 YRP17 DNA 的限制性酶切点图。



附图 14 YEP24 DNA 的限制性酶切点图。

酶	切点数	定	位	酶	切点数	定	位
Acl I	1	7695		Xho II	9	3177 6622	3784 5076 6525 8536 6634 7402 7419
Apo I	1	2873		BstN I	10	267 4468	1784 1878 2877 3539 4851 5811 6632 6845
BamH I	1	3784		Tth111 II	10	48 2583	828 1496 1821 2258 5331 6458 6491 6997
Bcl I	1	364		Ave II	12	33 4295 6915	1284 2179 2852 4280 4545 4898 4890 5109 7137
BspMI II	1	5073		Teq I	12	927 2764 5084	1009 1978 2267 2502 3748 4061 4588 4677 7428
Cla I	1	2266		Eco47 III	13	652 1589 3983	726 745 1384 1446 1571 1834 1818 3641 4184 5136
Eag I	1	4348		Hph I	13	3248 4937 7251	3535 3817 3862 4716 5494 5583 6628 6855 7477 7492
Esp I	1	2124		Not I	13	87 2200 3573	331 482 524 1916 2814 2747 2810 2998 5891 7256
Hpa I	1	1684		Bsp1286	14	2302 3888 5780	2711 2873 2966 3685 3894 3996 4583 4874 6198 7359 7444
Hse I	1	2784		Nci I	14	1734 3943 5564	3182 3378 3379 3579 4667 4893 5221 5529 6263 6959 7310
Hvu II	1	3638		Bln I	15	344 4508 6022	2132 3177 3784 3785 5076 6451 6525 6537 6635 7899 7482 7420
Sal I	1	4868		Hgo I	15	513 4858 5413	756 2881 3129 3799 4353 4385 4649 4739 5590 5986 6564 7314
Sma I	1	3378		Mae II	16	275 3245 5209 7696	685 1843 1745 2522 4310 4366 4955 4979 5637 6587 7883 7376 7696
SnaB I	1	1842		Mae I	17	328 1822 2539 6632	485 704 955 1286 2855 2872 2876 2219 2929 3639 4898 6379 6967
Sph I	1	3971		Sec I	17	261 2704 3538 4858	1734 1877 2236 2620 2876 3154 3378 3524 3937 3943 4576 4778 6044
Tth111 I	1	5628		Dde I	18	85 2125 4990 6734	1361 1485 1610 2185 2391 2992 3229 2185 5152 5694 6159 6568 7274 7700
Xba I	1	783					
Afl III	2	2678	5884				
Bcl I	2	264	4853				
Bam I	2	2855	4782				
BspMI	2	2475	4472				
EcoR I	2	1	2241				
EcoRV	2	2685	3594				
Fsp II	2	1977	2763				
HglE II	2	5784	6465				
Hru I	2	2999	4381				
Pf1M I	2	4724	4773				
Ppo I	2	3167	6844				
Pvu I	2	1919	7144				
Sce I	2	2809	7255				
Spe I	2	319	2218				
Ssp I	2	866	7579				
Stu I	2	2160	2933				
ApoL I	3	5780	6198 7444				
Ava I	3	1398	3378 4834				
Bgl I	3	4338	4572 5891				
Dra I	3	6641	6660 7352				
Dra III	3	2860	2862 3100				
Hind III	3	105	2272 3438				
Hda I	3	339	2411 5788				
Hel I	3	336	1396 3322				
Pat I	3	1996	2477 7818				
Sty I	3	2620	2784 4778				
Xan I	3	1889	5440 7372				
Ban II	4	2873	2966 3688 3894				
Not I	4	3810	4178 4338 4692				
Nar I	4	3822	3843 3957 4814				
Fin I	5	2653	3947 4297 4493 5170				
PpuMI	5	2178	2651 2651 4847 4689				
Aho II	6	3022	3843 3957 4614 7313				
Fsp I	6	7695 628 8997	1135 3669 4785 4863				
Gcl II	6	3287 7165	3784 3888 3948 4348				
Hinc II	6	217 7316	1684 3883 3139 4868				
NspB II	6	4548 7418	5475 5594 6224 6489				
Nsp7524 I	6	2678 5884	3324 3971 5225 5519				
Acc I	7	277 4888	2285 2214 2823 3328				
Bbv II	7	171 5809	753 1887 2766 4148				
EcoO109	7	284 4839	2178 2651 3932 4847				
Gau I	7	1782 5392	2831 2988 4220 4810				
Cfr10 I	8	35 4338	3569 3818 3819 4178 4892 6857				
Eae I	8	264 4348	3287 3784 3888 3948 4853 7165				
Bon I	9	3485 4175	3528 3822 3843 3957 4614 4698 8725				
HglA I	9	2382 5780	3685 3996 4583 4874 6198 7359 7444				

附图 15 YEP24 DNA 限制性酶切点的定位。



S0014753

收到期	90.9.21.
来源	西单书
书价	16.50.
单据号	159519
开票日期	90.9.20.

25438

58.17425

721

核酸研究技术 下册

1990

借者	还期	借者	还期
马林	90.11.26		
王芳	90.12.21		
	1.4.7		

58.17425

721

注 意

- 1 借书到期请即送还。
- 2 请勿在书上批改圈点，折角。
- 3 借去图书如有污损遗失等情形须照章赔偿。

25438

京卡 0701



ISBN 7-03-001650-5

Q·242

定 价：16.50 元

科技新书目215—082