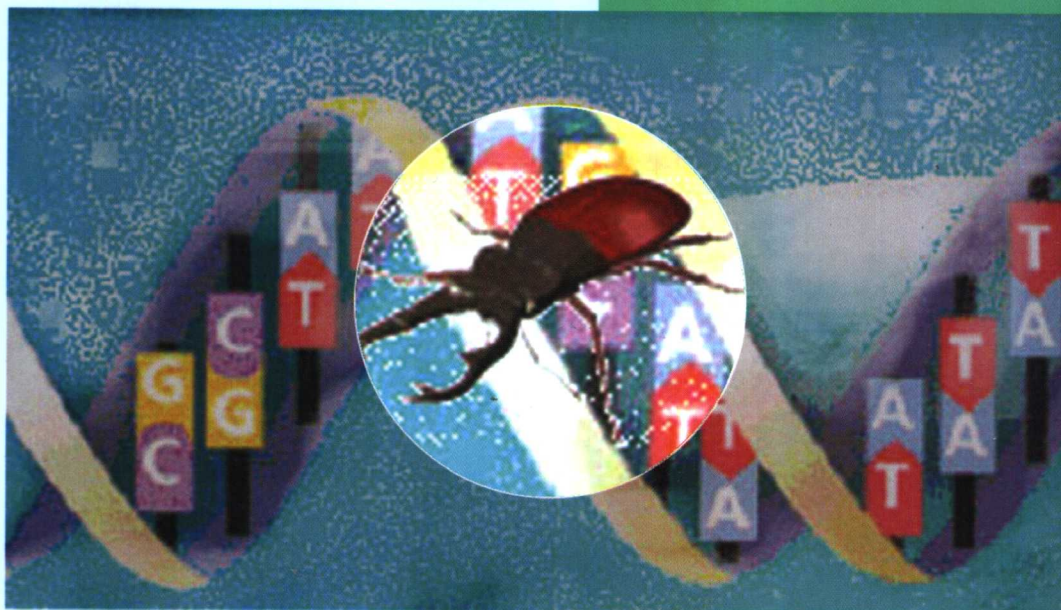




基因工程与害虫防治

JIYINGONGCHENGYUHAICHONGFANGZHI

主 编 林 同



东北林业大学出版社

基因工程与害虫防治

主编 林 同

东北林业大学出版社

图书在版编目 (CIP) 数据

基因工程与害虫防治/林同主编. —哈尔滨: 东北林业大学出版社, 2003.2

ISBN 7-81076-412-8

I. 基... II. 林... III. 基因—遗传工程—应用—害虫—防治 IV. S476

中国版本图书馆 CIP 数据核字 (2003) 第 016577 号

责任编辑: 刘学东

封面设计: 叶 方



NEFUP

基因工程与害虫防治

Jiyin Gongcheng Yu Haichong Fangzhi

主编 林 同

东北林业大学出版社出版发行

(哈尔滨市和兴路 26 号)

东北林业大学印刷厂印刷

开本 787 × 1092 1/16 印张 7.625 字数 176 千字

2003 年 2 月第 1 版 2003 年 2 月第 1 次印刷

印数 1—1 000 册

ISBN 7-81076-412-8

S·365 定价: 13.00 元

内容简介

基因工程技术是现代生物技术的核心，它正以日益显露的优势渗透入我们生活的方方面面；就害虫防治而言，基因工程技术给害虫的综合治理（IPM）注入了新鲜的血液，带来了深刻的变化。

本书对基因工程进行了概述，并分别论述了植物抗虫基因工程、杆状病毒重组杀虫剂、细菌和真菌工程菌的构建等基因工程技术在害虫防治领域中的应用，着重介绍了这些技术的基本原理、技术路线和方法以及最新进展；同时对由基因工程防治害虫而引发的生物安全性等当今热点问题进行了较详细的讨论。本书参考了大量的国内外最新文献，内容详实、全面，理论性强，具有系统性；列举了杨树转抗虫基因工程和舞毒蛾核型多角体病毒基因工程的实施方案等实例，具有实用性和可操作性；本书中的害虫防治方法是以前农林类昆虫学教科书中所没有介绍的，内容新颖，紧密跟踪现代基因工程防治害虫技术前沿领域。本书既可作为研究生使用的教材，又可供从事森保、植保或相关专业的教师、科研人员参考。

前 言

在现代社会中, 生物技术的发展日新月异。生物技术是利用生物的某些功能或潜能进行有效益的生产的工程技术。生物技术的生产过程主要有下列的步骤: 原始材料→上游过程→发酵或生物转化→下游过程→转化产品。生物技术可以分为传统生物技术、工业生物发酵技术和现代生物技术。现在人们常说的生物技术实际上就是现代生物技术。现代生物技术包括基因工程、蛋白质工程、细胞工程、酶工程和发酵工程等五大工程技术。其中基因工程技术是现代生物技术的核心。

基因工程正以其日益显露的优势渗透于我们生活的方方面面; 在农林业领域, 基因工程给人们带来的益处有目共睹。就害虫防治而言, 基因工程技术给害虫的综合治理(IPM) 注入了新鲜的血液, 带来了深刻的变化。如美国 1997 年种植了 100 万 hm^2 转苏云金杆菌毒蛋白基因的抗棉铃虫棉花, 基本上不使用或仅使用少量化学杀虫剂, 总共可减少杀虫剂用量近 40 万升, 并可使皮棉产量净增 7%。两项合计每公顷抗虫棉可增加净收益 83 美元, 总经济效益近 9 000 万美元。

众所周知, 为达到扩大杀虫谱, 提高杀虫效力, 或减少化学农药给环境带来的不利影响等诸多目的, 传统的害虫防治方法已显得无能为力, 基因工程技术的有效利用势在必行。鉴于此, 本书介绍了植物抗虫基因工程、杆状病毒重组杀虫剂、细菌和真菌工程菌等基因工程技术在害虫防治中的应用, 重点介绍了这些技术的基本原理、技术路线和方法以及最新进展, 旨在高效率地防治农林害虫提供理论铺垫和实用参考。

考虑到本书读者多具有植物保护和森林保护或相关专业背景, 故有关植保和森保的知识在本书中不再赘述。另外, 由于篇幅所限, 实施基因工程所需的生物化学和分子生物学知识和技术请参阅有关教科书和其他文献。从本书的系统性出发, 本书首先对基因工程进行了概述, 之后重点论述了植物抗虫基因工程和杆状病毒重组杀虫剂。

值得指出的是, 应该辩证地看待基因工程防治和传统防治方法的关系。无论是利用基因工程技术, 还是利用其他传统方法防治害虫, 都有其自己的优势和不足, 不能顾此失彼, 简单地屏弃某一种方法, 而应互相协调, 充分发挥各自的优势。比如, 由植物转基因技术和杆状重组病毒引发的生物安全性(biosafety) 问题已引起人们的广泛关注和争论, 本书就此问题进行了较详细的讨论。

本书引用了大量文献, 一次文献的作者和出处注在文中或参考文献中, 参考文献中一般不再列入二次文献的作者和出处, 在此向所有被引用文献的作者和版权单位一并表示真诚感谢, 对他们的辛勤劳动表示敬意!

在本书编写过程中, 东北林业大学刘宽余教授、王志英教授提出许多宝贵意见; 本

书由东北林业大学出版基金资助出版，并得到东北林业大学出版社的大力支持，在此一并致以衷心谢意。

希望本书对所有读者带来裨益，但由于水平有限，书中的不足和缺点在所难免，诚请诸位专家和广大读者批评指正。

林 同

2003.1 于哈尔滨

目 录

1 基因工程概论	(1)
1.1 基因工程的概念	(1)
1.2 基因工程操作中常用的工具酶	(2)
1.3 目的基因	(4)
1.4 基因工程载体	(6)
1.5 目的基因的分离	(10)
1.6 基因重组	(12)
1.7 目的基因导入受体细胞	(14)
1.8 克隆子的筛选和鉴定	(15)
2 植物抗虫基因工程	(19)
2.1 转基因植物	(20)
2.2 植物组织培养	(20)
2.3 外源基因导入植物	(23)
2.4 转基因植株鉴定	(26)
2.5 基因沉默	(28)
2.6 抗虫基因	(33)
2.7 我国农作物抗虫基因工程研究进展	(49)
2.8 国内外转基因林木研究进展及趋势	(50)
2.9 杨树抗虫基因工程	(51)
2.10 植物抗虫基因工程的安全性	(54)
2.11 抗虫基因工程实例 1: 抗 yellow stem borer (<i>Scirpophaga incertulas</i>) 的转 Bt CryIAC 基因印度稻	(58)
2.12 抗虫基因工程实例 2: 转 AaIT 基因抗虫杨树	(63)
3 重组杆状病毒杀虫剂	(69)
3.1 昆虫杆状病毒的分子生物学	(69)
3.2 昆虫杆状病毒表达系统	(72)
3.3 昆虫杆状病毒表达系统的优缺点	(78)
3.4 重组病毒杀虫剂	(78)
3.5 重组杆状病毒杀虫剂的生物安全性	(80)
4 细菌抗虫基因工程	(86)
4.1 苏云金芽胞杆菌基因工程	(86)
4.2 假单胞菌基因工程	(94)
4.3 其他细菌	(94)

4.4 展望	(97)
5 昆虫病原真菌基因工程	(98)
5.1 虫生真菌中已经克隆的基因	(98)
5.2 基因工程	(100)
参考文献	(101)

1 基因工程概论

1.1 基因工程的概念

基因工程 (gene engineering) 是 20 世纪 70 年代末、80 年代初在分子生物学、遗传学、生物化学、微生物学等学科的基础上建立起来的。一般来说, 基因工程是指在体外将核酸分子插入病毒、质粒或其他载体分子, 构成遗传物质的新组合, 并使之进入到以前没有这类分子的宿主细胞内, 而能持续稳定地繁殖。

由于基因工程是在分子水平上进行操作, 最终是为了创造出人们所需要的新品种, 因而它可以突破物种间的遗传障碍, 大跨度地超越物种间的不亲和性。如在基因工程中最常使用的大肠杆菌是一种原核生物, 但它却能大量表达来自于人类的某些基因; 再如各种人的多肽生长因子基因就可用大肠杆菌来生产。如果用常规的育种技术来做同一项工作, 那么成功的机会应为零。因此, 科学家们可以利用基因工程实现人类对各种物种改良的愿望。

现在生活在地球上的各种生物都是经过长期的生物进化演变而来的, 虽然不能说它们都很能适应现在的生态环境, 但至少可以说它们基本上都能适应当前的生态环境。也就是说, 每种生物体内或细胞内都处于精巧的调节控制和平衡之中。当用基因工程方法引入一段外源基因片段后, 原有的平衡可能被打破, 有可能导致细胞内的生物学功能发生紊乱, 最后可能导致细胞生长缓慢乃至细胞死亡。很显然, 开展基因工程研究既要使细胞像往常一样正常生长, 又要使细胞产生甚至大量产生人类所需要的外源基因表达产物。

概括起来, 基因工程包括以下内容或步骤:

- (1) 从复杂的生物有机体基因组中, 经过酶切消化或 PCR 扩增 (见 1.5.2) 等步骤, 分离出带有目的基因的 DNA 片段;
- (2) 在体外, 将带有目的基因的外源 DNA 片段连接到能够自我复制的并带有选择标记的载体分子上, 形成重组 DNA 分子 (见 1.6);
- (3) 将重组 DNA 分子转移到适当的受体细胞 (亦称宿主细胞), 并与之一起增殖 (见 1.7);
- (4) 从大量的细胞繁殖群体中, 筛选出获得了重组 DNA 分子的受体细胞克隆 (见 1.8);
- (5) 从这些筛选出来的受体细胞克隆中提取出已经得到扩增的目的基因, 供进一步分析、研究使用;
- (6) 把目的基因克隆到表达载体上, 导入宿主细胞, 使之在新的遗传背景下实现功能表达, 产生出人类所需要的物质。

由此可见,一个完整的基因工程包括基因的分离、重组、转移,基因在受体细胞内的保持、转录、翻译表达等全过程。基因工程的实施至少要有4个必要的条件:工具酶、基因、载体和受体细胞。

除基因工程以外,我们在文献中还见有遗传工程 (genetic engineering)、基因操作 (gene manipulation)、重组 DNA 技术 (recombinant DNA technique)、基因克隆 (gene cloning) 和分子克隆 (molecular cloning) 等。这些术语所代表的具体内容都是彼此相关的,在许多情况下被互相代替,很难严格区分,从某种意义讲,它们之间的差别只不过是各自考虑的角度和强调的侧重点不同罢了。

1.2 基因工程操作中常用的工具酶

基因工程的关键技术是 DNA 的连接重组,但 DNA 在连接之前必须进行加工,把 DNA 分子切割成所需的片段。有时为了便于 DNA 片段之间的连接,还需对片段末端进行修饰。基因工程的工具酶就其用途可分为3大类,即限制性内切酶、连接酶和修饰酶。

1.2.1 限制性内切酶

限制性内切酶 (restriction endonuclease) 是一类以环状或线形双链 DNA 为底物,能识别 DNA 中特定的核苷酸序列,并在合适反应条件下使每条链的一个磷酸二酯键断开,产生具有 3'-OH 和 5'-P 基团 DNA 片段的内脱氧核苷酸酶。至今发现的此类酶有3种类型,即 I 型酶、II 型酶和 III 型酶,目前在基因工程中真正有用的酶是 II 型酶,如果没有专门说明,通常所说的限制性内切酶是指 II 型酶。限制性内切酶根据其来源命名,如从 *Haemophilus influenzae* Rd 中提取的第三种限制性内切酶被命名为 *Hind* III。至今已发现上千种限制性内切酶,常用的有几十种。

1.2.1.1 限制性内切酶在 DNA 分子上的识别序列

限制性内切酶在双链 DNA 分子上能识别的特定核苷酸序列称之为识别序列,有时称之为识别位点。识别序列多数由4、5、6个核苷酸组成,如 *Dpn* I、*Apy* I 和 *EcoR* I 的识别序列分别是 GATC、CCTGG 和 GAATTC。有些识别序列超过6个核苷酸。有一些限制性内切酶虽来源不同,但有相同的识别序列,称之为同裂酶 (isoschizomer),如 *BamH* I 和 *Bst* I 具有相同的识别序列 GGATCC。

1.2.1.2 限制性内切酶的酶切位点

限制性内切酶具有特定的酶切位点,即限制性内切酶在其识别序列的特定位点对双链 DNA 进行切割,由此产生特定的酶切末端。一类是两条链断裂是交错对称的,产生的 DNA 的末端的一条链多出1至几个核苷酸,成为凸出末端,又称粘性末端;DNA 链 5'端凸出的称为 5'粘性末端,3'端凸出的称为 3'粘性末端。另一类是两条链上断裂的位置处在识别序列的对称结构中心,产生的 DNA 末端是平齐的,称为平末端。经限制性内切酶切割产生的 DNA 片段,不管是粘性末端还是平末端,5'端一定是 -P 基团,3'端一定是 -OH 基团。

有些限制性内切酶虽然识别序列不同,但切割 DNA 分子产生的 DNA 片段具有相同的粘性末端,这样的酶称为同尾酶(isocaudamer)。如 *Taq* I、*Cla* I 和 *Acc* I 是一组同尾酶,其中任何一种酶切割 DNA 分子都产生 5'端 CG 凸出的粘性末端。同尾酶在 DNA 片段重组时特别有用。

1.2.1.3 限制性内切酶反应系统

限制性内切酶反应底物是 DNA 分子,但只含识别序列的寡核苷酸是不会被切割的。如只含 GAATTC 的寡核苷酸不会被 *EcoR* I 切割,只有在识别序列两侧延伸 1 个或几个核苷酸后才能被 *EcoR* I 切割。所以人工合成的 *EcoR* I 连杆不是 GAATTC,而是 GGAAT-TCC。限制性内切酶反应系统的缓冲液因酶而异,常用的缓冲液有 A、B、L、M、H 等几种,厂家出售限制性内切酶时提供相应的缓冲液。限制性内切酶反应温度多数是 37℃,少数低于或高于 37℃。

1.2.2 DNA 连接酶

用于将两片段乃至数段 DNA 片段拼接起来的酶称为 DNA 连接酶。基因工程中最常用的酶是 T_4 DNA 连接酶,它催化 DNA 5'磷酸基与 3'羟基之间形成磷酸二酯键,其用途是:

- (1) 连接带匹配粘末端的 DNA 分子;
- (2) 使平端的双链 DNA 分子互相连接或与合成的寡核酸头相连接。这类反应要比粘末端之间的连接慢得多,但单价阳离子(150~200 mmol/L NaCl)或低浓度的聚乙二醇(PEG)可提高平端之间的连接效率。

除了 T_4 DNA 连接酶之外,还有大肠杆菌 DNA 连接酶,它所行使的催化反应基本和 T_4 DNA 连接酶相同,但只能连接具粘末端的 DNA 片段,而且催化反应需要有 NAD^+ 参与。

DNA 连接酶的最适反应温度是 37℃,但为了增强两 DNA 片段粘性末端互补碱基之间氢键的稳定性,实际反应温度是 4~15℃。

1.2.3 基因工程中的修饰酶

1.2.3.1 DNA 聚合酶

目前常用的 DNA 聚合酶是大肠杆菌 DNA 聚合酶 I、大肠杆菌 DNA 聚合酶 I 大片段(Klenow fragment)、 T_4 噬菌体 DNA 聚合酶、 T_7 噬菌体 DNA 聚合酶以及耐高温的 DNA 聚合酶(如 *Taq* DNA 聚合酶)等,不同来源的 DNA 聚合酶具有各自的酶学特性。大肠杆菌 DNA 聚合酶 I 具有 5'到 3' DNA 聚合酶活性、5'到 3'及 3'到 5'外切核酸酶活性,其大片段是经枯草杆菌蛋白酶裂解完整的 DNA 聚合酶 I 或用基因工程手段去除全酶中 5'到 3'外切核酸酶活性片段而得到的具有 5'到 3' DNA 聚合酶活性及 3'到 5'外切核酸酶活性的全酶大片段。 T_4 噬菌体 DNA 聚合酶的酶活性与上述大片段酶活性相似,然而其 3'到 5'外切核酸酶活性比 Klenow 大片段酶强近 200 倍。与 Klenow 大片段酶相比, T_4 噬菌体 DNA 聚合酶的最大特点是不从单链 DNA 模板上置换寡核苷酸引物,因此在体外突变反

应中, T_4 DNA 聚合酶比大片段酶有更高的效率。 T_7 噬菌体 DNA 聚合酶所催化合成的 DNA 的平均长度要比其他 DNA 聚合酶催化合成的 DNA 的平均长度大得多。耐高温的 DNA 聚合酶(如 *Taq* DNA 聚合酶)由于其最佳作用温度是 $75 \sim 80^\circ\text{C}$, 目前广泛用于多聚酶链反应(polymerase chain reaction, PCR)及 DNA 测序。

反转录酶是一种依赖于 RNA 的 DNA 聚合酶, 具有 5'到 3' DNA 聚合酶活性、5'到 3' 及 3'到 5'外切核糖核酸酶活性, 但不具备 3'到 5'外切脱氧核糖核酸酶活性。反转录酶在基因工程中的用途主要是以真核 mRNA 为模板合成 cDNA, 用于组建 cDNA 文库, 进而分离为特定蛋白质编码的基因。另外, 将 mRNA 与 PCR 偶联建立起来的反转 PCR(RT-PCR), 使真核基因的分离更加有效。

另一类型的 DNA 聚合酶是末端脱氧核苷酸转移酶, 简称末端转移酶。在二价阳离子存在下, 末端转移酶催化 dNTP 加于 DNA 分子 3'羟基端, 主要用于给载体或 cDNA 加上互补的同聚尾, 便于外源基因重组。

1.2.3.2 依赖于 DNA 的 RNA 聚合酶

SP_6 噬菌体 RNA 聚合酶及 T_7 和 T_3 噬菌体 RNA 聚合酶是目前常用的三种酶。这三种酶的共同特点是只识别各自噬菌体的 DNA 序列中特定的启动子区, 它们的主要用途是在离体的条件下, 合成单链 RNA 作为杂交探针, 完成体外翻译系统中功能性 mRNA 或体外剪接反应的底物的合成。

1.2.3.3 T_4 噬菌体多核苷酸激酶

该酶催化 ATP 的 γ -磷酸基转移至 DNA 或 RNA 片段的 5'末端。在基因工程中主要用于标记 DNA 片段的 5'末端, 制备杂交探针; 基因的化学合成中, 寡核苷酸片段 5'磷酸化; 测序引物的 5'磷酸标记。

1.2.3.4 碱性磷酸酶

碱性磷酸酶的功能是去除 DNA 或 RNA 5'末端的磷酸反应, 用于去除 DNA 片段 5'磷酸, 以防止在重组时自身环化, 从而提高重组效率; 在以 $[\gamma-^{32}\text{P}]$ ATP 标记 DNA 或 RNA 的 5'末端前, 去除 DNA 或 RNA 片段的非标记的 5'磷酸。

1.3 目的基因

基因工程的主要目的是通过优良性状相关基因的重组, 获得具有高度应用价值的新物种。为此必须从现有生物群体中, 根据需要分离出可用于克隆的此类基因, 这样的基因通常称之为目的基因, 目的基因主要是结构基因。

作为一个能转录和翻译的结构基因必须包括转录启动子、基因编码区和转录终止子。启动子是 DNA 上 RNA 聚合酶识别、结合和促使转录的一段核苷酸序列。转录 mRNA 的第一个碱基被定为转录起始位点, 这个碱基多数情况下是 A, 以此作为序列的 +1, 其上游的核苷酸序列为“-”, 下游的核苷酸序列为“+”。基因编码区包括起始密码子 ATG、开放阅读框(ORF)和终止密码子 TAA(或 TAG、TGA)。终止密码子提供转录停止信息。不同类型基因组的基因组成稍有不同, 因此必须根据基因组类型和实验需

要来分离含目的基因的 DNA 片段。

1.3.1 原核生物结构基因的组成

原核生物的基因多数以操纵子形式存在，完成同类功能的多个基因聚集在一起，处在同一个启动子的调控之下，下游同具一个终止子，但各基因分别有起始密码子 ATG 和终止密码子 TAA（或 TAG、TGA）。两个基因之间存在长度不等的间隔序列，但有例外。原核生物的启动子含 -35 序列保守区 5'TTGACA3'，提供 RNA 聚合酶识别的信号；-10 序列保守区 5'TATAAT3'，DNA 双链从此处解开转录。操纵子除启动子外，往往还有一些调控转录的其他因子，如乳糖操纵子除启动子（p）外，还有调节因子（i）和操纵因子（o）。

原核生物基因在起始密码子 ATG 上游约 10bp 处，有一个富含嘌呤核苷酸的 SD 保守序列区，一般为 AGGA，由此区转录的序列是翻译时核糖体识别、结合 mRNA 的位置。核糖体由此位置向前移动，寻找起始密码子 AUG。

原核生物基因转录终止之前有一段回文序列结构，称为终止子，使 RNA 聚合酶减缓移动或停止 RNA 的合成，但不是所有回文结构都是终止子。

1.3.2 真核生物结构基因的组成

真核生物染色体基因组的基因是单独存在的，并且两个基因之间有很长的间隔区。甚至起始密码子和终止密码子之间除编码序列区外还有一至多个非编码序列间隔区，称为内含子，编码序列区称为外显子。

真核生物结构基因的转录启动子通常包含 3 个保守序列区：在 -20 ~ -30 序列区有一个 TATA 框，DNA 双链在此处开始解开；在 -75 序列区有一个 CAAT 框，其作用与 RNA 聚合酶结合有关；在更上游处有一个 GC 框，是某些转录调控因子结合的序列。

对真核生物基因转录的终止信号和终止过程了解甚少。不过发现高等真核生物结构基因终止密码子下游一保守的 AATAAA 序列提供转录产物的释放信号。

此外，真核生物染色体基因组的结构基因不具 SD 序列区，核糖体与转录的 mRNA 的结合依赖于 mRNA5' 的帽结构。

1.3.3 其他类型基因组的基因组成

除上述原核生物基因组和真核生物染色体基因组外，还有自主复制的质粒基因组、病毒（噬菌体）基因组、线粒体基因组和叶绿体基因组，它们的基因组成类似于原核生物基因组。有的质粒 DNA 上有少量单独存在的结构基因，基因组成简单，含内含子、基因编码区和终止子。病毒（噬菌体）的基因因种类不同而不同，有的定位在 DNA 分子上，有的定位在 RNA 上；很多基因重叠排列，有的基因还含内含子。昆虫杆状病毒的基因组和基因见第三部分。

线粒体是真核生物的细胞器，其 DNA 上有一系列与呼吸作用相关的基因，这些基因不含 SD 序列；有些植物的线粒体基因有内含子。叶绿体是植物特有的细胞器，其 DNA 上有一系列与光合作用有关的基因。有些植物的叶绿体的基因不含 SD 序列；有的

叶绿体基因有内含子。

1.4 基因工程载体

如果只有基因而没有负责运载它的载体,则基因就不可能发挥作用。基因工程载体决定了外源基因的复制、扩增、传代乃至表达。基因工程载体应该满足以下条件:

- (1) 能自我复制并能带动插入的基因一起复制;
- (2) 有合适的限制性内切酶位点。在载体上单一的限制性内切酶位点越多越好,这样可以将不同限制性内切酶切割后的 DNA 片段方便地插入载体;
- (3) 具有合适的筛选标记;
- (4) 在细胞内拷贝数要多,这样才能使外源基因得以扩增;
- (5) 载体的相对分子质量要小,这样可以容纳较大的外源 DNA 插入片段。载体的相对分子量太大将影响重组体或载体本身的转化效率;
- (6) 在细胞内稳定性高,这样可以保证重组体稳定遗传而不易丢失。

现在有多种载体,它们分别由从细菌质粒、噬菌体 DNA、病毒 DNA 分离出的元件组装而成,可分为克隆载体和表达载体。表达载体又分胞内表达和分泌表达载体。从表达所用的受体细胞分,可分为原核细胞和真核细胞表达载体。从功能分,又可以分为测序载体、克隆-转录用载体、基因调控报告载体等多功能载体。

1.4.1 质粒载体

质粒载体是最早发展起来的一类载体,以质粒 DNA 为基础构建而成,是微生物和植物转基因研究的主要载体。

质粒是一种寓于宿主细胞中染色体外的裸露 DNA 分子,广泛存在于细菌之中,在某些蓝藻、绿藻和真菌细胞中也存在质粒。一个质粒就是一个 DNA 分子,多数以超螺旋环状共价双链分子形式存在,以 cccDNA 表示,体外在理化因子作用下可成为开环(ocDNA)或线形(l-DNA)分子。质粒 DNA 分子小的不足 1 500 bp (1 bp = 1 碱基对),大的可达 100 kb 以上 (1 kb = 1 000 bp)。质粒在宿主细胞内能自主复制,有的能复制几百上千个拷贝,称之为松弛型复制质粒;有的只能复制少数拷贝,称之为严紧型复制质粒。构建的质粒载体应具有复制起始点(ori),能在受体细胞内复制。此外,质粒载体还应有 1~2 种选择标记基因(*Amp^r*、*Cmp^r*、*Kan^r* 和 *lacZ'* 等)和允许外源 DNA 插入的克隆位点。

1.4.2 噬菌体和病毒载体

基因工程载体除质粒载体以外,还有一大类由噬菌体和病毒 DNA (或 RNA) 构建的载体。此类载体和外源 DNA 可以转染受体细胞,但通过噬菌体或病毒颗粒转导受体细胞必须用外壳蛋白进行体外包装。

1.4.2.1 λ 噬菌体载体

λ 噬菌体之所以可以作为基因工程的载体材料,是因为它有以下几方面的特性:

(1) λ 噬菌体含有线状双链 DNA 分子, 其长度为 48 502 bp, 两端各由 12 个核苷酸组成 5'端凸出的互补粘性末端, 当 λ DNA 进入宿主细胞后, 互补粘性末端连接成为环状 DNA 分子, 连接处称为 cos 位点;

(2) λ 噬菌体为温和噬菌体, λ DNA 可以整合到宿主细胞染色体上, 以溶源状态存在, 随染色体的复制而复制。当用紫外线或 45 °C 处理时, 释放出 λ DNA, 被壳蛋白包装, 成为溶菌态, 迅速增殖;

(3) λ 噬菌体能包装原 λ DNA 长度的 75% ~ 105%, 即使不对 λ DNA 进行改造, 也允许承载 5 kb 大小的 DNA 片段带入受体细胞。而且 λ DNA 分子中约有 20kb 的区域对 λ 噬菌体的生长不是绝对必需的, 允许替换和部分缺失;

(4) λ 噬菌体上的 D 基因和 E 基因对噬菌体的包装起决定性的作用, 缺少其中的任何一种基因都将导致噬菌体不能包装, 在宿主(受体)细胞中积累大量供包装用的其他壳蛋白;

(5) 在 λ DNA 分子上有多种限制性内切酶的识别序列, 便于用这些酶切割产生的外源 DNA 进行插入和置换。但是有的酶在 λ DNA 上有多个识别序列, 有的识别序列位于必需基因区域, 将影响外源 DNA 片段的插入和置换。

鉴于上述 λ 噬菌体的特性, 构建 λ 噬菌体载体的基本途径是:

(1) 去掉某种限制性内切酶在 λ DNA 分子上的一些识别序列, 只在非必需区保留 1~2 个识别序列。若保留一个识别序列, 可供外源 DNA 插入; 若保留两个识别序列, 则两个识别序列之间的区域可供外源 DNA 片段置换;

(2) 用合适的限制性内切酶切去部分非必需区, 但是由此构建的 λ DNA 载体不应小于 38 kb;

(3) 在 λ DNA 分子的合适区域插入可供选择的标记基因。

根据以上策略, 可以构建一系列利用不同限制性内切酶识别序列作为克隆位点的 λ 噬菌体载体。例如 gt₁₀es- λ B 是以 *EcoR* I 识别序列作为克隆位点的一种载体。*EcoR* I 在 λ DNA 上有 5 个识别序列, 若用此酶切割, 可得 6 个片段, 6 个片段大小分别是 21.6 kb、4.9 kb、5.5 kb、7.5 kb、5.9 kb 和 3.2 kb。片段 B 和 C 的完全缺失以及 E 片段中 2.6 kb 片段的缺失, 仍能维持噬菌体的溶菌生长。因此, 用 *EcoR* I 切去 C 片段, 保留 B 片段及其两侧的 *EcoR* I 识别序列, 并且使 E 片段中的 2.6 kb 缺失, 构建成 gt₁₀es- λ B 载体(图 1.1)。此载体分子大小为 48.5-5.5-2.6=40.4 kb。在保留的 2 个 *EcoR* I 识别序列处可插入约 13kb 或取代 18kb 具 *EcoR* I 切割末端的外源 DNA 片段, 包括选择标记基因。这种载体常用于构建 cDNA 文库见 1.5.1。

构建的 λ 噬菌体载体, 只要保留 λ DNA 两端含 cos 位点的短片段, 中间可以全是外源 DNA, 并且总长度大于 36 kb, 或小于 53 kb, 均可以按 λ 噬菌体进行体外包装后转导受体细胞。但这样包装的 λ 噬菌体不能增殖。由带 cos 位点的 λ DNA 片段与质粒构建的载体称为粘粒(cosmid)载体。这样的载体既可以按质粒载体的性质转化受体菌, 并在其中复制; 又可以按 λ 噬菌体进行体外包装转导受体细胞, 通过 cos 位点连接环化后, 按质粒复制的方式进行复制。一般构建的粘粒载体小于 20 kb, 可承载 30 kb 左右的外源 DNA 片段, 这种载体常用于构建真核生物基因组文库见 1.5.1。

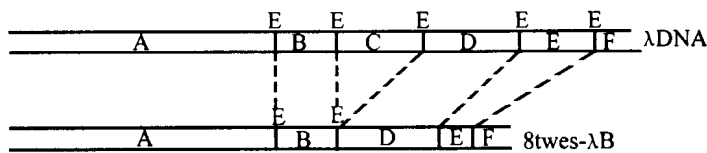


图 1.1 gt10es-λB 构建
(宋思杨等, 1999)

1.4.2.2 CaMV 载体

花椰菜花斑病病毒 (CaMV) 含一个约 8 kb 的环状双链 DNA 分子, 由此构建的载体可用于十字花科和少数非十字花科植物的基因转移。由于 CaMV 感染对植物是有害的, 因此不能直接用作载体, 并且 CaMV 最大只能包装大于 CaMV DNA 300 bp 的重组 DNA 分子, CaMV DNA 上的绝大多数基因都是不可缺少的, 不能被外源 DNA 所取代。构建 CaMV 载体主要有下面几种方法:

(1) 互补载体系统: 由缺陷性的 CaMV 同辅助病毒组成。缺陷性 CaMV 的 DNA 分子上仅具病毒少数几种功能相关的基因, 而辅助病毒的 DNA 上携带除缺陷性病毒具有的基因以外的 CaMV 的所有基因。两者之间的基因呈互补关系, 构成一个完整的基因组, 它们一起感染植物时, 可有效地将外源 DNA 转移到被感染的植物;

(2) 混合载体系统: 把组装了外源 DNA 的 CaMV DNA 插入 Ti 质粒载体, 通过根癌农杆菌介导, 转移到植物细胞内;

(3) CaMV35S 启动子融合基因载体系统: CaMV 的 DNA 在植物细胞中能高水平转录 35SRNA, 其启动子是一种强启动子。因此构建了含 35S 启动子的系列载体, 不含启动子的外源目的基因被组装在 35S 启动子下游, 组成融合基因, 可在植物细胞中高效表达。此类载体已被广泛用于植物基因工程 (见第二部分)。

1.4.2.3 SV40 载体

SV40 是一种含 5241 bp 环状双链 DNA 的猿猴空泡病毒 40, 其感染猿猴细胞后, 脱去蛋白外壳, DNA 进入细胞核, 开始基因早期转录和晚期转录。通过早期转录和翻译, 产生 T 抗原, 启动 DNA 复制。通过晚期转录和翻译, 合成壳蛋白质, 并与新复制的 DNA 一起组装成新的病毒颗粒, 使细胞破裂, 称为裂解感染。SV40 载体可用于哺乳动物, 并且对人类是安全的。由 SV40 的 DNA 构建的载体有以下两类:

(1) 取代型重组病毒载体: 由于 SV40 不能包装大于原 DNA 分子的重组 DNA, 所以采用外源 DNA 片段取代 SV40 DNA 中的部分区域, 如此构建的载体有晚期启动区取代载体和早期启动区取代载体。晚期启动区取代载体由于指令合成壳蛋白的基因被取代, 所以需要一种辅助病毒提供壳蛋白, 才能使 SV40DNA 包装重组成为有效的病毒颗粒。早期启动区取代载体由于转录 T 抗原的编码区被外源 DNA 所取代, 不能合成 T 抗原, 必须由一种辅助病毒来提供 T 抗原。后来建立了一种直接能提供 T 抗原的猿猴细胞系, 称为 COS 细胞系。早期启动区取代载体进入 COS 细胞后, 可组装成有效的病毒颗粒。

(2) 重组病毒质粒载体: 此类载体由 SV40 的部分 DNA 片段与质粒重组而成。如此

构建的载体中, 有的 DNA 重组分子内含 SV40 DNA 的完整早期转录区和复制起点, 有的只含 SV40 DNA 的复制起点。后者称为微型病毒复制子质粒载体。

1.4.2.4 反转录病毒载体

反转录病毒是一类含 RNA 的病毒, 进入受体细胞后, RNA 反转录成 DNA, 通过两端由数百碱基对组成的长末端重复序列 (LTR), 整合到染色体 DNA 上, 成为原病毒。

构建反转录病毒载体的途径是: 从感染的细胞中分离原病毒 DNA, 根据不同要求组装选择标记基因和启动子等元件, 与大肠杆菌质粒重组。这样的载体可转化大肠杆菌, 并在转化的细胞内转录重组的病毒 RNA, 但因其缺失包装所需的某些蛋白质的基因, 因此这些蛋白质需由辅助病毒提供, 此系统称为辅助病毒互补重组反转录病毒质粒载体系统。如果把反转录病毒包装必需的全部蛋白质的基因先整合到受体细胞基因组, 然后用重组病毒质粒载体直接转化这样的细胞, 可产生有效的反转录病毒颗粒。此系统称为不需辅助病毒的重组反转录病毒质粒载体系统。不管哪一种系统, 只要将外源目的基因插入原病毒 DNA 的合适克隆位点, 就能在被感染的动物细胞中表达。为使外源目的基因有效表达, 有的利用基因原有的启动子, 有的利用定位在反转录病毒长末端重复序列区的启动子。

除了以上几种病毒载体外, 还有昆虫杆状病毒载体等。关于昆虫杆状病毒载体, 将在第三部分详细介绍。

1.4.3 YAC 载体

目前能容纳最大外源 DNA 片段的载体是 YAC 载体。YAC 是酵母人工染色体 (yeast-artificial chromosome) 的缩写。人们发现, 真核生物的染色体有几个部分是最为关键的: 一是着丝粒 (centromere, CEN), 它主管染色体在细胞分裂过程中正确地分配到各子细胞中; 二是端粒 (telomere, TEL), 端粒位于染色体的末端, 它们对染色体末端的复制以及防止染色体被核酸外切酶切断具有重要的意义; 三是自主复制序列 (autonomously replicating sequence, ARS), 即在染色体上多处 DNA 复制起始的位点, 它和质粒的复制起始位点有些相似。了解了真核细胞染色体的重要结构以后, 人们试图用 DNA 重组技术分离开这些重要组分, 然后再将它们连接起来产生人工染色体。实验证明这条思路是可行的。由于酵母染色体 DNA 至少有几十万个碱基长, 因此人工染色体所能装进的外源 DNA 片段就要比其他类型载体装载得很多。简单地说, 酵母人工染色体载体有两个臂, 每个臂的末端有一个端粒, 而臂上则有着丝点等染色体必备元件。除此之外, 还有供选择的标记基因, 当外源 DNA 片段连接进两个臂中以后, 通过选择标记, 人们可以从酵母宿主细胞中筛选出重组的人工染色体。

图 1.2 是 YAC 的构建过程。目的核基因组经限制性内切酶酶解后变成大片段 DNA, 将这些大片段 DNA 连接进两个载体臂, 每个载体臂都以 TEL 结尾, 臂上有 CEN、ARS 以及选择标记 (TRP1 和 URA3)。有了这些元件, 带有外源 DNA 并在酵母中稳定存在的人工染色体都被筛选出来。为更好地进行重组染色体的选择, 有人还在 YAC 的载体臂加上了 tRNA 抑制子。另有人把 Pbr322 的单克隆位点引入载体臂上, 以便进行亚克隆构建和 YAC 克隆的限制性内切酶谱制定工作。也有人在 YAC 载体臂上加上了 T₃ 和 T₇ RNA