

生物技术和生物工程专业规划教材

# 基因工程实验指导

The Experimental Guide for Gene Engineering

朱旭芬 编著



高等教育出版社  
Higher Education Press

生物技术和生物工程专业规划教材

要参客

# 基因工程实验指导

The Experimental Guide for Gene Engineering

朱旭芬 编著

编著(11) 合编(11) 合编(11)

基因工程实验指导(第2版) 朱旭芬编著 北京: 高等教育出版社, 2008.1

ISBN 7-04-018364-1

I. 基... II. 朱... III. 基因工程—实验—教材 IV. Q78-33

高等教育出版社 林强—刘学高等

中国版本图书馆CIP数据核字(2007)第148238号

81118

策 划 编 辑 王 芳 策 划 编 辑 王 芳 策 划 编 辑 王 芳  
版 式设计 王 芳 版 式设计 王 芳 版 式设计 王 芳  
图 像 设计 王 芳 图 像 设计 王 芳 图 像 设计 王 芳

010-58581000

010-58581000

http://www.hep.com.cn

http://www.hep.com.cn

http://www.hep.com.cn

http://www.hep.com.cn

http://www.hep.com.cn

出版发行 高等教育出版社

社 址 北京市西城区德胜大街4号

邮 政 编 码 100011

总 机 010-58581000

经 销 蓝色畅想图书发行有限公司

印 刷 北京鑫海金泰印务有限公司

开 本 787×1092 1/16

印 张 13.75

字 数 330 000

次 2008年1月第1版

次 2008年11月第2次印刷

价 18.00元

本书如有缺页、倒页、脱页等质量问题, 可向出版单位联系调换。

版权所有 侵权必究

18384-1



高等教育出版社

Higher Education Press

## 内容提要

本书以基因克隆为主线,选编了35个实验,包含了从核酸与质粒的提取、基因文库构建与筛选、基因分离方法的确定、PCR法获得基因、克隆载体和表达载体的重组、重组子的鉴定,基因的表达到表达产物——蛋白的纯化、鉴定等相关的实验技术内容。同一实验项目提供了多种可供选择的实验方法,并强调指出了实验中的注意事项,还对与实验相关的内容进行了评议。

选编的实验内容衔接紧密,在实际教学时,可以针对不同层次、不同学时数的要求,既可以集中在两个星期连续进行,也可以分成十次实验阶段性完成。为提高教学效率,作者还将实验安排分段详细列于实验计划表中,便于教师组织实验教学。

## 图书在版编目(CIP)数据

基因工程实验指导/朱旭芬编著. —北京:高等教育出版社,2006.1  
ISBN 7-04-018364-1

I. 基... II. 朱... III. 基因-遗传工程-实验-高等学校-教材 IV. Q78-33

中国版本图书馆CIP数据核字(2005)第148526号

策划编辑 王莉 责任编辑 田军 封面设计 张楠 责任绘图 朱静  
版式设计 王莹 责任校对 王雨 责任印制 朱学忠

出版发行 高等教育出版社  
社 址 北京市西城区德外大街4号  
邮政编码 100011  
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司  
印 刷 北京鑫海金澳胶印有限公司

开 本 787×1092 1/16  
印 张 13.75  
字 数 330 000

购书热线 010-58581118  
免费咨询 800-810-0598  
网 址 <http://www.hep.edu.cn>  
<http://www.hep.com.cn>  
网上订购 <http://www.landaco.com>  
<http://www.landaco.com.cn>  
畅想教育 <http://www.widedu.com>

版 次 2006年1月第1版  
印 次 2006年1月第1次印刷  
定 价 18.00元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 18364-00

# 前 言

基因工程方法与技术随着分子生物学的迅猛发展,已渗透到现代生命科学的各个分支领域,成为生物工程的核心技术。

本实验教材的编写是基于自己多年来科研实践、教学内容与方法的探索,并汲取了国内外成熟的经验和信息资料,在2002年7月、2004年7月、2005年7月的实验讲义以及2004年9月全国生物技术实验师资培训教材的基础上进一步修改完善而成的。根据基因工程研究的发展趋势,结合基因操作技术和分子生物学研究中经常使用的基本方法以及从自然界中克隆基因的基本思路,本教材以基因克隆为主线,包含了从核酸与质粒的提取、基因文库构建与筛选,基因分离方法的确定、PCR法获得基因,与克隆载体和表达载体的重组、重组子的鉴定,基因的表达,到表达产物——蛋白质的纯化、鉴定等相关的实验技术内容。同一实验项目提供了多种可供选择的实验方法,并强调指出了实验中的注意事项,还对与实验相关的内容进行评议,试剂的配制等内容。为了能让实验有正确的理论指导,在每个实验前比较系统地介绍了与实验内容相关的基本原理,可为初学者提供有益的帮助,也为科研人员提供参考。

教材中的实验内容衔接紧密,可以根据各自的具体情况串联成一个综合性的大实验,针对不同层次、不同学时数要求的学生开设。为了尽可能全面、直观地了解基因工程的内涵,我们为本科生开设了以几丁质酶基因的克隆和表达为主题的96学时的系列大实验,这些实验既可以用10天时间连续进行,也可以分成10次实验间断完成,实验一环紧扣一环,在实验的间隙适当安排基因工程基本理论方面的专题讲座。旨在使学生对基因工程这一领域的主要技术有一个全方位、比较系统的了解,以便能将这些方法运用于今后的科研实践。为了提高教学效率,我们将实验安排分时段详细地列于实验计划表中,这样可以使教师与学生对实验每个环节都做到心中有数。

本实验课程既重视对学生基本实验操作技能的培养和训练,又注重对学生研究能力以及分析和解决问题能力的培养,以激发学生的学习和研究兴趣。通过系统的实验训练,培养学生独立实验、观察问题、分析问题和解决问题的能力,提高学生实验设计、综合运用能力,使学生具备从事分子生物学教学和科研的基本能力。

在编写本实验教材的过程中,受到了世界银行贷款的支助,得到了张铭、赵小立、史锋、朱成钢及金勇丰等老师的热情帮助,也得到了高等教育出版社生命科学分社的王莉老师的大力支持,在此对他们表示由衷的谢意。

限于本人的学识与水平,教材中难免有不妥和错漏之处,敬请大家随时赐教和指正,谢谢!

朱旭芬

2005年8月于浙江大学西溪校区北园

## 郑 重 声 明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

**反盗版举报电话：**(010) 58581897/58581896/58581879

**传 真：**(010) 82086060

**E - mail：**dd@hep.com.cn

**通信地址：**北京市西城区德外大街 4 号

高等教育出版社打击盗版办公室

**邮 编：**100011

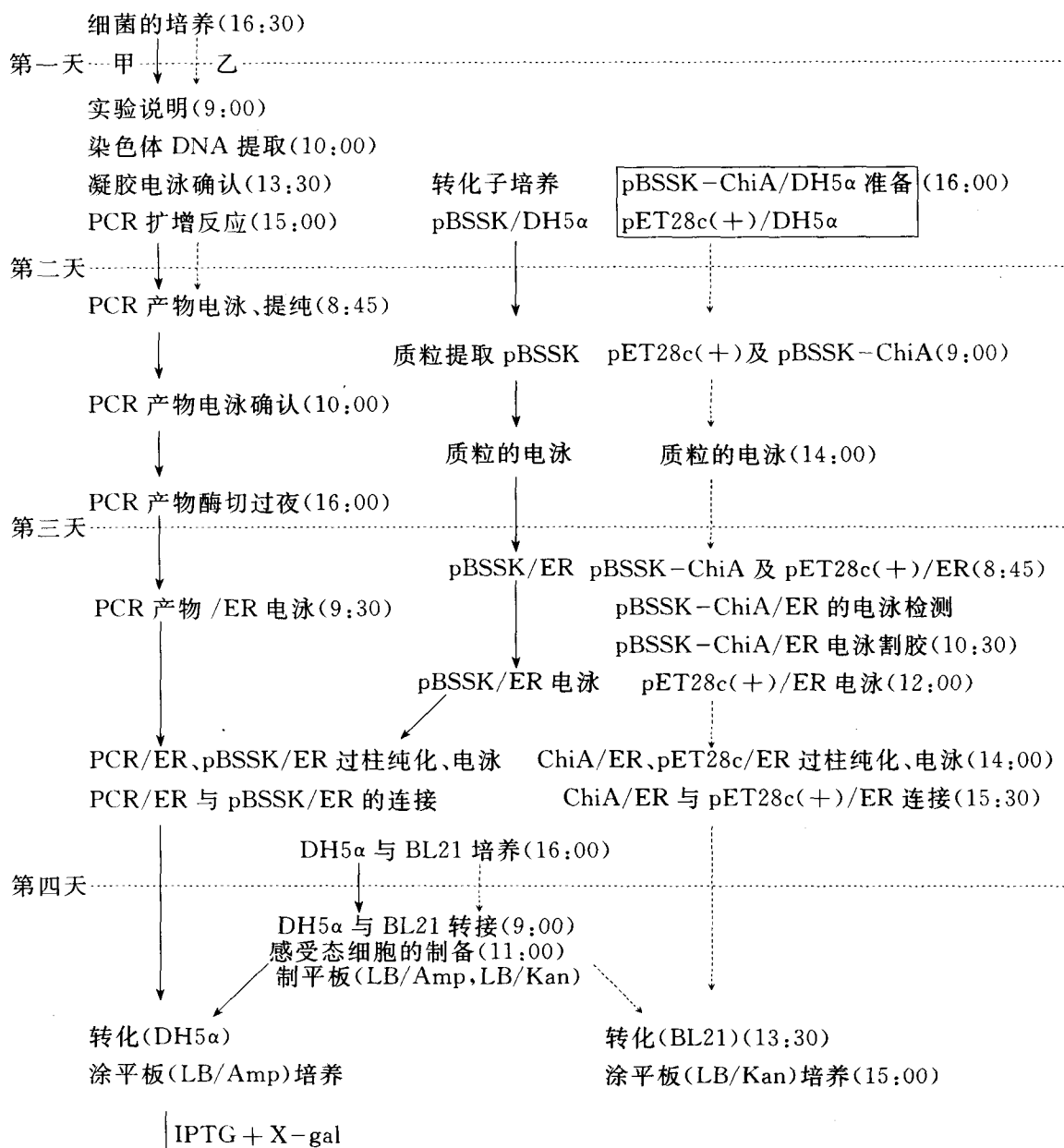
**购书请拨打电话：**(010)58581118

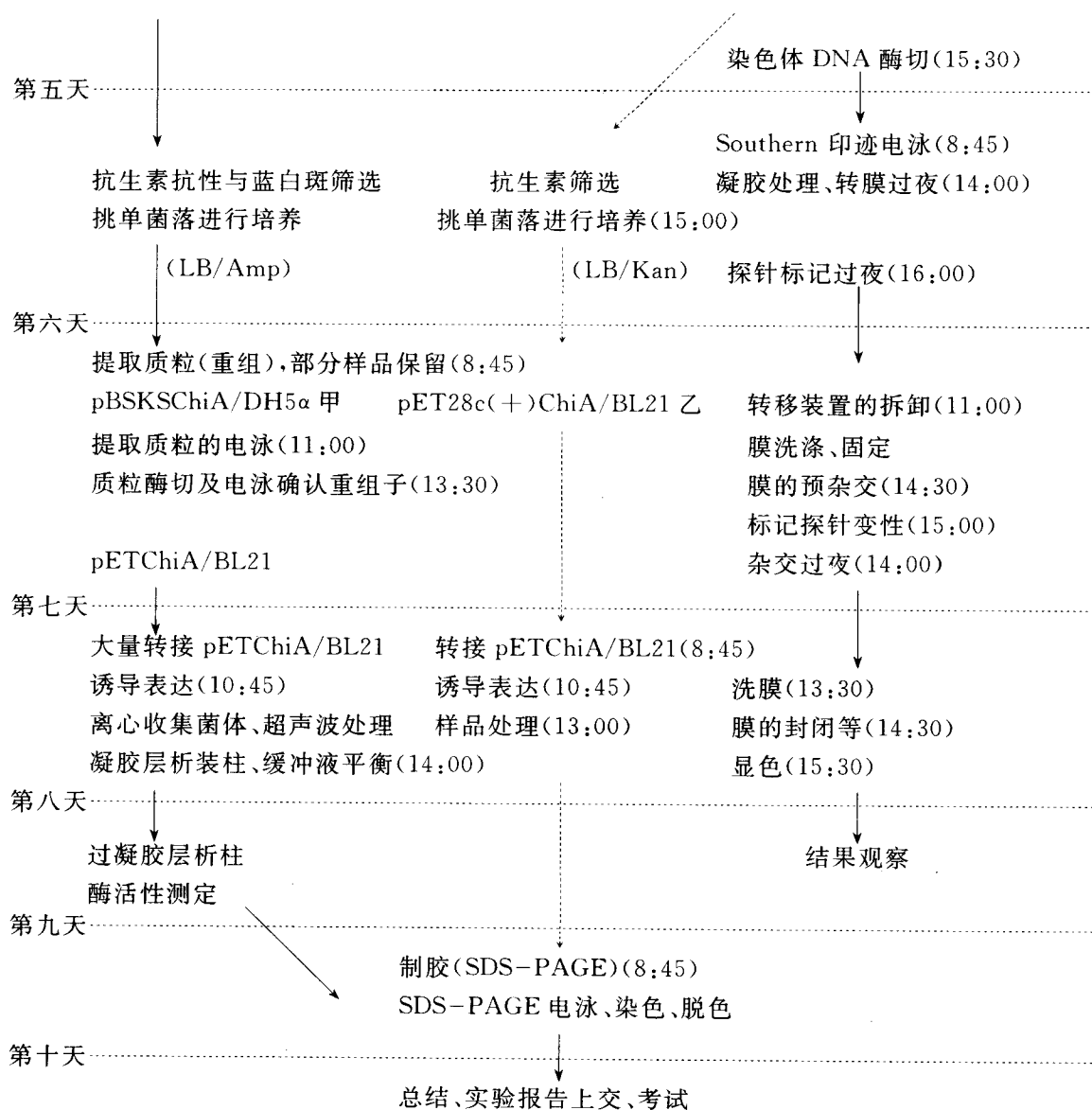
# 目 录

|                                 |     |                            |     |
|---------------------------------|-----|----------------------------|-----|
| 实验计划表 .....                     | 1   | 6-2 测序与序列同源性分析 .....       | 122 |
| 导论 .....                        | 4   | 6-3 Southern 印迹 .....      | 127 |
| 1 基因文库的构建 .....                 | 18  | 7 表达质粒的构建与诱导表达 .....       | 134 |
| 1-1 染色体 DNA 的提取 .....           | 18  | 7-1 表达质粒的构建 .....          | 134 |
| 1-2 总 RNA 和 mRNA 的提取 .....      | 25  | 7-2 Northern 印迹 .....      | 140 |
| 1-3 $\lambda$ 噬菌体 DNA 的提取 ..... | 33  | 7-3 外源基因的诱导表达 .....        | 145 |
| 1-4 染色体 DNA 的部分消化 .....         | 38  | 7-4 SDS-PAGE .....         | 149 |
| 1-5 基因组文库的构建 .....              | 40  | 7-5 Western 印迹 .....       | 155 |
| 1-6 cDNA 合成 .....               | 44  | 7-6 酶联免疫吸附测定 (ELISA) ..... | 160 |
| 1-7 cDNA 文库的构建 .....            | 49  | 7-7 蛋白质生物功能测定法 .....       | 162 |
| 2 目的基因的获得 .....                 | 55  | 8 工程菌的培养与目标产物分离 .....      | 164 |
| 2-1 PCR 扩增及 RT-PCR .....        | 55  | 8-1 发酵条件的优化 .....          | 164 |
| 2-2 核酸电泳 .....                  | 64  | 8-2 诱导表达蛋白质的分离沉淀 .....     | 166 |
| 2-3 核酸探针的标记及检测 .....            | 69  | 8-3 凝胶过滤层析 .....           | 172 |
| 2-4 核酸探针筛选基因文库 .....            | 77  | 8-4 离子交换层析 .....           | 180 |
| 3 载体质粒的制备 .....                 | 82  | 8-5 蛋白质浓度的测定 .....         | 186 |
| 3-1 质粒 DNA 的提取与纯化 .....         | 82  | 8-6 酶活性的测定 .....           | 189 |
| 3-2 核酸纯度、浓度与相对分子质量的测定 .....     | 90  | 8-7 分离纯化蛋白的收率计算与保存 .....   | 191 |
| 4 扩增质粒的构建 .....                 | 94  | 参考文献 .....                 | 193 |
| 4-1 限制性内切酶的酶切反应 .....           | 94  | 附录 1 简写 .....              | 194 |
| 4-2 凝胶电泳法进行 DNA 的分离、提纯 .....    | 100 | 附录 2 实验注意事项 .....          | 197 |
| 4-3 DNA 片段的体外连接 .....           | 103 | 附录 3 如何撰写实验报告 .....        | 199 |
| 5 重组 DNA 导入宿主细胞 .....           | 107 | 附录 4 常用的碱基、氨基酸符号及缓冲液 ..... | 201 |
| 5-1 感受态细胞的制备 .....              | 107 | 附录 5 大肠杆菌基因型 .....         | 204 |
| 5-2 重组质粒的转化 .....               | 111 | 附录 6 培养基与试剂的配制 .....       | 205 |
| 6 重组子的筛选与鉴定 .....               | 117 | 附录 7 实验仪器 .....            | 212 |
| 6-1 阳性克隆的筛选 .....               | 117 |                            |     |

# 实验计划表

## 方案一:连续 10 天的实验安排





实验结果汇总:

1. 染色体 DNA 提取
2. PCR
3. 质粒提取
4. 质粒的酶切
5. 割胶、过柱提纯
6. 连接、制备感受态细胞、转化、抗生素与蓝白斑筛选
7. 重组的扩增与表达质粒提取及酶切情况
8. 转化子的基因诱导表达

## 9. Southern 印迹

## 10. 凝胶层析

在有限的时间内(10 天为一周期),为了保证实验的完整性,两人一组(甲、乙)进行操作,其中一人构建扩增质粒(↓),另一人构建表达质粒(↓)。要求每个同学在实验过程中,在认真做好自己分内实验的同时,还应了解对方的实验内容。在实验完成后,能充分了解整个实验的设计方案。总之,通过完整的大实验体系,使同学们对基因工程中的分离、克隆、扩增、表达以及蛋白质的分离、纯化等一系列的实验过程加深认识,掌握实验方法,培养实验技巧,并对基因工程实验有一个全方位的了解。

### 方案二:分散实验的时间安排(如采用双休日的时间)

第一次实验:染色体 DNA 的提取(实验 1-1)、电泳确认(实验 2-2)、DNA 含量的测定(实验 3-2)

第二次实验:PCR 扩增目的基因(实验 2-1)、电泳确认(实验 2-2)、PCR 产物的纯化(实验 2-1)

第三次实验:质粒 pBSSK 及 pET28c 的提取(实验 3-1)、电泳确认(实验 2-2)

第四次实验:PCR 产物、pBSSK 及 pET28c 质粒的酶切(实验 4-1)、电泳确认、凝胶过柱纯化(实验 4-2)、电泳确认(实验 2-2)

第五次实验:连接反应(实验 4-3、7-1)、感受态细胞的制备(实验 5-1、7-1)、重组质粒的转化(实验 5-2、7-1)

第六次实验:抗生素与蓝白斑筛选、菌落的重组扩增质粒及表达质粒的提取(实验 6-1)、酶切、电泳确定(实验 7-1)

第七次实验:阳性重组表达质粒菌株的诱导表达(实验 7-3)、SDS-PAGE 检测(实验 7-4);大量诱导表达。

第八次实验:过凝胶层析柱(实验 8-3)、酶活的测定(实验 8-6);Southern 印迹(酶切)(实验 6-3)

第九次实验:Southern 电泳、凝胶、转膜(实验 2-3,实验 6-3)

第十次实验:杂交、显色(实验 6-3);总结、实验报告上交、考试

为了保证实验的完整性,实验也是分成甲乙两人一组进行,其中一人构建扩增质粒,另一人构建表达质粒。要求每个同学在实验过程中,在认真做好自己分内实验的同时,还应了解对方的实验内容。在实验完成后,能充分了解整个实验的设计方案。

虽然这个大实验的设计是分时间段进行的,但从染色体 DNA 的提取,到目的基因的扩增,从载体质粒的制备到扩增质粒、表达质粒的构建,再到重组子的导入鉴定、目的基因的诱导表达,最后进行基因的检测、蛋白质的分离,整个实验却是连续的、一环紧扣一环。所以每次实验结束后应注意妥善保管各自的实验产物,以便为下次实验提供材料保证。并且有些实验需要前一天进行接种等实验准备。而杂交实验也可以根据具体情况从 3 个杂交实验(Southern 印迹、Northern 印迹、Western 印迹)中任意选做一个。

# 导 论

我们生活的地球大约有 45 亿年的历史,而原始生物体(厌氧的异养生物)的诞生最早在 36 亿年前。从那以后,生物便以其顽强的活力世代繁衍,生生不息,发展成如今数以百万计、种类繁多、千姿百态的生命大千世界。尽管如此,生命活动的本质在不同生物体中,却是高度一致的,如构成 DNA 的 4 种核苷酸、组成蛋白质的 20 种氨基酸,以及核酸结构与蛋白质结构的对应关系,遗传密码等在整个生命世界中都是基本一致的,从而使不同生物的基因之间的转移和表达成为可能。

## 一、基因工程的研究进程

1860 至 1870 年孟德尔(G. J. Mendal)根据豌豆杂交实验提出遗传因子的概念,并总结出孟德尔遗传定律。

1909 年丹麦植物学家和遗传学家约翰逊(W. Johannsen)首次提出“基因(gene)”名词,用以表达孟德尔的遗传因子。

1944 年三位美国科学家埃弗利(O. T. Avery)以及麦克劳德(C. M. Macleod)、麦卡蒂(M. McCarty)进行了细菌转化的研究,证实生物体内遗传物质的基础就是 DNA。DNA 以其特定的结构与组织编排,以基因组的形式存在于细胞中。随后又发现一些基因组小、结构简单的生物,如病毒是以 RNA 作为遗传物质。

1953 年沃森(J. D. Watson)和克里克(F. H. C. Crick)在总结前人威尔金斯(M. Wilkins)等人的 X 射线工作的基础上,提出了 DNA 双螺旋结构的立体模型,并荣获 1962 年诺贝尔生理学医学奖。

1966 年尼伦伯格(Nirenberg)和科拉纳(Khorana)等破译了全部的遗传密码,编制了一本震惊世界的完整的遗传密码“辞典”——三联体密码,并进一步确认了遗传密码具有的特性:① 通用性;② 简并性;③ 偏倚性;④ 不重叠性和阅读方向性。

1967 年世界上 5 个实验室几乎同时发现了 DNA 的连接酶。

1970 年史密斯(H. O. Smith)、威尔考克斯(K. W. Wilcox)等从流感嗜血杆菌分离出第一种核酸内切酶 *Hind*Ⅲ,它可以在特定的位点将 DNA 分子切割开。

1972 年美国斯坦福大学的伯格(P. Berg)和杰克森(Jackson)等人将猴病毒基因组的 SV40 DNA 和  $\lambda$  噬菌体的 DNA 分别进行酶切消化,然后用 T4 连接酶进行连接,得到了第一个重组的 DNA 分子。1980 年伯格(P. Berg)获得了诺贝尔化学奖。

1973 年斯坦福大学的科恩(S. Cohen)等人在体外构建出含有四环素和链霉素两个抗性基因的重组质粒分子,并将其导入大肠杆菌中,获得了双抗性的大肠杆菌转化子(transformant)。完成了第一个细菌基因的克隆,标志着基因克隆技术的诞生。

从此基因工程的发展是日新月异,不仅在基础研究领域,还是在生产实际应用方面都取得了惊人的成绩。

基因工程(gene engineering)或称基因操作(gene manipulation)、重组 DNA 技术(recombinant DNA technology)是 20 世纪 70 年代从分子遗传学基础上发展起来的一门新技术,它综合采用了生物化学、遗传学和微生物学等学科的现代技术,在体外试管内对大分子 DNA 进行剪切加工,再与不同亲本 DNA 分子重组(recombination),并把它引入受体细胞中去,通过复制(replication)、转录(transcription)、翻译(translation)以及表达(expression),使生物获得新的遗传性状。可以说重组 DNA 技术实质上是人类对生物功能的模拟,是人们获取、整理、破译、编辑和表达遗传信息(基因)的一种操作平台与技术。基因工程的两个基本特点是分子水平的操作、细胞水平的表达。

重组 DNA 技术具有以下 3 个特点:① 可以打破生物界种的界限,突破亲缘关系的限制,加快变异的程度和速度;② 可以定向变异、育种,定性改造生物;③ 可以创造出自然界中原本没有的生物。

一个典型的 DNA 重组实验通常包含以下几个步骤,如图 1 所示。

1) 通过 PCR 扩增等方法获取供体(donor)生物的目的基因(target gene)或称外源基因,用限制性核酸内切酶(restriction endonucleases)分别将外源 DNA 和载体(vector)分子进行切割。

2) 利用 T4 DNA 连接酶(T4 DNA ligase)将含有目的基因的 DNA 片段连接到另一个称克隆载体(vector)的 DNA 分子上,构成一个新的重组 DNA 分子。

3) 借助于细胞转化(transformation)等手段,将重组 DNA 分子导入受体(host)细胞,并在受体细胞中复制保存。

4) 对那些吸收了重组 DNA 的受体细胞进行筛选和鉴定。

5) 对含有重组 DNA 的细胞进行大量培养,检测外源基因是否表达。

重组 DNA 技术的主题就是外源基因的分离、克隆、扩增和表达。其中最关键是基因的克隆。

## 二、基因工程的四大要素

基因工程的四大要素为工具酶、载体、受体(宿主)以及供体(外源基因)。

### 1. 工具酶(enzymes in molecular biology)

基因工程的迅猛发展,与工具酶的发现是紧密相关的。

1957 年 A. Kornberg 在大肠杆菌中发现了可在试管中合成 DNA 的第一个工具酶——DNA 聚合酶(polymerase)。

1962 年 W. Arber 等人发现存在着一种可使未甲基化的 DNA 分子断裂的限制性酶的存在。

1967 年世界上有 5 个实验室几乎同时发现了 DNA 连接酶(ligase)。

1970 年 H. O. Smith, K. W. Wilcox, J. Kelley 分离到第一种核酸内切限制酶。此后又发现了反转录酶等。

重组 DNA 中常用的工具酶包括:① 核酸限制性内切酶(restriction endonuclease)和甲基化酶(methylase)。② 核酸外切酶(exonuclease)和内切酶。③ DNA 和 RNA 聚合酶(polymerase)。④ 核酸末端修饰酶。⑤ DNA 连接酶(ligase)。这些工具酶种类不同,功能各异:有的是“手术刀”,专司切割之职,如限制性内切酶、核酸外切酶和内切酶;有的是“缝纫线”,具有连接之

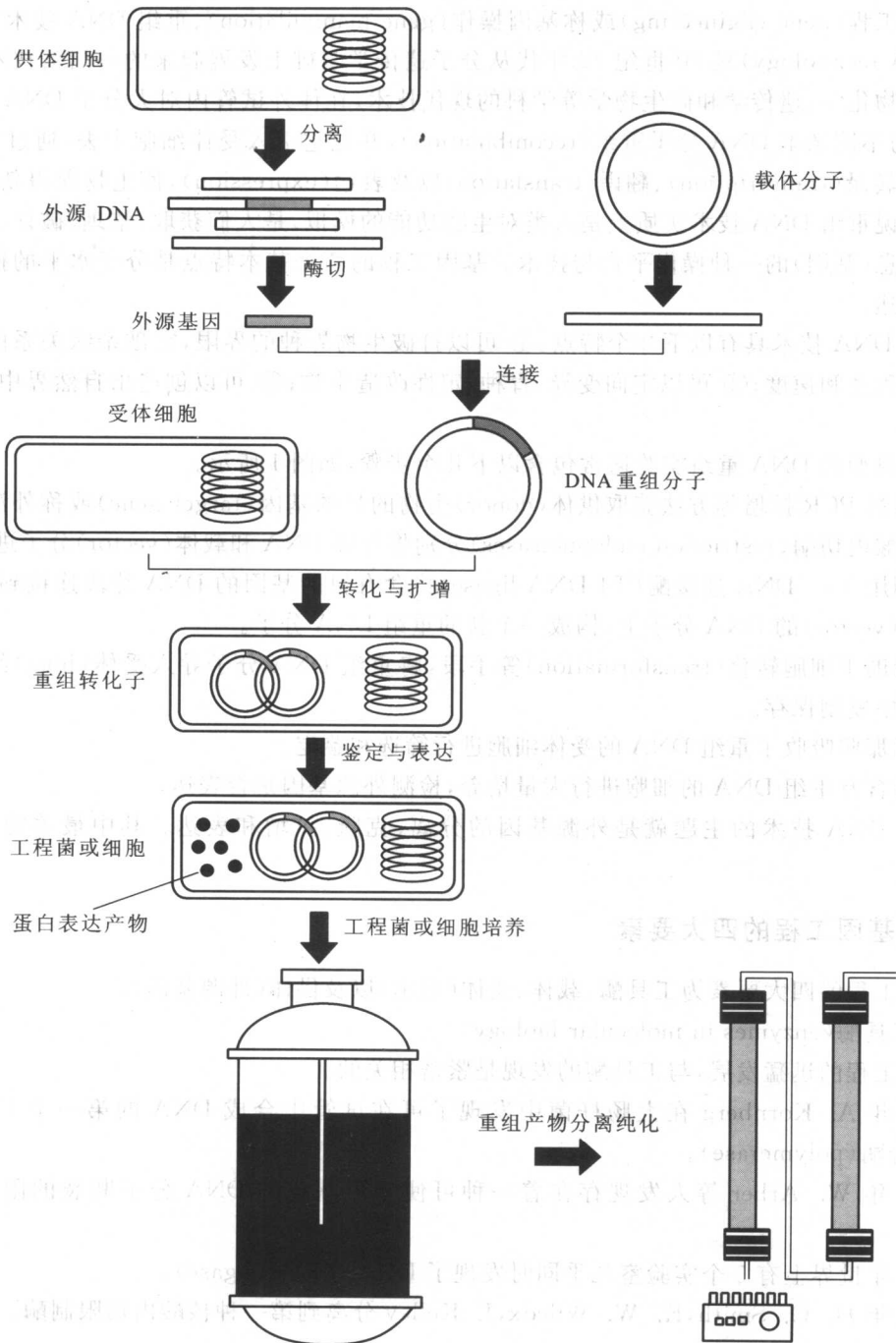


图 1 基因工程基本流程示意图(引自张惠展,1999)

功,如 DNA 连接酶;有的像“复印机”,行使复制之责,如 DNA 和 RNA 聚合酶;有的像“搬运工”,拥有末端转移之力,如核酸末端修饰酶等。其中大多数限制性内切酶和连接酶作用于双链

DNA,而不能作用于单链 DNA。在基因工程操作中最常用的工具酶是核酸限制性内切酶(restriction endonuclease)和 DNA 连接酶(ligase)。

## 2. 载体(vector)

外源基因往往不能独立地复制,需要借助某种运载工具将其引入宿主细胞中进行克隆、保存或表达。作为基因载体(vector)的必备条件:① 具有有效的运载能力和遗传标记基因;② 具有 DNA 复制起点(origin),携带外源性目的基因前后均能在宿主内自主复制;③ 在宿主内控制外源基因的表达活动;④ 具有多克隆位点 MCS(multiple cloning site),能携带大小不同的外源基因,与外源 DNA 连接时不影响载体的正常复制与扩增;⑤ 鉴定方便,装卸手续简便,安全可靠。

根据宿主细胞的不同,主要可以将载体分成五大类:细菌质粒(plasmid)载体(如 pUC、pET 系列等)、噬菌体(phage)载体、酵母穿梭载体(shuttle vector)、植物克隆载体和动物克隆载体。其中细菌质粒是基因工程中最常用的载体,大小为 1~200 kb,相对分子质量较小,因此易于在宿主间转移和迁移。但是,天然的质粒并不适宜作为载体,目前使用的质粒大多是在天然质粒的基础上进行改造,删除一些非必要的区段以及对宿主有不良影响的区段,削减载体的相对分子质量,使载体具有更大的容纳外源片段的能力。再加上容易选择或检测的标记,加上一些调控元件,有利于克隆基因的表达,并对限制性内切酶的酶切位点进行改造,便于外源基因插入到载体中的特定位置。

## 3. 受体(宿主,host)

为了保证外源基因在细胞中的大量扩增和表达,选择合适的克隆载体的宿主就成为基因工程重要问题之一。

一个理想宿主(host)的基本要求是:① 转化亲和性,受体细胞必须能够高效吸收外源重组 DNA 分子,即细胞易于形成感受态;② 具有使外源 DNA 进行高效复制的酶系统,能高水平表达目的基因;③ 遗传互补型,受体细胞必须具有与载体所携带的选择标记互补的遗传性状,方能使转化细胞的筛选成为可能;④ 限制缺陷型,为了避免重组 DNA 被摄入细胞后被菌体内的 DNA 限制酶的降解,保证重组质粒分子能够以它的原始状态存在,必须选用限制系统或限制与修饰系统均为缺陷型的宿主菌,即外切酶和内切酶活性缺陷型(RecB<sup>-</sup>, RecC<sup>-</sup>, HsdR<sup>-</sup>);⑤ 重组整合缺陷型,用于基因扩增或高效表达的受体细胞一般采用重组缺陷型(RecA<sup>-</sup>)菌株,使克隆载体 DNA 与宿主染色体之间不发生同源重组;⑥ 能有效地分泌和正确地修饰基因产物,使其具有正确的构象和活性;⑦ 培养方便,便于进行基因操作和筛选;⑧ 感染寄主缺陷型,从安全角度上考虑,受体细胞不能具有感染寄生性,即宿主细胞应该对人、畜、农作物无害或无致病性等,防止重组细菌扩散污染。

由于原核生物培养成本低,生长快,表达量高,基因操作方便,是目前多肽药物基因工程的主要表达系统。特别是大肠杆菌(*Escherichia coli*)DH5 $\alpha$ 、DH10B、BL21(DE3)等菌株,遗传背景清楚,载体受体系统完备,生长迅速,培养简单,重组子稳定,适用于外源 DNA 的扩增和克隆、原核生物基因的高效表达。然而原核表达系统也有问题,如原核生物中的表达的蛋白质不能糖基化,易形成包涵体(inclusion body),表达的产物易被蛋白水解酶降解,或不能正确加工折叠。

而能够糖基化的表达系统必须是真核细胞,其中两类系统很有希望:一是酵母表达系统,二是昆虫表达系统。酵母菌如毕赤酵母(*Pichia pastoris*)、酿酒酵母(*Saccharomyces cerevisiae*)具有真核生物的特征,遗传背景清楚,生长迅速,培养简单,外源基因表达系统完善,遗传稳定,内源

性蛋白产物种类繁多且含量高,能够糖基化,适用于真核生物基因高效表达、基因文库构建、真核生物基因表达调控的研究。但酵母的蛋白质糖基化的类型和程度与哺乳动物并不相同。而昆虫细胞具有真核生物的特征,外源基因表达量高,繁殖相对较快,培养比较容易,成本低廉,遗传稳定,糖基化的类型和程度则与哺乳动物更为相似,适用于真核生物基因的高效表达,目前备受人们关注。但其 DNA 重组操作系统还欠完善。

#### 4. 供体的目的基因

所谓目的基因(target gene, 又称靶基因)主要是指希望分离或克隆的基因。根据不同的要求与特性,人们所需的目的基因(基因源)是不同的,并且随着研究的进一步深入,对目的基因的认识无论从数量上还是从内涵上都在不断深化。目前世界面临的能源、粮食、人口、资源以及环境污染等严重问题,人们力图利用生物技术的方法加以解决。如干扰素基因、胰岛素基因、生长激素基因、动物抗体基因、固氮基因、光合作用的基因、抗病虫害基因、控制果实成熟的基因、改变花型花色的基因和氨基酸编码基因等。

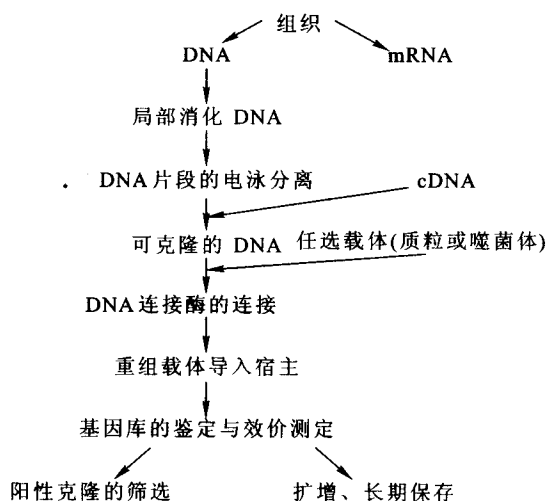


图 2 基因文库的构建

在重组 DNA 技术中,首先需要将某种生物的全部基因组的遗传信息,贮存在可以长期稳定保存的重组体中,以备需要时随时能够使用。基因文库(gene library 或 gene bank):是指通过克隆方法保存在适当宿主(如细菌或噬菌体)中的一群混合的 DNA 分子,所有这些分子中的插入片段的总和,可代表某种生物的全部基因组序列或 mRNA 序列。因此某一生物体的基因文库实质就是一个基因银行。

构建基因文库一般以改造的噬菌体作为载体(也有用细菌质粒作为载体的),具体步骤如下:① 载体 DNA 的制备;② 高相对分子质量染色体 DNA 及基因片段的制备,或者提取某一特定类型细胞表达的 mRNA 经反转录酶催化形成与之互补的 cDNA;③ 载体与 DNA 基因片段的体外连接;④ 包装蛋白的制备、重组体的体外包装;⑤ 将重组 DNA 分子导入宿主,重组克隆的筛选、鉴定;⑥ 文库的扩增、保存。(图 2)

构建的基因文库应具备代表性与完整性,基因文库的代表性是指文库中包含的重组 DNA 分子要全面地反映出供体细胞中表达的全部信息。文库的代表性可用一个量化的指标来衡量,即基因文库的库容量,它是指基因文库中所包含的独立的重组子克隆数。具备完好代表性的基因文库需要的库容量取决于供体细胞中基因总序列或表达出来的基因序列的总复杂度。

一般来说,一个基因文库应包含的克隆数与生物基因组的大小和被克隆 DNA 片段的长度有关。基因组越大,所需克隆数就越多;克隆时每个载体中允许插入的外源 DNA 片段越长,则所需克隆数越少。如果一个基因文库中的总的克隆较少,则从中筛选特异基因就比较容易,但由于插入片段较长,以后的分析就相对比较困难。

根据 Charke-Carbon 公式可以计算某一基因文库中应该包含的克隆(任何特定 DNA 序列)

数,也就是满足最低要求的基因文库的库容量。

$$N = \ln(1-P) / \ln(1-a/b)$$

$P$  是文库中包含供体细胞中任何 DNA 序列信息的概率(probability),即希望获得的概率(通常设为 99%),代表了基因文库的完备性。

$N$  是重组子数(基因文库大小、库容量),表示文库中以  $P$  概率出现在细胞中某段 DNA 序列理论上应具有的最少重组子克隆数。

$a$  是文库中每个重组子所含外源 DNA 片段的平均长度。

$b$  是生物基因组的大小,即单倍体基因组 DNA 的长度。

如要建立人的基因组( $3 \times 10^9$  碱基对)文库,以 20 kb 片段克隆进噬菌体中,要得到一个特定 DNA 序列的概率为 99%( $P=0.99$ )所需要的重组子数为:

$$N = \ln(1-0.99) / \ln(1-2 \times 10^4 / 3 \times 10^9) = 1.5 \times 10^5$$

即最少需要  $1.5 \times 10^5$  重组  $\lambda$  噬菌体。实际上将人的 DNA 片段随机组合进载体中确认 90%~95% 的 DNA 存在需要  $10^6$  个独立的重组  $\lambda$  噬菌体。

在构建 cDNA 文库时,要获得某一种低浓度特异 mRNA 的 cDNA 克隆,则可根据以下公式计算: $N = \ln(1-P) / \ln(1-1/n)$ 。其中的  $1/n$  是稀有 mRNA 在总 RNA 中所占的相对比例。如要获得成人纤维细胞中低丰度 mRNA(小于 14 拷贝/细胞),这种 mRNA 占 mRNA 总数的 30%,约有 11 000 种不同的 mRNA 属于这个范围。

$$N = \ln(1-P) / \ln(1-1/n) = \ln(1-0.99) / \ln(1-0.3/11\ 000) = 1.7 \times 10^5$$

则需要构建 17 万个不同重组克隆的基因文库。

在实际工作中,由于在基因文库的构建过程中存在多种操作和系统误差,一个具备完好代表性的人的基因文库至少应具备  $10^6$  个以上的库容量。

除了代表性外,文库中基因或 cDNA 片段的序列完整性(所含基因是否完整、是否是全长)也是反映文库质量的一个重要因素。如在细胞中表达出的各种 mRNA 都是由三个部分组成,即 5'端非翻译区,中间的编码序列和 3'端非翻译区。要从文库中分离获得目的基因完整的序列信息和功能信息,就要求文库中的重组 cDNA 片段应尽可能完整反映天然基因的结构。所以一个理想的基因文库应具备下列条件:① 重组克隆的总数不宜过大,以减轻筛选工作的压力;② 载体的容量必须大于绝大多数基因长度,以免基因被分隔在不同的克隆中;③ 含有相邻 DNA 片段的重组克隆之间,必须具有部分序列的重叠,以便于基因文库各克隆的排序;④ 克隆片段易从载体分子上完整卸下且最好不带有载体序列;⑤ 重组克隆应能稳定保存、扩增及筛选。

目前常用构建基因文库的载体为  $\lambda$  噬菌体和质粒。

1)  $\lambda$  噬菌体( $\lambda$  phage):构建基因文库中最早使用的载体,一般置换型噬菌体适用于基因组 DNA 的克隆,插入型噬菌体用于全长 cDNA 的克隆。重组子经专门的体外包装系统包装成具有感染能力的噬菌体颗粒,对宿主大肠杆菌的转染效率高,如通常以  $1\ \mu\text{g}$  cDNA 构建出的 cDNA 文库,转染宿主菌后,都可得到  $10^6 \sim 10^7$  以上的原始库容量,而是否含有 cDNA 文库可以通过对转染细胞的表型鉴定如蓝白斑分析,较为简单准确地反映出文库中含有插入片段的实际情况,有较好的质量控制指标。因此用这类载体系统构建出的基因文库质量高、代表性好。并且利用这

类载体构建的基因文库是以重组噬菌体颗粒形式存在,这些噬菌体颗粒的感染活性在 4℃ 环境中极其稳定,非常适合基因文库的长期保存。

2) 质粒(plasmid):使用质粒载体系统构建基因文库的最大特点是能够实现对基因文库(cDNA 文库)的功能性筛选。所谓功能性筛选是利用基因在体内体现其生物学活性所依据的生化基础,从文库中分离鉴定出目的基因的一种文库筛选策略。

利用质粒载体构建基因文库存在的主要缺点是插入片段的载体容量较小,含有大片段的重组克隆在文库的群体扩增过程中容易丢失,因此文库中包含全长 cDNA 或目的基因的克隆比例少。用质粒构建的基因文库常用转化方式导入宿主细胞中,转化效率普遍较低,如用 1 μg cDNA 构建出的质粒文库,经转化得到的原始克隆数在  $10^6$  以下。因此文库的库容量相对较小。此外质粒文库需要以活的转化菌形式存在和扩增,保存条件较为严格,而且保持过程中,部分转化菌会死亡,因此不适合文库的长期保存。

### 三、基因克隆的策略

虽然人们已经鉴定出生物的某些性状或疾病的基因,但常常不易获得目的基因。因为各种目的(靶)基因分布在生物的染色体上,染色体上基因存量非常庞大,原核生物的 DNA 分子平均有  $10^6$  个碱基对,基因多的可达数千种;真核生物的 DNA 分子可达  $10^9$  个碱基对,基因多至上万种。而基因工程中所需要的目的基因都仅有几千个或几百个碱基对。因此要在如此庞大的基因库中分离出绝对量极少、单拷贝的目的基因,难度不亚于大海捞针。但尽管如此,设计合适的基因克隆策略还是可以分离获得目的基因的。

要成功地分离目的基因必须具备以下条件:① 目的基因必须表型明确;② 单个目的基因编码优良性状基因或疾病基因;③ 目的基因是一对等位基因控制的显性基因;④ 容易获得目的基因的突变体和重组体。

基因克隆的方法:包括① PCR(polymerase chain reaction)扩增法;② 核酸探针筛选法;③ 免疫反应筛选法;④ 通过酶活性筛选;⑤ 差示杂交法;⑥ DNA 标签法;⑦ 染色体步移法(chromosome walking);⑧ 作图克隆法;⑨ 酵母双杂体系等。

#### 1. PCR(polymerase chain reaction)合成法

1) 已知某一基因的序列或基因 5' 和 3' 两端的序列,则可以设计引物,进行基因的 PCR 合成(实验 2-1)。

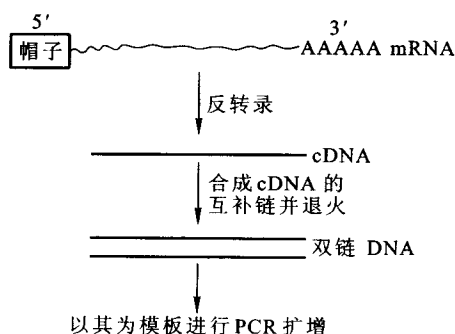


图 3 反转录 PCR

2) 反转录 PCR(reverse transcription PCR, RT-PCR):这种方法指的是以 RNA 为模板,经反转录获得与 RNA 互补的 DNA 即 cDNA,然后以 cDNA 链为模板进行 PCR 反应(图 3)。

3) 反向 PCR(inverse PCR):特别适用于扩增已知序列两端的未知序列。用相应的限制性内切酶水解后,将酶切的片段在连接酶的作用下环化,使得已知序列位于环状分子上。根据已知序列的两端序列设计两个引物,以环状分子为模板进行 PCR,就可以扩增出已知序列的两侧的未知序列(图 4)。

4) RACE 法(rapid amplification of cDNA ends):特别适用于扩增那些已知部分序列的目的基因。RACE 法又可分为 5'RACE 和 3'RACE 法。如果已知部分 cDNA 序列,就可以以已知的碱基序列为基础进行 PCR,扩增至 cDNA 两末端。如果未知区域是 mRNA 的上游区域即 5',称为 5'RACE。如果是 mRNA 下游的 3'端,则为 3'RACE。

3'RACE 法:此方法中需要设计 4 种两对引物,即 5'端带有 oligo dT 的适合引物 AP1,5'端与 AP1 相同但不含有 oligo dT 的适合引物 AP2 以及利用已知序列设计的引物 GSP1(gene specific primer)、GSP1 下游的已知序列设计的引物 GSP2。

在设计 5'端带有 oligo dT 的适合引物 AP1 时应注意这种引物的 3'端是 oligo dT,与模板结合状态是非常不稳定的,所以设计时应考虑到  $T_m$  值要保持在 60℃ 以上较好。此外还应在引物的靠近 oligo dT 部位上设计适当的酶切位点。

如 Pharmacia 的 Ready-To-Go 的 cDNA 第一链合成试剂盒中具有 *Not* I 的适合引物:

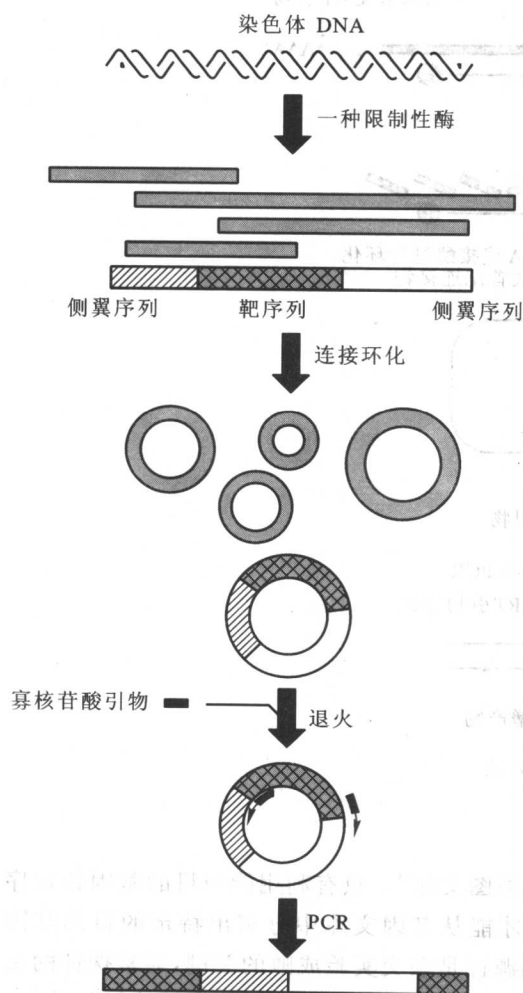


图4 反向 PCR

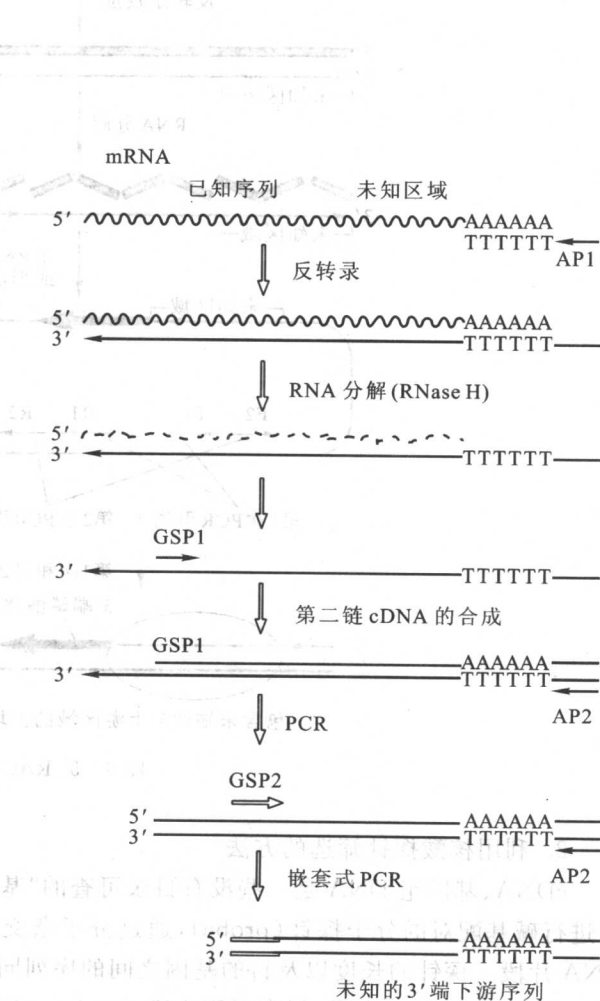


图5 3'RACE 法