

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

C12N 15/67

C12N 15/55 C12P 21/02

C12N 9/16 C12N 1/21

/(C12N21/02, C12R1 : 19)

(C12N9/16,

C12R1 : 19)(C12N1/21,

C12R1 : 19)

# [12] 发 明 专 利 说 明 书

[21] ZL 专利号 97190500.2



[45] 授权公告日 2003 年 12 月 24 日

[11] 授权公告号 CN 1131889C

[22] 申请日 1997.3.10 [21] 申请号 97190500.2

[30] 优先权

[32] 1996.3.13 [33] JP [31] 85801/1996

[32] 1996.7.18 [33] JP [31] 208897/1996

[86] 国际申请 PCT/JP97/00748 1997.3.10

[87] 国际公布 WO97/34006 日 1997.9.18

[85] 进入国家阶段日期 1998.1.8

[71] 专利权人 宝酒造株式会社

地址 日本京都府京都市

[72] 发明人 佐川裕章 上野春美 大岛淳

加藤郁之进

[56] 参考文献

US-5137823-A 1992.08.11 C12N9/22

审查员 李 博

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 王景朝

权利要求书 2 页 说明书 33 页 附图 26 页

[54] 发明名称 质粒

[57] 摘要

一种质粒载体, 其特征在于其具有能被非宿主内在的 RNA 聚合酶识别并控制目的基因表达的启动子序列和具有能在外源因子的诱导下增加拷贝数的复制起点; 利用该载体表达和分离目的基因的方法, 具有 AccIII 限制性内切核酸酶活性的多肽以及编码该多肽的 DNA。 本发明首次提供了导入的外源目的基因所编码蛋白质对所说宿主是致死性的或有毒性的质粒载体, 利用该载体有效表达所述蛋白质的方法。 而且提供了对现有技术来说是困难的, 在不存在修饰酶基因的情况下分离构成限制修饰系统的限制性内切核酸酶基因的方法。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种质粒载体, 其特征在于其具有能被非宿主内在的 RNA 聚合酶识别并控制目的基因表达的启动子序列和在外源因子的诱导下拷贝数增加的复制起点。
- 5       2. 权利要求 1 所述的质粒载体, 其中启动子序列可被来自噬菌体的 RNA 聚合酶识别。
3. 权利要求 2 所述的质粒载体, 其中的启动子序列能被来自 SP6 噬菌体的 RNA 聚合酶识别。
4. 权利要求 3 所述的质粒载体, 其中启动子序列含序列表中序列号  
10 30 所示的碱基序列。
5. 权利要求 1 所述的质粒载体, 其中复制起点在启动子的控制之下。
6. 权利要求 5 所述的质粒载体, 其中复制起点在 lac 启动子的控制之下。
- 15       7. 权利要求 1 所述的质粒载体, 其中含有作为选择标记的药物抗性基因。
8. 权利要求 1 所述的质粒载体, 其选自 pACE601、pACE611、pACE701、pACE702。
9. 一种质粒载体, 其中待表达的目的基因被整合进权利要求 1 的质  
20 粒载体中。
10. 一种目的基因的表达方法, 其特征在于向宿主中导入一种质粒载体和识别该质粒载体上的启动子序列的 RNA 聚合酶基因, 在所述质粒载体中待表达的目的基因被整合进权利要求 1 的质粒载体中; 并用外源因子诱导质粒载体拷贝数的增加和 RNA 聚合酶的表达以进行目的基因  
25 的转录和翻译。
11. 权利要求 10 所述的目的基因的表达方法, 其特征在于质粒载体拷贝数的增加和 RNA 聚合酶的表达是由不同的外源因子诱导的。
12. 权利要求 10 所述的目的基因的表达方法, 其特征在于质粒载体拷贝数的增加和 RNA 聚合酶的表达是由同一外源因子诱导的。
- 30       13. 权利要求 10 所述的目的基因的表达方法, 其中诱导质粒载体拷贝数增加的外源因子选自异丙基- $\beta$ -D-硫代半乳糖苷 (IPTG) 的添加, 乳糖的添加, 半乳糖的添加, 阿拉伯糖的添加, 色氨酸浓度的降低

及转化子培养温度的调整其中之一。

14. 权利要求 10 所述的目的基因的表达方法，其中诱导 RNA 聚合酶表达的外源因子选自如下之一：异丙基- $\beta$ -D-硫代半乳糖苷 (IPTG) 的添加，乳糖的添加，半乳糖的添加，色氨酸浓度的降低以及
- 5 转化子培养温度的调整。

15. 权利要求 10 所述的目的基因的表达方法，其特征在于通过其它的质粒载体或噬菌体载体将 RNA 聚合酶基因导入到了宿主中。

16. 权利要求 10 所述的目的基因的表达方法，其特征在于 RNA 聚合酶基因被整合进了宿主的染色体中。

- 10 17. 权利要求 15 或 16 所述的目的基因的表达方法，其特征在于 RNA 聚合酶基因来自 SP6 噬菌体。

18. 权利要求 10 所述的目的基因的表达方法，其中目的基因编码对宿主为致死性的或有害的蛋白质。

- 15 19. 权利要求 10 的目的基因的表达方法，其特征在于用大肠杆菌作为宿主。

20. 一种分离目的基因的方法，其特征在于在目的基因的分离方法中应用了权利要求 1 的质粒载体。

21. 权利要求 20 所述的目的基因的分离方法，其中目的基因编码对宿主为可致死或有害的蛋白质。

- 20 22. 权利要求 21 所述的目的基因的分离方法，其中编码对宿主可致死或有害的蛋白质的基因是限制性内切核酸酶的基因。

## 质粒

## 技术领域

- 5 本发明涉及可在基因重组技术中使用的质粒载体，以及用该质粒载体表达基因的方法。本发明还涉及用该质粒载体分离目的基因的方法。另外，本发明还涉及可用作基因工程试剂的限制性内切核酸酶及其基因，更具体地说是 AccIII 限制性内切核酸酶及其 DNA。

## 背景技术

- 10 在用基因工程技术构建目的基因的表达系统时，将其放在被所用宿主 RNA 聚合酶识别的启动子的支配之下，以控制该基因的表达。但是，在编码对所述宿主有害的蛋白质的情况下，由于不能严格控制所用启动子的表达，有时质粒本身的构建由于基因产物的表达而变得不可能。

- 作为解决上述问题的表达系统，开发了有利用大肠杆菌作为宿主和感染大肠杆菌的 T7 噬菌体的 RNA 聚合酶的 PET 系统（Novagen 制）[分子生物学杂志(J.Mol.Biol.)189 卷，113 - 130 页（1986）；基因（Gene），56 卷，125 - 135 页（1987）]。pET 系统使启动子识别特异性及转录活性高的 T7 RNA 聚合酶在大肠杆菌中表达，并且该 T7 RNA 聚合酶使表达载体上 T7 启动子下游配置的目的基因转录和高表  
15 达。由于是在 T7 RNA 聚合酶的存在下进行目的基因的转录，只要该宿主不产生该聚合酶，不表达目的基因的情况下也可在宿主中进行质粒构建；如构建目的基因被置于由宿主 RNA 聚合酶所识别的启动子的控制之下这样的表达系统，则不同质粒构建本身不受到阻碍。

- 但是，由于 T7 RNA 聚合酶基因已被克隆进入噬菌体载体并溶原进  
25 表达宿主，宿主就不能自由选择，宿主的更换需要烦杂的操作。另外，由于宿主中 T7 RNA 聚合酶表达不受严格控制，非诱导状态的宿主内 T7 RNA 聚合酶也表达，结果引起配置在表达载体上 T7 启动子下游的目的基因在非诱导状态表达。为了使这种非诱导状态下目的基因的表达降低，使用 T7 RNA 聚合酶抑制剂即 T7 溶菌酶抑制 T7 RNA 聚合酶活性[分子生物学杂志(J.Mol.Biol)219 卷，37 - 44 页（1991）]；或在 T7 启动子  
30 的下游放置乳糖操纵子以防止 T7 RNA 聚合酶接触到 T7 启动子[分子生

物学杂志(J.Mol.Biol)219 卷, 45 - 59 页(1991)]。

但是, 这些措施对转录活性高的 T7 RNA 聚合酶不是很满意, 不能完全抑制非诱导状态下 T7 RNA 聚合酶的活性。因此, 在目的基因产物对宿主致死的情况下, 即使能构建质粒, 也不能制备用于表达该基因的转化子。总之, PET 系统中有两个问题需要解决, 即宿主不能自由变更, 和 T7 RNA 聚合酶的表达控制不准确。

另一方面, 当 T7 噬菌体性质类似的噬菌体有感染沙门氏菌属细菌(鼠伤寒沙门氏菌)的 SP6 噬菌体[科学(Science, 133 卷, 2069 - 2070 页(1961))]。SP6 噬菌体生产的 RNA 聚合酶是分子量约 10 万的单一肽, 由于其启动子识别特异性强转录活性也高, 常将其用于体外 RNA 合成[生物化学杂志(J.Biol.Chem)257 卷, 5772 - 5778 页(1982)。生物化学杂志(J.Biol.Chem)257 卷, 5779 - 5788 页(1982)]。另外, 已将 SP6 RNA 聚合酶基因克隆进了大肠杆菌内, 进行大量的表达[核酸研究(Nucl.Acids.Res)15 卷, 2653 - 2664 页(1987 页)]。

对宿主来说表达产物为致死性的基因的例子有限制性内切核酸酶基因。基本上, 限制性内切核酸酶被用于通过对进入产限制性内切核酸酶的微生物细胞中的噬菌体和其它外源 DNA 进行切割以进行自身防御。另一方面, 多数产限制性内切核酸酶的微生物产修饰酶, 该修饰酶如这些限制性内切核酸酶一样识别相同的碱基序列, 以保护其自身的 DNA 不受这些限制性内切核酸酶的切割。具体地, 修饰酶通过向所识别的碱基序列中的一个或多个碱基上加上甲基基团修饰 DNA, 使与识别与修饰酶相同序列的限制性内切核酸酶结合并切割该 DNA 成为不可能。该机制称为限制修饰系统, 构成限制修饰系统的限制性内切核酸酶和修饰酶基因对称为限制修饰系统基因。因此, 当限制性内切核酸酶基因在缺少限制修饰系统基因中的修饰酶基因的微生物中表达时, 微生物 DNA 被切割, 导致细胞死亡。实际上, 在 Mob I 限制修饰系统基因中有两个修饰酶基因; 已有报道, 由于在另一修饰酶基因的共存下的不完全甲基化使宿主 DNA 修饰不完全, 就不可能对限制性内切核酸酶基因进行克隆[核酸研究(Nucl.Acids.Res)21 卷, 2309 - 2313 页(1993)]。

还有, 如果在携带限制修饰系统基因的细胞中缺失限制修饰系统基因, 细胞中修饰活性的缺失导致基因组 DNA 的不完全甲基化, 如果是残留的非常小量的限制性内切核酸酶导致其自身基因组 DNA 的致死性断

裂〔科学 ( Science ) 267 卷, 897 - 899 页 ( 1995 ) 〕。总之, 在缺少构成限制修饰系统的修饰酶的情况下, 修饰酶代表对细胞非常有害的蛋白质; 用现有技术分别克隆和表达其基因还是不可能的。

就限制性内切核酸酶来说, 可根据酶特征分为三种类型: I、II 和 III。具体地, II 型限制性内切核酸酶, 每一种识别特定的 DNA 碱基序列并在特定的位点或附近切割该序列, 该型酶被广泛地用于基因工程领域, 并且有各种特征的限制性内切核酸酶已从各种微生物中分离出来〔核酸研究 ( Nucl.Acids.Res ) 24 卷, 223 - 235 ( 1996 ) 〕。在本说明书中, 下文中将 II 型限制性内切核酸酶称为“限制性内切核酸酶”。但是应该指出的是, 一些微生物产生很少量的限制性内切核酸酶, 而其它的微生物产生大量的限制性内切核酸酶。例如, 限制性内切核酸酶 Acc III 由乙酸钙不动杆菌 ( 下方称为 Acc 细菌 ) 产生, 该菌已于昭和 60 年 11 月 9 日 ( 最初保藏日期 ) 被以保藏号 FERM BP - 935 保藏于通商产业省工业技术院微生物工业技术研究所〔地址: 日本茨城县筑城市谷田部镇东 1 丁目 1 番 3 号 ( 邮区号 305 ) 〕, 但是产生的酶量少且该微生物还同时产生限制性内切核酸酶 AccI 和 AccII。因此, 需要更先进的技术用该微生物高纯度和低花费的提供作为试剂的限制性内切核酸酶 AccIII。在提供高纯度和低花费的作为试剂的限制性内切核酸酶时, 可有效地利用基因工程技术分离目的限制性内切核酸酶基因和选择性大量生产目的限制性内切核酸酶。为此目的, 已报道了一些分离限制性内切核酸酶基因的方法。

首先, 要提到“鸟枪”法, 其中将产限制性内切核酸酶微生物的基因组 DNA 用合适的限制性内切核酸酶切割, 将得到的片段插入到合适的质粒载体中, 并筛选一个表达该限制性内切核酸酶基因的克隆。筛选目的基因的方法可例举出这样一个方法, 其中以导入了限制修饰系统基因的宿主所需要的自身防御功能为基础, 以对噬菌体感染的抗性为指标, 分离了限制修饰系统的基因〔美国国家科学院院刊 (Proc.Natl.Acad.Sci) 78 卷, 1503 - 1507 页 ( 1981 ) 〕。但是, 在本方法中, 限制修饰系统基因的大小必须在允许分离的范围之内, 并且所分离的限制修饰系统基因的表达表现出对噬菌体的足够的抗性, 以允许宿主的选择性存活。另一方面, 作为限制修饰系统基因的一般特征, 应该提到在基因组上限制性内切核酸酶和修饰酶基因的较近的定位; 实际上, 在已获得的许多限制

修饰系统基因中得到了确认〔核酸研究 ( Nucl.Acids.Res ) 19 卷, 2539 - 2566 ( 1991 ) 〕。因此, 有这样一种方法, 其中在上述特征的基础上, 以修饰酶基因的表达作为指标筛选限制修饰系统基因〔特开昭 63 - 87982 核酸研究 ( Nucleic.Acids.Res ) 19 卷, 1831 - 1835 页 ( 1991 ) 〕。  
5 但是, 当限制性内切核酸酶基因不与修饰酶基因接近时, 该方法不能产生限制性内切核酸酶基因。

而且, 上述的“鸟枪”法有一个与基因组 DNA 来源生物体与宿主之间转录 - 译差异相关的基本问题。例如, 由于伴随限制修饰系统基因启动子和核糖体的结合位点在宿主中很好的起作用的失败使基因不能充分  
10 表达时, 即使得了也需要大量的劳动选择转化子。为了避免该缺点, 有这样一种方法, 其中分析了限制性内切核酸酶蛋白质的氨基酸序列, 在所得序列的基础上以 PCR 为基础的 DNA 扩增方法, 限制性内切核酸酶基因从产限制性内切核酸酶的微生物的基因组 DNA 得到, 并且其中利用了已知的蛋白表达系统〔特开平 6 - 277070 〕。由于在常规的蛋白表达  
15 系统中限制性内切核酸酶的存在对宿主是致死的, 就需要通过例如让构成限制修饰系统的修饰酶与该限制性内切核酸酶共存在以保护宿主。

尽管在上述的限制性内切核酸酶基因的分离方法中, 在分离该基因的同时, 还必须分离构成限制修饰系统基因的修饰酶基因, 也已报道了只分离限制性内切核酸酶基因的分离方法〔核酸研究  
20 ( Nucleic.Acids.Res ) 22 卷, 2399 - 2403 页 ( 1994 ) 〕。但是, 在该方法中, 为了分离编码的限制性内切核酸酶, 其酶活性的最适温度为约 70 °C 的基因, 当要分离的编码在接近培养温度时有高的比活性的限制性内切核酸酶的基因时, 修饰酶基因的共存是必需的。

例外情况, 象限制性内切核酸酶 DpnI 一样, 除非其 DNA 识别序列  
25 中的特定的核酸碱基没有被甲基化修饰, 其限制性内切核酸酶不表现出切割活性。具有该特性的限制性内切核酸酶基因被认为是例外是因为, 即使在缺少另一个特定的基因时, 通过选择合适的宿主生物体也可被分离。实际上, 已分离了编码 Mrr 蛋白质的 Mrr 基因, 该 Mrr 蛋白质不是 II 型限制性内切核酸酶并识别含甲基化核酸碱基的特定的碱基序列并具  
30 有 DNA 切割活性〔细菌学杂志 ( J.Bacteriol ) 173 卷, 5627 - 5219 页 ( 1991 ) 〕。

如上所述, 除了特殊情况, 用现有技术分离限制性内切核酸酶基因,

必须同时表达该基因和构成限制修饰系统基因的修饰酶基因。

#### 发明公开

因此，本发明的第一个目的是提供能分离此类基因并能导入宿主中有效表达蛋白质的质粒载体，因为该基因产物对宿主是致死的或有害的，因此该基因的分离或表达系统的构建在现有技术中是有困难的。本发明的第二个目的是提供用该质粒载体表达目的基因的方法。本发明的第三个目的是提供用该质料载体分离目的基因的方法，特别是在不需要构成限制修饰系统的修饰酶基因的共存下分离限制性内切核酸酶基因。本发明的第四个目的是提供具有 AccIII 限制性内切核酸酶活性的多肽。另外，本发明的第五个目的是提供编码具有 AccIII 限制性内切核酸酶活性多肽的基因。

首先，为了解决 pET 系统中的问题，本发明人已构建了用 SP6 RNA 聚合酶作为新的精确的表达控制系统的表达系统并评估了该系统。

由于该表达载体是用将 SP6 RNA 聚合酶基因导入了 miniF 质粒的系统质粒构建的，因此可通过向宿主中同时导入携带有克隆在 SP6 启动子下游的目的基因的表达载体和该系统质粒而将目的基因的表达系统构建在各种大肠杆菌菌株中。另外，应用 lac 启动子和反义技术，本发明人使在系统质粒上精确控制 SP6 RNA 聚合酶基因的表达成为可能。用  $\beta$ -半乳糖苷酶基因作为报告基因评估该表达系统表明在该系统中 SP6 RNA 聚合酶基因受到精确控制，并且在非诱导条件下几乎没有表达，并且在诱导之后有足够量的 SP6 RNA 聚合酶，以有效地表达目的基因和随后高水平地表达该目的基因。

但是，已证明当所用的宿主是大肠杆菌时，由于 SP6 启动子非常弱但可被大肠杆菌 RNA 聚合酶精确识别，在非诱导条件下 SP6 启动子下游的目的基因表达量非常小。因此证明当基因产物对大肠杆菌非常有害和可致死时，表达系统的构建是不可能的，因此本发明的目的无法很好的实现。

有了这些发现，本发明进行了更广泛的调研并意外地发现(i) 非诱导条件下目的基因的表达可被抑制到不可检测的水平，以及(ii) 表达诱导增加了含目的基因的质粒的拷贝数并引起 RNA 聚合酶的表达，导致置于可被该 RNA 聚合酶识别的启动子下游的目的基因的转录和翻译，这是通过结合两种控制方法用新的控制目的基因的表达的系统进行的，即基因

拷贝数的控制和通过启动子对转录进行控制。换句话说,本发明的质粒载体是具有独特的解决上述基因工程领域中问题的质粒载体。本发明人  
在此发现的基础上进行了更进一步的研究,并开发了用该载体表达目的  
基因的方法。

- 5 本发明人还开发了应用上述质粒载体分离其表达产物对宿主致死的  
基因的方法,更具体地说是一种分离限制性内切核酸酶基因的方法,但  
无现有技术缺点。

本发明人还发现在没有 AccIII 修饰酶的共存下,能够分离以前尚未  
得到过的编码 AccIII 限制性内切核酸酶的 DNA,并能用上述的方法大量  
10 表达 AccIII 限制性内切核酸酶。

即本发明的要点有:

(1) 一种质粒载体,其特征在于包括一种控制目的基因表达的启动子  
序列,该启动子序列可被非宿主内在的 RNA 聚合酶识别,和一个在外源  
因子的诱导下可增加拷贝数的复制起点;

- 15 (2) 上述(1)中所述的质粒载体,其中启动子序列由来源于噬菌体的  
RNA 聚合酶所识别;

(3) 上述(2)中所述的质粒载体,其中启动子序列被来源于 SP6 噬菌  
体的 RNA 聚合酶识别;

- 20 (4) 上述(3)中所述的质粒载体,其中启动子序列含序列表中 SEQ  
ID NO: 30 所示的碱基序列;

(5) 上述(1)到(4)任一项所述的质粒载体,其中复制起点在启动子的  
控制之下;

(6) 上述(1)-(5)任一项所述的质粒载体,其中复制起点在 lac 启动子  
的控制之下;

- 25 (7) 上述(1)-(6)任一项所述的质粒载体,包括作为选择标记的药物  
抗性基因;

(8) 上述(7)中所述的质粒载体,其选自 pACE601、pACE611、  
pACE701、pACE702;

- 30 (9) 一种质粒载体,其中氨被表达的目的基因被掺入到上述(1)至(8)  
任一项的质粒载体中;

(10) 一种表达目的基因的方法,其特征在于在宿主中导入了质粒载  
体,其中目的基因被掺入到了上述(1)至(8)任一项所述的质粒载体中,和

识别质粒载体中的启动子序列的 RNA 聚合酶基因,并用外源因子诱导质粒载体拷贝数的增加和 RNA 聚合酶的表达以转录和翻译目的基因;

(11) 上述(10)中的表达目的基因的方法,其特征不在于质粒载体的拷贝数的增加和 RNA 聚合酶的表达是由不同的外源因子诱导的;

5 (12) 上述(10)中所述的表达目的基因的方法,其特征不在于质粒载体拷贝数的增加和 RNA 聚合酶的表达是由相同的外源因子诱导的;

(13) 上述(10)-(12)任一项所述的目的基因的表达方法,其中诱导质粒载体拷贝数增加的外源因子是选自添加异丙基- $\beta$ -D-硫代半乳糖苷(IPTG),添加乳糖,添加半乳糖,阿拉伯糖的添加,色氨酸浓度的降低和转换子培养温度的调整中的一种或几种;

(14) 上述(10)-(12)任一项的目的基因的表达方法,其中诱导 RNA 聚合酶表达的外源因子是选自异丙基- $\beta$ -D-硫代半乳糖苷(IPTG)的添加,乳糖的添加,半乳糖的添加,阿拉伯糖的添加,色氨酸浓度的降低和转换子培养温度的调整中的一种或几种;

15 (15) 上述(10)所述目的基因的表达方法,其特征不在于其它的质粒载体或噬菌体将 RNA 聚合酶基因导入宿主中;

(16) 上述(10)中所述的目的基因的表达方法,其特征不在于 RNA 聚合酶基因被整合到了宿主的染色体中;

(17) 上述(15)或(16)中所述目的基因的表达方法,其特征不在于 RNA 聚合酶基因来源于 SP6 噬菌体;

(18) 上述(10)到(17)任一项所述的目的基因的表达方法,其中目的基因编码对宿主致死或有害的蛋白质;

(19) 上述(10)-(18)任一项所述的目的基因的表达方法,其特征不在于大肠杆菌被用作宿主;

25 (20) 一种分离目的基因的方法,其特征不在于上述(1)到(8)任一项所述的质粒载体被应用于目的基因的分离方法中;

(21) 上述(20)中所述的目的基因的分离方法,其中目的基因编码一种对宿主致死或有害的蛋白质;

(22) 上述(21)中所述的目的基因的分离方法,其中编码一种对宿主致死或有害的蛋白质的基因是一种限制性内切核酸酶基因;

30 (23) 一种多肽,其含有序列表中序列号 1 所示的氨基酸序列的全部或部分并具有 AccIII 限制性内切核酸酶活性;

(24) 一种多肽, 其具有由序列表中序列号 1 所示的氨基酸序列或其部分中一个或多个氨基酸残基的缺失, 添加插入或置换之一得到的氨基酸序列, 并具有 AccIII 限制性内切核酸酶活性;

5 (25) 一种编码多肽的 RNA, 该多肽含有序列表中序列号 1 所示的氨基酸序列的全部或部分, 并具有 AccIII 限制性内切核酸酶活性;

(26) 一种 DNA, 其有序列表中序列号 2 所示的 DNA 的全部或部分的 DNA, 其中该 DNA 的表达产物具有 AccIII 限制性内切核酸酶活性;

10 (27) 一种编码多肽的 DNA, 该多肽是由序列表中序列号 1 中所示的氨基酸序列或其部分中一个或多个氨基酸残基的缺失, 添加, 插入或替代至少之一而得到的, 并具有 AccIII 限制性内切核酸酶活性;

(28) 一种能与上述(25)到(27)中任一项所述的 DNA 杂交并编码一种具有 AccIII 限制性内切核酸酶活性多肽的 DNA。

#### 附图简述

图 1 示模板表达质粒 pMSP6L 和 pMSP6F 的组成。

15 图 2 示模板表达质粒 pMSP60 的组成。

图 3(A), (B)示失控质粒 pHS2870 的构建程序。

图 4 示失控质粒 pHS2870 的组成。

图 5(A),(B)示质粒 pCRS04 的构建程序。

图 6 示质粒 pCRS04 的组成。

20 图 7 示质粒 pCRA19 的组成。

图 8(A),(B),(C),(D),(E)示系统质粒 pFSP6 的构建程序。

图 9(A),(B)示质粒 pACE601 的构建程序。

图 10 示质粒 pACE611 的组成。

图 11(A),(B)示质粒 pACE611 的构建程序。

25 图 12(A),(B)示质粒 pCRS70 的构建程序。

图 13(A),(B)示质粒 pACE701 和 pACE702 的构建程序。

图 14 示在将 pFSP6 和 pACE601 或 pFSP6 和 pACE601-NspIII 导入大肠杆菌 HB101 中之后, 根据 IPTG 诱导或非诱导通过  $\lambda$  DNA 降解反应溶液琼脂糖凝胶电泳所示出的 Nsp7524III 限制性内切核酸酶的活性。在该  
30 图中, M 指  $\lambda$  - EcoRT141 大小标准; 1 指  $\lambda$  - DNA; 2 指诱导后的 HB101/pFSP6/pACE601; 3 指诱导前的 HB101/pFSP6/pACE601; 4 指诱导后的 HB101/pFSP601/pACE601 -

NspIII； 5 指诱导前的 HB101/pFSP6/pACE601 - NspIII。

图 15 示在将 pFSP6 和 pACE601 或 pSFP6 和 pACE601 - NspIII 导入大肠杆菌 HB101 之后，根据 IPTG 的诱导或非诱导对  $\lambda$  - DNA 延长降解 2 和 3 小时后的琼脂糖凝胶电泳的结果。在该图中，M 指  $\lambda$  - EcoT141 大小标准； 1 指 Aval 消化的  $\lambda$  - DNA； 2 指 1 小时（诱导）后的 HB101/pFSP6/pACE601； 3 指 2 小时后（诱导）的 HB101/pFSP6/pACE601； 4 指 3 小时后（诱导）的 HB101/pFSP6/pACE601； 5 指 1 小时后（非诱导）的 HB101/pFSP6/pACE601； 6 指 2 小时（非诱导）后的 HB101/pFSP6/pACE601； 7 指 3 小时后（非诱导）的 HB101/pFSP6/pACE601； 8 指 1 小时后（诱导）的 HB101/pFSP6/pACE601-NspIII； 9 指 2 小时后（诱导）的 HB101/pFSP6/pACE601-NspIII； 10 指 3 小时后（诱导）的 HB101/pFSP6/pACE601-NspIII； 11 指 1 小时后（非诱导）的 HB101/pFSP6/pACE601-NspIII； 12 指 2 小时后（非诱导）的 HB101/pFSP6/pACE601-NspIII； 以及 13 指 3 后（非诱导）的 HB101/pFSP6/pACE601-NspIII。

图 16 示 AccIII 限制修饰系统基因的构建。

#### 实施本发明的最佳方式

本发明的第一个方面涉及质粒载体。更具体地说，本发明涉及一种质粒载体，其特征在于该质粒载体包含一个启动子序列，该启动子可被非宿主内在的 RNA 聚合酶识别并可控制目的基因的表达和在外源因子的诱导下可增加拷贝数的复制起点。质粒载体中所含可以控制目的基因的表达的启动子序列，它可以是任何启动子序列；可使用被特定的 RNA 聚合酶，例如来源于 T7，T3，SP6 和其它噬菌体的 RNA 聚合酶识别的启动子序列，优选被来源 SP6 噬菌体的 RNA 聚合酶所识别启动子序列。可被来源于 SP6 噬菌体的 RNA 噬菌体的 RNA 聚合酶识别的启动子即 SP6 启动子最小区域的碱基序列如序列表中序列号 30 所示。而且，该质粒载体还可含有作为选择标记的药物抗性基因。

本发明控制质粒载体拷贝数即作用于质粒载体复制起点的外源因子包括但不限于，例如添加各种化学物质如乳糖或结构相关物质象异丙基 -  $\beta$  - D - 硫代半乳糖苷（IPTG）半乳糖或结构相关物质，以及阿拉

伯糖或结构相关物质；色氨酸浓度的降低；和转化子培养温度的调整。

通过将复制起点置于可被这些因子控制的启动子的控制之下，使拷贝数的控制成为可能。可用于此目的启动子只要能达到此目的即可无特别限制，例如当用大肠杆菌作宿主时可用 lac、trp、tac、gal、ara、和 P<sub>L</sub>

5 启动子等。这些启动子也可用于诱导 RNA 聚合酶基因的表达。

本发明质粒载体复制起点的例子包括但不限于如在 lac 启动子的控制之下用作失控质粒复制起点的质粒载体。在这种情况下，拷贝数的控制是通过添加乳糖和结构相关物质，特别是优选添加异丙基-β-D-硫代半乳糖苷（IPTG）实现的。

10 本发明的第二方面涉及表达目的基因的方法，其特征在于将本发明的由整合目的基因进入质粒而制备的质粒载体和识别质粒载体上的启动子序列的 RNA 聚合酶基因导入宿主中，并用外源因子诱导质粒载体拷贝数的增加和 RNA 聚合酶的表达以转录和翻译目的基因。识别质粒载体上启动子序列的 RNA 聚合酶基因包括但不限于来源于上述噬菌体的 RNA 聚合酶基因，优选来源于 SP6 噬菌体的 RNA 聚合酶基因。诱导此类 RNA 聚合酶表达或质粒载体拷贝数增加的外源因子包括但不限于例如添加各种化学物质，如乳糖或结构相关物质异丙基-β-D-硫代半乳糖苷（IPTG），半乳糖或结构相关物质，以及阿拉伯糖或结构相关物质；色氨酸浓度的降低；以及转化子培养温度的调整。

20 质粒载体拷贝数的增加和 RNA 聚合酶表达的诱导由不同的外源因子或相同的外源因子的作用实现。而且，在这种情况下，编码 RNA 聚合酶的基因可被整合到宿主的染色体上，或由本发明的质粒载体以外的其它质粒载体或噬菌体载体导入宿主中。在后者，可根据目的适当的改变宿主。在该情况下，含 RNA 聚合酶基因的载体向宿主中的导入可在导入 25 本发明的质粒载体之前或之后进行。

尽管本发明的目的基因不受限制，但只有在其编码的蛋白对宿主致死或有害时，才有意义。此类蛋白质包括，例如上述的限制性内切核酸酶，其它核酸酶，核酸结合蛋白和蛋白酶。

30 本发明提供了一种通过两种控制方法联合控制目的基因表达的系系统，即对基因拷贝数的控制和如上所述通过启动子对转录的控制，不同于常规方法，对克隆至本发明的质粒中的目的基因表达的控制是非常严紧的，在非诱导条件下其表达可被抑制至不可检测的水平。通过表达诱

导当质粒的拷贝数增加时,即目的基因的拷贝数增加时,随着RNA聚合酶的同时表达,通过RNA聚合酶的作用,目的基因即在启动子的控制下转录和翻译。

因此可用本发明的质粒制备含对宿主致死或有害的基因的转化子并  
5 表达该基因,这是现有技术难以实现的。

例如,当用本发明的质粒载体尝试克隆编码Nsp7524III限制性内切核酸酶的基因时,在没有对应的修饰酶基因共存下,该基因被成功的保持在大肠杆菌中,还可通过将含RNA聚合酶基因的系统质粒导入该转化子中诱导Nsp7524III限制性内切核酸酶的表达,从而可在得到转化子的  
10 细胞中表达Nsp7524III限制性内切核酸酶。

另外,用本发明的质粒载体作为克隆载体制备基因文库,就可分离用常规方法不能分离的基因,并确定在单一宿主中其表达产物的活性。当然,本发明的质粒载体的用途不限于克隆编码对宿主有害产物的基因,而且可用作普通目的的质粒载体。

可用,例如,质粒pFSP6和噬菌体M13sp6将本发明的RNA聚合酶  
15 基因导入宿主中。质粒pFSP6构建的细书如参考实施例(2)所示。用该质粒转化的大肠杆菌HB101命为HPB101/pFSP6,已于平成7年12月22日(原保藏日)保藏于通商产业省工业技术院生命工学工业技术研究所〔地名:日本茨城县筑城市东1丁目1番3号(邮区号:305)〕保藏  
20 号FERM BP-5742。噬菌体M13sp6的构建方法见实施例2(b)中所述。

下面以质粒pFSP6为例,更详细地说明本发明。

首先描述的是用已构建的多拷贝模板表述质粒pMSP6L, pMSP6F(图1)和pMSP60(图2)进行实验的结果,所有的模板表述质粒含有可被SP6 RNA聚合酶识别的启动子序列和作为报告基因的 $\beta$ -半乳糖苷酶基因,并且都可用上述的系统质粒pFSP6表达目的基因。  
25

模板表述质粒的构建实例参考实施例(3)所示。这些质粒分别整合有SP6启动子的最小区域(pMSP6L),内在的SP6启动子区域(pMSP6F),或SP6启动子最小区域和lac操纵子区域(pMSP60),作为SP6质粒序列。

当质粒pMSP6L或pMSP6F被单独导入如表1中所示的大肠杆菌菌株中时,如表2中所示,尽管宿主大肠杆菌不表达SP6 RNA聚合酶,观察到了 $\beta$ -半乳糖苷酶活性,其量比用不含SP6启动子的质粒pMS434  
30

得到的要高。简而言之，可以推测 SP6 启动子被来源于大肠杆菌的 RNA 聚合酶识别，并且其下游的β - 半乳糖苷酶基因被转录和翻译，还有，当大肠杆菌 MRi80 被用作宿主时，β - 半乳糖苷酶的活性也降至 MRi7 的 1/7 到 1/14，其中象上述的表达质粒模板一样，该菌株在 pcnB 基因中有突变并且其中具有来自 ColEI 的复制起点的质粒的拷贝数减少至本菌株 MRi7 的 1/6 到 1/13。这提示了β - 半乳糖苷酶基因表达水平和所导入的质粒的拷贝数之间的相关。

表 1

| 菌株     | 基因型                                                                        | 起源/参考文献                                |
|--------|----------------------------------------------------------------------------|----------------------------------------|
| MC4100 | F-,araD139,Δ(argF-lac)U169, thiA,rpsL150,relA1,flbB5301, deoCl,ptsF25,rbsR | Casadaban 等, 分子生物学杂志 104(1984) 541-555 |
| MRi7   | MC4100Δrbs-7                                                               | Lopilato 等. 细菌学杂志.158(1984)665-673     |
| MRi80  | MRi7 pcnB80                                                                | Lopilato 等.分子和普通微生物学.205(1986) 285-290 |

表 2

| 菌株     | β - 半乳糖苷酶活性 |        |        |
|--------|-------------|--------|--------|
|        | 质粒          |        |        |
|        | pMSP6F      | pMSP6L | pMS434 |
| MC4100 | 202         | 242    | 30     |
| MRi7   | 234         | 287    | 23     |
| MRi80  | 17          | 27     | 4      |

接下来，如表 3 所示，当将同样的模板表达质粒当上述的系统质粒 pFSP6 一起导入大肠杆菌中并在 SP6 RNA 聚合酶基因未被诱导表达时测定β - 半乳糖苷酶活性，发现半乳糖苷酶的活性和在不含系统质粒的大肠杆菌中得到的一样高。这表明在非诱导条件下，质粒 pFSP6 上 SP6 RNA 聚合酶基因的表达几乎被完全抑制。另一方面，当 SP6 RNA 聚合酶基因

的表达被所添加的异丙基 $\beta$  - D - 硫代半乳糖苷 ( IPTG ) 所诱导时,  $\beta$  - 半乳糖苷酶活性增加 18 到 32 倍, 表明 $\beta$  - 半乳糖苷酶基因可通过 SP6 启动子表达。具体地, 当所用的宿主是大肠杆菌 MRi80 时, 诱导 4 小时后 $\beta$  - 半乳糖苷酶活性比非诱导条件下得到的增加 25 到 32 倍, 尽管非诱导条件下的活性保持在非常低的水平。

表 3

| 菌株            | $\beta$ - 半乳糖苷酶活性 ( 非诱导/诱导 ) |      |        |      |        |    |
|---------------|------------------------------|------|--------|------|--------|----|
|               | pMSP6F                       |      | pMSP6L |      | pMS434 |    |
|               | 非诱导                          | 诱导   | 非诱导    | 诱导   | 非诱导    | 诱导 |
| MC4100(pFSP6) | 221                          | 4875 | 234    | 5256 | 25     | 19 |
| MRi7(pFSP6)   | 220                          | 3958 | 242    | 4243 | 25     | 21 |
| MRi80(pFSP6)  | 17                           | 425  | 30     | 970  | 3      | ND |

在 SP6 启动子的下游携带有 lac 操纵子序列的质粒 pMSP60 中, 非诱导条件下  $\beta$  - 半乳糖苷酶活性比其它模板表达质粒中的要低; 但是, 其表达不能被完全抑制。这些结果提示, 在使用多拷贝质粒的表达系统中, 很难完全抑制目的基因的表达, 并且在解决这个问题方面控制含目的基因的质粒的拷贝数比控制系统质粒中 RNA 聚合酶基因的表达更有效。

因此本发明人在这些结果的基础上作了更进一步的调查研究, 以探索一种新的表达载体, 并开发了本发明质粒载体。

具体地说, 本发明人构建了一种能在非诱导条件下降低表达水平的失控质粒, 以确保在诱导之后的足量表达。首先, 考虑到 SP6 RNA 聚合酶在系统质粒中的表达是由 IPTG 诱导的, 根据图 3 所示的程序构建了含有允许在 IPTG 诱导时增加拷贝数的控复制起点的质粒。首先, 通过从质粒 pUC106A 去除含 lac 启动子和操纵子的 PvuII 片段区构建了质粒 pUC106AdPO, 质粒 pUC106A 是通过将碱基序列如序列表中序列号 3 和序列号 4 所示的两条 DNA 链制备而来的 106 Nde I DNA 片段导入质粒载体 pUC19(宝酒造社制)的 NdeI 位点而得到的。

其次, 可通过将 RNA 1870 片段插入到质粒 pSTV28(宝酒造社制)的

XbaI 位点而得到质粒 pHS2870，其中 RNA1870 片段是用靠近 RNAII 区排列的(质粒的复制起点)RNAIIA 引物(碱基序列如序列表中序列号 5 所示)和靠近 106NdeI 序列的末端的 1870 引物(碱基序列如序列表中序列号 6 所示)和用该质粒为模板通过 PCR 得到的。如此构建的质粒 pHS2870 如图 4 所示。通常，该质粒以每个大肠杆菌细胞约 30 个拷贝存在，但是 IPTG 的添加使该拷贝数增加到几百个。

然后从该质粒中去除含有衍生自质粒 pSTV28 的 P15A 复制起点的 AccI-NspI 片段(质粒 pCRS01)，接着进一步删去含非必需的限制性内切酶位点的 EcoRI - XbaI 片段(质粒 pCRS02)，之后再导入来源于质粒 pMJR1560 [基因(gene) 51 卷, 225 - 267 页(1987)] 的含乳糖阻遏子(lac Iq)基因的 DNA 片段以产生质粒 pCRS04。从质粒 pHS2870 构建质粒 pCRS04 的流程图如图 5 所示。如此构建的质粒 pCRS04 如图 6 所示。该质粒是一个被 IPTG 诱导的失控质粒，该质粒在每个大肠杆菌细胞中的拷贝数是 1 - 2 个拷贝，但是 IPTG 的添加使该拷贝数增加至几百个。

为了控制已被插入至质粒中的目的基因的表达，Psp6 - Olac EX 接头，即含 SP6 启动子序列和 lac 操纵子序列的双链寡核苷酸，可被导入到质粒的 Nhe I 位点以构建质粒 pACE601。Psp6 - Olac 接头是从其碱基序列如序列表中序列号 7 和 8 所示的两个 DNA 链制备而来的。质粒 pACE601 是含有作为选择标记的氯霉素抗性基因的失控质粒载体。

并且，通过将含有来源于质粒体 pUC118 (宝酒造社制)的 $\beta$  - 半乳糖苷酶基因的 BspHI 片段导入至上述质粒 pCRS04 的 NheI 位点(质粒 pCRS70)，接着去除含氯霉素抗性基因的 NcoI - BsaAI 片段，并用上述的 Psp6 - Olac 接头置换，构建了质粒 pACE701 和 pACE702，在这两个质粒中接头以相互相反的方向插入。这些质粒是含有作为选择标记的氨苄青霉素抗性基因的失控质粒载体。

如上所述，可用 $\beta$  - 半乳糖苷酶基因作为报告基因测定上述得到的质粒 pACE601、pACE701 和 pACE702 控制目的基因表达潜力。可通过将 $\beta$  - 半乳糖苷酶基因的 DNA 片段导入上述三质粒的每一种的 SP6 启动子下游的 NcoI 位点而构建模板表达质粒，该 DNA 片段是用引物 trpA-N-NcoI 和 lacZ-C-NcoI 以质粒 pMS434 [基因(gene), 57 卷, 89 - 99 页(1987)] 为模板通过 PCR 扩增得到的。引物 trpA-N-NcoI 和 lacZ-

C-NcoI 的碱基序列分别如序列表中序列号 9 和 10 所示。由此得到质粒被分别命名为 pACE601Z, pACE701Z 和 pACE702Z。在用这些模板质粒转化的大肠杆菌中, 绝对未检测到 $\beta$ -半乳糖苷酶的活性, 表明了对 $\beta$ -半乳糖苷基因表达的精确控制。

- 5 使用 Nsp7524III 限制性内切酶基因, 可以证明可用本发明的质粒分离和表达其表达产物对宿主致死的基因。质粒 pBRN3 含有 Nsp7524III 限制修饰系统基因。用该质粒转化的大肠杆菌 MC1061, 命名为大肠杆菌 MC1061/pBRN3, 已于平成 7 年 9 月 28 日 (原保藏日) 保藏于通商产业省工业技术院生命工学工业技术研究所 [地址: 日本茨城县筑城市东 1 丁目 1 番 3 号 (邮区号 305)], 保藏号 FERM BP - 5741。以引物对 L - ORF 种 NspR - ORF 和从转化子制备的质粒 pBRN3 模板通过 PCR 为基础的 DNA 扩增, 可得到只含 Nsp7524III 限制性内切酶基因的 DNA 片段。引物 L - ORF 和 NspR-ORF 的碱基序列分别如序列表中序列号 11 和 12 所示。通过向大肠杆菌 HB101 中导入质粒 pACE601-
- 10 NspIII, 可得到转化子大肠杆菌 HB101/pACE601 - NspIII, 其中的质粒是通过将得到的 DNA 片段插入到上述质粒 pACE601 SP6 启动子下游的 NcoI 位点而得到的。尽管缺乏 Nsp7524III 修饰酶基因, 转化子也可使宿主稳定的保持有 Nsp7524III 限制性内切酶基因。
- 15

- 而且, 可以通过将上述的系统质粒 pFSP6 导入该转化子从而制备转化子大肠杆菌 HB101/pFSP6/pACE601 - NspIII。通过培养转化子和在合适的时间添加 IPTG 诱导表达, 在培养中产生 Nsp7524III 限制性内切酶的可能性即可被证实。
- 20

- 本发明的第三个方面提供了分离目的基因的方法, 其特征在于使用了上述的质粒载体。本发明的分离方法特别适用于其表达产物是对宿主致死或有害的基因的分离, 特别是限制性内切酶基因。使用本发明的方法, 可以分离以前尚未分离的限制性内切酶基因, 例如在甚至缺少 AccIII 修饰酶基因时分离 AccIII 限制性内切酶。限制性内切酶基因的分离不仅可使用“鸟枪法”而且可以使用产限制性内切酶微生物基因组 DNA 的盒式文库; 而所谓的“鸟枪法”是产目的酶的微生物的基因组 DNA 被合适的限制性内切酶切割并且得到的 DNA 片段被有接插入到例如质粒 pACE611 中。通过用盒式文库为模板以 DNA 扩增获得一个基因, 就可能表达一个从其转录机制不同于宿主的微生物中制备而来的基因。用盒
- 25
- 30

式文库分离限制性内切酶基因的方法的描述，可参照一个包括 AccIII 限制性内切酶的实施例。

产 AccIII 限制性内切酶的微生物即 Acc 细菌基因组 DNA 盒式文库的制备如下：从 Acc 细菌培养物中提取基因组 DNA，并用合适的限制性内  
5 酶消化，之后将得到的 DNA 片段当含有与该片段有互补突出末端盒式文库。通常将这些文库称为 Acc 基因组盒式文库。其次，从 Acc 细菌纯化得到的目的限制性内切酶，其氨基酸序列被部分确定，并在此序列的基础上合成了一个引物。使用该引物和盒式引物，以每个盒式文库为模  
10 板进行以 PCR 为基础的 DNA 扩增反应，以得到含 AccIII 限制性内切酶基因片段的 DNA 片段。可通过，例如直接对 PCR 产物测序确定所得到的 DNA 片段的碱基序列，以确定 AccIII 限制性内切酶基因的全长碱基序列。

通过从该碱基序列设计一个能扩增 AccIII 限制性内切酶基因全长序列的引物，并使用该引物以 Acc 细菌基因组 DNA 为模板进行基于 PCR  
15 的 DNA 扩增反应，即可获得含 AccIII 限制性内切酶基因全长序列的 DNA 片段。获得的 DNA 片段被插入到质粒 pACE611 SP6 启动子的下游使密码子框被调整，并将得到的重组质粒导入例如大肠杆菌 JM109 中以产生转化子。通过培养各转化子，用本发明的基因表达方法诱导基因表  
20 达，并测定得到的各培养物的 AccIII 限制性内切酶活性，以及画出被各转化子携带的质粒 DNA 的限制性内切酶图谱，即可在可表达条件下对含 AccIII 限制性内切酶基因的转化子进行筛选。

由此得到的 AccIII 限制性内切酶基因被准确地插入到质粒 pACE611 中，并将得到的质粒命名为 pCRA19。该质粒的限制性内切酶图谱如图 7  
25 所示，其中粗实线指含 AccIII 限制性内切酶基因的 DNA 片段。整合有质粒 pCRA19 的大肠杆菌 JM109，被命名为大肠杆菌 JM109/pCRA19，已于平成 8 年 5 月 28 日（原保藏日）保藏于通商产业省工业技术院生命工学工业研究（地址：日本茨城县城市东 1 丁目 1 番 3 号（邮区号 305）），保藏登记号 FERM BP - 5743。尽管该转化子中缺少 - AccIII 修饰酶  
30 基因，它稳定地保持有该限制性内切酶基因。通过上述的方法，不需要构成限制修饰系统基因的修饰酶基因的共存，即可分离限制性内切酶基因。

本发明的第四个方面提供了具有 AccIII 限制性内切酶活性的多肽。

在本说明书中，在一些情况下术语 AccIII 限制性内切酶可作为具有 AccIII 限制性内切酶活性多肽的通用术语。例如，本发明的限制性内切酶包括序列表中序列号 1 所示的氨基酸序列。

5 不局限于上述，本发明的限制性内切酶还包括含序列表中序列号 1 所述全部氨基酸序列或部分序列并有 AccIII 限制性内切酶活性的多肽。而且，具有由序列表中序列号 1 所示氨基酸序列的全部或部分缺失，添加，插入或取代一个或多个氨基酸残基中的至少之一而得到并有 AccIII 限制性内切酶活性的多肽也在本发明的范围之内。

10 通常，由于体内或纯化中的蛋白质的修饰以及由于编码该蛋白质的基因的多型性或变异天然存在的蛋白质在其氨基酸序列中有氨基酸残基的缺失，插入，添加，取代或其它变异。但无论如何，已知有一些多肽在生理或生物活性上与无变异的蛋白质基本等价。因此，这些与相应的蛋白质结构上不同，但与该蛋白质在功能或活性上无显著差异的多肽也  
15 在本发明的范围之内。而且当人工向该蛋白质的氨基酸序列中导入变异时，也是相同的，并且可产生更多样化的变异体。例如，据报道在于大肠杆菌中表达的蛋白质的 N 末端的甲硫氨酸常常由于甲硫氨酸氨肽酶的作用而被去除，但依据不同的蛋白质类型并不是都去除，由此一些表达蛋白有甲硫氨酸残基而其它的蛋白质却没有。但是，在大多数情况下，  
20 甲硫氨酸残基和存在与否不影响蛋白质的活性，还已知在人白介素 2 (IL - 2) 的氨基酸序列中用特定的丝氨酸取代半胱氨酸得到的多肽仍保持着 IL - 2 的活性〔科学 (Science, 224 卷, 1431 页(1984))〕。

另外，在用基因工程技术生产目的蛋白时经常将目的蛋白表达为融合蛋白。例如，来源于另一个蛋白质的 N 末端肽链被添加到目的蛋白的  
25 N 末端以增强目的蛋白表达，或者通过向目的蛋白的 N 或 C 末端添加合适的肽链表达该蛋白使用时添加的蛋白有亲合性的载体以易化目的蛋白的纯化。因此，尽管该多肽的氨基酸序列与本发明的 AccIII 限制性内切酶的有部分不同，只要其具有与本发明的限制性内切酶基本等价的活性，该多肽就属于本发明的范畴。

30 AccIII 限制性内切酶的获得可以通过，例如，培养上述含 AccIII 限制性内切酶基因的转化子大肠杆菌 JM109/pCRA19，在培养合适时间添加诱导剂 IPTG 以增加 pCRA19 的拷贝数，随后用噬菌体 M13sp6 感染

进行。

从转化子培养物中收获 AccIII 限制性内切酶可以通过, 例如, 从培养物中收集细胞, 随后通过超声破碎, 超速离心等提取酶, 并接着结合核酸去除, 盐析, 亲和层析, 凝胶过滤, 离子交换层析等纯化该酶。由于作为该纯化起始物质的上述培养中不含有其它的限制性内切酶如 AccI 和 AccII, 可比其它常规纯化方法更容易得到目的纯度的酶制品。

本发明的第五个方面提供了编码具有 AccIII 限制性内切酶活性的多肽的 DNA。

本发明的 DNA 编码具有限制性内切酶活性的多肽, 包括例如编码序列列表中序列号 1 所示氨基酸序列的 DNA 还包括但不局限于包括序列列表中序列 2 中所述碱基序列的 DNA。具体地说, 下述的 DNA 在本发明的范围之内:

- (1) 编码一种多肽的 DNA, 该多肽含有序列表中序列号 1 所述的氨基酸序列的全部或部分并具有 AccIII 限制性内切酶活性;
- (2) 含有序列表中序列号 2 中所示的 DNA 的全部或部分, 其中该 DNA 的表达产物具有 AccIII 限制性内切酶活性;
- (3) 编码一种多肽的 DNA, 该多肽是由序列列表中序列号 1 的氨基酸序列或其部分中一个或多个氨基酸残基的缺失, 添加, 插入或取代而得到的并具有 AccIII 限制性内切酶活性; 以及
- (4) 能与上述(1)-(3)中的 DNA 杂交并编码具有 AccIII 限制性内切酶活性等的 DNA。

另外, 如果用上述得到的 DNA 作为探针在严紧的条件下进行杂交, 可得到与以上所得到的 DNA 序列号 2)类似但都有些不同但又编码有相同酶活性的多肽的 DNA。这样的 DNA 也包括在本发明的范围之内。

此类严紧条件指例如, 固定有 DNA 的膜被在含有  $6 \times \text{SSC}$  ( $1 \times \text{SSC}$  是在 1 升水中含 8.76g NaCl 和 4.41g 柠檬酸钠的溶液), 1% SOS, 100 $\mu\text{g/ml}$  醋精 DNA 0.1% 牛血清白蛋白, 0.1% 聚乙烯吡咯烷酮和 0.1% Ficoll 的溶液中, 在 65 $^{\circ}\text{C}$  温育 20 小时。

通过杂交获得编码 AccIII 限制性内切酶的类似的 DNA 的各种方法包括, 例如如下的方法。

首先, 通过常规方法将从合适的基因源得到的 DNA 与质粒或噬菌体载体连接以产生 DNA 文库。该文库被导入合适的宿主中; 得到的转化子

培养在平板上；长出的克隆或噬菌斑被转移至硝酸纤维素或尼龙膜上并变性，之后将 DNA 固定到膜上。在上述的条件下，将这些膜在含预先用  $^{32}\text{P}$  标记的探针等组成的上述溶液中温育杂交（所用探针可以是编码序列表中序列号 1 所示氨基酸序列的全部或部分的任意多核苷酸，代表性的多核苷酸由序列表中序列号 2 的碱基序列的全部或部分组成，或包含之）。杂交结束后，洗掉非特异性结合的探针，之后进行放射自显影等，以识别已与探针杂交的克隆。重复该方法直至目的杂交克隆被分离。由此得到的克隆保持有编码具有目的酶活性的多肽的 DNA。

通过例如，下述方法确定得到的 DNA 的碱基序列以确证其编码目的酶蛋白质。

为了碱基测序，将转化子培养在试管中，当宿主是大肠杆菌，并且用常规方法制备了质粒，前提是转化子是用质粒制备的，当用得到的质粒为模板或在将插入片段取出并亚克隆入 M13 噬菌体载体中之后，用双脱氧法确定碱基序列，在用噬菌体载体制备转化子的情况下，也可用基本相同的方法确定碱基序列。

这些从培养到碱基测序的基本实验方法在例如，《分子克隆：实验手册》（Molecular Cloning: A Laboratory Manual, 1982, T. Maniatis 等冷泉港实验室出版社出版）中有描述。

不管如何，通过将所确定的碱基序列与序列表中序列号 2 所示的碱基序列相比较，或将由所确定的碱基序列推导而来的氨基酸序列与序列表中序列号 1 所示的氨基酸序列相比较，可以证实所得到的 DNA 与编码目的 AccIII 限制性内切酶的 DNA 类似。

当得到的 DNA 不含目的限制性内切酶编码区的全部时，可以通过在所得 DNA 碱基序列的基础上合成引物，并用该引物通过 PCR 扩增缺失区域，或者用得到的 DNA 片段为探针重复筛选 DNA 文库而得到整个编码区。

可以制备含由此所得编码 AccIII 限制性内切酶的类似 DNA 的转化子，使转化子表达该 DNA 所编码的酶蛋白，并纯化所表达的酶蛋白。转化子的制备以及酶蛋白的表达和纯化都可用本发明的质粒完成。这样得到的酶蛋白保留有 AccIII 限制性内切酶活性。

另外，AccIII 修饰酶以及编码该酶的 DNA，此二者在目前尚未得到，也可用上述的编码 AccIII 限制性内切酶基因的 DNA 得到。

例如，根据在许多情况下限制性内切酶基因与修饰酶基因位置靠近，该目的可如下实现：用盒式文库为模板通过 DNA 扩增反应得编码靠近限制性内切酶基因的蛋白质的基因区，将其导入合适的表达载体使基因表达。并用合适的方法确定 AccIII 修饰酶活性。还可根据，例如从该  
5 转化子制备的 DNA 对 AccIII 限制性内切酶断裂活性的抗性，确认 AccIII 修饰酶活性。只要测定了上述基因区的碱基序列以证实它与已知的限制修饰系统的修饰酶基因之间保守区同源[分子生物学杂志 (J.Mol.Biol) 206 卷, 305 - 321 页 (1989)]，就可基因表达之前在一定程度上预料到该基因是修饰酶基因。

10 通过使用上述的 Acc 细菌基因组 DNA 的盒式文库，可得到 AccIII 修饰酶及编码它的 DNA。其氨基酸序列和碱基序列分别如序列表中序列号 13 和 14 所示。

本文的 AccIII 修饰酶不仅限于如上所述。如在 AccIII 限制性内切酶说明中所指出的，它是一个含有序列表中序列号 13 所述的氨基酸序列的  
15 全部或部分并有 AccIII 修饰酶活性的多肽。而且，由序列表中序列号 13 的氨基酸序列或其部分的一个或多个氨基酸残基的缺失，添加，插入或取代至少之一而得到并具有 AccIII 修饰酶活性的多肽也在本发明的范畴之内。

本发明的 AccIII 修饰酶编码 DNA 包括编码序列表中序列号 13 所述  
20 氨基酸序列的 DNA，还包括，但不局限于，例如，包含序列表中序列号 14 碱基序列的 DNA。具体地说，下述的 DNA 在本发明的范畴之内。

(1) 编码一种多肽的 DNA，该多肽含有序列表中序列号 13 中所述的氨基酸序列的全部或部分并具有 AccIII 修饰酶活性；

(2) 含有序列表中序列号 14 所示 DNA 的全部或部分的 DNA，其中  
25 该 DNA 的表达产物具有 Acc III 修饰酶活性；

(3) 编码一种多肽的 DNA，该多肽是由序列表中序列号 13 的氨基酸序列或其部分中的一个或多个氨基酸残基的缺失，添加，插入或取代至少之一而得到并具有 Acc III 修饰酶活性；以及

(4) 能与上述(1)-(3)中所述的 DNA 杂交并编码具有 Acc III 修饰酶  
30 活性等多肽的 DNA。

下文将通过参考实施例和工作实施例更详细地描述本发明，这些实施例并不是限制性的。在本文所述的方法中，一些基本的如质粒制备，

限制性内切酶消化等是根据《分子克隆：实验手册》，第二版(Molcuelar Cloning. A Laboratory Manual 2nd ed.), T. Maniatis 等编，冷泉港实验室出版社出版中所述的方法实现的。

### 参考实施例

#### 5 (1) 培养基和培养条件

将大肠杆菌用 LB 培养基(1%蛋白胨, 0.5%酵母提取物, 0.5% NaCl, pH7.0)在 37 °C 通气培养。根据大肠杆菌所携带的质粒向培养基中各自加入各种抗生素: pFSP6 加入卡那霉素 25µg/ml, pXX325 加入氨苄 20µg/ml, 对氨苄抗性的 Col E1 型质粒加入 50µg/ml 氨苄, 以及对氨苄  
10 抗性的 pUC 型质粒加入 100 µg/ml 氨苄。在表达诱导实验中, 将培养至平台期的培养肉汤以 1%接种到新鲜培养基中, 然后于 37 °C 通气培养, 之后在达到 OD<sub>600</sub> 0.6 时( $6 \times 10^8$  细胞/ml)加入异丙基-β-D 半乳糖苷 (IPTG)至终浓度 0.2mM, 接着进一步培养。

#### (2) 系统质粒 pFSP6 的构建

15 根据图 8 所示的方法, 用大肠杆菌 JM109 为宿主构建系统质粒 pFSP6。用 Hind III(宝酒造社制)消化质粒 pSP6 - 2[核酸研究, 15 卷, 1653 - 2664 页(1987)]并使两个末端平齐, 之后进一步用 Bam HI(宝酒造社制)消化以产生一个约 2.8kb 的含 SP6 RNA 聚合酶基因的 DNA 片段, 将该片段与预先用 Bam HI 和 HincII(宝酒造社制)消化的质粒载体使  
20 pUC18(宝酒造社制)混合以便连接构建质粒 pUCSP。其次, 将质粒 pMJR1560[基因(gene), 51 卷, 225 - 267(1987)]用 KpnI 消化并使其两个末端平齐化, 之后进一步用 PstI(由宝酒造社制)消化以生产一个约 1.3kb 的含 lacIq 基因的 DNA 片段, 然后将其分离。将上述的质粒 pUCSP 用 Hind III 消化并使两个末端平齐, 之后进一步用 Pst I 消化以产生一个  
25 DNA 片段, 将该 DNA 片段与上述的约 1.3kb DNA 片段混合连接以构建质粒 pUCSPlac。

而且, 将一个由用 PstI 消化 pUCKm[分子生物学杂志(J.mol.Biol)47 卷, 217 - 226 页(1981)]得到的含有卡那霉素抗性基因的约 1.5kb 的 DNA 片段, 导入上述质粒 pUCSPlac 的 PstI 位点以构建质粒  
30 pUCSPlacKm。用 AatII(东洋纺社制)消化该质粒, 之后用 Nsp I (宝酒造社制)部分消化以分离一个约 6.5kb 的 DNA 片段, 然后使其两端平齐。

另一方面, 用 Hind III 消化质粒 pXX325[美国图象科学院院刊

(Proc.Natl.Acad.Sci.USA),80 卷, 4784 - 4788 页(1983)]并使两端平齐,之后进一步用 Sal I 消化以产生一个含 miniF 质粒复制起点的约 6.8kb 的 DNA 片段,将该片段与上述的约 6.5kb 的 DNA 片段混合连接以构建质粒 pFSP6。用该质粒转化大肠杆菌 HB101 以形成大肠杆菌 HB101/pFSP6。

### 5 (3) 多拷贝模板表达质粒的构建

从其碱基序列如分别从序列表中序列号 15 和 16 所示的两条 DNA 链制备了 Psp6 接头,它含 SP6 启动子最小区的双链寡核苷酸,将它们与事先用 XhoI(宝酒造社制)和 Hind III 消化的质粒 pMS434[基因, (gene), 57 卷, 89 - 99 页(1987)]混合连接以构建模板表达质粒 pMSP6L,该质粒中  
10 在 $\beta$ -半乳糖苷酶基因的上游插入了上述的启动子序列。

另一方面,可通过将含有用 Acc II 和 Hind III 消化 pSP64[核酸研究, 12 卷, 7035 - 7056(1984)]而得到的内在 SP6 启动子序列的 DNA 片段导入至上述质粒 pMS434 的 Xho I - Hind III 位点,从而构建了模板表达质粒 pMSP6F(图 1)。而且,含有 SP6 启动子最小区和 lac 操纵子区的  $P_{sp6}$  -  $O_{lac}$  接头被插入到上述质粒 pMS434 的 Xho I - Hind III 位点,从而构建  
15 了模板表达质粒 pMSP60(图 2),其中,  $P_{sp6}$  -  $O_{Lac}$  接头是从其碱基序列分别如序列表中序列号 17 和 18 所示的两条 DNA 链制备而来的。

### (4) SP6 启动子对大肠杆菌 RNA 聚合酶的启动子活性

参考实施例(3)中构建的质粒 pMSP6L 和 pMSP6F 和作为对照的不含  
20 SP6 启动子序列的质粒 pMS434,被各自导入了表 1 中所示的大肠杆菌中。用参考实施例(1)的方法分别培养得到的转化子,之后各自收集对数生长期的细胞培养肉汤。用一部分测定收集的各培养肉汤的  $OD_{600}$  值,而将剩下的部分转移到预冷的试管中并补加氯霉素至终浓度为  $100\mu\text{g/ml}$ 。用该培养肉汤作样品,用 J.H.Miller 所编的《分子遗传学实验》(Experiments in Molecular Genetics)冷泉港实验室出版社出版, 1972  
25 中所述的方法,测定 $\beta$ -半乳糖苷酶活性。

如表 2 所示,整合有模板表达质粒 pMSP6L 或 pMSP6F 的大肠杆菌 MC 4100,其中上述二质粒都指带有含 SP6 启动子的序列,表现出相似的低 $\beta$ -半乳糖苷酶活性,尽管其中插入启动子序列。但是,整合的不  
30 含 SP6 启动子的质粒 pMS434 表现出更低的酶活性。另外,对比作为宿主的大肠杆菌 MRi 7 和 MRi 80 发现 MRi80 中 $\beta$ -半乳糖苷酶活性降低了,其中该质粒的拷贝数也减少了。

### (5)系统质粒 pFSP6 的评价

将参考实施例(4)中制备的其中进行导入了系统质粒 PFSP6 的转化子, 用参考实施例(1)中所述的方法培养, 测定 IPTG 加入前后非诱导条件下和诱导条件下的 $\beta$  - 半乳糖苷酶活性。结果如表 3 所示。在非诱导条件下, 所有的转化子所表现出的 $\beta$  - 半乳糖苷酶活性都与表 2 中所示的不存在质粒 pFSP6 时得到的水平相似。

另一方面, 随着 IPTG 添加的诱导, 与非诱导条件下的相比, 整合有质粒 pMSP6L 和 pMSP6F 的转化子的 $\beta$  - 半乳糖苷酶活性是 18 到 32 倍, 而在整合了质粒 pMS434 的转化子中无活性的增加。

### 10 (b)Lac 操纵子的表达控制作用

用参考实施例(1)所述的方法培养由向整合有质粒 pFSP6 的大肠杆菌 MC4100 中分别导入质粒 pMSP6L, pMSP6F, pMSP60 和 pMS434 而得到的转化子, 测定了非诱导条件 $\beta$  - 半乳糖苷酶的活性。结果如表 4 所示。尽管活性水平比用质粒 pMSP6L 和 pMSP6F 得到的转化子低, 但是整合了含 lac 操纵子序列的质粒 pMSP60 的转化子仍表现出一些 $\beta$  - 半乳糖苷酶活性。简而言之, 通过简单地导入 lac 操纵子序列不能完全抑制非诱导条件下的表达。

| 菌株     | $\beta$ - 半乳糖苷酶活性(非诱导条件) |        |        |        |
|--------|--------------------------|--------|--------|--------|
|        | pMSP6F                   | pMSP6L | pMSP60 | pMS434 |
| MC4100 | 221                      | 234    | 30     | 25     |

### 实施例 1

#### (1)培养基和培养条件

20 用 LB 培养基(1%蛋白胨, 0.5%酵母提取物, 0.5%NaCl,pH7.0)于 37℃ 通气培养大肠杆菌。根据大肠杆菌所含质粒向培养基中各自加入各种浓度的抗生素, 具体如下: pFSP60 加入卡那霉素 25 $\mu$ g/ml, 氨苄抗性 Col E1 型质粒加入氨苄 50 $\mu$ g/ml, 氨苄抗性 pUC 型质粒加入氨苄 100 $\mu$ g/ml, 氯霉素抗性失控质粒加入氯霉素 30 $\mu$ g/ml, 和氨苄抗性失控质粒加入氨苄 30 $\mu$ g/ml。在表达诱导实验中, 培养至平台期得到的培养肉汤以 1%加入到新鲜培养基中, 然后在 37℃ 通气培养, 之后在达到 OD<sub>600</sub> 值为 0.6( $6 \times 10^8$  细胞/ml)时添加异丙基 -  $\beta$  - D - 硫代半乳糖苷(IPTG)至终浓度为 0.2mM; 接着进一步进行培养。

#### (2)失控质粒 pHS2870 的构建

表达质粒构建基础上形成的失控质粒 pHS2870，以图 3 中所示的方法构建。将从其碱基序列分别如序列表中序列号 3 和 4 所示的两条 DNA 链制备而来的 106 Nde I DNA 片段与预先用 Nde I 消化的质粒载体 pUC19 混合以便连接。将由此得到的质粒 pUC106A 用 PvuII 消化并进行自身连接以产生质粒 pUC106AdPO。用该质粒为模板，使用 RNA IIA 引物（其碱基序列如序列表中序列号 5 所示）和 1870 引物（其碱基序列如序列表中序列号 6 所示）进行 PCR 以产生一个扩增的 DNA 片段，然后将其用 XbaI 消化并与预先用 XbaI 消化的质粒 pSTV28 混合连接以产生质粒 pHS2870。质粒 pHS2870 的构建如图 4 所示。

#### 10 (3) 构建失控制质粒 pCRS04

在用 AccI（宝酒造社制）和 NspI 消化之后，将上述的质粒 pHS2870 两端平齐并进行自身连接以产生质粒 pCRS01，该质粒缺失 P15A 复制起点。然后用 EcoRI（宝酒造社制）和 XbaI 消化该质粒，之后用如上同样的方式使其两端平齐并进行自身连接以产生质粒 pCRS02。在将用 KpnI（宝酒造社制）和 PstI 消化质粒 pMJR1560 得到的约 1.2kb DNA 片段两末端平齐之后，将其与预先用 NspV 和 VspI（均由宝酒造社造）消化并且两端平齐的上述质粒 pCRS02 混和，连接以产生质粒 pCRS04。构建质粒 pCRS04 的流程图如图 5 所示，而构建质粒 pCRS04 的构建如图 6 所示。

#### 20 (4) 失控表达载体 pACE601 的构建

将上述的预先用 NheI（宝酒造社造）消化并随后使两末端平齐的质粒 pCRS04 与  $P_{sp6}$ -O<sub>lac</sub> EX 接头混合，连接以构建质粒 pACE601，其中  $P_{sp6}$ -O<sub>lac</sub> EX 接头由两条其碱基序列如序列表中序列号 7 和 8 所示的 DNA 链制备而来。质粒 pACE601 的构建流程图如图 9 所示。这些质粒的构建选用大肠杆菌 JM109 为宿主的。

#### 25 (5) 失控表达载体 pACE611 的构建

将失控表达载体 pACE601 用 XhoI - HindIII 消化，并将由序列表中序列号 15 和 16 所示的合成 Oligo-DNA 组成的  $P_{sp6}$  接头插入到其中的位点中以构建失控表达载体 pACE611，pACE611 构建物如图 10 所示。质粒 pACE611 的构建流程图如图 11 所示。

#### 30 (6) 失控表达载体 pACE701 和 pACE702 的构建

用 BspHI(由 NEB 产)消化质粒载体 pUC118（宝酒造社制）并使其

末端平齐而得到的一个约 1kb 的 DNA 片段，之后将其与上述的预先用 Nhe I 消化并使末端平齐的质粒 pCRS04 混合，以连接构建质粒 pCRS70。质粒 pCRS70 构建的流程图如图 12 所示。其次，用 Nco I（宝酒造社制）和 BsaA I（NEB 制）消化该质粒，之后使两末端平齐并与上述的  $P_{sp6}$ - $O_{lac}$  EX 接头混合，以连接构建两个质粒 pACE701 和 pACE702，该两个质粒以相互相反的方向整合了所插入的接头。质粒 pACE701 和 pACE702 的构建流程图如图 13 所示。使用了大肠杆菌作为这些质粒构建的宿主。

#### (7) 模板失控表达质粒的构建

以上述的质粒 pMS434 为模板，使用引物 trpA-N-NcoI 和 lacZ-C-NcoI 进行 PCR，以产生含  $\beta$  - 半乳糖苷酶基因的 DNA 片段。引物 trpA - N - NcoI 和 lac Z - C - Nco I 的碱基序列分别如序列表中序列号 9 和 10 所示。在用 NcoI 消化之后，将该片段与所有却用 Nco I 消化的上述质粒 pACE601，pACE701 和 pACE702 各自混合，连接产生模板失控质粒 pACE601Z，pACE701Z,和 pACE702Z。所有的都整合]被导入  $P_{sp6}$  -  $O_{lac}$  序列下游的  $\beta$  - 半乳糖苷酶基因。这些质粒的构建是以大肠杆菌 JM109 为宿主的。

#### (8) 非诱导条件下失控表达质粒表达水平的评价

将分别向整合有上述的系统质粒 pFSP6 的大肠杆菌 MC4100（下方作为 MC4100/PFSP6）导入了上述的模板失控表达质粒 pACE601Z，pACE701Z 和 pACE702Z 而得到的转化子 MC4100/pFSP6/pACE601Z，MC4100/pFSP6/pACE701,和 MC4100/pFSP6/pACE702Z 分别培养在实施例(1)所述的条件下；用上述参考实施例(4)所述的方法分别分析在对数生长期收集的各培养肉汤的  $\beta$  - 半乳糖苷酶活性。在所有的检测的转化子中， $\beta$  - 半乳糖苷酶活性低于检测限度，表明了比参考实施例(4)中用多拷贝质粒得到的转化子的更大的表达抑制作用。

#### (9) Nsp7524 III 限制性内切酶基因表达质粒的构建

含有 Nsp7524 III 限制修饰系统基因的质粒 pBRN3 是从大肠杆菌 MC1061/pBRN3（FERM BP - 5741）制备而来的。用该质粒为模板使用引物 L - ORF 和 NspR - ORF3 进行 PCR 以产生一个约 1kb 的仅含 Nsp7524 III 限制性内切酶基因的 DNA 片段。引物 L - ORF 和 NspR - ORF3 的碱基序列分别如序列表中序列号 11 和 12 所示。在用 NcoI 消化

之后,将该片段与上述的预先用 NcoI 消化的质粒 pACE60 混合,以连接构建质粒 pACE601-Nsp III,该质粒整合有仅在  $P_{sp6} - O_{lac}$  序列的下游导入了 Nsp7524 III 限制性内切核酸酶基因。该质粒被稳定的保持在不合 Nsp7524 III 修饰酶基因的大肠杆菌 JM109 中。

#### 5 (10) Nsp7524III 限制性内切酶基因表达系统的构建

将在向整合有上述的系统质粒 pFSP6 的大肠杆菌 HB101 (下文称为 HB101/pFSP6) 中导入了上述的含 Nsp7524 III 限制性内切酶基因的质粒 pACE601 - Nsp III 以及质粒 pACE601 而得到的转化子 HB101/pFSP6/pACE601-Nsp III,和 HB101/pFSP6/pACE601 分别培养于 LB 培养基中直至平台期;将收集的各培养肉汤以 1 % 接种到两试管新鲜基  
10 培养基中并于 37 °C 通气培养。在  $OP_{600}$  值达到 0.6 之后,向每个试管中加入 IPTG 至终浓度 0.2mM,之后继续培养。培养结束后,收集细胞,悬浮于细胞破碎缓冲液 A 中 (20mM Tris - HCl, pH7.5 10mM2 - 巯基乙醇),并超声破碎,接着离心产生粗提取物。

15 将 1 $\mu$ l 该粗提物加入到 30 $\mu$ l 反应混合物中 (10mM Tris - HCl, pH7.5, 10mM  $MgCl_2$ , 1mM DTT, 50mM NaCl, 1 $\mu$ g  $\lambda$ -DNA),之后将反应混合物与琼脂糖凝胶电泳以检查  $\lambda$  - DNA 的消化和证实限制性内切酶的活性。如图 14 所示,只有在加入 IPTG 诱导表达的 HB101/pFSP6/pACE601 - NspIII 中,观察到了,限制性内切酶活性,并且  
20 其  $\lambda$  - DNA 断裂带型同 Aval,即 Nsp7524 III 的一个同裂酶的带型一致。

上述  $\lambda$  - DNA 消化反应延长至 2 和 3 小时的结果如图 15 所示。这时,由于本实验所用的粗提取物所含有来自大肠杆菌核酸酶,由 Nsp7524III 活性得到的 DNA 片段由大肠杆菌核酸酶进行了进一步的降解,在泳道 8 到泳道 10 得到随时间增加而其密度降低的带,而泳道 10  
25 没有带。由于 Nsp7524 III 诱导的缺乏;甚至在延长反应时间时也在泳道 11 到 13 上没有产生 Nsp7524 III 消化片段,表明由于大肠杆菌核酸的直接作用, $\lambda$ -DNA 被变成了低分子量的片段。在来自 HB101/pFSP6/pACE601 的粗提物中没有检测到限制性内切酶活性。

#### 实施例 2 AccIII 限制性内切酶基因的分离和表达

30 (1) 测定 AccIII 限制性内切酶蛋白质 N - 末端氨基酸序列以及合成对应于该氨基酸序列的引物 DNA

将大约 2ku 的 AccIII 限制性内切酶的商业产品 (宝酒造社制) 上

Sephacryl 5300 柱 ( Pharmacia 造 ) 进行凝胶过滤以测定 AccIII 限制性内切酶的分子重。从检测到活性的洗脱位置判断, 表明 AccIII 限制性内切酶的分子量约为 70,000 。由于大多数的限制性内切酶的蛋白质是二聚体, 预测 AccIII 限制性内切酶蛋白在 SDS 聚丙烯酰胺凝胶电泳上会迁移到分子量 35,000 的位置。

接下来, 为了得到 AccIII 限制性内切酶蛋白质 N 末端氨基酸测序用的样品, 将大约 2ku 的 AccIII 商业产品与一个蛋白质分子量标准一起进行 SDS 聚丙烯酰胺凝胶电泳, 之后将蛋白质转移到 PVDF 膜上并用溴酚蓝染色以确认蛋白质的位置。在用 10 % 乙酸 - 50 % 甲醇脱色之后, 切下印迹有约 35,000 分子量蛋白质位置的 PVDF 膜, 并用蛋白质测序仪 G1000A ( Hewlett - Packard ) 进行自动 Edman 降解以确定如序列表中序列号 19 所示的 N 末端氨基酸序列。以该序列为基础, 然后合成了分别如序列表中序列号 20 和 21 所示的 AccIII 引物 1 和 2 以用作一对盒式引物。

#### 15 (2) Acc 细菌基因组 DNA 的制备

根据喜多等在核酸研究 ( Nucleic Acid Research ) 13 卷, 8685 - 8694 ( 1985 ) 中所述的方法, 培养 Acc 细菌并得到湿菌体。将得到的 2g 湿菌体悬浮于 10ml 缓冲液 B[25mM Tris-HCl(pH8.0), 50mM 葡萄糖, 10mM EDTA], 在以 2mg/ml 于缓冲液 B 中存在的 1ml 溶菌酶的存在下搅拌, 并于 37 °C 放置 20 分钟。接着, 向该溶液中加入 28ml 缓冲液 C[100mM NaCl, 100mM Tris-HCl(pH8.0)], 随后搅拌。进一步加入 1ml 以 20mg/ml 溶于 TE[10mM Tris-HCl, 1mM EDTA(pH8.0)] 中的蛋白酶 K 和 4ml 10 % SDS 溶液, 随后搅拌, 之后将混合溶液于 37 °C 放置 1 小时。

向该溶液中加入 6ml 5M 氯化钠水溶液和 6ml 缓冲液 D[CTAB ( 十六烷基三甲基溴化铵 ), 0.7M NaCl ], 随后搅拌, 接着将混合液于 60 °C 放置 20 分钟。将该溶液用苯酚和随后用氯仿处理, 之后分离水相。向该水层中加入在 TE 中配制的 10mg/ml 的 RNAaseA ( Sigma 产 ) 50 $\mu$ l, 随后搅拌, 之后将混合液在 37 °C 放置 40 分钟。溶液静置后, 用苯酚处理, 随后用氯仿处理, 之后分离水相。向水相中加入等体积的冷乙醇, 通过在一玻璃毛细管上缠绕回收所沉淀的 DNA。用 70 % 乙醇洗该 DNA 并在 3ml TE 中溶解以产生约 200 $\mu$ g 的基因组 DNA。

#### (3) Acc 基因组盒式文库的制备

下面的方法基本上根据宝酒造社在《基因工程指南》(1995 - 1996) F16 - F17 页所述的方法进行实施。

将(2)中所得的基因组 DNA 用 EcoRI (宝酒造社制) 完全消化; 将得到的 DNA 片段当互补具有与其的粘末端的 EcoRI 盒酒造社制) 连接, 产生 EcoRI 盒式文库。类似地, 将基因组 DNA 分别用限制性内切酶 Bgl II, EcoT14 I, EcoT22 I, Hind III, Pst I, Sal I 和 Xba I (均由宝酒造社制) 完全消化。将得到的每一种基因组 DNA 片段分别与几个具有互补突出主要的盒(宝酒造社制)中的一种结合, 以分别产生 Bal II, EcoT14 I, EcoT22 I, Hind III, Pst I, Sal I 和 Xba I 盒式文库。这些盒式文库被通称为 AccIII 基因组盒式文库。

#### (4) AccIII 限制性内切酶基因的分析

用(3)中所得到的 EcoRI 盒式文库为模板, 以 AccIII 引物 1 和盒式引物 C<sub>1</sub> (宝酒造社制) 作为引物时进行基于 PCR 的 DNA 扩增反应。为了有效和特异性地扩增目的区, 用 AccIII 引物 2 和盒式引物 C<sub>2</sub> (宝酒造社制) 为引物对以所得的反应液的一部分为模板进行了第二个基于 PCR 的 DNA 扩增反应, 以产生扩增的 DNA 片段。类似地, 分别以 Bgl II, EcoT14I, EcoT22I, Hind, PstI, SalI 和 XbaI 盒式文库为模板, 分别实施上述的两步基于 PCR 的 DNA 扩增反应, 以得到扩增的 DNA 片段。用琼脂糖凝胶电泳分析每个扩增的 DNA 片段; 来源于 XbaI 盒式文库的扩增的 DNA 片段表现出特别高的扩增效率。

因此, 用 XbaI 盒式文库为模板以 AccIII 引物 1 和盒式引物 C<sub>1</sub> 为引物对, 再次进行基于 PCR 的 DNA 扩增反应。将反应混合物进行琼脂糖凝胶电泳; 从凝胶中回收了一个约 0.5kb 的扩增的 DNA 片段。以所得的 DNA 片段为模板, 以 AccIII 引物 2 和盒式引物 C<sub>2</sub> 为引物对进行了第二个基于 PCR 的 DNA 扩增反应。所得的约 0.5kb 的扩增的 DNA 片段的碱基测序表明该片段编码子(1)中测定了其 N 末端氨基酸序列的蛋白质的一部分。另一方面, 推测的 AccIII 限制性内切酶蛋白质的分子量约为 35,000 因此, 估计编码该蛋白的基因文库为约 1kb。因此除了上述的约 0.5kb 碱基序列的信息是之外, 只要有了与 AccIII 引物 1 和 2 退火的 5' - 末端区域准确碱基序列的信息, 和剩下的 3 区约 0.5kb 碱基序列的信息就有希望确定 AccIII 限制性内切酶基因的碱基序列。

因此, 为了确定 AccIII 限制性内切酶 5' 末端区域的准确的碱基序列,

- 合成了 AccIII 3'物了（如序列表中序列 22 所示）和 AccIII 引物 4（如序列表中序列 23 所示）。接下来，以(3)中所得的 EcoRI 盒式文库为模板用 AccIII 引物 4 和盒式引物 C<sub>1</sub> 为引物对进行了基于 PCR 的 DNA 扩增。用一部分该反应液为模板，使用 AccIII 引物 3 和盒式引物 C<sub>2</sub> 的引物对进行了基于 PCR 的 DNA 扩增反应。类似地，分别用 BglII，EcoT14I；EcoT22I，HindIII，PstI，Sall 和 XbaI 盒式文库为模板，分别实施了上述的两步基于 PCR 的 DNA 扩增反应。在所扩增的 DNA 片段中，确定了最短的，即约 1kb 的用 EcoT14I 盒式文库为模板得到的 DNA 片段的碱基序列。
- 而且，为了确定假定的编码 AccIII 限制性内切酶蛋白的基因区的引端的约 0.5kb 的序列，合成了如序列表序列号 24 所示的 AccIII 引物 5，和如序列表序列号序列 26 所示的引物 6。以(3)中所得的 EcoRI 盒式文库为模板，使 AccIII 引物 5 和盒式引物 C1 引物对进行基于 PCR 的 DNA 扩增反应。用该反应混合物的一部分当为模板，使用 AccIII 引物 6 和盒式引物 C2 引物对进行了基于 PCR 的 DNA 扩增反应。类似地，分别以 BglII，EcoT14I，EcoT22I，HindIII，PstI，Sall 和 XbaI 盒式文库为模板，进行了上面两步基于 PCR 的 DNA 扩增反应。

- 在所扩增的 DNA 片段中，分别以 EcoT22I 和 HindIII 盒式文库为模板得到约 0.5kb 和 0.8kb 的 DNA 片段，这两个片段进行了碱基测序。结合所确定的三个碱基序列的结果，得到了一个约 1.6kb DNA 片段的碱基序列信息。其碱基序列如序列表中序列号 26 所示。而且，能够编码该蛋白质的开放阅读框的查找结果表明；在碱基编号 558 到 1442 为 ORF1，在碱基编号 1588 到 1434 被认为是 ORF2，以及在碱基编号 1 到 535 是 ORF3。由于其编码(1)中 N 末端氨基酸序列确定的蛋白质，认为 ORF1 是 AccIII 限制性内切酶基因。AccIII 限制性内切酶基因的碱基序列如序列表中序列号 2 所示，其中从 5'端数第 4 个碱基是 C。从该碱基序列推导而来的氨基酸序列如序列表中序列号 1 所示，其中从 N 末端数第 2 个氨基酸是 Leu。ORF2 的翻译方向与 ORF1 和 ORF3 相反。

#### (5) 质粒 pCRA19 的构建

- 接下来，构建了表达 AccIII 限制性内切酶基因的质粒。首先，为了得到该基因，合成了如序列表中序列号 27 所示的引物 Acc - RL 和如序列表中序列号>8 所示的引物 Acc - RR。在两个引物中导入了限制性内

切酶 NcoI 识别序列位点。用上述的引物 sh3，就可能用限制性内切酶 NcoI 从一个扩增的 DNA 片段切出 AccIII 限制性内切酶基因，而该 DNA 片段是用 Acc 细菌的基因组 DNA 为模板通过 PCR 扩增得到的，并且当该基因被插入到 pACE611 载体的 NcoI 位点时，翻译密码子框在 SP6 启动子的一致控制之下。另一个方面，该引物对的使用引起从 AccIII 限制性内切酶基因 5' 端数第四个碱基的取代即 C 到 G，以及从该基因所编码的蛋白质 N 端数第 2 个氨基酸的取代基 Leu 到 Val。使用上述的引物时并以 (2) 所得 Acc 细菌的基因组 DNA 为模板，进行了基于 PCR 的碱基扩增。用限制性内切酶 NcoI（宝酒造社制）完全消化该 DNA 片段，之后进行琼脂糖凝胶电泳；收回了一个约 900bp 大小的 DNA 片段。接下来，将该 DNA 片段插入到 pACE611 载体的 NcoI 位点，从而使其定位于 SP6 启动子的下游。

将重组 DNA 导入大肠杆菌 JM109 中产生转化子。随后挑选 30 个转化子，并将每个转化子接种到 5ml 含 30 $\mu$ g/ml 氯霉素的 LB 培养基并于 37 $^{\circ}$ C 培养。当细胞培养肉汤 OD<sub>600</sub> 达到 0.6 小时，加入 IPTG 至终浓度 2mM 以增加质粒拷贝数，接着于 37 $^{\circ}$ C 培养 2 小时，之后从各培养物分别制备质粒 DNA，所获得的每种质粒 DNA 同时用限制性内切酶 HindIII 和 XbaI 切割，接着用琼脂糖凝胶电泳确定得到的 DNA 片段的长度；检测到了一个以正确的方向插入的含基因区的质粒。被命名为 pCRA19 的质粒被导入至大肠杆菌 JM109 中以产生转化子，该转化子被称为大肠杆菌 JM109/pCRA19。

#### (6) AccIII 限制性内切酶基因表达系统的构建

首先，将 SP6 RNA 聚合酶基因导入大肠杆菌 JM109 中以构建一个可进一步表达该基因的噬菌体基因。具体地说，通过将由用 BamHI-HindIII 消化 pSP6 - 2 而得到的 SP6 RNA 聚合酶基因片段插入至商业产品噬菌体载体 M13mp18（宝酒造社制）的多克隆位点中的 BamHI-HindIII 位点，从而构建了 SP6 RNA 聚合酶表达基因噬菌体 M13sp6。

接着，转化子大肠杆菌 JM109/pCRA19 被接种到 5ml 含 30 $\mu$ g/ml 氯霉素的 LB 培养基中并于 37 $^{\circ}$ C 培养。当细胞培养肉汤的 OD<sub>600</sub> 达到 0.6 时，加入 IPTG 至终浓度 2mM 以增加拷贝数，随后在 37 $^{\circ}$ C 进一步培养 2 小时。然后用噬菌体 M13sp6 感染这些细胞以表达由 AccIII 限制性内切酶基因所表达的蛋白质，随后于 37 $^{\circ}$ C 进一步培养 16 小时。将得到的湿

细胞中的一部分 11mg 悬浮于 180 $\mu$ l 的细胞破碎缓冲液 E 中〔 20mM Tris-HCl(pH7.5), 10mM 2-巯基乙醇〕, 之后超声破碎细胞, 然后离心 ( 18000g, 10 分钟) 将固体与液体分离。

在宝酒造社所产生 AccIII 限制性内切酶所附资料页所示的活性测定条件下测定上清的活性, 测定表明产量为每克湿细胞产生约 8000 单位的 AccIII 限制性内切酶, 约是同每单位重量湿 Acc 细胞得到活性的 16 倍。在所得上清既没有看到 AccIII 之外其它限制性内切酶的活性, 也没有看到 AccIII 修饰酶活性。对导入到质粒 pCRA19 中的 AccIII 限制性内切酶基因来说, 表明从翻译起点碱基开始数, 第 4 个碱基在用 PCR 扩增 DNA 时被由 C 置换成了 G, 结果导致从所翻译的蛋白质的 N - 末端数第 2 个氨基酸被取代, 即由 Leu 到 Val, 并且该蛋白质具有 AccIII 限制性内切酶活性。

由此开发了一个不存在 AccIII 修饰酶的 AccIII 限制性内切酶基因的分离/物质生产系统。

#### 15 (7) AccIII 修饰酶基因的分离

修饰酶基因的限制性内切酶基因定位经常很靠近。因此, 为了确定上述(4)中所推测的 ORF2 区域, 从而测定了一个约 1.4kb DNA 片段的碱基序列, 该 DNA 片段是以用 EcoRI 盒式文库为模板所制备的用于测定上述(4)AccIII 限制性内切酶基因 3'区的扩增的 DNA 片段中得到的。为了测定上述(4)中所推测的 ORF3 区域, 从用 BglII 盒式文库为模板制备的用于测定上述(4)中的 AccIII 限制性内切酶基因 5 - 末端区的扩增的 DNA 片段中得到一个约 2.2kb 的 DNA 片段, 测定了其碱基序列。结合这些碱基序列和上述(4)中测定的含有 AccIII 限制性内切酶基因的约 1.6kb DNA 片段的碱基序列, 从而得到了如序列表中序列号 29 所示的约 4.2kb 碱基序列的信息。AccIII 限制性内切酶基因定位于碱基编号 1913 到 2797, ORF2 位于碱基编号 3712 到 2789, 而 ORF3 位于碱基编号的 691 到 1890。在这些 ORF 中, 证明 ORF2 含有一个与已知的各种限制修饰系统的修饰酶的保守区高度同源的部分。

接着, 为了得到 ORF3 区, 用 EcoT22I 和 HpaI (均由宝酒造社产) 从扩增的 DNA 片段切下了一个约 1.1kb 的含 ORF<sub>2</sub> 的部分, 其中扩增的 DNA 片段是在各自的步骤中分别使用 Acc 引物 6 和盒式引物 C1 引物对及另一引物对 Acc 引物 5 和盒式引物 C2 以 EcoRI 盒式文库作模板通过两

从扩增的 DNA 片段切下了一个约 1.1kb 的含 ORF2 的部分，其中扩增的 DNA 片段是在各自的步骤中分别使用 Acc 引物 6 和盒式引物 C1 引物对及另一引物对 Acc 引物 5 和盒式引物 C2 以 EcoRI 盒式文库作模板通过两步 PCR 得到的，并将该 1.1kb 片段插入到了 pUC118 载体 lac 启动子下游的 SmaI 位点。如果该重组质粒含 AccIII 修饰酶基因，并且如果 AccIII 修饰酶可在大肠杆菌中表达，则由通过将该重组质粒导入大肠杆菌 JM109 中而得到的转化子培养物中的 DNA 将会被 AccIII 修饰酶甲基化并获得对 AccIII 限制性内切核酸酶切割的抗性。

但是由于用于构建该重组质粒的 pUC118 载体不含有 AccIII 限制性内切核酸酶识别序列，用该重组质粒本身去确证 AccIII 修饰酶活性就不太适合了。另一方面，在质粒 pSTV29 中有 AccIII 限制性内切核酸酶识别序列，它可与该重组质粒在大肠杆菌 JM109 中共存。因此，用 pSTV29 作为 AccIII 修饰酶表达的指示，将上述的重组质粒和 pSTV29 同时导入大肠杆菌 JM109 中以产生转化子。将 3 个转化子各自于 37℃ 在含 100μg/ml 氨苄，30μg/ml 氯霉素和 2mM IPTG 的 LB 培养基中培养 16 小时，之后从各培养物中制备质粒 DNA。

由此制备的 DNA 是质粒 pSTV29 和上述重组质粒的混合物。当用 AccIII 限制性内切核酸酶消化每种 DNA 样品时，所有样品中的 pSTV29 都表现出对 AccIII 限制性内切核酸酶的抗性，表明 AccIII 修饰酶基因被插入到了 DNA 样品中所含的重组质粒中。而且，当用限制性内切核酸酶 Hind III 和 XbaI 同时消化该 DNA 样品并随后用琼脂糖凝胶电泳分析所得到的 DNA 片段的长度时，证实了重组质粒中 ORF2 的存在。从而证明 ORF2 是 AccIII 修饰酶基因。所得的 AccIII 修饰酶基因和由其推导而得的氨基酸序列分别如序列表中序列号 14 和序列号 13 所示。

根据本发明所证明的 AccIII 限制修饰系统基因的结构如图 16 所示，其中 M，R 和箭头分别代表修饰酶基因，限制性内切核酸酶基因和 ORF 的方向。

#### 工业应用性

本发明首次提供了一种能将其编码的蛋白质对宿主致死或有害的外源基因导入宿主中的质粒，以及一种能用该质粒载体有效表达该蛋白质的方法。还提供一种在现有技术中是困难的在无修饰酶基因共存时分离构成限制修饰系统的限制性内切核酸酶基因的方法。而且，本发明分离

了 AccIII 限制性内切核酸酶基因和 AccIII 修饰酶基因，并且与现有的纯化方法相比用含该基因的质粒转化的大肠杆菌，可以容易的得到目的纯度的在基因工程中有用的 AccIII 限制性内切核酸酶或 AccIII 修饰酶的酶制品。

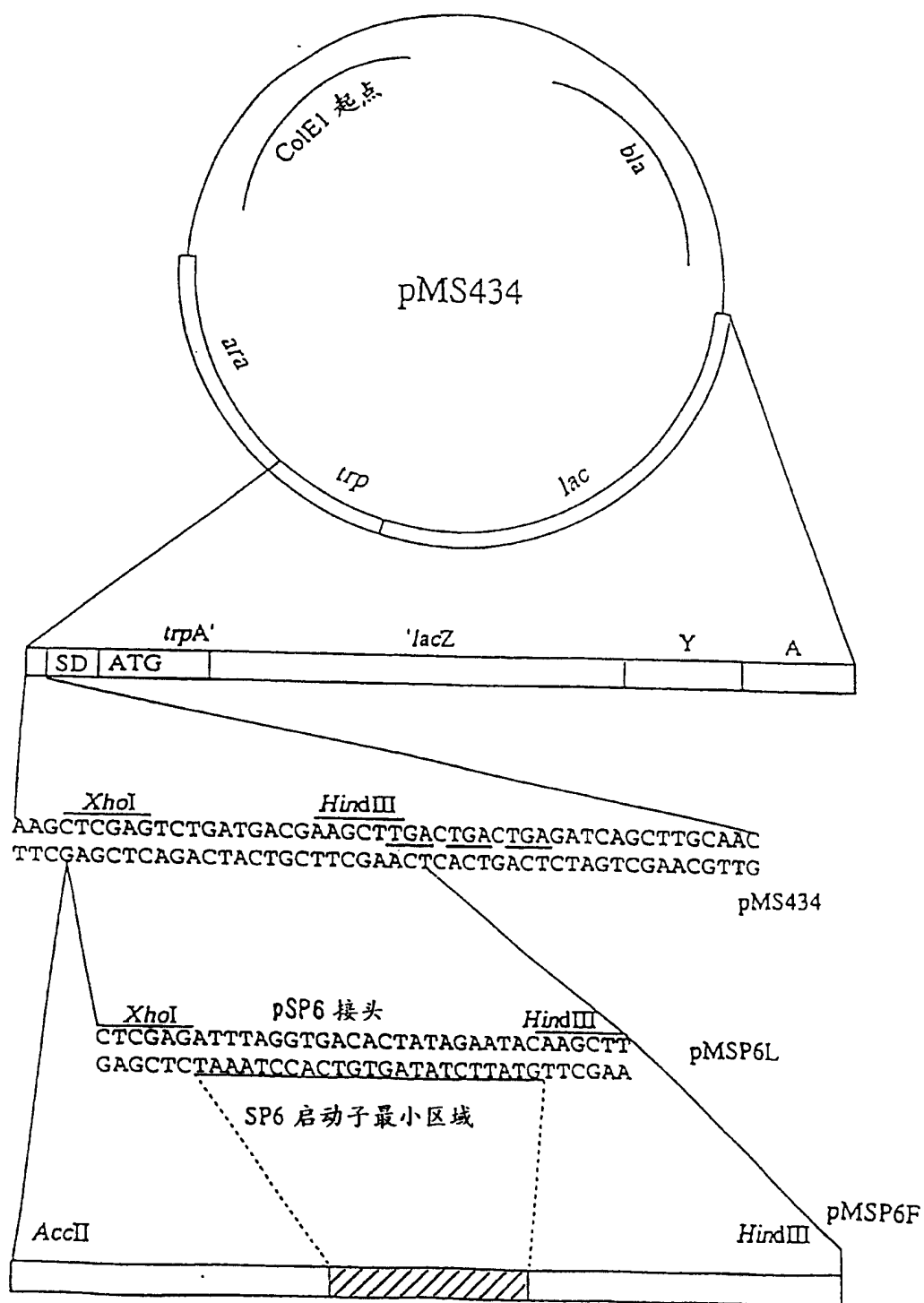


图 1

pSP64 由来的 Acc II-Hind III 片段

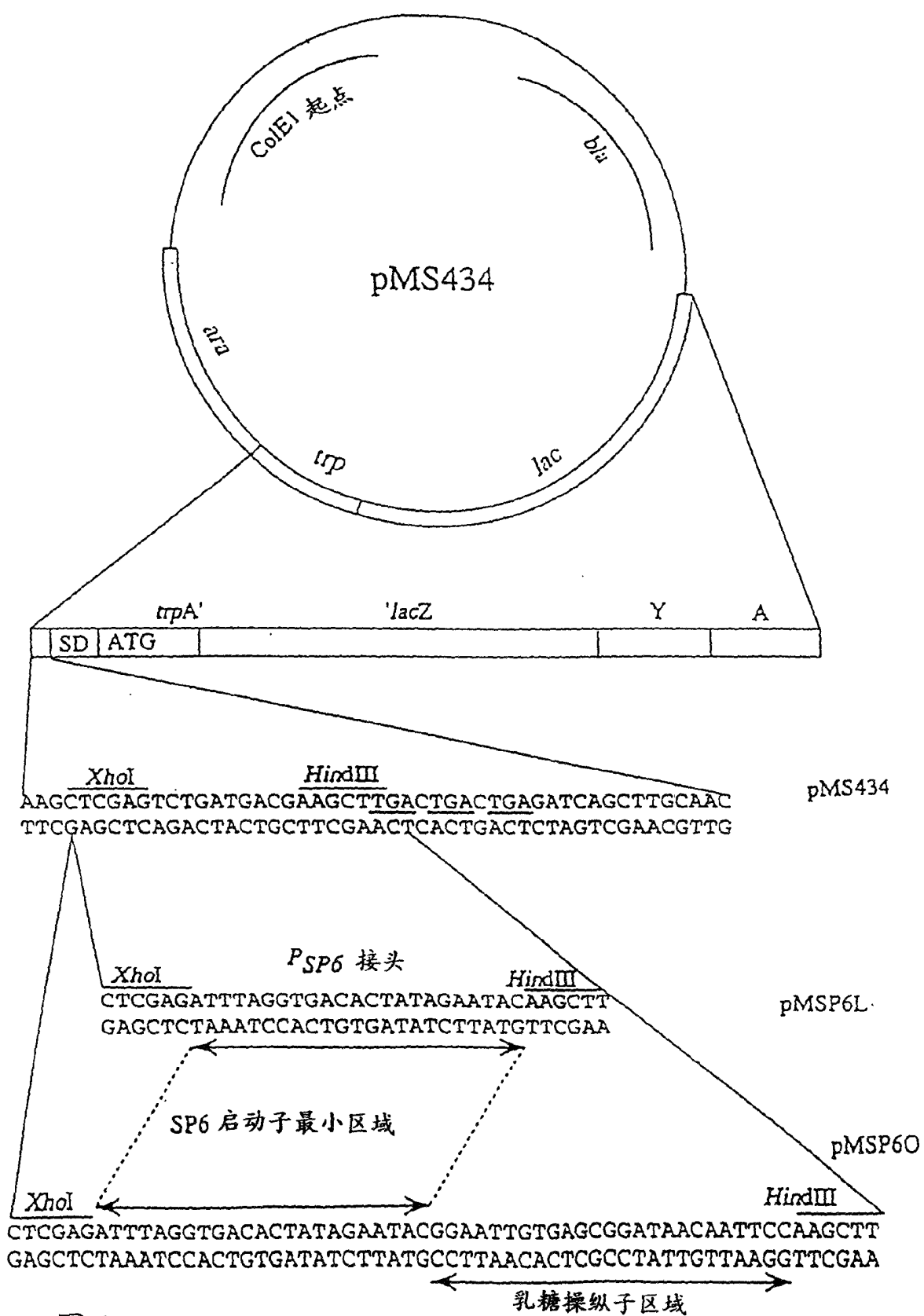
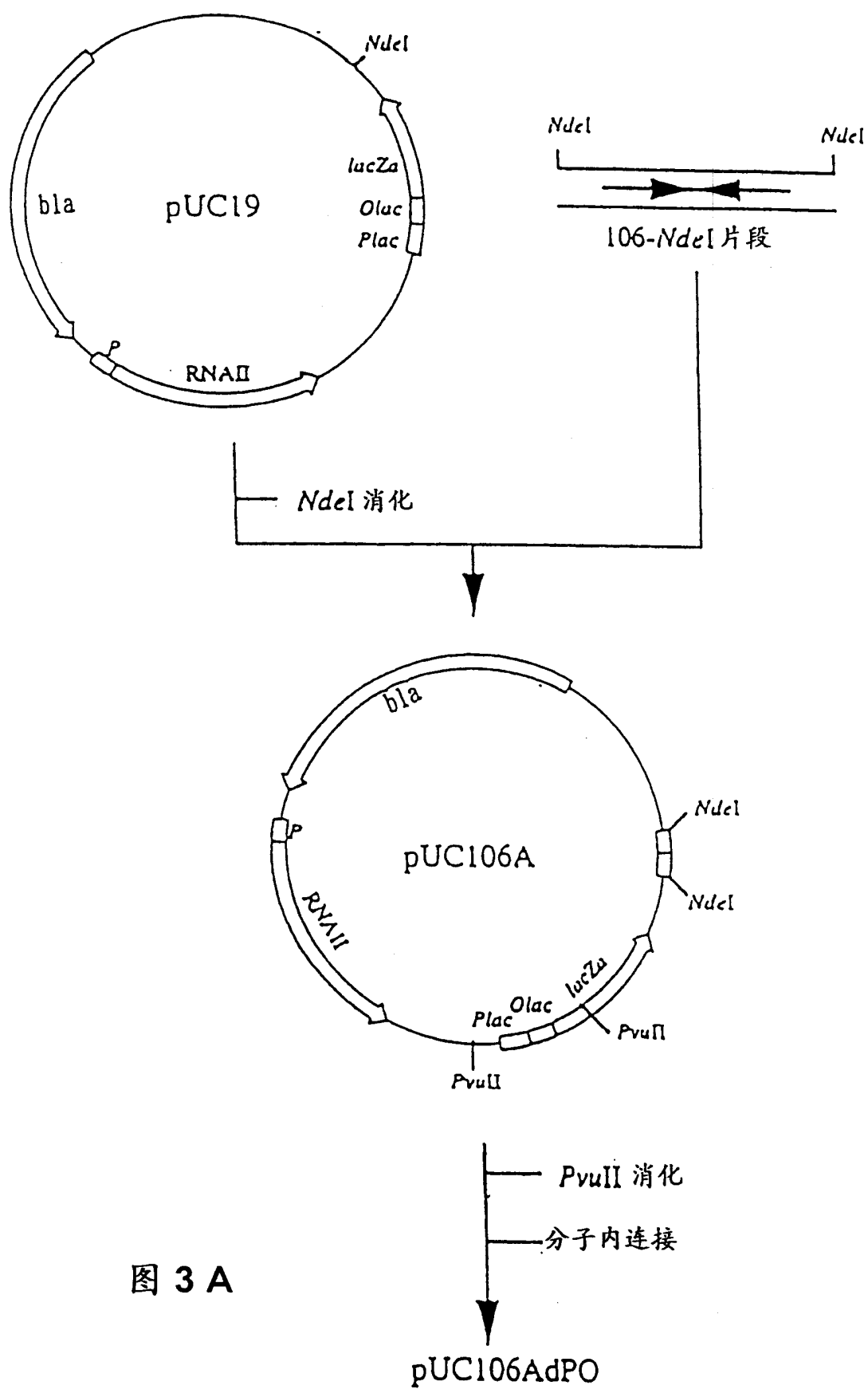


图 2

 $P_{SP6}-O_{lac}$  接头



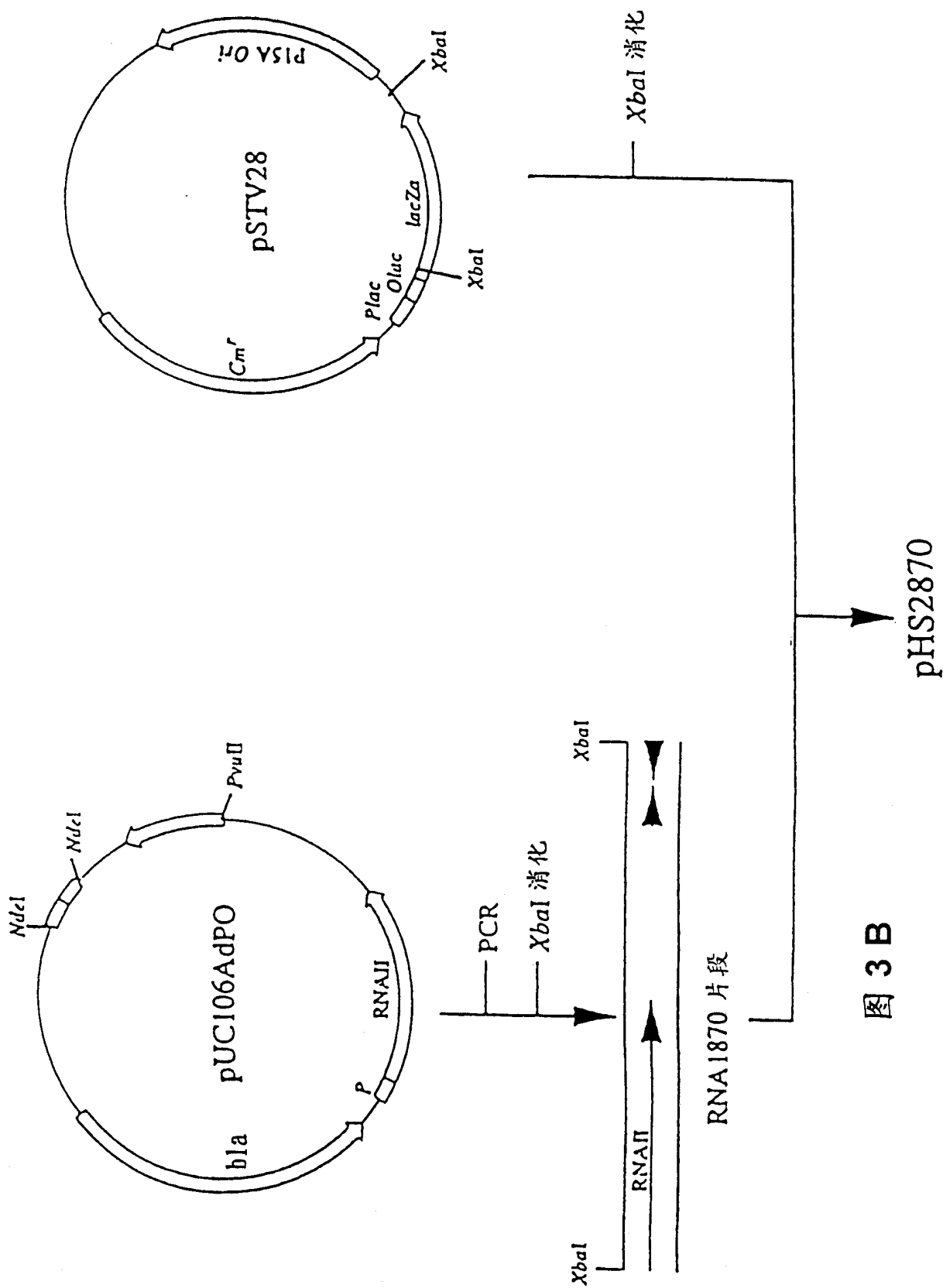


图 3 B

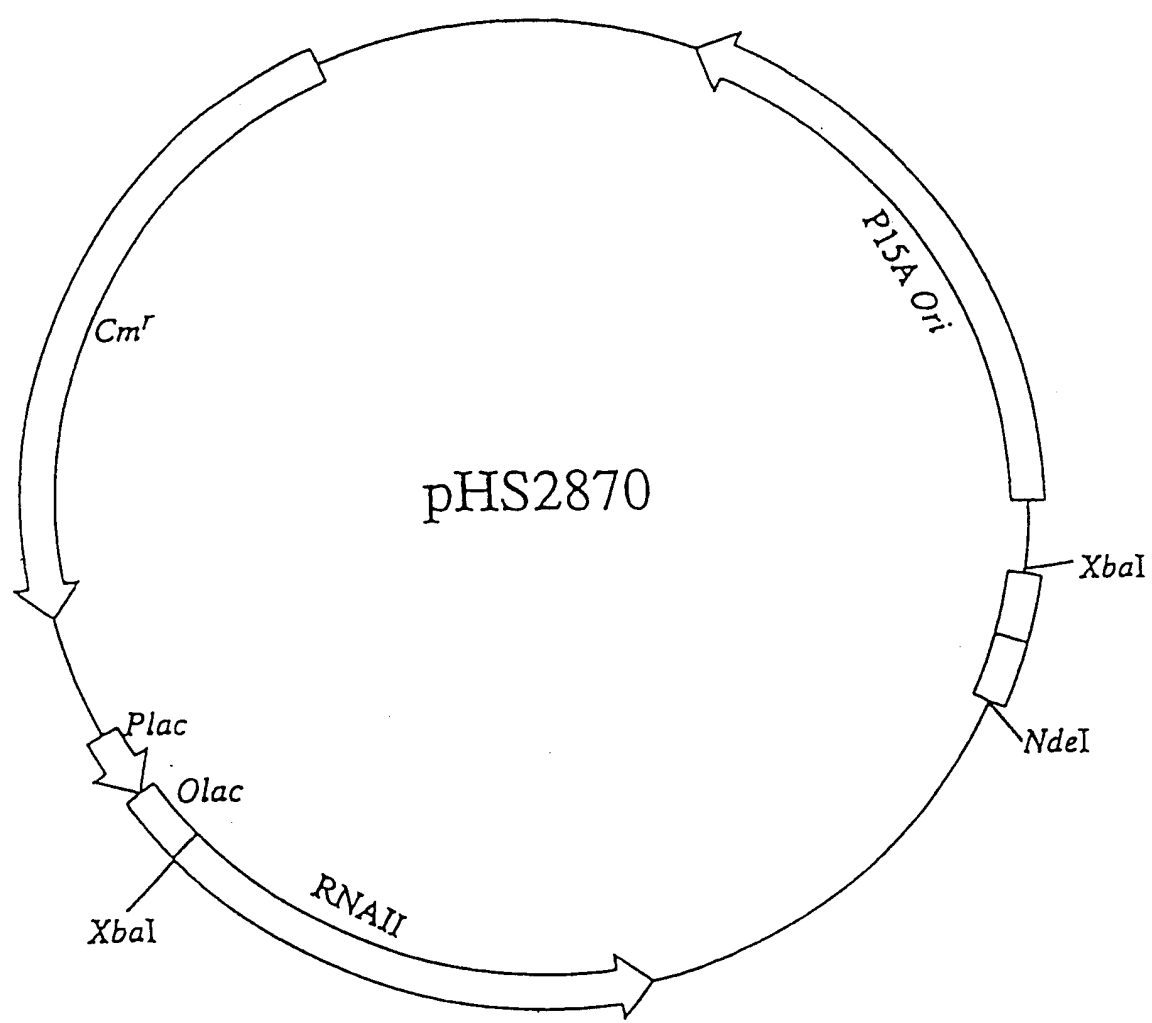


图 4

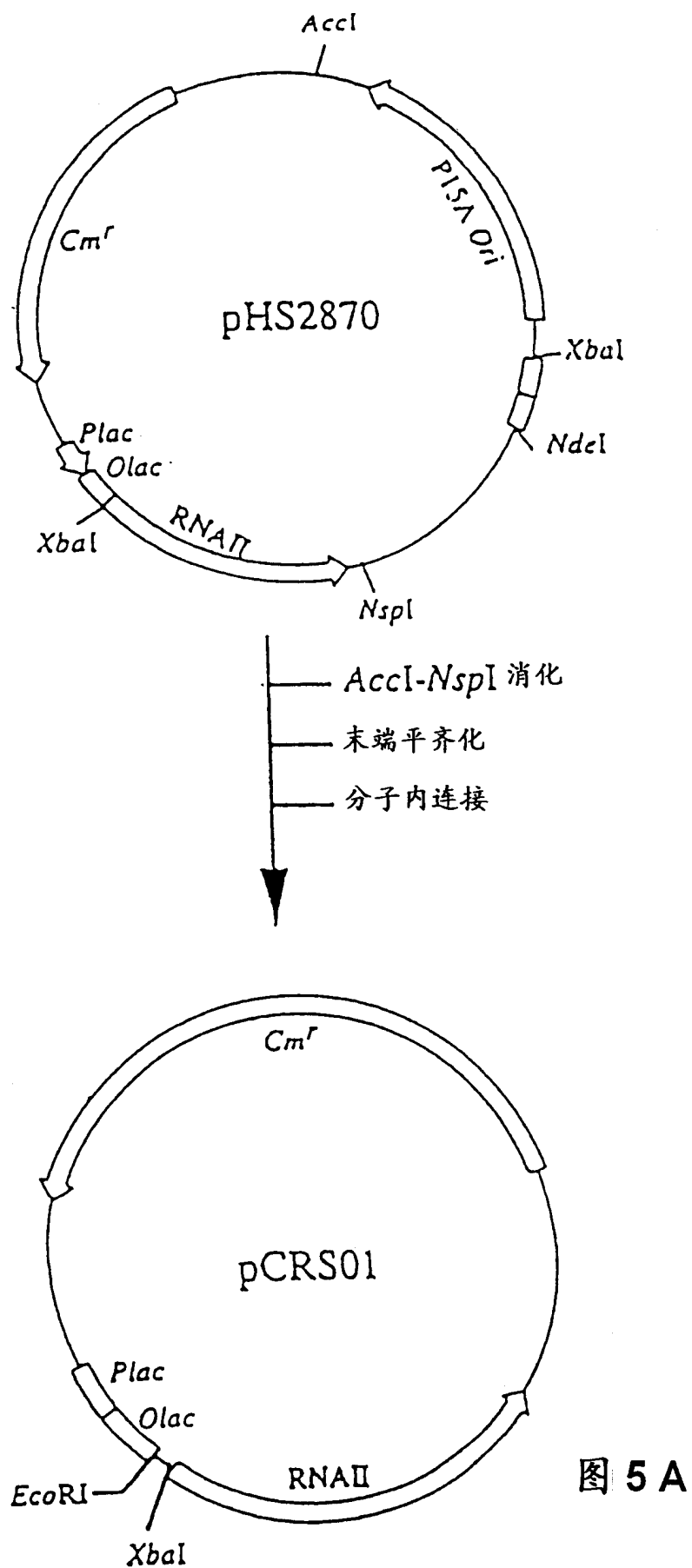
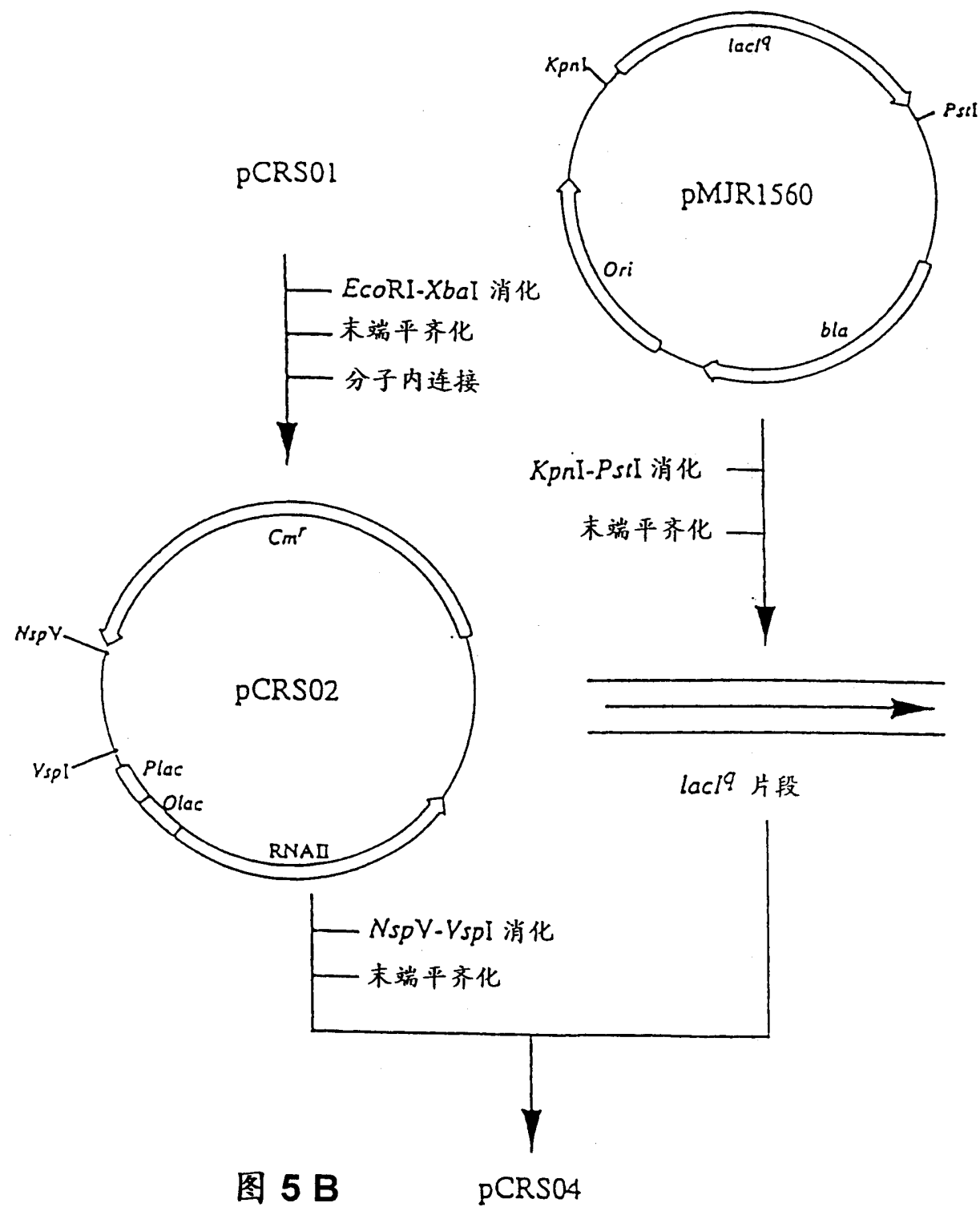


图 5 A



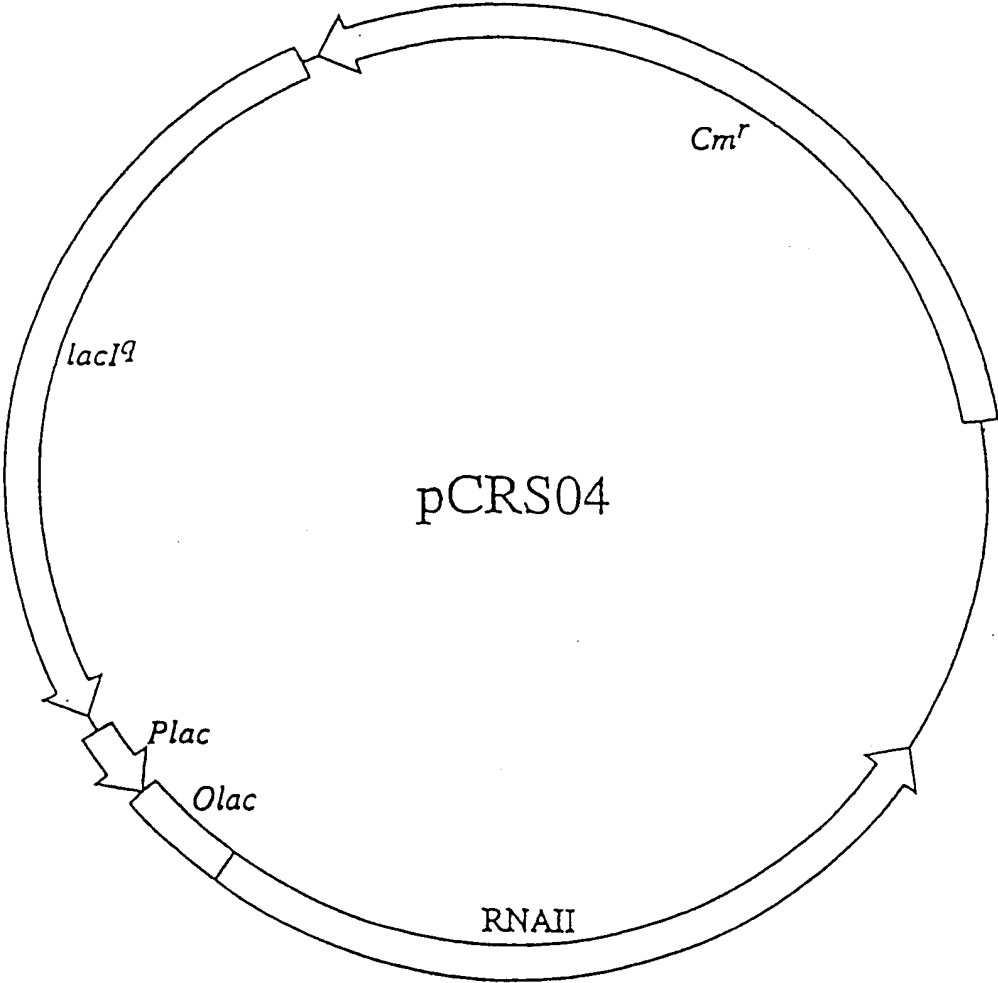


图 6

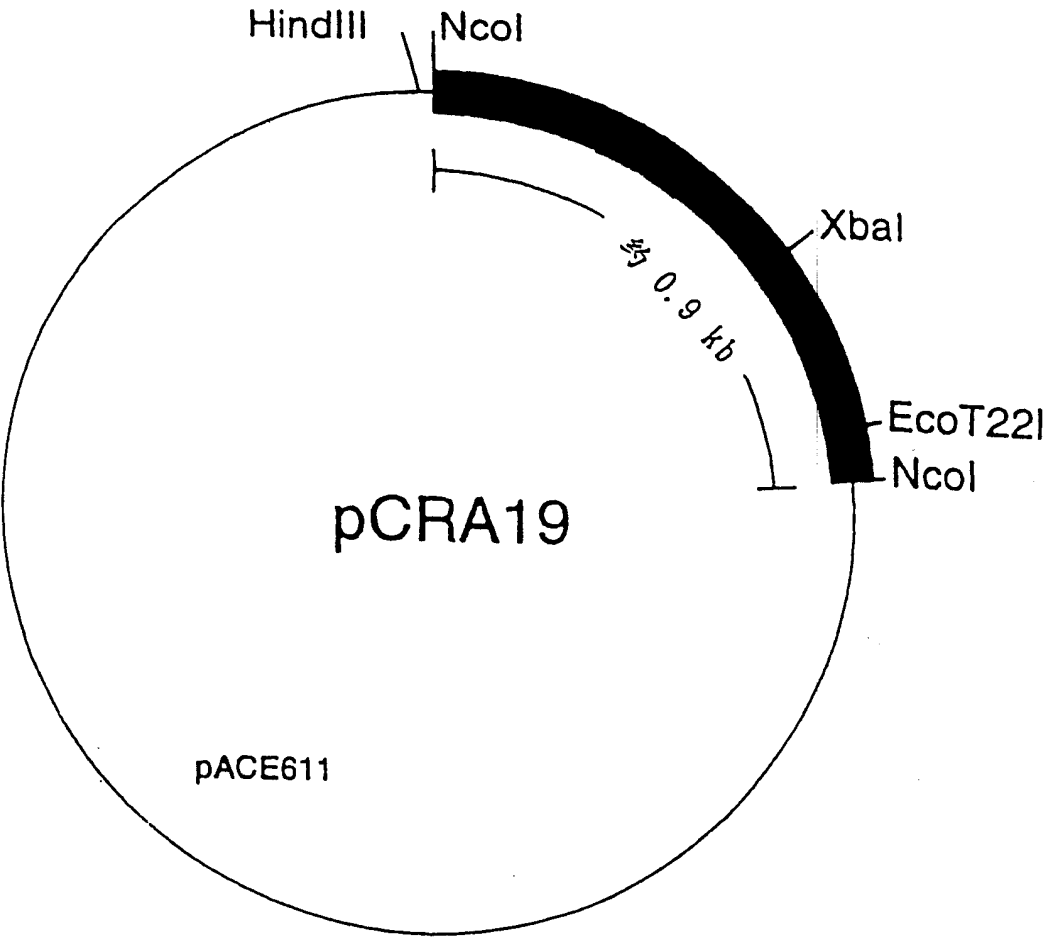


图 7

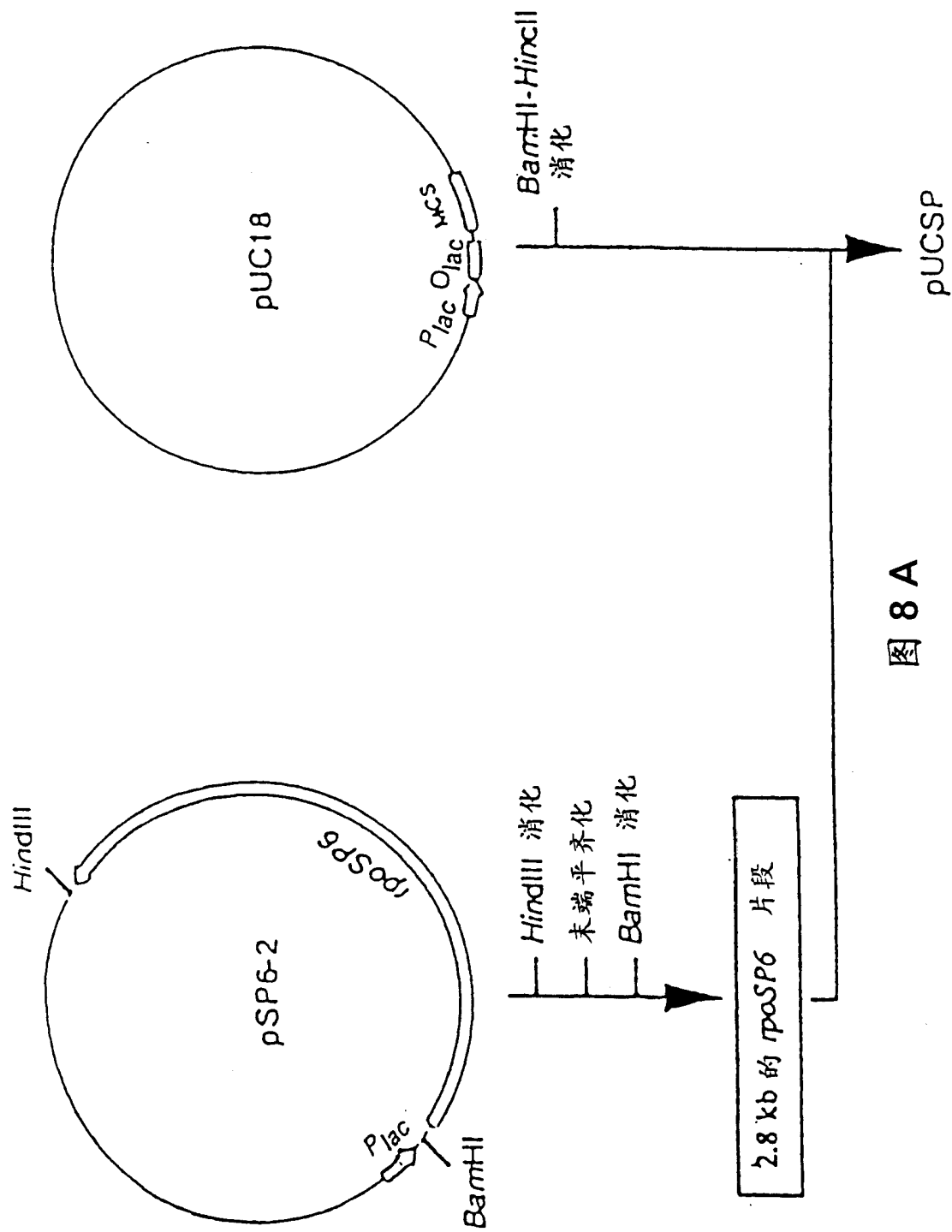
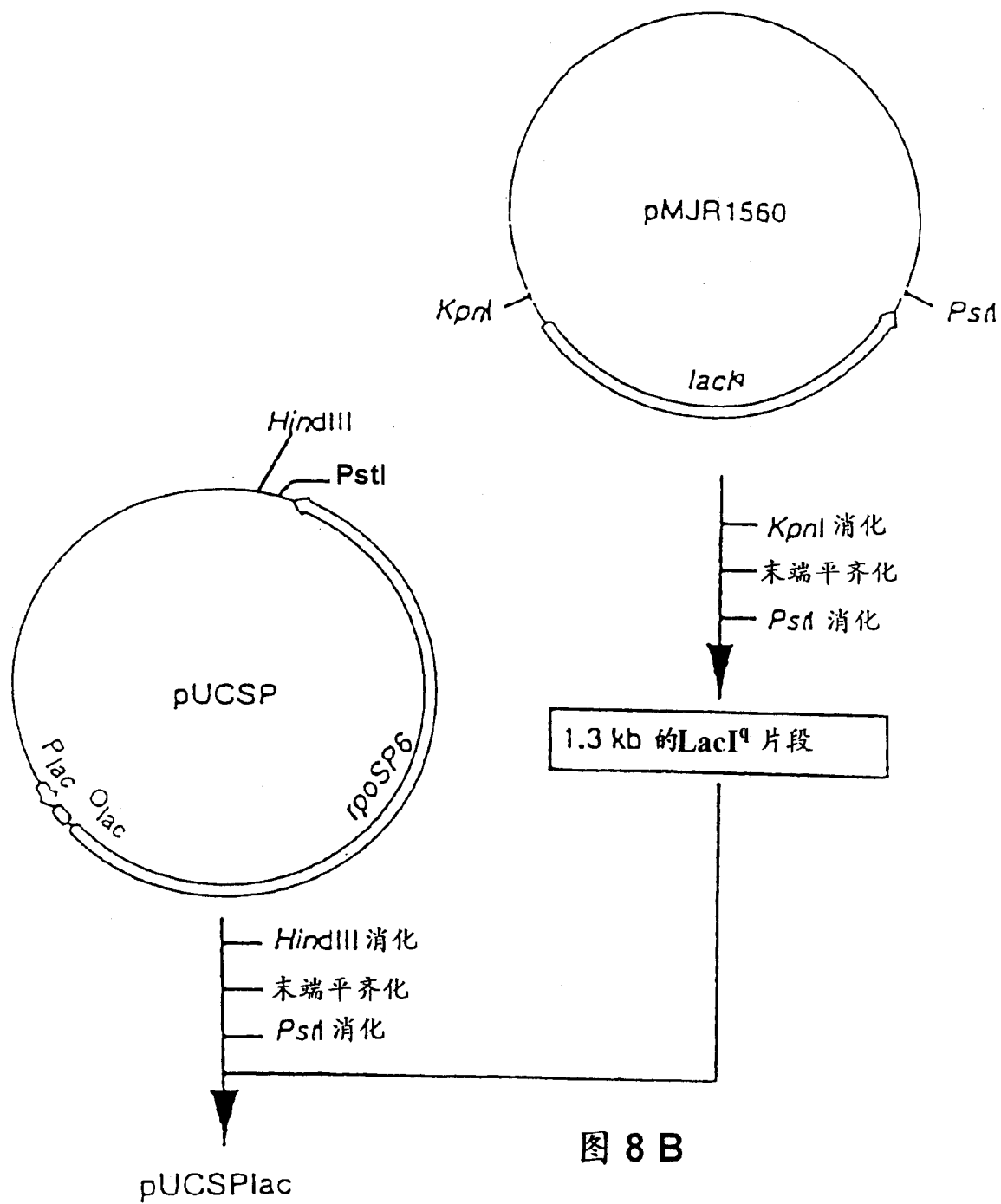


图 8 A



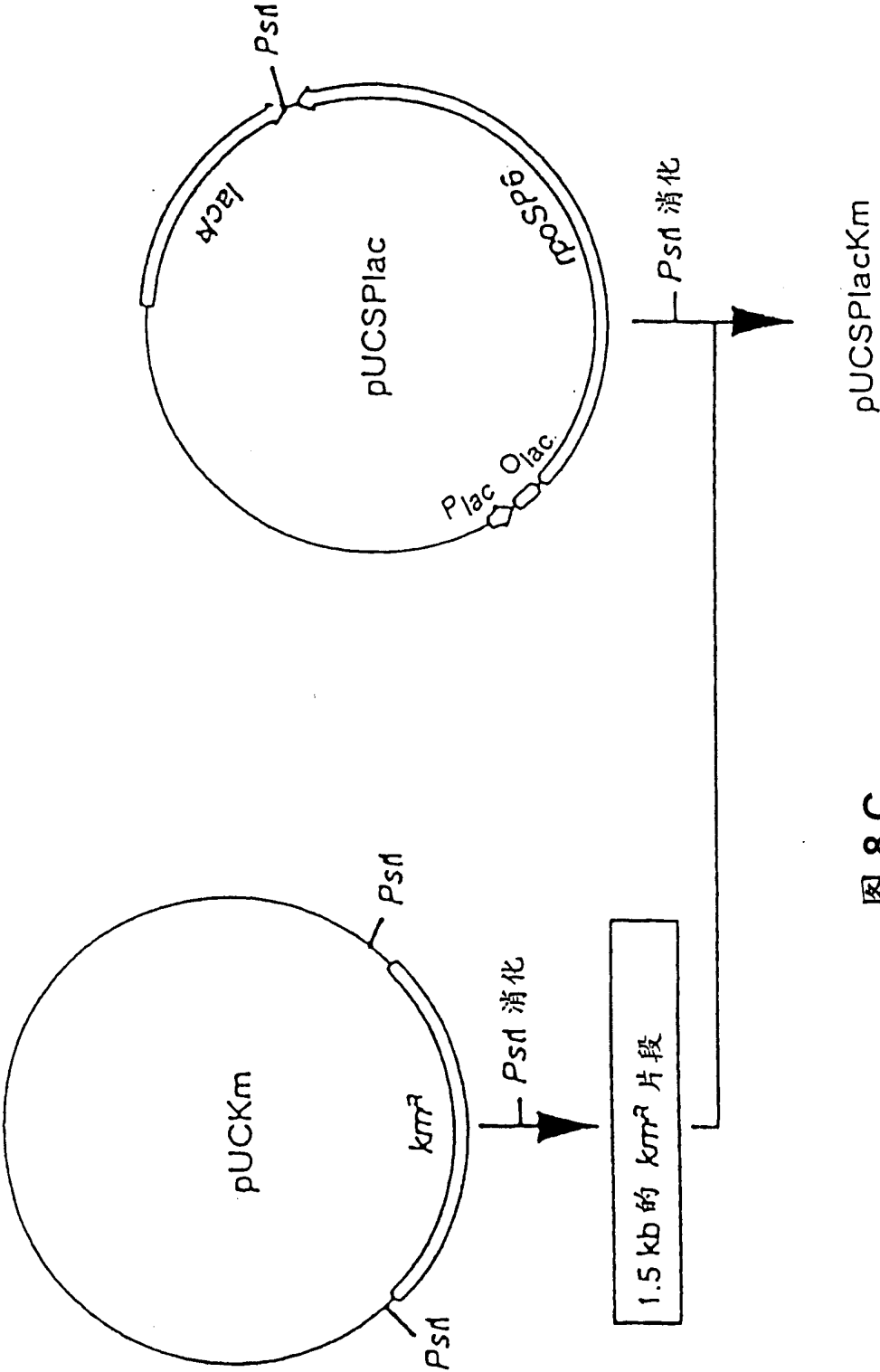


图 8 C

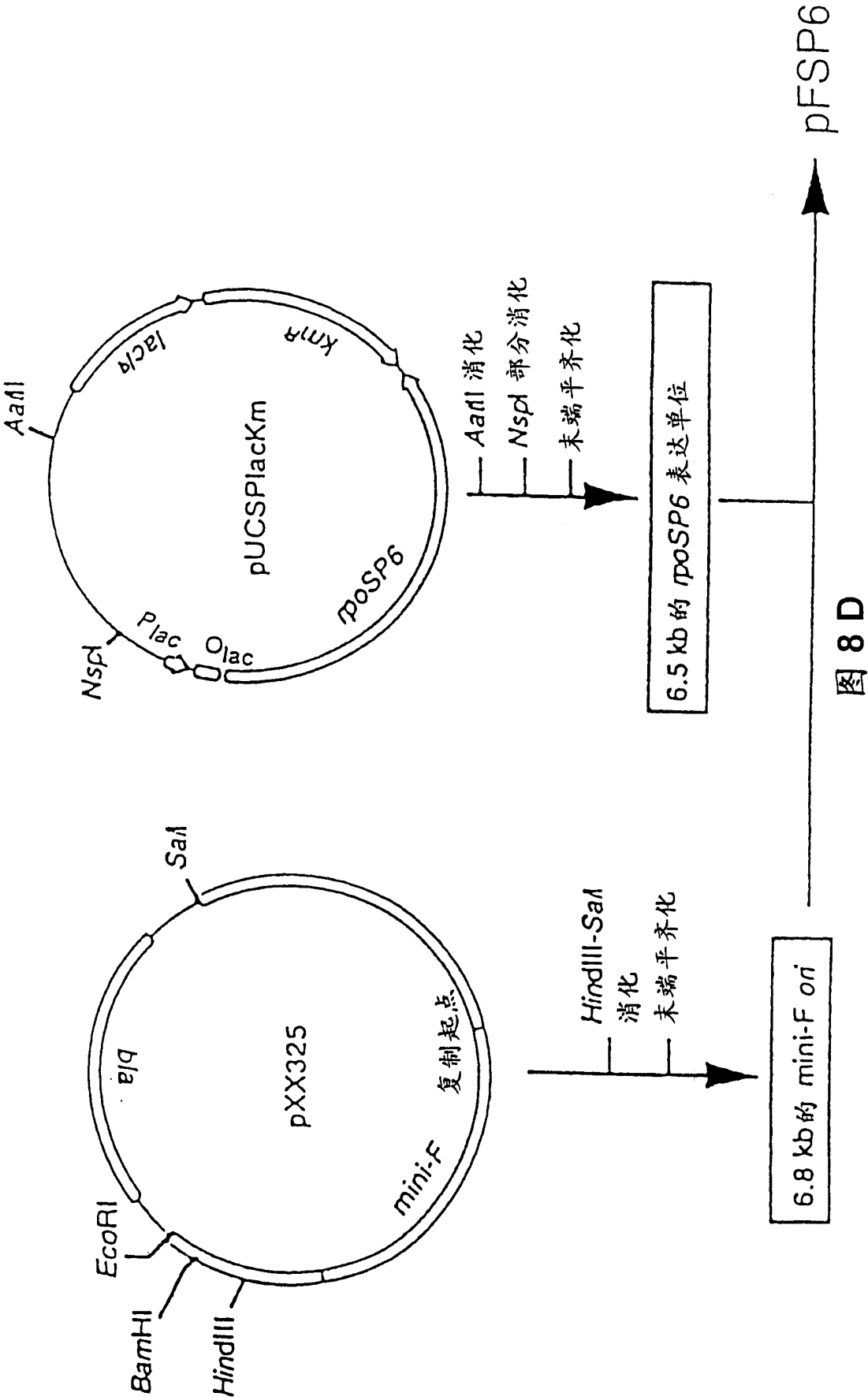


图 8 D

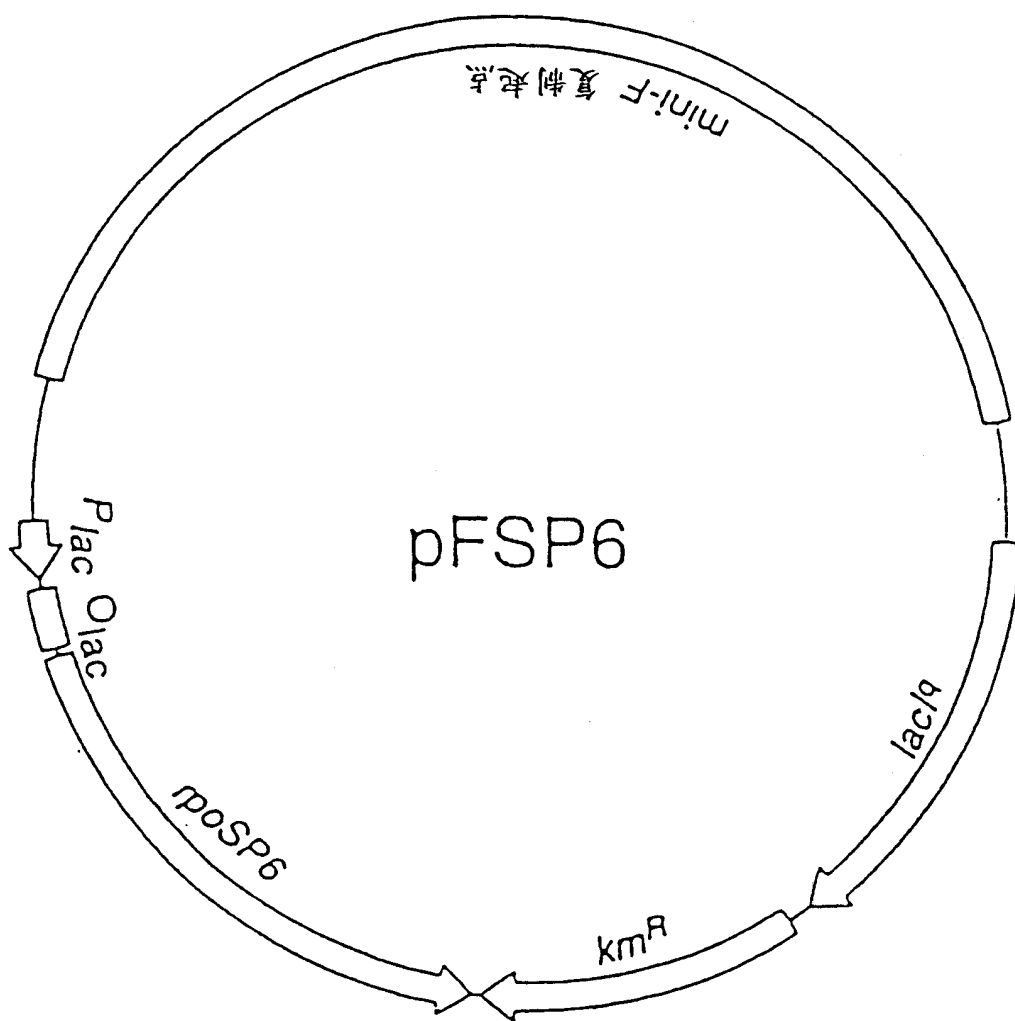


图 8 E

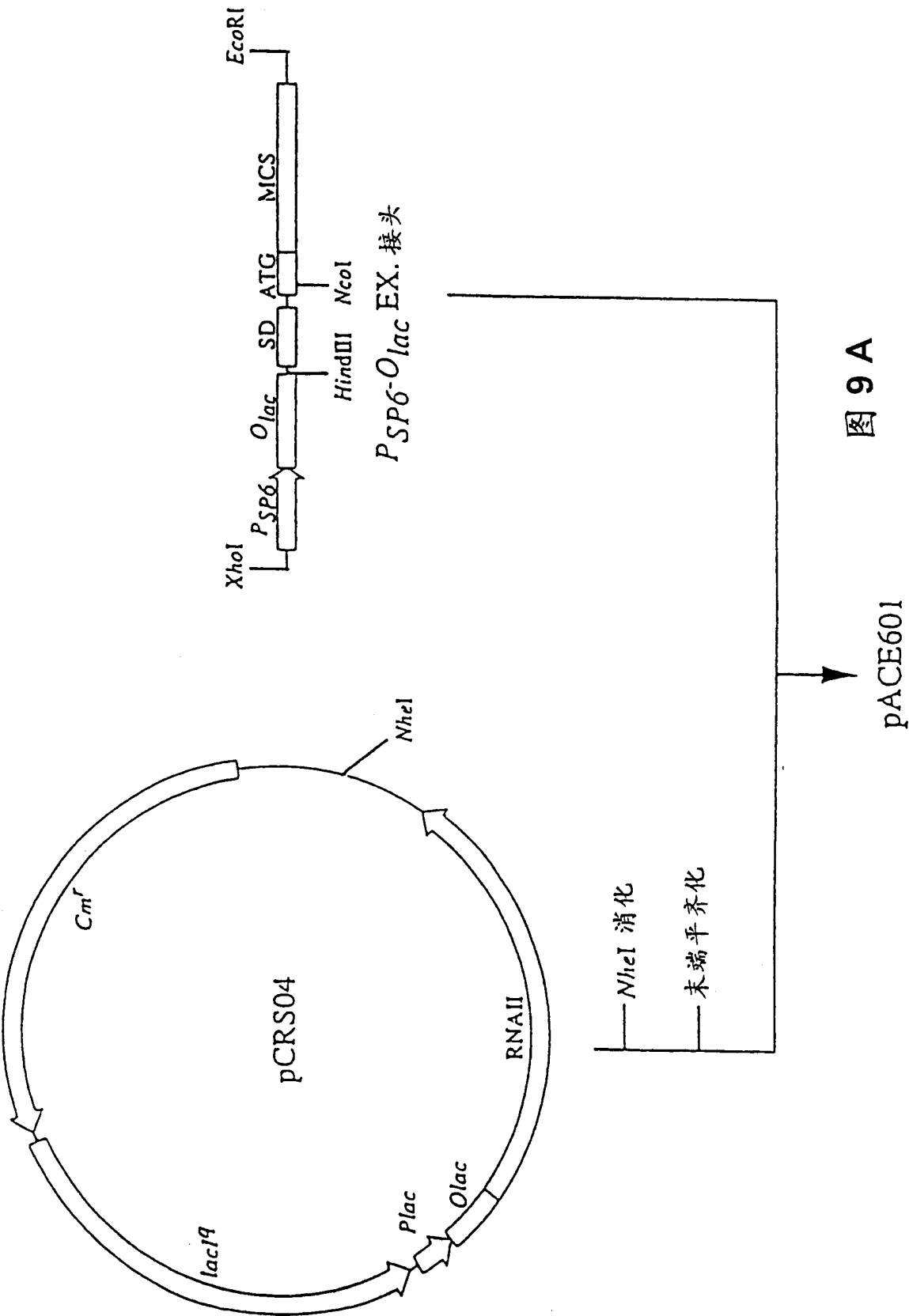


图 9 A

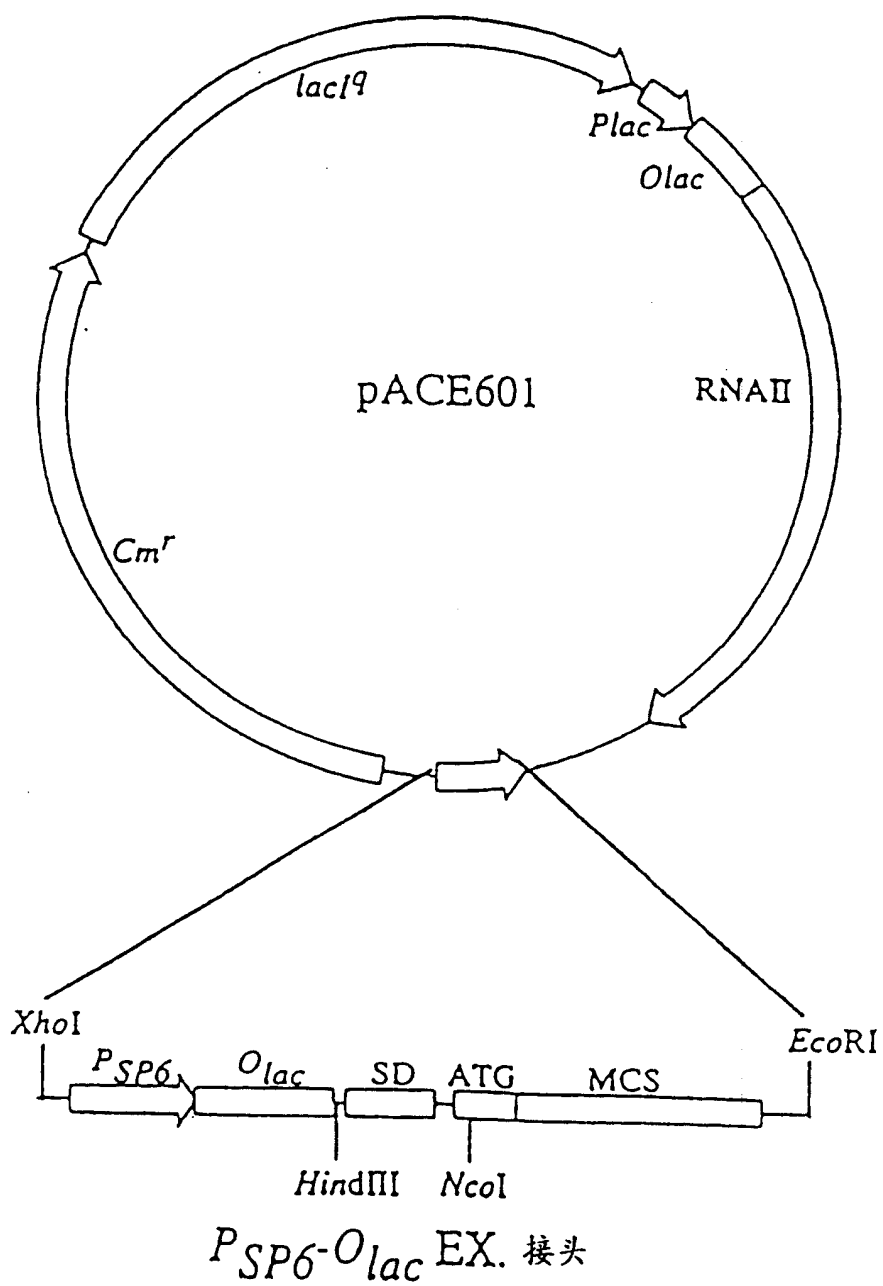


图 9 B

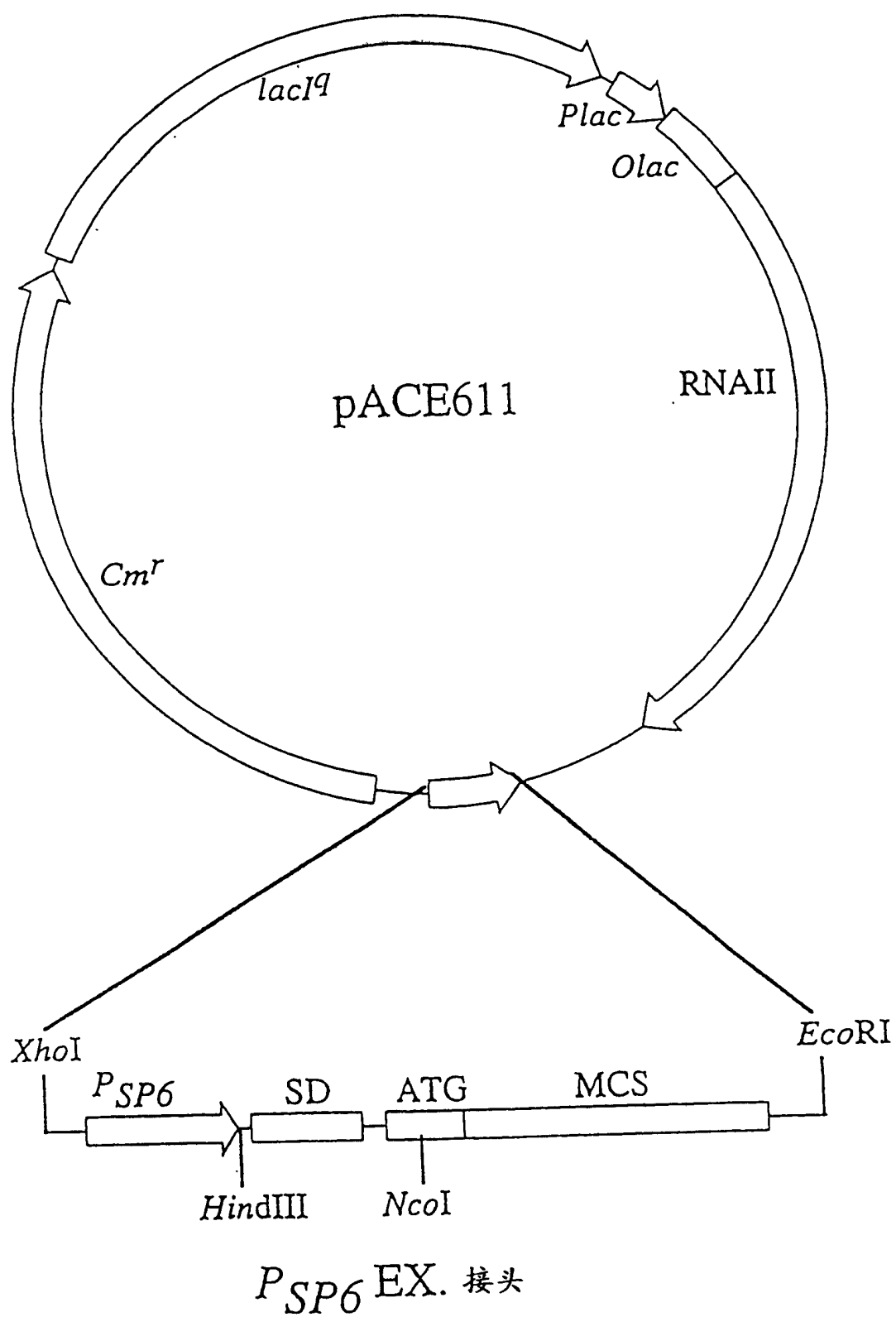
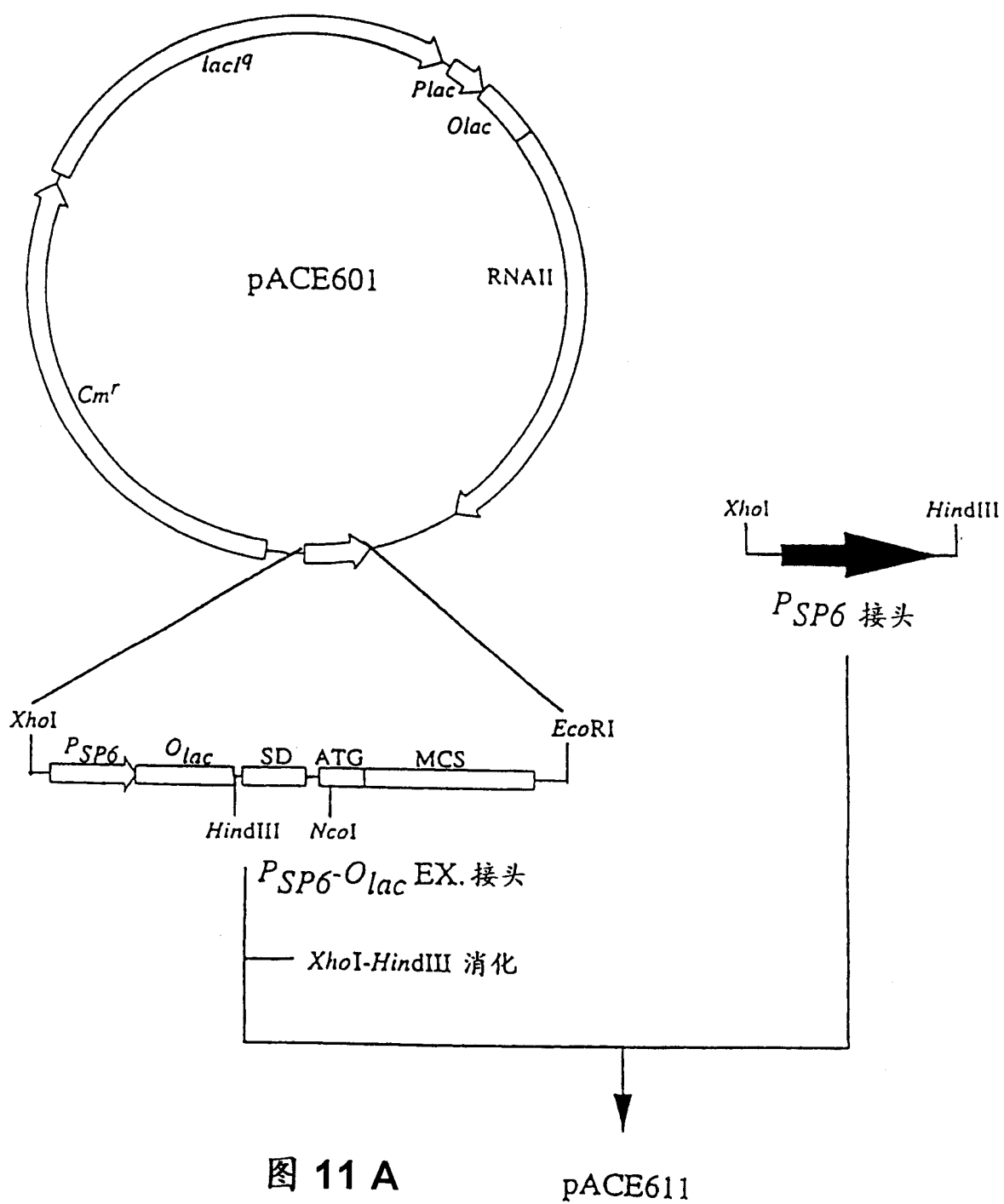


图 10



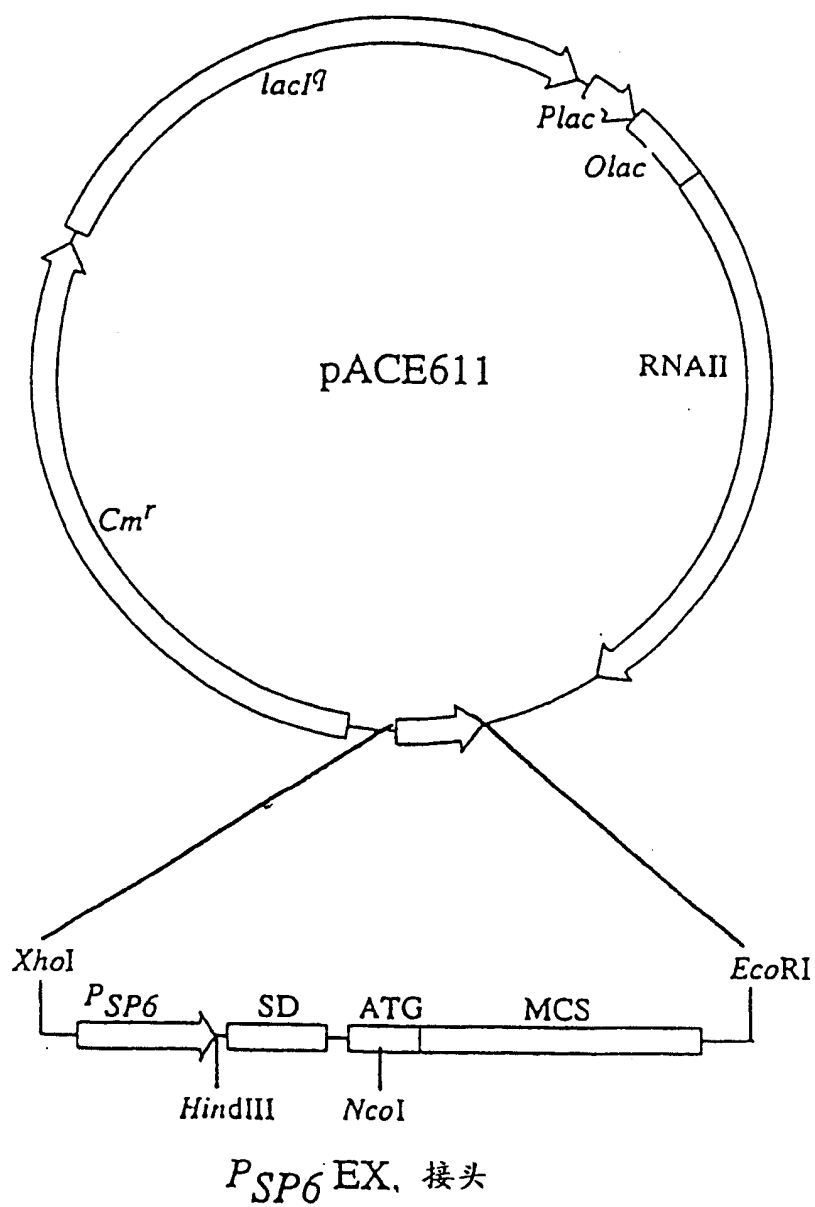


图 11 B

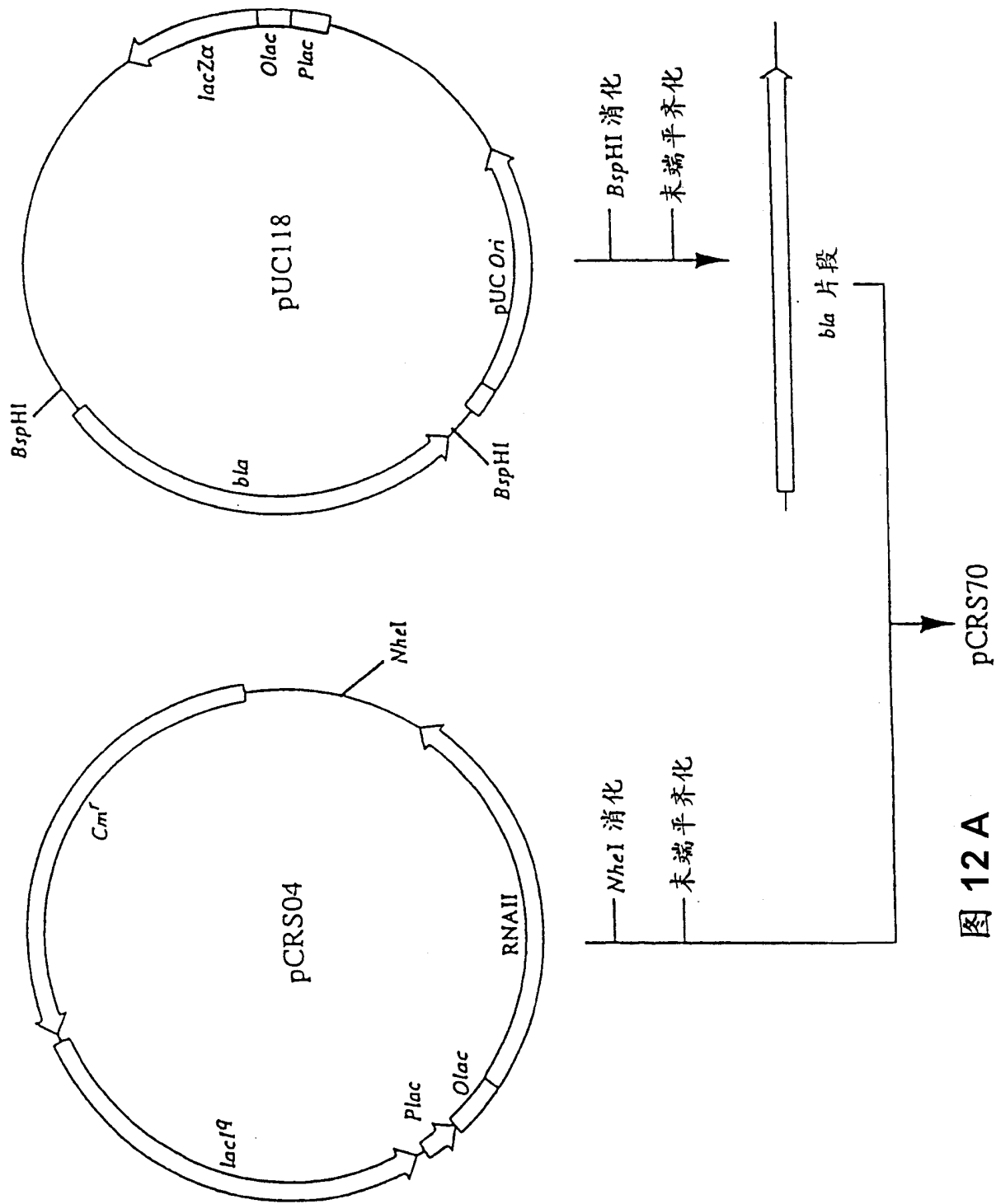


图 12 A

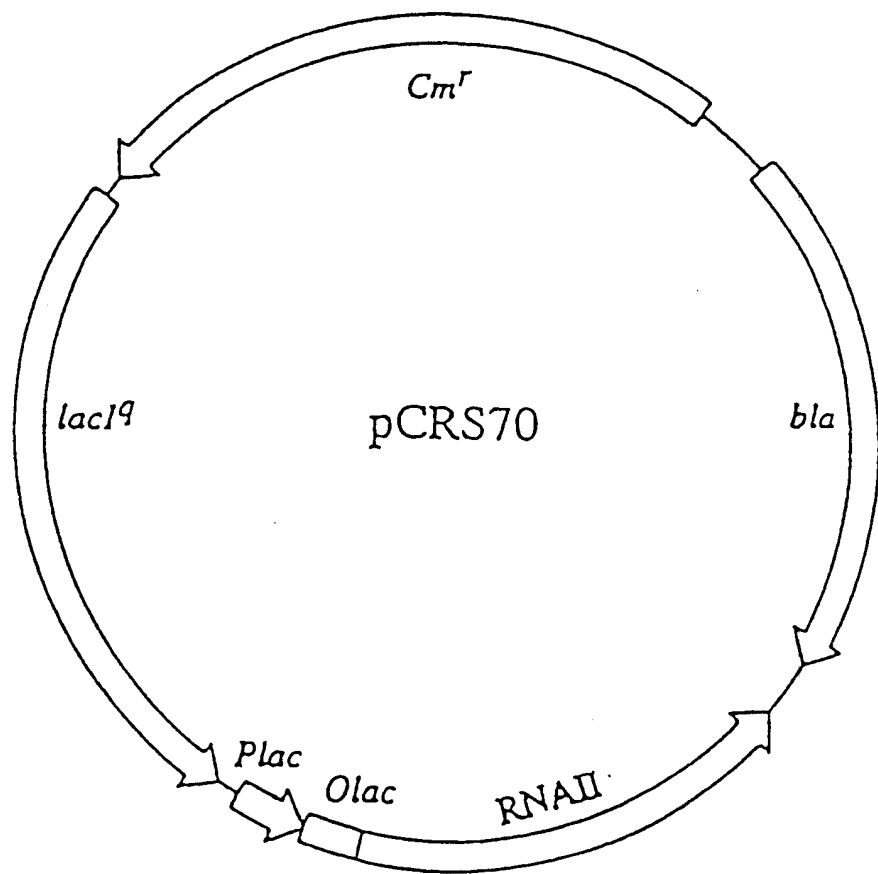


图 12 B

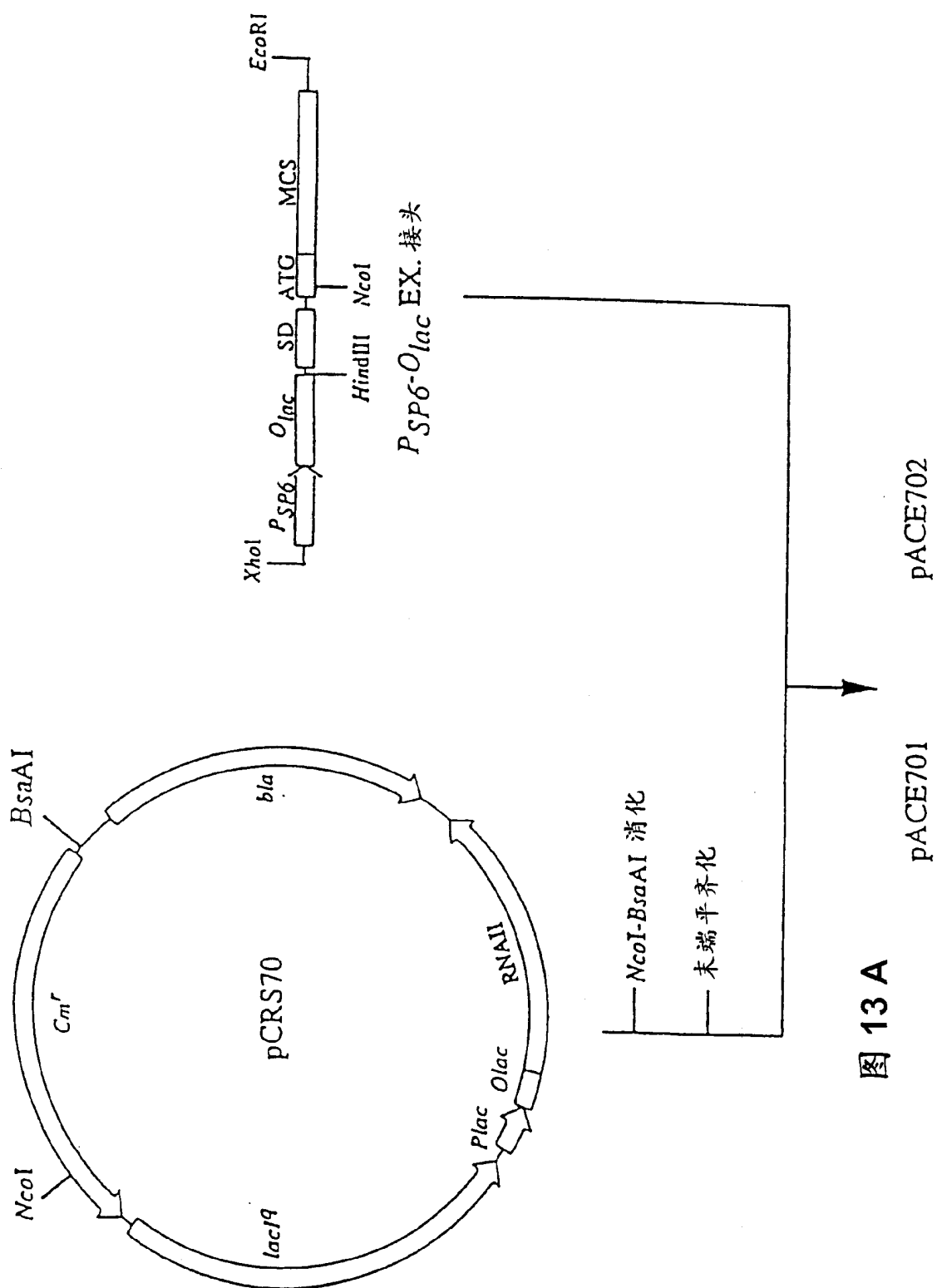


图 13 A

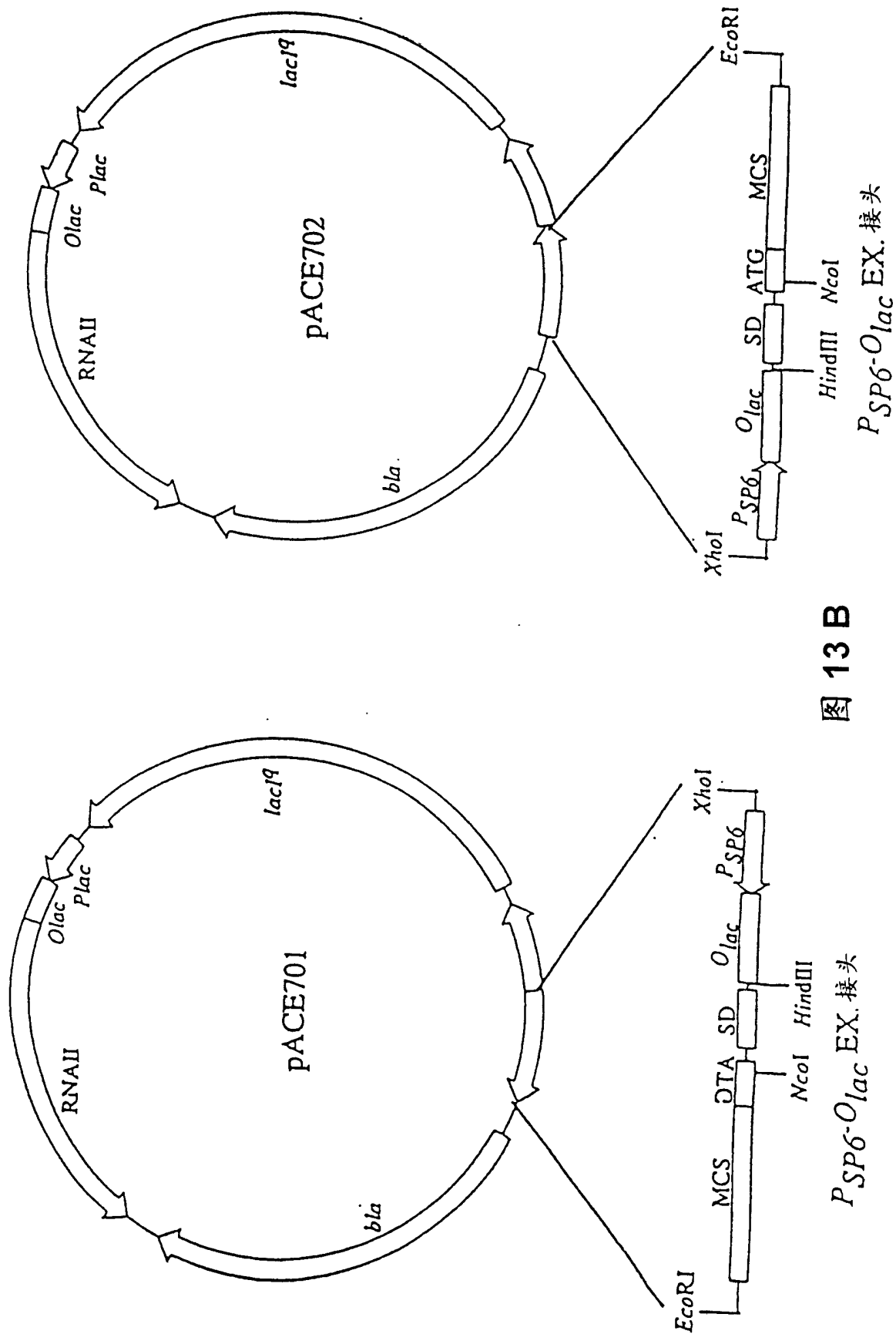


图 13 B

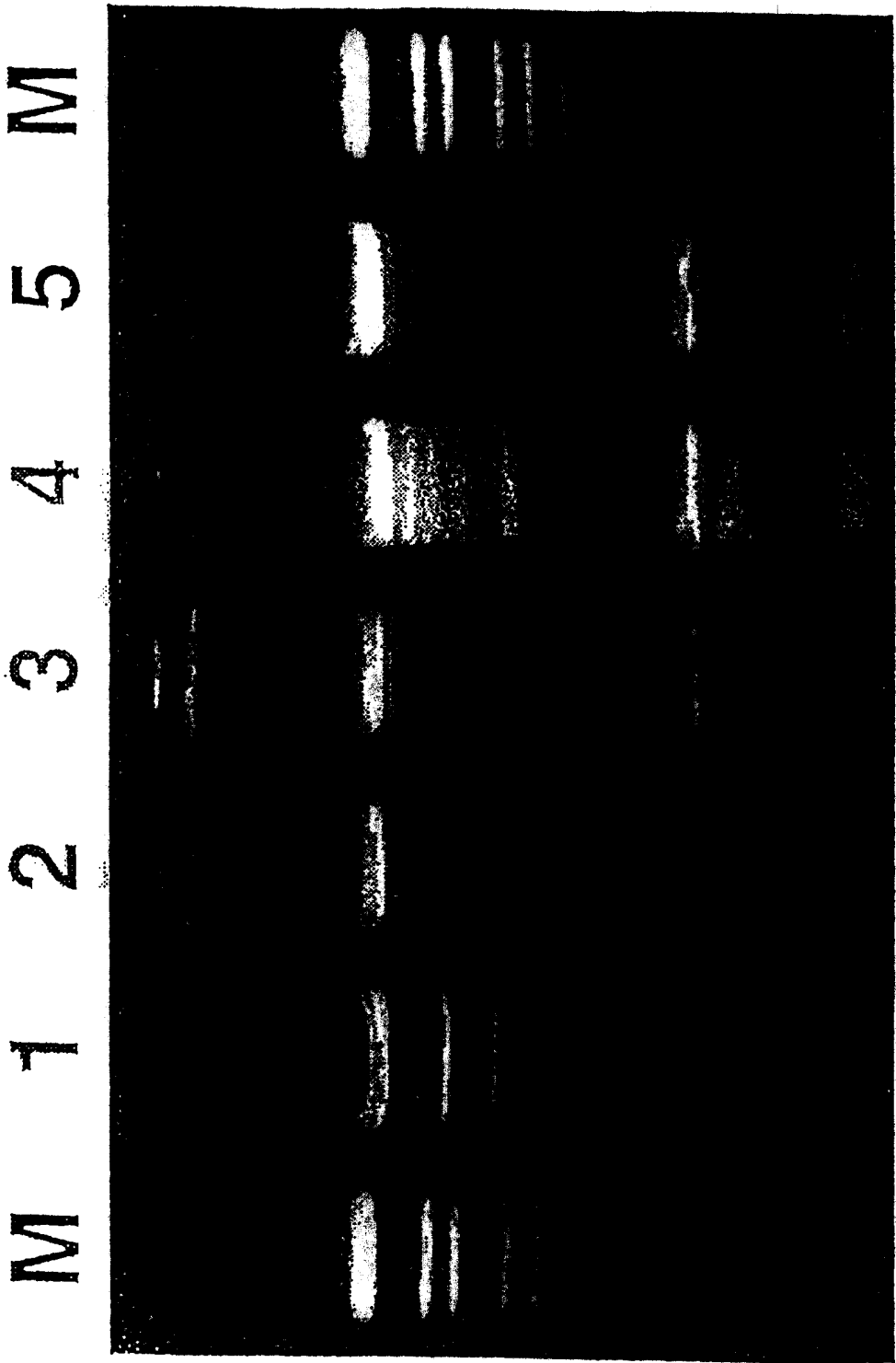


图 14

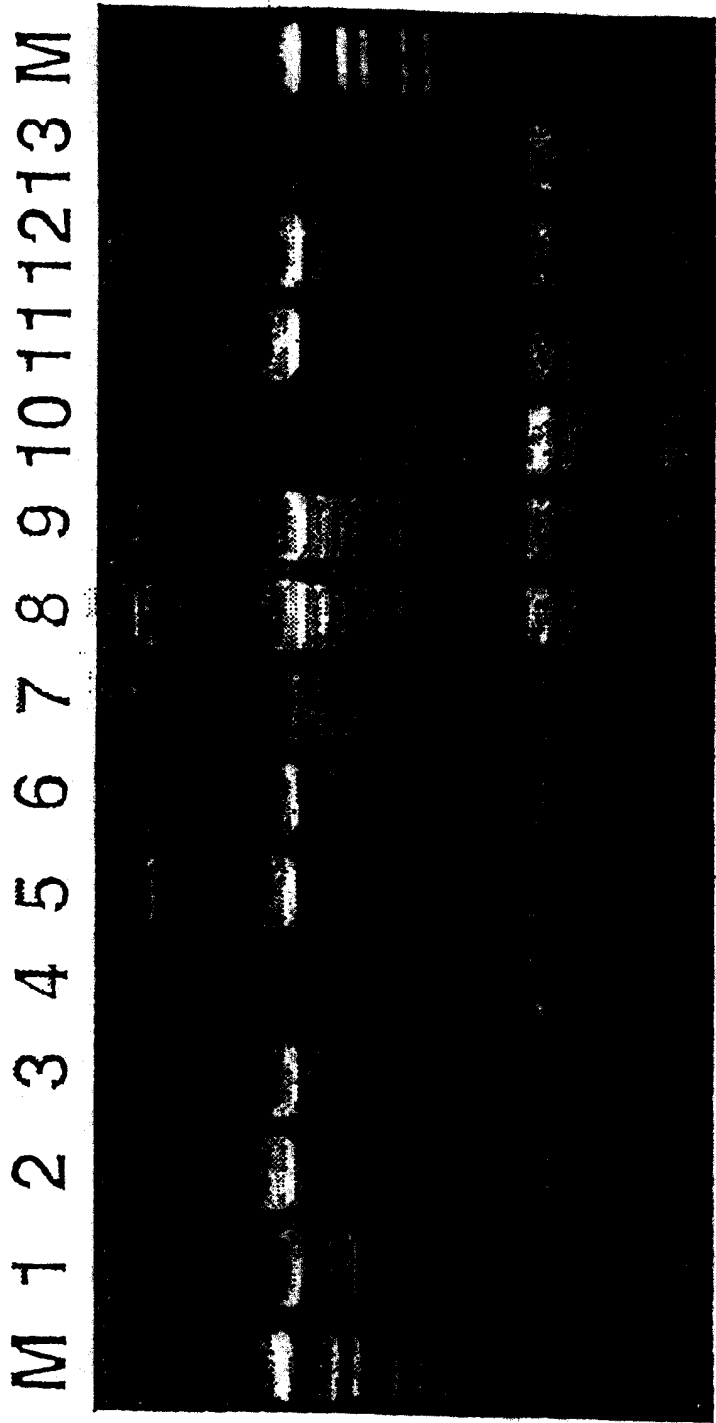


图 15

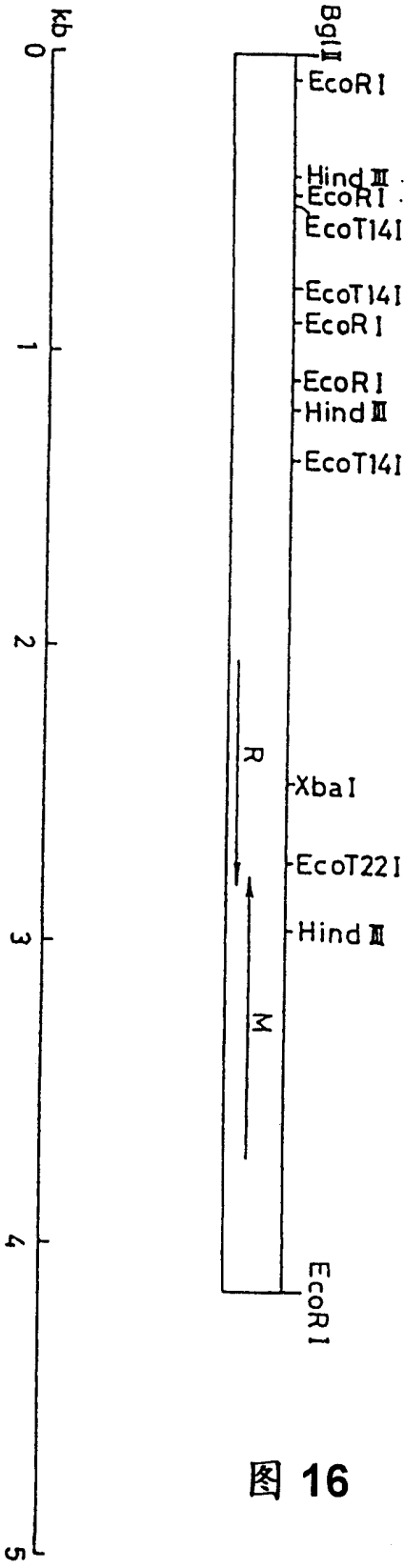


图 16