



(12) 发明专利

(10) 授权公告号 CN 1890373 B

(45) 授权公告日 2011. 12. 07

(21) 申请号 200480036792. 2  
 (22) 申请日 2004. 05. 04  
 (30) 优先权数据  
 10/682, 764 2003. 10. 09 US  
 (85) PCT申请进入国家阶段日  
 2006. 06. 09  
 (86) PCT申请的申请数据  
 PCT/US2004/013517 2004. 05. 04  
 (87) PCT申请的公布数据  
 W02005/040336 EN 2005. 05. 06  
 (73) 专利权人 英特尔克松公司  
 地址 美国弗吉尼亚州  
 (72) 发明人 托马斯·D·里德  
 (74) 专利代理机构 北京市柳沈律师事务所  
 11105  
 代理人 巫肖南 封新琴  
 (51) Int. Cl.  
 C12N 15/63 (2006. 01)  
 C12P 19/34 (2006. 01)  
 (56) 对比文件  
 Li Lin et al. Efficient linking and

transfer of multiple genes by a multigene assembly and transformation vector system. PNAS. 2003, 100(10), 5962-5967.

Inge J.W.M. Goderis et al. A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. Plant Molecular Biology. 2002, 5017-27.

J. MICHAEL THOMSON et al. ARTIFICIAL GENE-CLUSTERS ENGINEERED INTO PLANTS USING A VECTOR SYSTEM BASED ON INTRON- AND INTEIN-ENCODED ENDONUCLEASES. In Vitro Cell. Dev. Biol.—Plant. 2002, 38537-542.

审查员 王慧

权利要求书 2 页 说明书 13 页  
序列表 11 页 附图 11 页

(54) 发明名称  
DNA 克隆载体质粒及其使用方法

(57) 摘要  
出于基因表达或基因表达分析的目的, 本发明涉及用于构建 DNA 分子 ( 诸如转基因 ) 的一组克隆载体质粒。本发明还涉及用于在一系列可变的克隆步骤中使用克隆载体质粒生成最终转基因产物的方法。质粒克隆载体经过改造将载体的最终用户及其使用方法对 DNA 片段成分所需的操作量降至最低。使用本发明生成的转基因可用于具有少许或没有其它修饰的单一生物体或多种生物体, 包括细菌、酵母、小鼠和其它真核细胞。

1. 一种克隆载体质粒,其中以所列顺序,编码模块结构的序列元件包括:
  - a) 三个不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
  - b) 5' 寡核苷酸引物位点,
  - c) 正向取向的唯一 HE 位点,
  - d) 随机核苷酸序列侧翼的一对不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
  - e) 限定启动子模块 5' 部分的一组固定的不可变的超过 6 个核苷酸的限制位点,
  - f) 随机核苷酸序列,
  - g) 限定相对于启动子 / 内含子模块的 3' 位置和相对于表达模块的 5' 位置之间共有接点的一组固定的不可变的超过 6 个核苷酸的限制位点,
  - h) 随机核苷酸序列,
  - i) 限定相对于表达模块的 3' 位置和相对于 3' 调控模块的 5' 位置之间接点的一组固定的不可变的超过 6 个核苷酸的限制位点,
  - j) 随机核苷酸序列,
  - k) 限定相对于 3' 调控模块的 3' 位置的一组固定的不可变的超过 6 个核苷酸的限制位点,
  - l) 随机核苷酸序列侧翼的一对不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
  - m) 反向取向的唯一 HE 位点,与位于 5' 寡核苷酸引物位点 3' 的 HE 位点相同,
  - n) 反向取向的 3' 寡核苷酸引物位点,和
  - o) 限定 3' 插入位点的四个不可变且唯一的 6 个或不到 6 个核苷酸的限制位点。
2. 权利要求 1 的克隆载体质粒,其中序列元件的模块结构的安排容许导入三个不连续的核苷酸序列区域,使得如此导入的核苷酸序列取代所述启动子模块、表达模块和 3' 调控模块的随机核苷酸序列。
3. 一种克隆载体质粒,其中以所列顺序,所编码的序列元件包括:
  - a) 限定 5' 插入位点的两个不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
  - b) 寡核苷酸引物位点,
  - c) 随机核苷酸序列侧翼的一对相反取向的唯一 HE 位点,
  - d) 容许在所述一对唯一 HE 位点下游克隆穿梭载体模块的不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
  - e) 一组固定的不可变的超过 6 个核苷酸的限制位点,
  - f) 随机核苷酸序列,
  - g) 一组固定的不可变的超过 6 个核苷酸的限制位点,
  - h) 正向取向的唯一 HE 位点,
  - i) 随机核苷酸序列侧翼的一对不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
  - j) 寡核苷酸引物位点,
  - k) 一对相反取向的唯一 BstXI 位点,其中 BstXI 识别位点中的可变核苷酸区限定为与通过反向互补取向安排的两个相同 HE 识别位点生成的非互补尾相同的核苷酸,
  - l) 随机核苷酸序列侧翼的一对相反取向的唯一 HE 位点,
  - m) 反向取向的寡核苷酸引物位点,
  - n) 随机核苷酸序列侧翼的一对不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,

- o) 反向取向的唯一 HE 位点, HE 位点与正向取向的 HE 位点相同,
- p) 一组固定的不可变的超过 6 个核苷酸的限制位点,
- q) 随机核苷酸序列,
- r) 一组固定的不可变的超过 6 个核苷酸的限制位点,
- s) 不可变且唯一的 6 个或不到 6 个核苷酸的限制位点,
- t) 随机核苷酸序列侧翼的一对相反取向的唯一 HE 位点,
- u) 反向取向的寡核苷酸引物, 和
- v) 三个不可变且唯一的 6 个或不到 6 个核苷酸的限制位点。

4. 权利要求 3 的克隆载体质粒, 其中序列元件的模块结构的安排容许导入两个不连续的转基因, 使得如此导入的核苷酸序列取代随机核苷酸序列。

5. 权利要求 3 的克隆载体质粒, 其中序列元件的模块结构的安排容许导入两个不连续的核苷酸序列区域以诱导转基因进行同源重组以实现基因靶向突变。

6. 权利要求 3 的克隆载体质粒, 其中序列元件的模块结构的安排容许导入两个不同的正或负选择元件。

7. 一种用于构建转基因的方法, 包括下列步骤:

- a) 提供包含启动子模块、表达模块和 3' 调控模块的顺序排列的依照权利要求 1-2 任一项的克隆载体质粒,
- b) 将转基因中包含的第一段核苷酸序列导入穿梭载体,
- c) 将第一段核苷酸序列由穿梭载体转移至包含模块元件的依照权利要求 1-2 任一项的克隆载体质粒,
- d) 将任何其它所需核苷酸序列顺序导入另一个穿梭载体, 和
- e) 将其它所需核苷酸序列转移至已经接受了第一段核苷酸序列的依照权利要求 1-2 任一项的克隆载体质粒。

## DNA 克隆载体质粒及其使用方法

### 发明领域

[0001] 本发明涉及克隆载体质粒领域,以及克隆载体质粒用于构建 DNA 构建物或转基因的用途。

### [0002] 发明背景

[0003] 重组 DNA 技术是分子生物学的基础,在这里可以概括为出于研究核酸及其蛋白质产物的结构和功能的目的而修饰和扩增核酸。

[0004] 单个的基因、基因调控区、基因子集、和事实上包含它们的整个基因组都是由反向平行的双链核苷酸序列组成的,核苷酸常规标示为 A、T、G、和 C。可以将这些 DNA 序列以及由 mRNA 分子衍生的 cDNA 序列切割成不同的片段,分离、并插入载体诸如细菌质粒以研究基因产物。出于研究或生产基因产物的目的,可以对质粒即最初衍生自细菌的染色体外 DNA 部件进行操作或重新导入宿主细菌。质粒的 DNA 与所有染色体 DNA 类似,即它也是由相同的 A、T、G、和 C 组成来编码基因和基因调控区的,但是它是由少于大约 30,000 碱基对或 30 千碱基 (kb) 组成的较小分子。另外,双链质粒的核苷酸碱基对形成连续环状分子,这也是质粒 DNA 与染色体 DNA 的区别。

[0005] 质粒增强了细菌生物体之间遗传物质的快速交换,而且容许细菌生物体快速适应环境的变化,诸如温度、食物供应、或其它挑战。所获得的任何质粒都必须表达有助于宿主存活的一种或多种基因,否则它将被生物体破坏或丢弃,因为维持不必要的质粒将浪费资源。细胞的克隆群含有相同的遗传材料,包括它可能含有的任何质粒。在这样的宿主细胞克隆群中使用含有 DNA 插入片段的克隆载体质粒将扩增目的 DNA 的可用数量。然后可以分离和回收如此克隆的 DNA,用于构建 DNA 构建物所需步骤的后续操作。由此,应当认识到克隆载体质粒是研究基因功能的有用工具。

[0006] 尽管在质粒中发现的有些元件是天然存在的,然而其它元件经过改造增强了质粒作为 DNA 载体的效用。这些元件包括抗生素或化学药品抗性基因和多克隆位点 (MCS),等等。这些元件中的每一种在本发明以及现有技术中都是有作用的。关于每一种元件所发挥的作用的描述将强调现有技术的限制并展示本发明的效用。

[0007] 宿主可以通过质粒获得的一种特别有用的基因是将赋予宿主抗生素抗性的基因。在重组 DNA 技术的日常实践中,将抗生素抗性基因用作正或负选择元件,从而相对于其它质粒优先增强期望质粒的培养和扩增。

[0008] 为了获得宿主细菌的维持,质粒还必需包含指导宿主复制质粒的序列区段。称为复制起点 (ORI) 元件的序列指导宿主使用它的细胞酶来生成质粒的拷贝。当这样的细胞分裂时,子细胞将各自保留任何这种质粒的一个或几个拷贝。派生得到了将这种复制最大化的某些大肠杆菌细菌菌株,每个细菌可生成多达 300 个拷贝。如此可以增强期望质粒的培养。

[0009] 任何克隆载体中的另一种必需元件是用于插入目的遗传材料的位置。这是插入“野生型”质粒从而赋予克隆载体功效的合成元件。任何典型的商品化克隆载体质粒都包含至少一个这样的区域,称为多克隆位点 (MCS)。MCS 通常包含可以由一种或一系列限制

性核酸内切酶（下文称为“限制酶”）切割的核苷酸序列，每一种限制酶都具有独特的识别序列和切割模式。DNA 分子中编码的所谓识别序列（称为限制酶“位点”）包含双链回文序列。对于有些限制酶，少至 4-6 个核苷酸就足以提供识别位点，而有些限制酶需要 8 个或更多核苷酸的序列。例如 EcoRI 酶识别六核苷酸序列：<sup>5'</sup>G-A-A-T-T-C<sup>3'</sup>，其中 5' 指通常称为“上游”末端的分子末端，而 3' 同样指“下游”末端。识别序列的互补链是它的反向平行链，<sup>3'</sup>G-A-A-T-T-C<sup>5'</sup>。由此，双链识别位点可以在包含它的较大双链分子中如下展现：

[0010] <sup>5'</sup> .....G-A-A-T-T-C.....<sup>3'</sup>

[0011] <sup>3'</sup> .....C-T-T-A-A-G.....<sup>5'</sup>。

[0012] 正如许多其它限制酶，EcoRI 并非精确的在二分对称轴处进行切割，而是在两条 DNA 链中相距四个核苷酸的位置，由“/”指示的核苷酸之间：

[0013] <sup>5'</sup> .....G/A-A-T-T-C.....<sup>3'</sup>

[0014] <sup>3'</sup> .....C-T-T-A-A-/G.....<sup>5'</sup>，

[0015] 使得双链 DNA 分子断裂，并在新形成的“末端”具有由此产生的核苷酸结构：

[0016] <sup>5'</sup> .....G<sup>3'</sup>， <sup>5'</sup> A-A-T-T-C.....<sup>3'</sup>

[0017] <sup>3'</sup> .....C-T-T-A-A<sup>5'</sup>， <sup>3'</sup> G.....<sup>5'</sup>。

[0018] 这种交错切割产生具有 5' 突出末端的 DNA 片段。因为 A-T 和 G-C 配对在互相接近时自发形成，所以将诸如这样的突出末端称为粘端。这些末端中的任何一个能够与由相同限制酶切割的任何其它互补末端形成氢键。由于含有特异识别序列的任何 DNA 将与含有相同序列的任何其它 DNA 以相同方式受到切割，因此那些切割产生的末端将是互补的。因此，由相同限制酶切割的任何 DNA 分子的末端就像相邻拼图碎片的“匹配”一般互相“匹配”，而且能够由酶连接起来。正是这种特性容许形成重组 DNA 分子，并且容许将外源 DNA 片段导入细菌质粒，或是任何其它 DNA 分子。

[0019] 在构建重组 DNA 分子时需要考虑的另一项基本原则是分子中存在的所有限制位点都将受到特定限制酶的切割，而非只是目的位点。DNA 分子越大，重新出现任何限制位点的可能性就越大。假设任何限制位点沿着 DNA 分子随机分布，那么平均每  $4^4$ （即 256）个核苷酸将出现一个四核苷酸位点，每  $4^6$ （即 4096）个核苷酸将出现一个六核苷酸位点，而每  $4^8$ （即 114,688）个核苷酸将出现一个八核苷酸位点。由此，容易认识到较短的识别序列将频繁出现，而较长的识别序列将难得出现。在设计转基因或其它重组 DNA 分子的构建时，这是一个至关重要的问题，因为这样的计划常常需要装配不同大小的几段 DNA。这些片段越大，几段 DNA 成分中出现希望使用的位点的可能性就越大，最多使得操作变得困难。

[0020] 频繁出现的限制酶在本文中称为常见限制酶（common restriction enzyme），而它们的相关（cognate）序列称为常见限制位点。相关序列超过 6 个核苷酸的限制酶称为罕见限制酶（rare restriction enzyme），而它们的相关位点称为罕见限制位点。由此，称谓“罕见”和“常见”并不是指任何具体限制酶的相对丰度或利用率，而是指构成其相关识别位点的核苷酸序列在任何 DNA 分子或 DNA 分子分离片段或是任何基因或其 DNA 序列中的出现频率。

[0021] 最近分离得到第二类限制性核酸内切酶，称为归巢核酸内切酶（homing endonuclease, HE）。HE 酶具有较长的、不对称的识别位点（12-40 个碱基对）。HE 识别位点极其罕见。例如，称为 I-SceI 的 HE 具有 18bp 的识别位点（5' ... TAGGGATAACAGGGTAAT ... 3'），根据预测，在每  $7 \times 10^{10}$

个碱基对的随机序列中只出现一次。这种出现频率相当于 20 个哺乳动物大小的基因组中只有一个位点。HE 位点的罕见本质大大增加了遗传工程师能够切割最终转基因产物而不破坏转基因完整性的可能性,如果在克隆载体质粒的适当位置含有 HE 位点的话。

[0022] 由于来自任何生物体来源的 DNA 分子将以相同方式受到其相关限制酶的切割,因此可以用限制酶切割来自任何物种的外源 DNA 片段,插入已经用相同限制酶切割过的细菌质粒载体,并在合适的宿主细胞中进行扩增。例如,可以用 EcoRI 在两个位置切割人类基因,分离具有 EcoRI 末端的片段,并在通常称为连接反应或连接混合液的体系中与同样经过 EcoRI 切割的质粒混合。在连接混合液中,在合适的条件下,有些分离的人类基因片段将与质粒分子的末端匹配。这些重新接合的末端能够连接起来,并用酶将质粒再环化,质粒此时含有它的新 DNA 插入片段。然后将连接混合液导入大肠杆菌或其它合适的宿主,而重新改造的质粒将随细菌的分裂而扩增。如此可以由细菌获得和收获相对较大数目的人类基因拷贝。然后可以出于研究、分析、或生产其基因产物蛋白质的目的而进一步操作这些基因拷贝。

[0023] 重组 DNA 技术在所谓“转基因”的生成中得到频繁体现。转基因通常包含衍生自一种或多种供体生物体且导入宿主生物体的多种遗传材料。通常,使用克隆载体作为计划的起点或“主链”来构建转基因,而且设计一系列复杂的克隆步骤在该载体中装配最终产物。包含核苷酸序列的转基因的元件包括但不限于 1) 调控启动子和 / 或增强子元件, 2) 将表达为 mRNA 分子的基因, 3) 为 mRNA 信息提供稳定性的 DNA 元件, 4) 模仿哺乳动物内含子基因区域的核苷酸序列, 和 5) mRNA 加工信号, 诸如添加在天然存在的 mRNA 末端的 polyA 尾。在有些情况中,实验设计可能需要添加定位信号,从而将基因产物运输至特定的亚细胞位置。这些元件中的每一种都是切自供体基因组的较大 DNA 分子的片段,或者在有些情况中是在实验室中合成的。以精确的顺序和 5'-3' 取向将每一种片段与其它片段一起连接到克隆载体质粒中。

[0024] 可以以 DNA 片段的形式分离任何基因的启动子,并置于合成分子诸如质粒内,从而指导期望基因的表达,假设可以提供刺激目的启动子的必要条件的话。例如,可以分离胰岛素基因的启动子序列,与报道基因一起置于克隆载体质粒中,并用于研究在适当细胞类型中表达胰岛素基因所需要的条件。或者,可以在克隆载体质粒中将胰岛素基因启动子与任何目的基因的蛋白质编码序列相连,并用于驱动目的基因在表达胰岛素的细胞中表达,假设如此构建的 DNA 转基因中存在所有必需元件的话。

[0025] 报道基因是某些转基因类型的特别有用的成员。报道基因所包含的核苷酸序列编码的蛋白质将在转基因中与它相连的特定目的启动子的指导下表达,从而为启动子活性提供可测量的生化应答。相对于内源细胞蛋白质的背景,报道基因通常易于检测或测量。常用的报道基因包括但不限于 LacZ、绿色荧光蛋白、和荧光素酶,以及其它报道基因,其中许多对于本领域技术人员而言是众所周知的。

[0026] 在细菌基因组中没有发现内含子,但是它是在哺乳动物细胞中正确形成 mRNA 分子所必需的。因此,在哺乳动物系统中使用的任何 DNA 构建物必须具有至少一个内含子。可以由任何哺乳动物基因分离内含子,并与容许哺乳动物细胞切除内含子并将剩余 mRNA 末端剪接到一起的适当剪接信号一起插入 DNA 构建物。

[0027] mRNA 稳定元件指受到保护某些 mRNA 免于降解的结合蛋白识别的 DNA 序列。包含

mRNA 稳定元件通常将提高 mRNA 在某些哺乳动物细胞类型中的基因表达水平,因而可用于某些 DNA 构建物或转基因。可以由天然存在的 DNA 或 RNA 分离 mRNA 稳定元件,或者人工合成并包含在 DNA 构建物中。

[0028] 定位信号指编码目的蛋白亚细胞行程的蛋白质信号的 DNA 序列。例如,核定位信号将蛋白质引向细胞核;质膜定位信号将它引向质膜,等等。由此可以将定位信号掺入 DNA 构建物以促进其蛋白质产物转移至期望的亚细胞位置。

[0029] DNA 构建物中可能编码有标签序列,使得蛋白质产物将具有独特的附加区域。这个附加区域作为蛋白质标签,能够将它与它的内源对应物区分开。或者,它可以担当可以通过本领域众所周知的多种技术检测的鉴别物,包括但不限于 RT-PCR、免疫组化、或原位杂交。

[0030] 对于复杂的转基因或是包含特别大的 DNA 区域的转基因而言,这些 DNA 片段中存在多个识别位点的可能性增加了。想起每 4096bp 出现一个编码任何一种六核苷酸位点的识别序列。如果启动子序列是 3000bp,而且要将 1500bp 的目的基因装配到 3000bp 的克隆载体中,许多 6 个或更少核苷酸的位点将不能使用在统计学上是很有可能的,因为任何可用的位点只能存在于两种片段中。此外,位点必须存在于将要装配的适当分子的适当区域。另外,大多数克隆计划将需要添加额外的 DNA 元件,从而增加增长中分子的复杂性和任何特定位点不合时宜重复的可能性。由于任何限制酶切割它在分子中的所有位点,如果限制酶位点重复出现的话,那么所有不合适的位点将与期望位点一起受到切割,从而破坏分子的完整性。由此必须小心设计每一个克隆步骤,不会因为早已用于掺入前述元件的限制酶对它的切割而破坏增长中的分子。最后,如果研究人员希望将完成后的转基因导入哺乳动物生物体,那么通常必须将完整装配的转基因构建物在转基因的至少一个末端在唯一的限制位点处线性化,由此还要求构建物的任何其它区域不存在另一个唯一位点。由于大多数 DNA 构建物设计用于单一目的,因此对于可能需要进行的任何未来修改没有太多考虑,从而进一步增加了未来实验改变的难度。

[0031] 出于包括下列各项在内的几个原因,转基因的设计和构建通常耗费大量的时间和精力:

[0032] 1. 多种限制酶和 HE 酶可用于生成一批末端,然而其中大多数彼此不相容。许多限制酶(诸如 EcoRI)生成具有 5' 突出粘端或“尾”的 DNA 片段;其它酶(例如 PstI)生成具有 3' 突出尾的 DNA 片段;而还有一些酶(例如 BalI)在对称轴处切割而生成平端片段。其中一些将与由其它限制酶和 HE 酶切割而形成的末端相容,但是大多数有用的末端并非如此。在设计 DNA 构建物时必须仔细考虑每一次 DNA 片段分离中可能生成的末端。

[0033] 2. 必须首先由它们的基因组来源分离装配 DNA 构建物或转基因所需要的 DNA 片段,置于质粒克隆载体中,并扩增以获得有用数量。可以使用任何数目的商品化或个别改造的克隆载体来进行上述步骤。每一种不同的商品化克隆载体质粒在极大程度上是独立开发的,因此含有不同的序列和用于插入目的基因或遗传元件的 DNA 片段的限制位点。因此必须根据任何指定的一组实验的需要个别裁剪基因以适应每一种这样的载体。通常需要改变相同的 DNA 片段,进一步用于后续实验或克隆到新 DNA 构建物或转基因的其它组合中。由于每一种 DNA 构建物或转基因定制用于一种具体的应用而没有考虑下次将如何使用它,因此通常必须将它“翻新(retro-fitted)”以用于后续应用。

[0034] 3. 另外,任何指定基因或遗传元件的 DNA 序列是不同的,而且可能含有使之与现

有载体不相容的内部限制位点,从而使操作变得复杂。在将几个 DNA 片段装配成单一 DNA 构建物或转基因时尤其如此。

[0035] 由此,需要容许用户将许多 DNA 片段快速装配成一个分子的系统,尽管在 DNA 片段的末端或内部发现有限制位点的冗余。这样的系统可能还提供用于快速改变片段的末端从而向其添加其它限制序列的简单方法。含有单一或相对的 HE 限制位点对将增加具有可用于克隆的唯一位点的可能性。还容许容易的替代或清除一个或多个片段的系统将提高用户目前无法获得的多功能性 (versatility) 水平。因此,容许将 DNA 片段插入或移出克隆载体中侧翼为罕见限制位点的“盒 (cassette)”区域的“模块系统 (modular system)”在重组 DNA 技术领域将是特别有用的,且是受欢迎的。

[0036] 发明概述

[0037] 本发明涉及一种克隆载体质粒,其中编码模块结构的序列元件包括:三个不可变 (non-variable) 且唯一的常见限制位点、5' 寡核苷酸引物位点、正向取向的唯一 HE 位点、随机核苷酸序列侧翼的一对不可变且唯一的 (non-variable and unique) 常见限制位点、限定启动子模块 5' 部分的一组固定的不可变罕见限制位点 (fixed grouping of non-variable rare restriction site)、随机核苷酸序列、限定相对于启动子 / 内含子模块的 3' 位置和相对于表达模块的 5' 位置之间共享接点的一组固定的不可变罕见限制位点、随机核苷酸序列、限定相对于表达模块的 3' 位置和相对于 3' 调控模块的 5' 位置之间接点的一组固定的不可变罕见限制位点、随机核苷酸序列、限定相对于 3' 调控模块的 3' 位置的一组固定的不可变罕见限制位点、随机核苷酸序列侧翼的一对不可变且唯一的常见限制位点、反向取向的唯一 HE 位点 (与位于 5' 寡核苷酸引物位点 3' 的 HE 位点相同)、反向取向的 3' 寡核苷酸引物位点、和限定 3' 插入位点的四个不可变且唯一的常见限制位点。

[0038] 本发明涉及一种克隆载体质粒,其中编码分子结构的序列元件包括:限定 5' 插入位点的两个不可变且唯一的常见限制位点、寡核苷酸引物位点、随机核苷酸序列侧翼的一对相反取向的唯一 HE 位点、容许在一对唯一 HE 位点下游克隆穿梭载体模块的不可变且唯一的常见限制位点、一组固定的不可变罕见限制位点、随机核苷酸序列、一组固定的不可变罕见限制位点、正向取向的唯一 HE 位点、随机核苷酸序列侧翼的一对不可变且唯一的常见限制位点、寡核苷酸引物位点、一对相反取向的唯一 BstXI 位点 (其中 BstXI 识别位点中的可变核苷酸区限定为与通过反向互补取向安排的两个相同 HE 识别位点生成的非互补尾相同的核苷酸)、随机核苷酸序列侧翼的一对相反取向的唯一 HE 位点、反向取向的寡核苷酸引物位点、随机核苷酸序列侧翼的一对不可变且唯一的常见限制位点、反向取向的唯一 HE 位点 (HE 位点与正向取向的 HE 位点相同)、一组固定的不可变罕见限制位点、随机核苷酸序列、一组固定的不可变罕见限制位点、不可变且唯一的常见限制位点、随机核苷酸序列侧翼的一对相反取向的唯一 HE 位点、反向取向的寡核苷酸引物、和三个不可变且唯一的常见限制位点。

[0039] 本发明还涉及一种用于构建转基因的方法,包括下列步骤:提供包含启动子模块、表达模块、和 3' 调控模块的顺序排列的克隆载体质粒;将转基因中包含的第一段核苷酸序列导入穿梭载体;将第一段核苷酸序列由穿梭载体转移至包含模块元件的克隆载体质粒;将任何其它所需核苷酸序列顺序导入另外的穿梭载体;并将所述其它所需核苷酸序列转移至所述已经接受了第一段核苷酸序列的克隆载体质粒。

[0040] 附图描述

[0041] 图 1 是本发明模块概念的线性图谱。

[0042] 图 2 是停泊质粒 (Docking Plasmid) 图谱。

[0043] 图 3 的线性限制图谱图示了停泊质粒 MCS 中可以包含的限制酶位点的实例。

[0044] 图 4 是初级停泊质粒 (Primary Docking Plasmid) 图谱。

[0045] 图 5 的线性限制图谱图示了初级停泊质粒 MCS 中可以包含的限制酶位点的实例。

[0046] 图 6 是穿梭载体 P(SVP) 质粒图谱。

[0047] 图 7 的线性限制位点图谱图示了 SVP MCS 中可以包含的限制酶位点的实例。

[0048] 图 8 是穿梭载体 E(SVE) 质粒图谱。

[0049] 图 9 的线性限制位点图谱图示了 SVE MCS 中可以包含的限制酶位点的实例。

[0050] 图 10 是穿梭载体 3(SV3) 图谱。

[0051] 图 11 的线性限制位点图谱图示了 SV3MCS 中可以包含的限制酶位点的实例。

[0052] 序列表简述

[0053] SEQ:ID 01 是 PE3 停泊质粒 MCS 的核苷酸序列实例。

[0054] SEQ:ID 02 是 PE3 停泊质粒的核苷酸序列实例。

[0055] SEQ:ID 03 是初级停泊质粒 MCS 的核苷酸序列实例。

[0056] SEQ:ID 04 是初级停泊质粒的核苷酸序列实例。

[0057] SEQ:ID 05 是 SVP 质粒 MCS 的核苷酸序列实例。

[0058] SEQ:ID 06 是 SVP 质粒的核苷酸序列实例。

[0059] SEQ:ID 07 是 SVE 质粒 MCS 的核苷酸序列实例。

[0060] SEQ:ID 08 是 SVE 质粒的核苷酸序列实例。

[0061] SEQ:ID 09 是 SV3 质粒 MCS 的核苷酸序列实例。

[0062] SEQ:ID 10 是 SV3 质粒的核苷酸序列实例。

[0063] 用于描述本发明的术语的定义

[0064] 在用于本文时,术语“克隆载体”和“克隆载体质粒”可以互换使用,指最低限度包含复制起点、用于对容纳质粒的宿主细胞进行正选择的手段诸如抗生素抗性基因、和多克隆位点的环状 DNA 分子。

[0065] 在用于本文时,术语“复制起点”(ORI)指指导质粒在宿主细胞中复制的核苷酸序列。

[0066] 在用于本文时,术语“多克隆位点”指出于将 DNA 片段克隆到克隆载体质粒中的目的而包含限制位点的核苷酸序列。

[0067] 在用于本文时,术语“克隆”指将 DNA 分子连接到质粒中并将其转移到适当宿主细胞中用于在宿主繁殖过程中复制的过程。

[0068] 在用于本文时,术语“DNA 构建物”指通过克隆载体质粒中的连续克隆步骤而合成的,且常用于在任何适当细胞宿主(诸如体外的培养细胞或体内的转基因小鼠)中指导基因表达的 DNA 分子。用于生成这种小鼠的转基因也可以称为 DNA 构建物,尤其是在设计和合成转基因的那段时间里。

[0069] 在用于本文时,术语“穿梭载体”指本发明中用于生成中间分子的专用(specialized)克隆载体质粒,所述中间分子将更改 DNA 片段的末端。

[0070] 在用于本文时,术语“停泊质粒”指本发明中用于将 DNA 片段装配到 DNA 构建物中的专用克隆载体质粒。

[0071] 在用于本文时,“限制性核酸内切酶”或“限制酶”指与相关 DNA 序列结合并在该序列内的精确位置切割 DNA 分子的一类催化分子的一个或多个成员。

[0072] 在用于本文时,术语“相关序列 (cognate sequence)”指限制酶结合并切割 DNA 分子或基因所需要的最小限度的一段核苷酸。

[0073] 在用于本文时,术语“DNA 片段”指任何分离的 DNA 分子,包括但不限于编码蛋白质的序列、报道基因、启动子、增强子、内含子、外显子、polyA 尾、多克隆位点、核定位信号、或 mRNA 稳定信号,或是任何其它天然存在的或合成的 DNA 分子。或者, DNA 片段可以完全是合成起源的,在体外生成的。另外, DNA 片段可以包含分离的天然存在和 / 或合成片段的任意组合。

[0074] 在用于本文时,术语“基因启动子”或“启动子”(P)指基因表达所必需的核苷酸序列。

[0075] 在用于本文时,术语“增强子区域”指不是靶基因表达所必需的但在适当条件下将提高基因表达水平的核苷酸序列。

[0076] 在用于本文时,术语“报道基因”指所编码的蛋白质可用于监测特定目的启动子活性的核苷酸序列。

[0077] 在用于本文时,术语“染色质修饰结构域”(CMD)指与维持和 / 或改变染色质结构相关的多种蛋白质相互作用的核苷酸序列。

[0078] 在用于本文时,术语“polyA 尾”指通常在信使 RNA (mRNA) 分子末端处发现的一串腺嘌呤核苷酸。将 polyA 尾信号掺入 DNA 构建物或转基因的 3' 末端有助于表达目的基因。

[0079] 在用于本文时,术语“内含子”指基因中在两个蛋白质编码区或外显子之间发现的非蛋白质编码区的核苷酸序列。

[0080] 在用于本文时,术语“非翻译区”(UTR)指涵盖 mRNA 分子中的非蛋白质编码区的核苷酸序列。这些非翻译区可以位于 mRNA 分子的 5' 端 (5' UTR) 或 3' 端 (3' UTR)。

[0081] 在用于本文时,术语“mRNA 稳定元件”指受到某些结合蛋白识别的 DNA 序列,所述结合蛋白被认为是保护某些 mRNA 免于降解。

[0082] 在用于本文时,术语“定位信号”(LOC)指编码目的蛋白亚细胞行程的信号的核苷酸序列。

[0083] 在用于本文时,术语“标签序列”(TAG)指编码容许进行检测或在有些情况中与任何内源对应物进行区分的独特蛋白质区域的核苷酸序列。

[0084] 在用于本文时,术语“引物位点”指 DNA 模板中单链 DNA 寡核苷酸能够与它退火从而启动 DNA 测序、PCR 扩增、和 / 或 RNA 转录的核苷酸序列。

[0085] 在用于本文时,术语“基因表达宿主选择基因”(GEH-S)指在用适当抗生素或化学药品进行处理时能够赋予细胞或生物体以抗性或毒性的遗传元件。

[0086] 在用于本文时,术语“重组臂”指促进转基因 DNA 与基因组 DNA 之间进行同源重组的核苷酸序列。成功的重组要求在要通过同源重组掺入宿主基因组的转基因 DNA 区的侧翼存在左重组臂 (LRA) 和右重组臂 (RRA)。

[0087] 在用于本文时,术语“pUC19”指本领域技术人员熟知的质粒克隆载体,它在 NCBI

基因库数据库中的编号是 L09137。

[0088] 在用于本文时,术语“随机核苷酸序列”指与编码相同分子之成分的其它元件之序列不同的核苷酸序列的任意组合。

[0089] 在用于本文时,术语“唯一 (unique)”指在 DNA 分子中的其它区域没有发现的任何限制性核酸内切酶或 HE 位点。

[0090] 发明详述

[0091] 本发明是经过优化降低了将多个 DNA 片段重新装配成 DNA 构建物或转基因通常所需的操作量的一组克隆载体。初级载体(本文中称为停泊质粒)所包含的多克隆位点(MCS)具有以线性模式排列的三套罕见限制位点和/或 HE 位点。这种排列限定了这样的模块体系结构,它容许用户将多个插入片段装配成一个转基因构建物,但不会打乱在先前的克隆步骤中早已掺入停泊质粒中的 DNA 元件的完整性。

[0092] 将至少三种 HE 的两个识别位点以相反取向置于三个模块区的侧翼,用于生成不能自我退火的基因盒接收位点。因为 HE 位点是不对称的且非回文的,所以有可能通过以相反取向安置两个 HE 识别位点而生成不互补的 3' 突出粘性尾。由此,HE I-SceI 如“/”所示切割其相关识别位点:

[0093] 5'...TAGGGATAA/CAGGGTAAT...3',

[0094] 3'...ATCCCTATTGTCCCAATTA...5'。

[0095] MCS 中反向安置的第二个位点将生成两个不互补的粘性突出尾:

[0096] 5'...TAGGGATAA CCCTA...3'

[0097] 3'...ATCCC AATAGGGAT...5'。

[0098] 这在必须将较大转基因亚克隆到载体中时是特别有用的。由于插入片段的大小,载体在热力学上更加有利于自身退火而非接受大插入片段。通过这样安排限制位点而生成的非互补尾的存在为抵消自我连接的热力学倾向提供了化学力。

[0099] 大多数 HE 突出尾的不对称本质在联合 BstXI 限制酶位点(5' CCANNNN/NTGG3')使用时还产生了强大的克隆工具。在排除自我退火的同时,BstXI 的序列中性区域(sequence-neutral domain)可用于生成与两个相反取向的 HE 突出尾相容的粘端。

[0100] BstXI(正向 I-SceI) 正向 I-SceI 反向 I-SceI BstXI(反向 I-SceI)

[0101] 5' -CCAGATAA CAGGGTAAT//ATTACCCTGTTAT GTGG-3'

[0102] 3' -GGTC TATTGTCCCATTA//TAATGGGAC AATACACC-5'

[0103] 本发明的次级载体(本文中称为穿梭载体)所包含的多克隆位点具有侧翼为罕见限制位点和/或 HE 位点的常见限制位点。穿梭载体设计用于将 DNA 片段克隆到罕见位点之间的常见限制位点中。随后可以通过罕见限制位点或 HE 位点处的切割释放所克隆的片段,并使用在穿梭载体中发现的相同罕见限制位点和/或 HE 位点掺入停泊质粒。

[0104] 由此,与常规克隆载体不同的是,MCS 的设计容许将 DNA 片段的“盒”或模块插入停泊质粒的模块区域。同样,可以使用相同的罕见限制酶和/或 HE 酶容易的将每一个切除,并用任何其它目的 DNA 片段取代。这种特征容许用户快速且容易的改变实验计划的方向,无需重建整个 DNA 构建物。由此,本发明的克隆载体质粒容许用户使用常见限制位点将 DNA 片段克隆到中间载体中,产生接受盒的模块(cassette-accepting module),然后借助罕见限制位点将该片段转移至最终构建物的期望模块处。另外,它容许将来改变分子,用其它盒

模块取代停泊质粒中的个别模块。下面的描述强调了本发明相对于现有技术的区别。

[0105] 可以作为由穿梭载体转移至 PE3 停泊位置质粒的模块来装配转基因的各个成分（启动子增强子 P、表达的蛋白质 E、和 / 或 3' 调控区 3）。如果需要更高级别的复杂性，那么可以将装配好的转基因或其它核苷酸序列转移到初级停泊质粒中。由更加复杂的 PE3 停泊位置质粒和初级 Dosing 质粒开始，下文将更为详细的说明五类克隆载体质粒中的每一种，从而能够理解掺入每一种质粒的成分。

[0106] PE3 停泊质粒（图 2）包含 pUC19 的主链但具有下列修改，其中所述序列是依照编号 L09137 的 pUC19 基因库序列文件编号的：

[0107] 1. 停泊质粒中只使用了 806 至 2617 (Af13-Aat2) 的序列，

[0108] 2. pUC19 中 1729 处的 BspHI 位点由 TCATGA 突变成 GCATGA，

[0109] 3. pUC19 中 1493 处的 AcII 位点由 AACGTT 突变成 AACGCT，

[0110] 4. pUC19 中 1120 处的 AcII 位点由 AACGTT 突变成 CACGCT，

[0111] 5. pUC19 中的 AhdI 位点由 GACNNNNNGTC 突变成 CACNNNNNGTC，

[0112] 6. 在上表突变步骤 2 之后，将编码 BspHI/I-PpoI/BspHI 的序列插入 pUC 19 中唯一剩余的 BspHI 位点。

[0113] PE3 停泊质粒的多克隆位点 (MCS) (图 3) 以所列顺序包含下列序列元件：

[0114] 1. 限定突变后 pUC19 载体的 5' 插入位点的三个不可变且唯一的常见限制位点（例如 AatII、BlnI、和 Eco0109I），

[0115] 2. 一个 T7 引物位点，

[0116] 3. 一个唯一的 HE 位点（例如 I-SceI (正向取向)），

[0117] 4. 可以担当染色质修饰结构域接受模块 (RNAS-CMD-1) 的随机核苷酸序列侧翼的一对不可变且唯一的常见限制位点（例如 KpnI 和 AvrII），5. 限定启动子模块 5' 部分的一组固定的不可变罕见限制位点（例如 AsiSI 和 SgrAI），

[0118] 6. 可以担当启动子 / 内含子接受模块的随机核苷酸序列 (RNAS-P)，

[0119] 7. 限定启动子 / 内含子模块的 3' 部分和表达模块的 5' 部分之间共有接点的一组固定的不可变罕见限制位点（例如 PacI、AscI、和 MluI），

[0120] 8. 可以担当表达接受模块的随机核苷酸序列 (RNAS-E)，

[0121] 9. 限定表达模块的 3' 部分和 3' 调控模块的 5' 部分之间接点的一组固定的不可变罕见限制位点（例如 SnaBI、NotI、和 SalI），

[0122] 10. 可以担当 3' 调控结构域接受模块的随机核苷酸序列 (RNAS-3)，

[0123] 11. 限定 3' 调控模块的 3' 部分的一组固定的不可变罕见限制位点（例如 SwaI、RsrII、和 BsiWI），

[0124] 12. 可以担当染色质修饰结构域接受模块的 DNA 随机核苷酸序列

[0125] (RNAS-CMD-2) 侧翼的一对不可变且唯一的常见限制位点（例如 XhoI 和 NheI），

[0126] 13. 反向取向的唯一 HE 位点，与上文第 3 项的 HE 位点相同，

[0127] 14. 反向取向的 T3 引物位点，和

[0128] 15. 限定上文所述突变后 pUC19 载体的 3' 插入位点的四个不可变且唯一的常见限制位点（例如 BspEI、PmeI、SapI、和 BspHI）。

[0129] 初级停泊质粒（图 4）可用于装配最初在 PE3 停泊位置质粒 (DockingStation

Plasmid) 中构建的两个完整转基因或构建基因靶向转基因所需要的两个同源臂,或是用于导入两类正或负选择元件的。初级停泊质粒的多克隆位点 (MCS) (图 5) 以所列顺序包含下列序列元件:

[0130] 1. 限定上文所述突变后 pUC19 载体的 5' 插入位点的两个不可变且唯一的常见限制位点 (例如 AatII 和 B1pI),

[0131] 2. M13 反向引物位点,

[0132] 3. 可以担当基因组表达宿主选择基因接受模块 (RNAS-GEH-S1) 的 DNA 随机核苷酸序列侧翼的一对唯一 HE 位点 (例如 PI-SceI (正向取向) 和 PI-SceI (反向取向)),

[0133] 4. 容许在所述 HE 对位点下游克隆穿梭载体模块的不可变且唯一的常见限制位点 (例如 Eco0109I),

[0134] 5. 限定左重组臂模块的 5' 部分的一组固定的不可变罕见限制位点 (例如 SgrAI 和 AsiSI),

[0135] 6. 可以担当左重组臂接受模块的随机核苷酸序列 (RNAS-LRA),

[0136] 7. 限定左重组臂接受模块的 3' 部分的一组固定的不可变罕见限制位点 (例如 PacI、MluI、和 AscI),

[0137] 8. 唯一的 HE 位点 (例如 I-CeuI (正向取向)),

[0138] 9. 可以担当染色质修饰结构域接受模块 (chromatin modification domainacceptor module) (RNAS-CMD-1) 的 DNA 随机核苷酸序列侧翼的一对不可变且唯一的常见限制位点 (例如 KpnI 和 AvrII),

[0139] 10. T7 引物位点,

[0140] 11. 可以担当复杂转基因接受模块 (complex transgene acceptor module) 的 DNA 随机核苷酸序列 (RNAS-PE3-1) 侧翼的一对相反取向的唯一 BstXI 位点 (其中 BstXI 识别位点中的可变核苷酸区限定为与通过反向互补取向安排的两个相同 HE 识别位点生成的非互补尾相同的核苷酸,例如 PI-SceI (正向取向) 和 PI-SceI (反向取向)),

[0141] 12. 可以担当复杂转基因模块的 DNA 随机核苷酸序列 (RNAS-PE3-2) 侧翼的一对相反取向的唯一 HE 位点 (例如 I-SceI (正向取向) 和 I-SceI (反向取向)),

[0142] 13. 反向取向的 T3 引物位点,

[0143] 14. 可以担当染色质修饰结构域接受模块的 DNA 随机核苷酸序列 (RNAS-CMD-2) 侧翼的一对不可变且唯一的常见限制位点 (例如 XhoI 和 NheI),

[0144] 15. 反向取向的唯一 HE 位点,与上述第 8 项的 HE 位点相同,

[0145] 16. 限定右重组臂模块的 5' 部分的一组固定的不可变罕见限制位点 (例如 SnaBI、SalI、和 NotI),

[0146] 17. 可以担当右重组臂接受模块的随机核苷酸序列 (RNAS-RRA),

[0147] 18. 限定右重组臂接受模块的 3' 部分的一组固定的不可变罕见限制位点 (例如 RsrII、SwaI、和 BsiWI),

[0148] 19. 容许在一对 HE 位点上游克隆穿梭载体模块的不可变且唯一的常见限制位点 (例如 BspEI),

[0149] 20. 可以担当基因组表达宿主选择基因接受模块的 DNA 随机核苷酸序列 (RNAS-GEH-S2) 侧翼的一对相反取向的唯一 HE 位点 (例如 PI-PspI (正向取向) 和

PI-PspI(反向取向)),

[0150] 21. 以反向取向安置的 M13 正向引物位点,

[0151] 22. 限制上文所述突变后 pUC19 载体的 3' 插入位点的三个不可变且唯一的常见限制位点(例如 PmeI、SapI、和 BspHI)。

[0152] 本发明的三种克隆载体质粒称为穿梭载体。穿梭载体,像 PE3 和初级停泊质粒,也是由 pUC19 主链构建而成的。正如 PE3 和初级停泊质粒,每一种穿梭载体对 pUC19 主链进行了上文所列 1-6 的相同修改。各种穿梭载体(SV)鉴别为启动子/内含子穿梭载体(P)、表达穿梭载体(E)、和 3' 调控穿梭载体(3);自此以后,分别以 SVP、SVE、SV3 表示。下文更加详尽的描述了每一种穿梭载体。

[0153] 穿梭载体 P(SVP):

[0154] SVP 是可用于制备将装配到转基因构建物中的启动子和内含子序列的克隆载体质粒(图 6)。SVP 质粒的实例可以以所列顺序在 MCS 中包含下列序列元件(图 7):

[0155] 1. 限定上文所述突变后 pUC19 载体的 5' 插入位点的两个不可变且唯一的常见限制位点(例如 AatII 和 B1pI),

[0156] 2. T7 引物位点,

[0157] 3. 容许在 T7 引物位点下游有效克隆穿梭载体模块的不可变且唯一的常见限制位点(例如 Eco0109I),

[0158] 4. 限定启动子模块的 5' 部分的一组固定的不可变罕见限制位点(a fixed grouping of non-variable rare restriction sites)(例如 AsiSI 和 SgrAI),

[0159] 5. 包含穿梭载体中唯一的任何分组的常见或罕见限制位点的可变 MCS(例如图 7 所示一系列限制位点),

[0160] 6. 限定启动子模块的 3' 部分的一组固定的不可变罕见限制位点(例如 PacI、AscI、和 MluI),

[0161] 7. 容许在 T3 引物位点上游有效克隆穿梭载体模块的不可变且唯一的常见限制位点(例如 BspEI),

[0162] 8. 反向取向的 T3 引物位点,和

[0163] 9. 限定上文所述突变后 pUC19 载体的 3' 插入位点的两个不可变且唯一的常见限制位点(例如 PmeI 和 SapI)。

[0164] 穿梭载体 E(SVE):

[0165] 这是可用于制备将装配到转基因构建物中的、由转基因表达的序列的克隆载体质粒(图 8)。SVE 质粒的实例可以以所列顺序在 MCS 中包含下列序列元件(图 9):

[0166] 1. 限定上文所述突变后 pUC19 载体的 5' 插入位点的两个不可变且唯一的常见限制位点(例如 AatII 和 B1pI),

[0167] 2. T7 引物位点,

[0168] 3. 容许在 T7 引物位点下游有效克隆穿梭载体模块的不可变且唯一的常见限制位点(例如 Eco0109I),

[0169] 4. 限定表达模块的 5' 部分的一组固定的不可变罕见限制位点(例如 PacI、AscI、和 MluI),

[0170] 5. 由穿梭载体中唯一的任何分组的常见或罕见限制位点组成的可变 MCS(例如图

9 所示一系列限制位点),

[0171] 6. 限定表达模块的 3' 部分的一组固定的不可变罕见限制位点 (例如 *Sna*BI、*Not*I、和 *Sal*I),

[0172] 7. 容许在 T3 引物位点上游有效克隆穿梭载体模块的不可变且唯一的常见限制位点 (例如 *Bsp*EI),

[0173] 8. 反向取向的 T3 引物位点, 和

[0174] 9. 限定上文所述突变后 pUC19 载体的 3' 插入位点的两个不可变且唯一的常见限制位点 (例如 *Pme*I 和 *Sap*I)。

[0175] 穿梭载体 3(SV3) :

[0176] 这是可用于制备将装配到转基因构建物中的 3' 调控序列的克隆载体质粒 (图 10)。SV3 质粒的实例可以以所列顺序在 MCS 中包含下列序列元件 (图 11) :

[0177] 1. 限定上文所述突变后 pUC19 载体的 5' 插入位点的两个不可变且唯一的常见限制位点 (例如 *Aat*II 和 *B1p*I),

[0178] 2. T7 引物位点,

[0179] 3. 容许在 T7 引物位点下游有效克隆穿梭载体模块的不可变且唯一的常见限制位点 (例如 *Eco*0109I),

[0180] 4. 限定 3' 调控模块的 5' 部分的一组固定的不可变罕见限制位点 (例如 *Sna*BI、*Not*I、和 *Sal*I),

[0181] 5. 由穿梭载体中唯一的任何分组的常见或罕见限制位点组成的可变 MCS (例如图 11 所示一系列限制位点),

[0182] 6. 限定 3' 调控模块的 3' 部分的一组固定的不可变罕见限制位点 (例如 *Swa*I、*Rsr*II、和 *Bsi*WI),

[0183] 7. 容许在 T3 引物位点上游有效克隆穿梭载体模块的不可变且唯一的常见限制位点 (例如 *Bsp*EI),

[0184] 8. 反向取向的 T3 引物位点, 和

[0185] 9. 限定上文所述突变后 pUC19 载体的 3' 插入位点的两个不可变且唯一的常见限制位点 (例如 *Pme*I 和 *Sap*I)。

[0186] 尽管本发明公布了用于在质粒克隆载体中构建转基因的方法, 然而可以使用类似方法在较大染色体外 DNA 分子中构建转基因, 诸如粘粒或人工染色体, 包括细菌人工染色体 (BAC)。可以掺入质粒克隆载体的多种遗传元件还容许通过少许操作或无需更多操作就将最终的转基因产物转移到多种宿主生物体中。

[0187] 作为实践本发明的方法的实例, 可以构建包含这些元件的转基因 :

[0188] 1. 表面活性蛋白 C(SP-C) 的人类启动子的核苷酸序列,

[0189] 2. 编码小鼠基因粒细胞-巨噬细胞集落刺激因子受体  $\beta$  c (GMR  $\beta$  c) 的蛋白质产物的序列,

[0190] 3. 兔  $\beta$  球蛋白内含子序列, 和

[0191] 4. SV40polyA 信号。

[0192] SP-C 序列包含内部 *Bam*HI 位点, 而且只能用 *Not*I 和 *Eco*RI 由其亲本质粒中释放出来。GMR  $\beta$  c 具有内部 *Not*I 位点, 而且可以用 *Bam*HI 和 *Xho*I 由其亲本质粒中切割出来。兔

$\beta$  球蛋白内含子序列可以用 EcoRI 由其亲本质粒中切割出来。SV-40polyA 尾可以用 XhoI 和 SacI 由其亲本质粒中切割出来。因为几种限制位点的冗余,这些亲本质粒都不能用于装配所有的所需片段。

[0193] 用于在 PE3 停泊质粒发明中构建期望转基因的步骤如下:

[0194] 1. 由于 NotI 和 PspOM1 生成相容粘端,因此用 NotI 和 EcoRI 切出人类 SP-C 启动子序列,并克隆到穿梭载体 P 的 PspOM1 和 EcoRI 位点中。将该反应的产物称为 pSVP-SPC。

[0195] 2. 在本领域技术人员熟知的扩增和回收步骤之后,将兔  $\beta$  球蛋白内含子序列克隆到 pSVP-SPC 的 EcoRI 位点中。通过对称为 pSVP-SPC-r  $\beta$  G 的由此产生的中间构建物进行测序来确认内含子在所述中间构建物中的取向。

[0196] 3. 使用 AsiSI 和 AscI 切出并克隆启动子和内含子作为来自 pSVP-SPC-r  $\beta$  G 的一个相连片段。同时,用 AsiSI 和 AscI 切割 PE3 停泊质粒,准备与启动子 / 内含子区段进行连接。将启动子 / 内含子片段连接到停泊质粒中,扩增,并回收。

[0197] 4. 使用本领域技术人员熟知的技术补平 GMR  $\beta$  c 片段的 XhoI 位点以产生 3' 平端。然后将它克隆到 pSVP-SPC-r  $\beta$  G 的 BamHI 位点和末端补平的 Pvu2 位点。扩增并回收由此产生的质粒 (pDP-SPC-GMR  $\beta$  c-r  $\beta$  G)。

[0198] 5. 最后的克隆步骤是添加 SV-40 的 polyA 尾。用 XhoI 和 SacI 切出 SV-40polyA 片段,正如受体载体 pDS1-SPC-GMR  $\beta$  c-r  $\beta$  G。将这两种 DNA 片段进行凝胶纯化和回收。以 SV-40polyA 与 pDS1-SPC-GMR  $\beta$  c-r  $\beta$  G 以 10 : 1 摩尔比制备连接混合液。扩增并回收连接产物。新质粒 pDS1-SPC-GMR  $\beta$  c-r  $\beta$  G-pA 含有转基因所需要的所有元件,包括 3' 末端的唯一限制位点,可用于将整个 pDS1-SPC-GMR  $\beta$  c-r  $\beta$  G-pA 质粒线性化,以转染到真核细胞中或显微注射到受精卵的原核中。

[0001]

序 列 表

<110> 英特里克松公司(Intrexon Corporation)

<120> DNA 克隆载体质粒及其使用方法

<130> 60/417,282

<140> US 10/682,764

<141> 2003-10-09

<150> 60/417,282

<151> 2002-10-09

<160> 10

<170> PatentIn version 3.3

<210> 1

<211> 3452

<212> DNA

<213> 人工合成的 DNA 分子

<220>

<223> 合成的构建物

<220>

<221> misc\_feature

<222> (98)..(347)

<223> "n" 是 a, c, g, 或 t

<220>

<221> misc\_feature

<222> (411)..(430)

<223> "n" 是 a, c, g, 或 t

<220>

<221> misc\_feature

<222> (530)..(779)

<223> "n" 是 a, c, g, 或 t

<220>

<221> misc\_feature

<222> (810)..(1059)

<223> "n" 是 a, c, g, 或 t

<220>

<221> misc\_feature

<222> (1163)..(1182)

<223> "n" 是 a, c, g, 或 t

<220>

<221> misc\_feature

<222> (1240)..(1489)

<223> "n" 是 a, c, g, 或 t

<400> 1

tgagcagcgg ataacaattt cacacaggaa acagctatga ccatgattac tctgtagcat	60
ctatgtcggg tgcggagaaa gagglaatga aatggcannn nnnnnnnnnn nnnnnnnnnn	120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnntgc catttcatta	360
cctctttctc cgcacccgac atagataggc cctgcgccgg cggcgatcgc nnnnnnnnnn	420
nnnnnnnnnn ttaattaaac gcgtggcgcg cctaactata acggtcctaa ggtagcggag	480
taccgctggc cctagggtaa tacgactcac tataggcca cataagtggg nnnnnnnnnn	540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	600

[0002]

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	780
cacttaigtg gtagggataa cagggtaatn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna ttacctgtt atccctatcc	1080
ctttagtgag ggittaaltct cgaggcagga gctagctcgc taccttagga ccgttatagt	1140
tatacgtagt cgacgcggcc gcnnnnnnnn nnnnnnnnnn nncgggtccga tttaaatcgt	1200
acgtccggat ggcaaacagc tattatgggt attatgggt nnnnnnnnnn nnnnnnnnnn	1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna cccataatac	1500
ccataatagc tgtttgccag ctacagagtt tactggccgt cgttttaca cgtcgtgact	1560
gggaaaacc tgccggitta aacgctcttc cgtctccttc atgtgagcaa aaggccagca	1620
aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc	1680
tgacgagcat cacaaaaatc gacgctcaag tcagagggtg cgaaaccgga caggactata	1740
aaataaccag gcgtttcccc ctggaagctc cctcgtcgc tctcctgttc cgaccctgcc	1800
gcttaccgga tacctgtccg cctttctccc ttccggaagc gtggcgcttt ctcatagctc	1860
acgctgtagg taictcagtt cggigttagt cgttcgtccc aagctgggct gigtgcacga	1920
acccccggt cagccccgacc gctgcgcctt atccggtaac taicgtcttg agtccaacc	1980
ggaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	2040
gtatgtaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct aactagaag	2100
gacagtattt ggtatctcgc ctctgctgaa gccagttacc ttccgaaaa gagttgtag	2160
ctcttgatcc ggcaaaaaa ccaccgctgg tagcggtagt tttttgttt gcaagcagca	2220
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	2280
cgctcagtg aacgaaaact cacgttaagg gatlttggtc atgataacta tgactctctt	2340
aaggtagcca aatcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa	2400
aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa	2460
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcadc catagttgcc	2520
tgactccccg tgggtgtagat aactacgata cgggagggct taccatctgg cccagtgct	2580
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca	2640
gccggaaggc ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt	2700
aatgtgtgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgtggtt	2760
gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc	2820
ggttccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcggtttag	2880

[0003]

```

tcttccggtc ctccgatcgt tgcagaagt aagttggccg cagtggtatc actcatggtt 2940
atggcagcac tgcataatc tcttactgtc atgcatccg taagatgctt ttctgtgact 3000
ggtagtact caaccaagtc attctgagaa tagtgatgc ggcgaccgag ttgctcttgc 3060
ccggcgtaaa tacgggataa taccgcgcca catagcagaa cttttaaagt gctcatcatt 3120
ggaaagcgtt cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg 3180
atgtaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct 3240
gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa 3300
tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt 3360
ctcatgcgcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc 3420
acatttcccc gaaaagtgcc acctgacgtc gc 3452

```

```

<210> 2
<211> 792
<212> DNA
<213> 人工序列

```

```

<220>
<223> PE3 停滞质粒 MCS

```

```

<220>
<221> misc-feature
<222> (86)..(185)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc-feature
<222> (208)..(307)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc-feature
<222> (330)..(429)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc-feature
<222> (451)..(550)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc-feature
<222> (593)..(692)
<223> "n" 是 a, c, g, 或 t

```

```

<400> 2
gagagagaga cgtcgtgag caggccctgt aatacgactc actatagggg gcgccggagc 60
ttagggataa cagggtaatg gtaccnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
nnnnncctag ggcgatcgcc gccggcggnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300
nnnnnnntta attaaggcgc gccacgcgtn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 420
nnnnnnnnnt acgtagcggc cgcgtcgacg nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 480
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 540
nnnnnnnnnn atttaaatcg gtccgcgtac gcatatagct aacagcctcg agnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660

```

[0004]

```

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngctagcat tacctgttta tcctagccg 720
ctggcgcttc cctttagtga gggtaattt cgggagtta aacgctcttc cgcttccttc 780
atgagagaga ga 792

<210> 3
<211> 3452
<212> DNA
<213> 人工序列

<220>
<223> 初级停滞质粒

<220>
<221> misc_feature
<222> (98)..(347)
<223> "n" 是 a, c, g, 或 t

<220>
<221> misc_feature
<222> (411)..(430)
<223> "n" 是 a, c, g, 或 t

<220>
<221> misc_feature
<222> (530)..(779)
<223> "n" 是 a, c, g, 或 t

<220>
<221> misc_feature
<222> (810)..(1059)
<223> "n" 是 a, c, g, 或 t

<220>
<221> misc_feature
<222> (1163)..(1182)
<223> "n" 是 a, c, g, 或 t

<220>
<221> misc_feature
<222> (1240)..(1489)
<223> "n" 是 a, c, g, 或 t

<400> 3
tgagcagcgg ataacaattt cacacaggaa acagctatga ccatgattac tctgtagcat 60
ctatgtcggg tgcggagaaa gaggtaatga aatggcannn nnnnnnnnnn nnnnnnnnnn 120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnntgc catttcatta 360
cctctttctc cgcacccgac atagataggc cctgcgccgg cggcgatgc nnnnnnnnnn 420
nnnnnnnnnn ttaaitaaac gcgtggcgcg cctaactata acggtcctaa ggtagcggagg 480
taccgctggc cctagggtaa tacgactcac tatagggccca cataagtggn nnnnnnnnnn 540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 780
cacttatgtg gtagggataa cagggtaatn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 900

```

[0005]

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna ttacctgtt atccctatcc	1080
ctttagtgag ggttaattct cgaggcagga gctagctcgc taccttagga cegttagt	1140
tatacgtagt cgacgcggcc gcnnnnnnnn nnnnnnnnnn nncggtcgca tttaaatcgt	1200
acgtccggat ggcaaacagc tattatgggt attatgggtn nnnnnnnnnn nnnnnnnnnn	1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna cccataatac	1500
ccataatagc tgtttgccag ctacagagtt tactggccgt cgttttaca cgtcgtgact	1560
gggaaaacc tggcggttta aacgctcttc cgcttccttc atgtgagcaa aaggccagca	1620
aaaggccagg aaccgtaaaa aggccgcgtt gcigggcttt ttccataggc tccgcccccc	1680
tgacgagcat cacaaaaatc gacgctcaag tcagaggtag cgaaacccga caggactata	1740
aagataccag gcgtttcccc ctggaagctc cctcgtcgc tctcctgttc cgacctgcc	1800
gcttaccgga tacctgtccg cttttctccc ttcgggaagc gtggcgcttt ctcatagctc	1860
acgctgtagg tatctcagtt cggtagtagt cgttcgctcc aagctgggct ggtgacaga	1920
acccccggt cagccccacc gctgcgcctt atccgtaac tatcgtcttg agtccaacc	1980
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	2040
gtatgtaggc ggtgctacag agttctttaa gtggtggcct aactacggct aactagaag	2100
gacagtattt ggtatctcgc ctctgctgaa gccagttacc ttcggaaaaa gaggtagtag	2160
ctcttgatcc ggcaaaaaa ccaccgctgg tagcggtagt tttttgttt gcaagcagca	2220
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	2280
cgctcagtag aacgaaaact cacgttaagg gatitggtc atgataacta tgactctctt	2340
aaggtagcca aattcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa	2400
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttagtctga cagttaccaa	2460
tgcttaalca gtgaggcacc tatctcagcg atctgtctat ttcgttcac catagttgcc	2520
tgactccccg tggigtatag aactacgata cgggagggct taccatctgg cccagtgct	2580
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca	2640
gccggaagg cgcgagcagc aagtggctct gcaactttat ccgcctccat ccagctctat	2700
aattgttggc ggggaagctag agtaagtagt tcgccagtta atagtttgcg caacgtggtt	2760
gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtagggcttc atcagctcc	2820
ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcggtagc	2880
tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttate actcatggtt	2940
atggcagcac tgcataattc tcttactgtc atgccatccg taagaigctt tctgtgact	3000
ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc	3060
ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt	3120
ggaaagcgtt ctccggggcg aaaactctca aggatcttac cgtgttgag atccagttcg	3180

[0006]

```

atgtaaccca ctctgacacc caactgatct tcagcatctt ttactttcac cagcgtttct 3240
gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa 3300
tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca gggltatigt 3360
ctcatgctcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc 3420
acatttcccc gaaaagtgcc acctgacgtc gc 3452

```

```

<210> 4
<211> 1628
<212> DNA
<213> 人工序列

```

```

<220>
<223> 初级停泊质粒 MCS

```

```

<220>
<221> misc.feature
<222> (114)..(363)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc.feature
<222> (427)..(446)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc.feature
<222> (546)..(795)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc.feature
<222> (826)..(1075)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc.feature
<222> (1179)..(1198)
<223> "n" 是 a, c, g, 或 t

```

```

<220>
<221> misc.feature
<222> (1256)..(1505)
<223> "n" 是 a, c, g, 或 t

```

```

<400> 4
gagagagaga cgctcgtgag cagcggataa caatttcaca caggaacag ctatgacat 60
gattactctg tagcatctat gtcgggtgcg gagaaagagg taatgaaatg gcannnnnnn 120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 360
nnntgccatt tcattacctc ttctctcgca cccgacatag ataggccctg cgccggcggc 420
gatcgcnnnn nnnnnnnnnn nnnnnntaa ttaaacgcgt ggcgcgccta actataacgg 480
tcctaaggta gcgagglacc gctggccta gggtaaatag actcactata gggccacata 540
agtggnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 780

```

[0007]

```

nnnnnnnnnn nnnnccact tatgtggtag ggataacagg gtaatnnnnn nnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnattac 1080
cctgittatcc ctatcccttt agtgagggtt aattctcgag gcaggagcta gctcgcctacc 1140
ttaggaccgt tatagttata cgtagtcgac gcggcgcgenn nnnnnnnnnn nnnnnnnncg 1200
gtccgattta aatcgtagct ccggatggca aacagctatt atgggtatta tgggtnnnnn 1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1500
nnnnnaccca taatacccat aatagctggt tgccagctac agagtttact gcccgctggt 1560
ttacaacgct gtgactggga aaaccctggc ggtttaaacg ctcttcegct tccttcatga 1620
gagagaga 1628

<210> 5
<211> 2092
<212> DNA
<213> 人工序列

<220>
<223> SVP 质粒

<400> 5
tgagcgaat acgactcact atagggaggc cctgcgatcg ccgccggcgg atatcggagc 60
tgctgggccc agggagcttc tagaggagct ggatccgctg gagaattcgg agctggaaag 120
cttggagctg ctctgcaggg agctgcatgc gctggcgcac agctgtaat taaggcgcgc 180
cacgcgttcc ggattccctt tagtgagggt taattgttta aacgctcttc cgttcccttc 240
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcggt gctggcggtt 300
ttccataggc tccgcccccc tgacgagcat cacaaaaac gacgctcaag tcagagggtg 360
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 420
tctcctgttc cgacctgcc gcttaccgga tacctgtccg cctttctccc itcgggaagc 480
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggttaggt cgttcgctcc 540
aagctgggct gtgtgcacga acccccgtt cagcccagacc gctgcgcctt atccggtaac 600
tatcgtcttg agtccaacc ggtaagacac gacttatcgc cactggcagc agccactggt 660
aacaggatta gcagagcgag giatgtaggc ggtgctacag agttcttgaa gtgggtggcct 720
aactacggct aactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 780
ttcgaaaaa gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgt 840
ttttttgtt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 900
atcttttcta cggggtctga cgctcagttg aacgaaaact cacgttaagg gattttggtc 960
atgataacta tgactctctt aaggtagcca aattcatgag attatcaaaa aggatcttca 1020
cctagatcct tttaaattaa aaatgaagtt ttaaatacat ctaaaagtata tatgagtaaa 1080
cttggctcga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 1140

```

[0008]

ttcgttcac catagttgcc tgactccccg tggigtatag aactacgata cgggagggct 1200  
 taccatctgg cccagtgct gcaatgatac cgcgagacc acgtcaccg gctccagatt 1260  
 tatcagcaat aaaccagcca gccggaagg cggagcgcag aagtggtcct gcaacittat 1320  
 ccgcctccat ccagtcctatt aatgtttgcc gggagctag agtaagtagt tcgccagita 1380  
 atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg 1440  
 giatggcttc atcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt 1500  
 tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg 1560  
 cagtggtatc actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg 1620  
 taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgaatgc 1680  
 ggcgaccgag ttgctcttgc ccggcgicaa tacgggataa taccgcgcca catagcagaa 1740  
 ctttaaaagt gctcatcatt gaaagcggt cttcggggcg aaaacttca aggatcttac 1800  
 cgctgttgag atccagttcg atgtaaccca ctctgtcacc caactgatct tcagcatctt 1860  
 ttactttcac cagcgtttct gggtagcaaa aaacaggbaag gcaaaatgcc gcaaaaaagg 1920  
 gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattatgaa 1980  
 gcatttatca gggttattgt ctcatgcgcg gatacatatt tgaatgtatt tagaaaaata 2040  
 aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc gc 2092

<210> 6  
 <211> 234  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> SVP 质粒 MCS

<400> 6  
 cgctgagcgt aatacagctc actatagga ggccctgcga tcgcccccgg cggatatcgg 60  
 agctgctggg cccaggagc ttctagagga gctggatccg ctggagaait cggagctgga 120  
 aagcttgagg ctgctctgca gggagctgca tgcctggcg cacagctgtt aattaaggcg 180  
 cgccacgcgt tccggattcc ctttagtgag ggttaattgt ttaaacgctc ttcc 234

<210> 7  
 <211> 2097  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> SVE 质粒

<400> 7  
 tgagcgtaat acgactcact ataggagggc cctgttaatt aaggcgcgcc acgcgtgata 60  
 tcggagctgc tgggcccagg gagcttctag aggagctgga tccgctggag aattcggagc 120  
 tggaaaagctt ggagctgctc tgcaggagc tgcattgcgt ggcgcacagc tgtacgtagc 180  
 ggccgcgtcg actccggatt ccttttagtg agggtaatt gtttaaacgc tcttccgctt 240  
 ccttcatgig agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 300  
 cgtttttcca taggtccgc cccctgacg agcatcacia aaatcgacgc tcaagtcaga 360  
 ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 420  
 tgcctctcc tgttccgacc ctccgctta ccggatacct gtccgccttt ctcccttcgg 480

[0009]

```

gaagcgtggc gctttctcat agctcacgct gtaggiatct cagttcggtg taggtcgttc 540
gctccaagct gggctgtgtg cacgaacccc cgttcagcc cgaccgctgc gccttatccg 600
gtaactatcg tcttgagtcc aacccggtaa gacacgacti atcggcactg gcagcagcca 660
ctggtaacag gattagcaga gcgaggtatg taggcgggtc tacagagttc ttgaagtggg 720
ggcctaacta cggctacact agaaggacag tatttgggat ctgcgctctg ctgaagccag 780
ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 840
gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 900
ctttgatctt tcttacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 960
tggatcatgat aactatgact ctcttaaggt agccaaattc atgagattat caaaaaggat 1020
cttcacctag atccttttaa attaaaaatg aagttttaa tcaatctaaa gtaatatga 1080
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 1140
tctatttcgt tcatccatag ttgcctgact ccccgtagtg tagataacta cgatcgggga 1200
gggcttacca tctggcccca gtgctgcaat gataccgca gaccacgct caccggctcc 1260
agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtctgcaac 1320
tttatccgcc tccatccagt ctattaattg ttccgggaa gctagagtaa gtagttcgcc 1380
agttaatagt ttgcgcaacg tggttgcat tgctacaggc atcgtggtgt cacgctcgtc 1440
gtttggatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 1500
catgtttgtc aaaaaagcgg ttagctcctt cggctcctcg atcgttgtca gaagtaagtt 1560
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 1620
atccgtaaga tgctttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 1680
tatcggcgca ccgagttgct ctgcccggc gtcaatcagg gataataccg cgccacatag 1740
cagaacttta aaagtgtca tcattggaaa gcgttcttcg gggcgaaaac tctcaaggat 1800
cttacgctg ttgagatcca gttcgaigta acccactcgt gcaccaact gatcttcagc 1860
atcttttact ttcaccagcg tttctgggtg agcaaaaaa ggaaggcaaa atgccgcaaa 1920
aaagggcaata agggcgacac ggaatgttg aatactcata ctcttcttt ttcaatatta 1980
ttgaagcatt tatcagggtt attgtctcat gcgcggatac atatttgaat gtaattagaa 2040
aaataacaa ataggggttc cgcgcacatt tccccgaaa gtgccacctg acgtcgc 2097

```

<210> 8  
<211> 239  
<212> DNA  
<213> 人工序列

<220>  
<223> SVE 质粒 MCS

```

<400> 8
cgctgagcgt aatacgactc actatagggg gggcctgtta attaaggcgc gccacgcgtg 60
atatacggagc tgctgggccc agggagcttc tagaggagct ggatccgctg gagaattcgg 120
agcttgaaaag cttggagctg ctctgcaggg agctgcatgc gctggcgcac agctgtacgt 180
agcggccgcg tcgactccgg attcccttta gtaggggta attgtttaaa cgtctctcc 239

```

<210> 9  
<211> 2096  
<212> DNA  
<213> 合成 DNA 分子

[0010]

<400> 9	
tgagcgtaat acgactcact atagggaggc ccgttacgta gcggccgct cgacgatac	60
ggagctgctg ggcccagggg gcctctagag gagctggatc cgctggagaa ttcggagctg	120
gaaagcttgg agctgctctg cagggagctg catgcgctgg cgcacagctg atttaaatcg	180
gtccgcgtac gtccggattc cctttagtga gggtaattg tttaaacgct ctccgcttc	240
cttcatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc	300
gtttttccat aggtccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag	360
gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa gctccctcgt	420
gcgctctcct gttccgacct tggccttac cggataccct gccgcctttc tcccttcggg	480
aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg agglcgttcg	540
ctccaagctg ggctgigtgc acgaaccccc cgttcagccc gaccgctcgc cttatccgg	600
taacatctgt cttgagtcca acccgtaag acacgactta tgcacctgg cagcagccac	660
tggtaacagg attagcagag cgaggtaigt aggcgggtgt acagagttct tgaagtgtg	720
gcctaacact ggctacacta gaaggacagt atttggatc tgcgctctgc tgaagccagt	780
taccttcgga aaaagagttg gtagctctg atccggcaaa caaacaccg ctggtagcgg	840
tggtttttt gtttgaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc	900
tttgatctt tctacgggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt	960
ggatcatgata actatgactc tcttaaggta gccaaattca tgagattatc aaaaaggatc	1020
ttcacctaga tccctttaaa ttaaaaatga agtttttaaat caatctaaag tataatgag	1080
taaaccttgg ctgacagtta ccaatgctta atcagtggag cacctatctc agcgtctgt	1140
ctatttcgtt catccatagt tgcctgactc cccgtgggtg agataactac gatacgggag	1200
ggcttaccat ctggccccag tgcctgcaatg ataccggag acccacgctc accggctcca	1260
gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcttgcact	1320
ttatccgctt ccatccagtc tattaattgt tggcgggaag cttagagtaag tagttcgcca	1380
gttaatagtt tgcgcaactg ggttgccatt gctacaggca tctgtggtgc acgctcgtcg	1440
tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc	1500
atgttgtgca aaaaagcggg tagctccttc ggtccctccga tctgtgtcag aagtaagttg	1560
gccgcagttg taccactcat ggttatggca gcaclgata attctcttac tgtcatgcca	1620
tccgtaagal gcttttctgt gactggtgag tactcaacca agtcatctg agaatagttg	1680
atgcggcgac cgagttgctc ttgccggcg tcaatacggg ataataccgc gccacatagc	1740
agaactttaa aagtgctcat cattggaaag cgttcttcgg ggcgaaaact ctcaaggatc	1800
ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca	1860
tcttttactt tcaccagcgt tcttgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa	1920
aagggataa gggcgacacg gaaatgttga atactcatac tcttctttt tcaatattat	1980
tgaagcattt atcagggtta ttgtctcatg cgcggataca tatttgaatg tatttagaaa	2040
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtcgc	2096
<210> 10	
<211> 238	
<212> DNA	
<213> 合成 DNA 分子	

[0011]

---

<400> 10	
cgctgagcgt aatacgaactc actatagggga ggccctgtac gtagcggccg cgtcgacgat	60
atcggagctg ctgggcccag ggagcttcta gaggagctgg atccgctgga gaattcggag	120
ctggaagct tggagctgct ctgcagggag ctgcatgcgc tggcgcacag ctgatttaaa	180
tcggiccgcg tacgtccgga ticccttag tgagggttaa ttgtttaaac gctcttcc	238

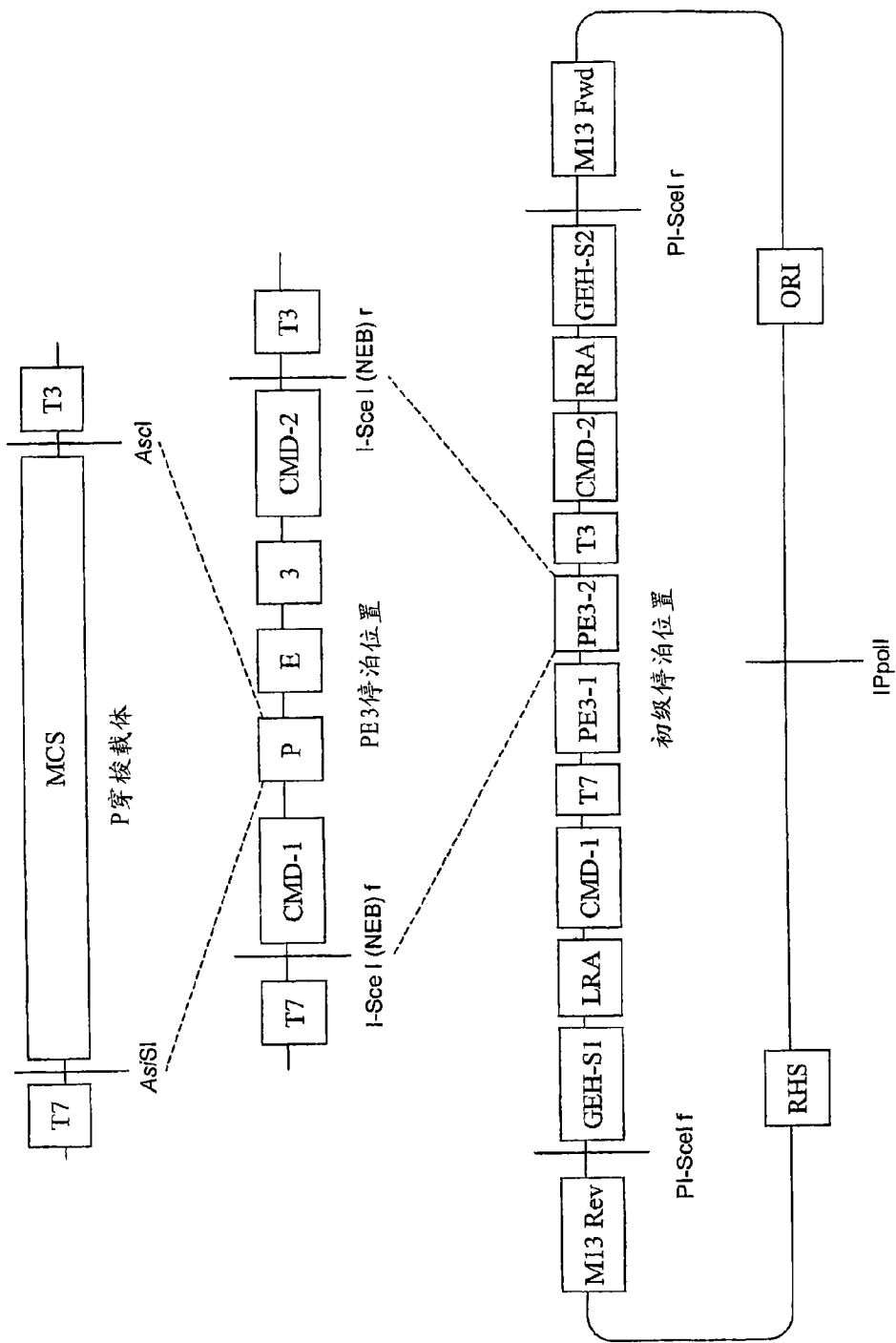


图 1

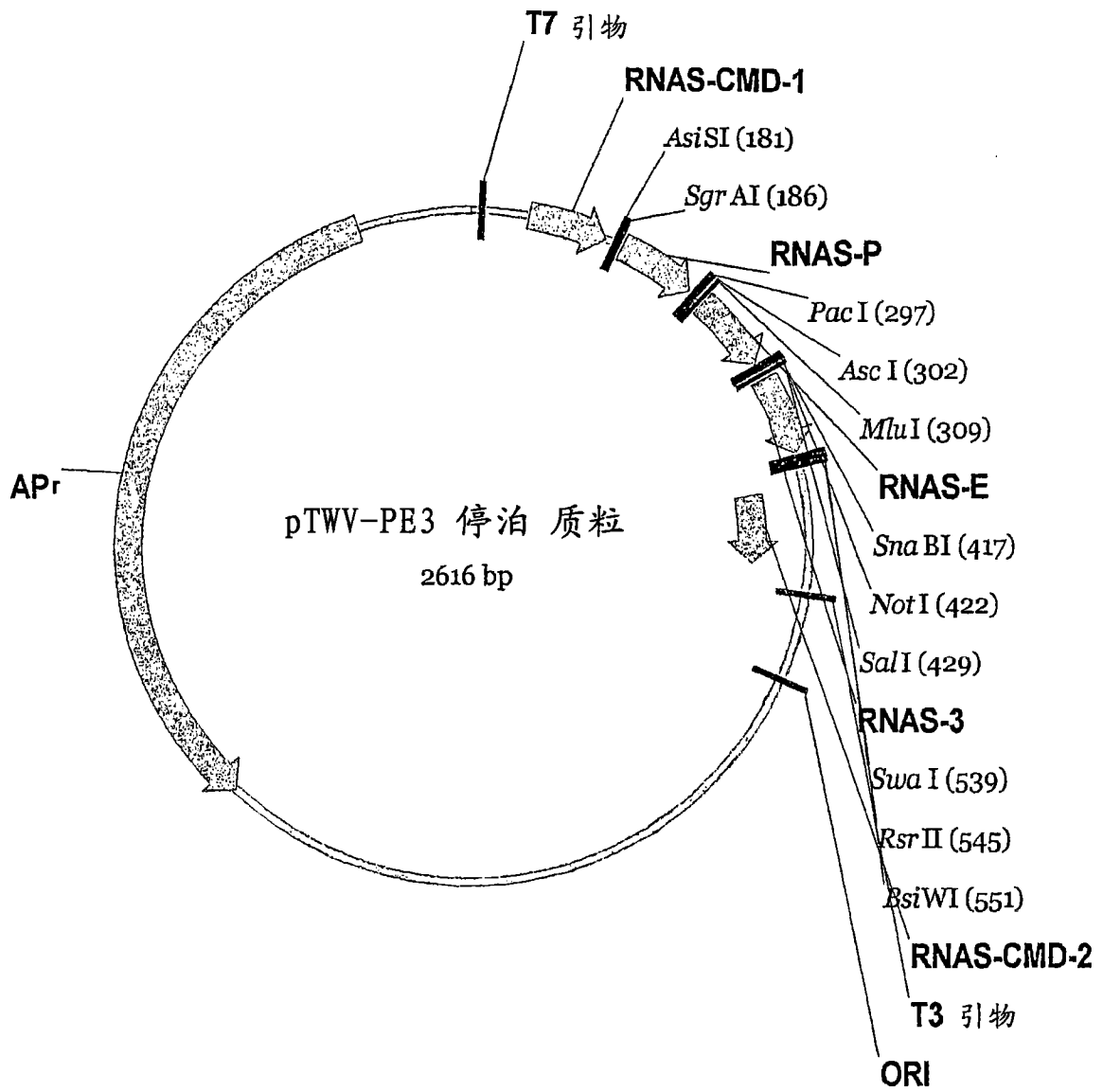


图 2

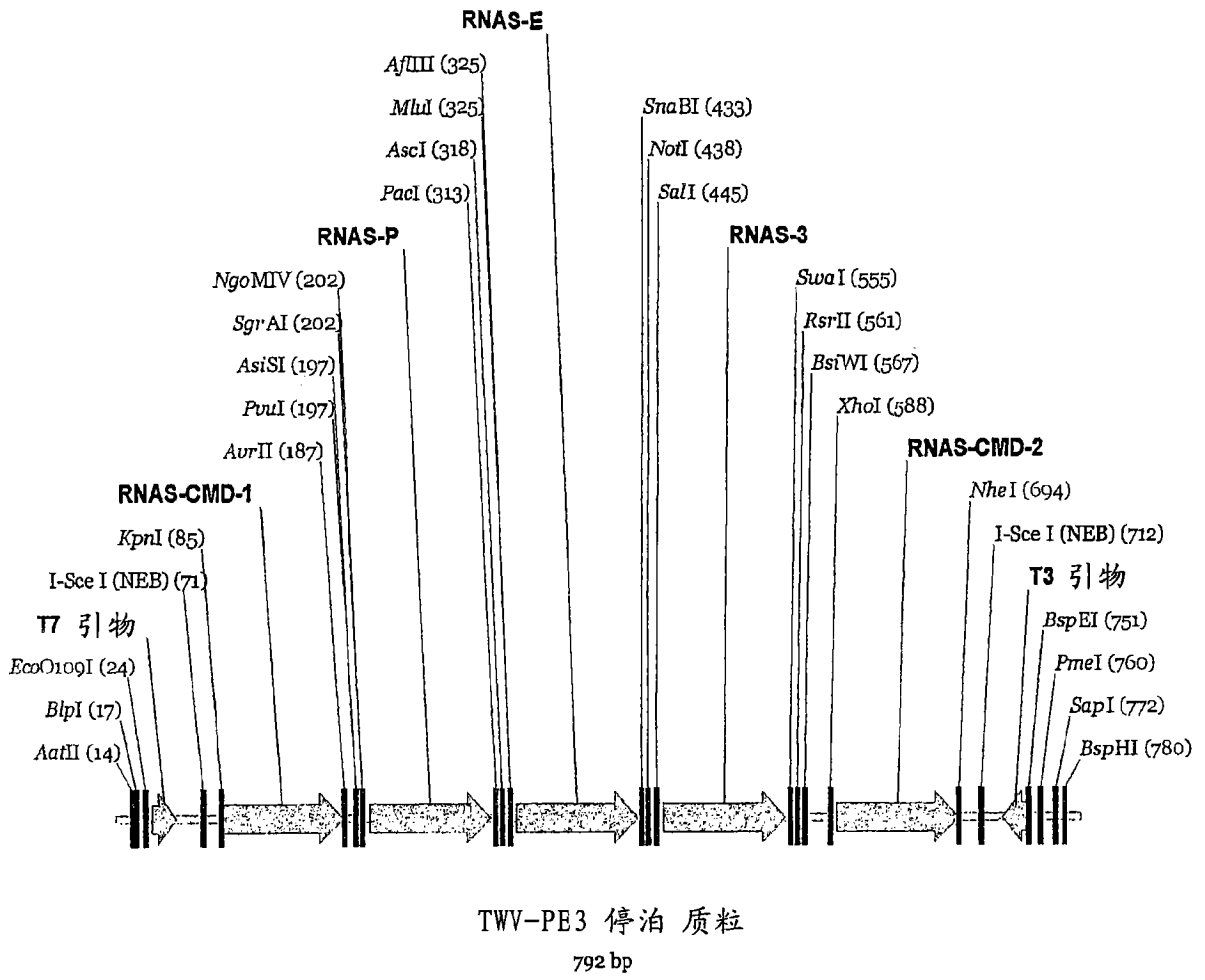


图 3

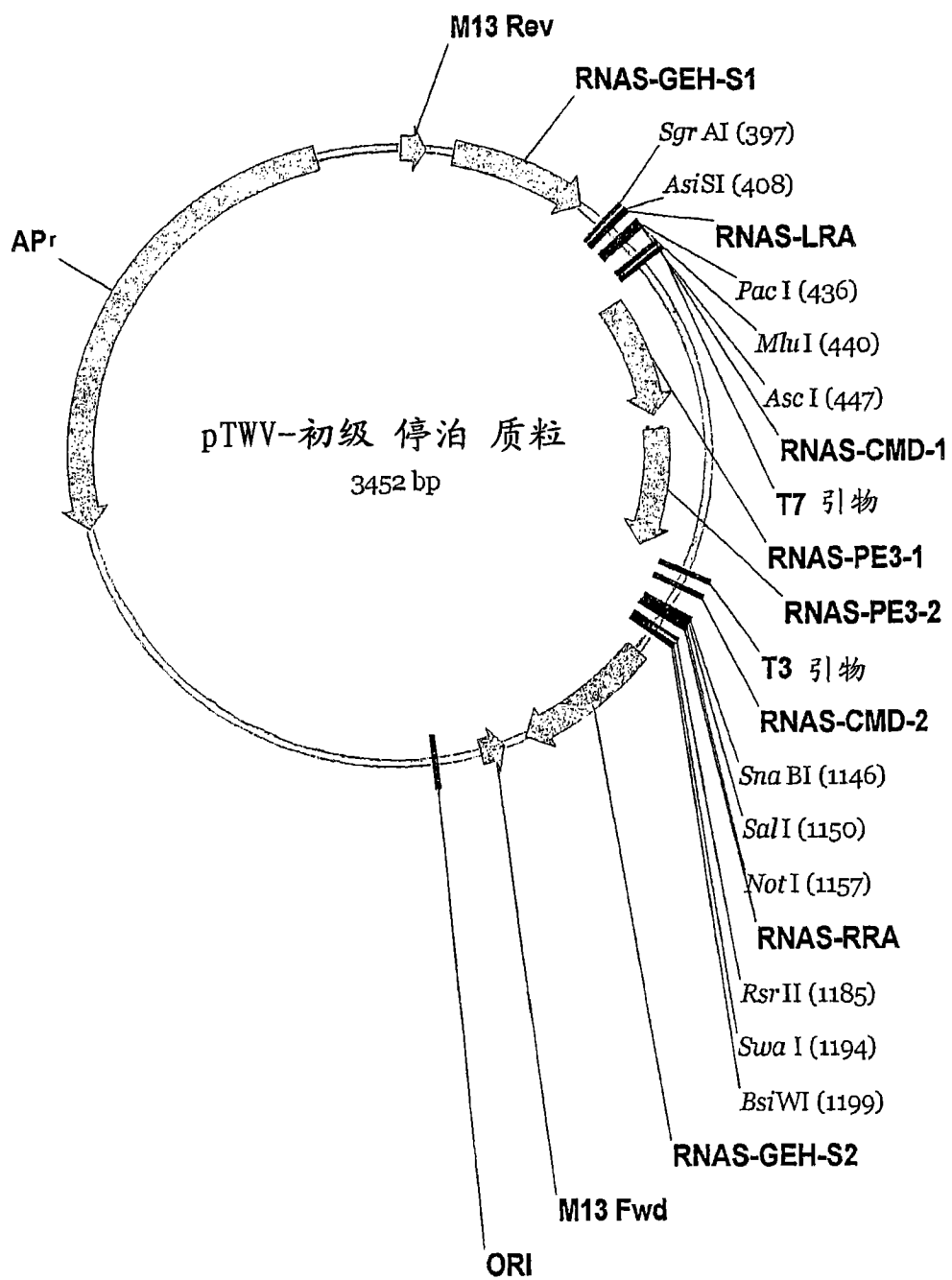
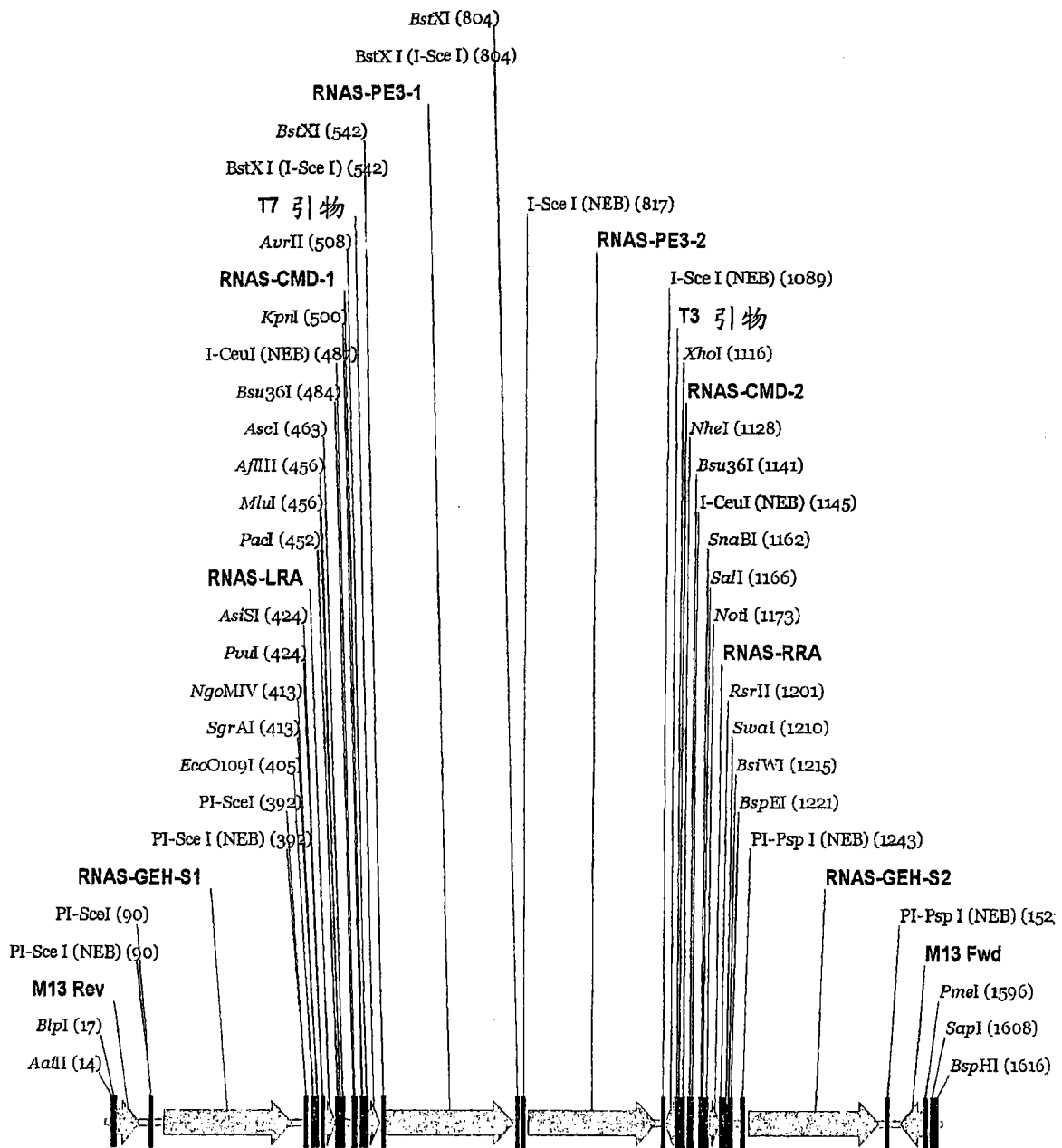


图 4



TWV-初级停泊质粒  
1628 bp

图 5

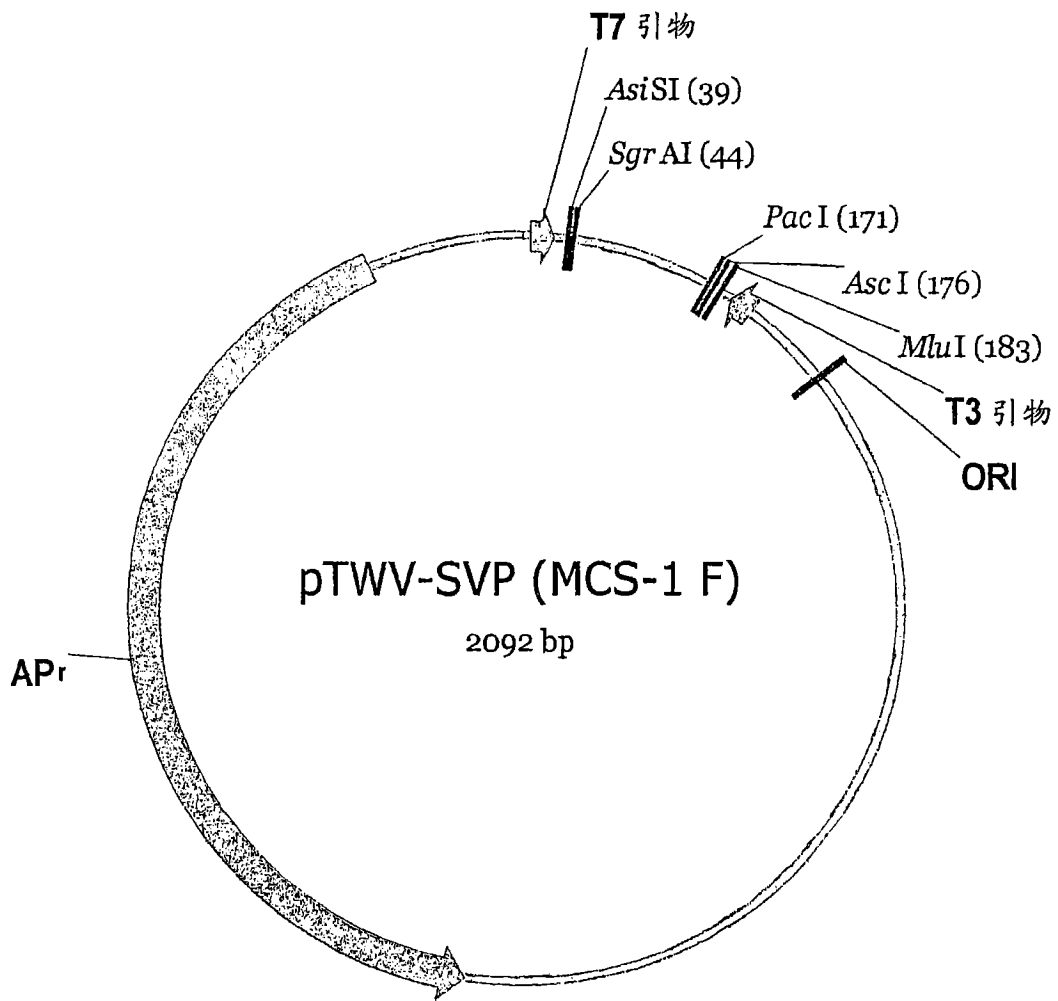


图 6

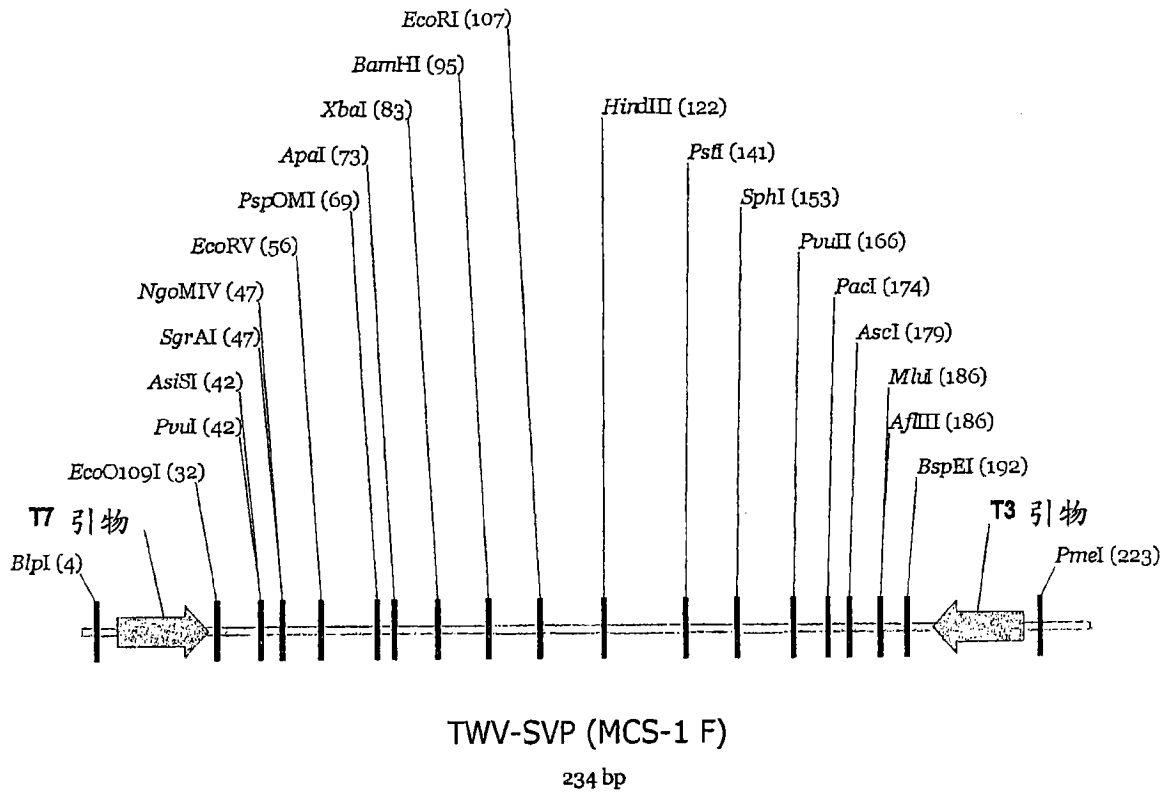


图 7

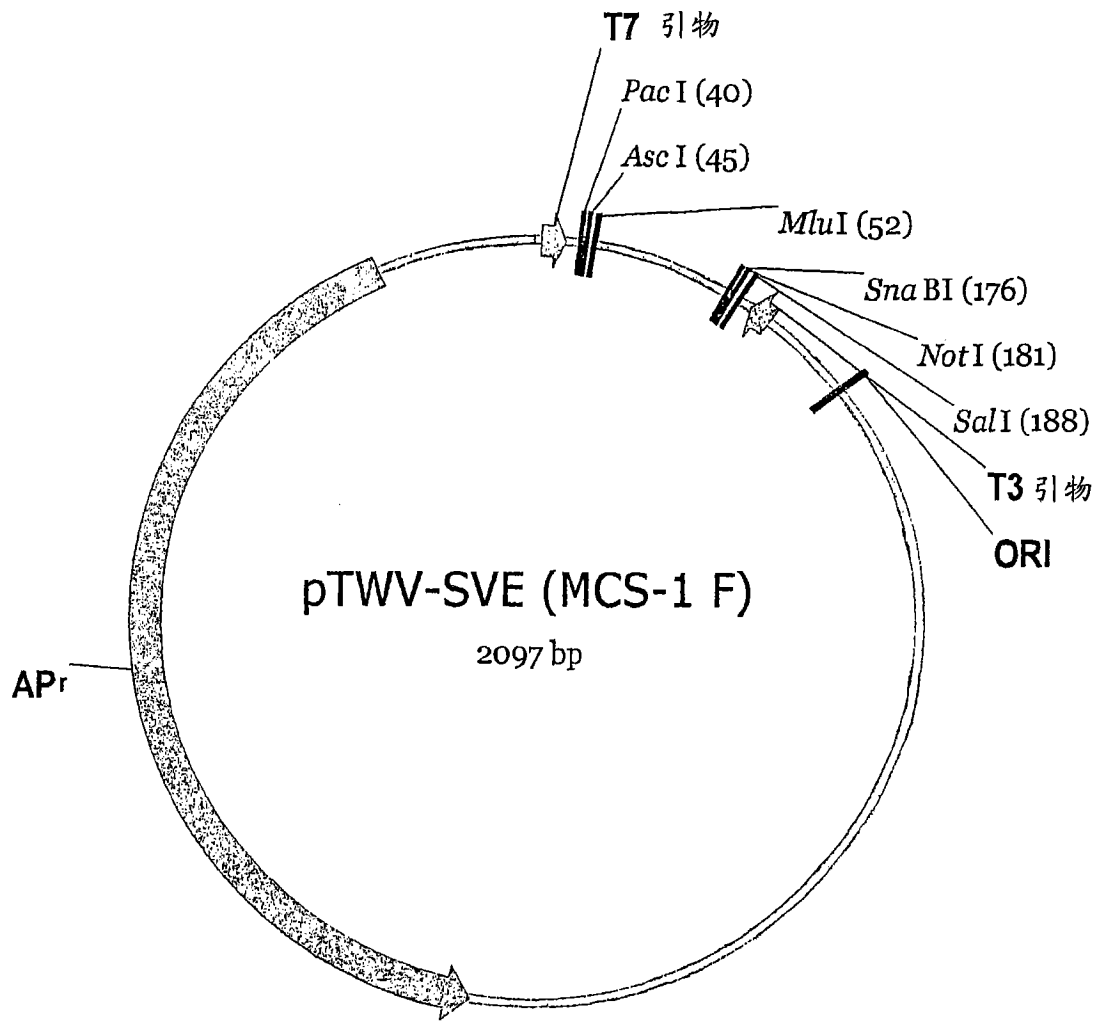


图 8

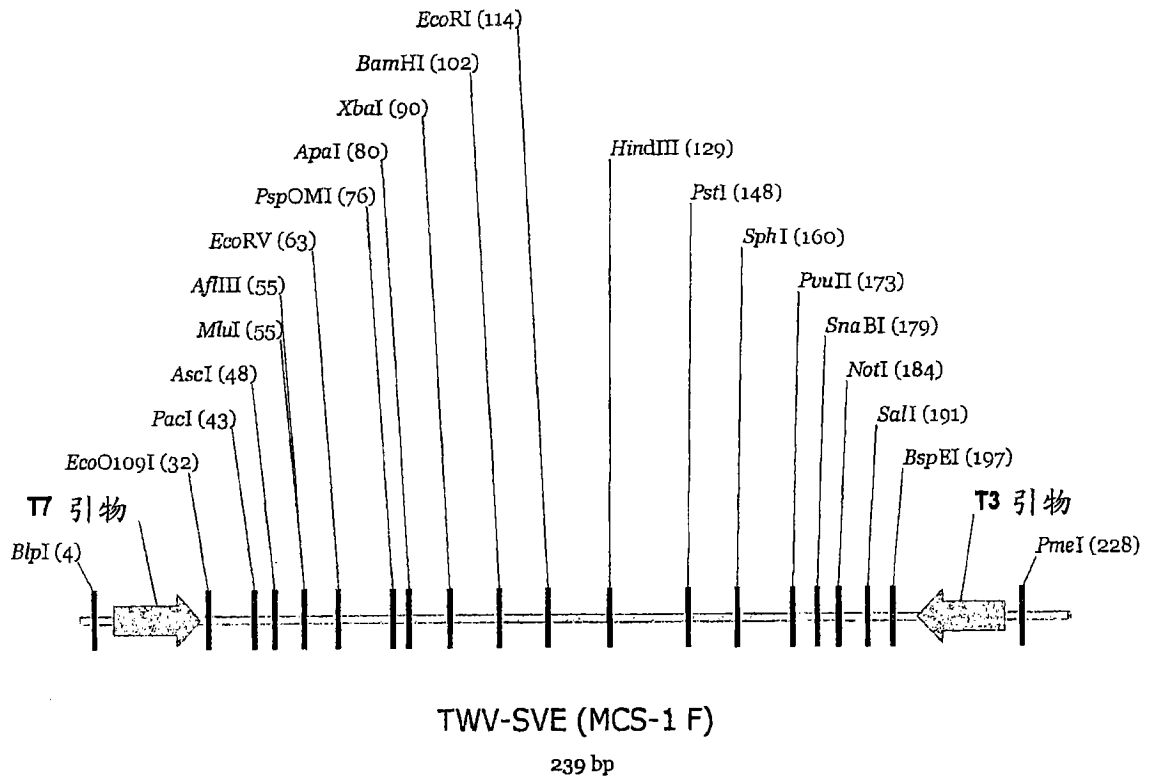


图 9

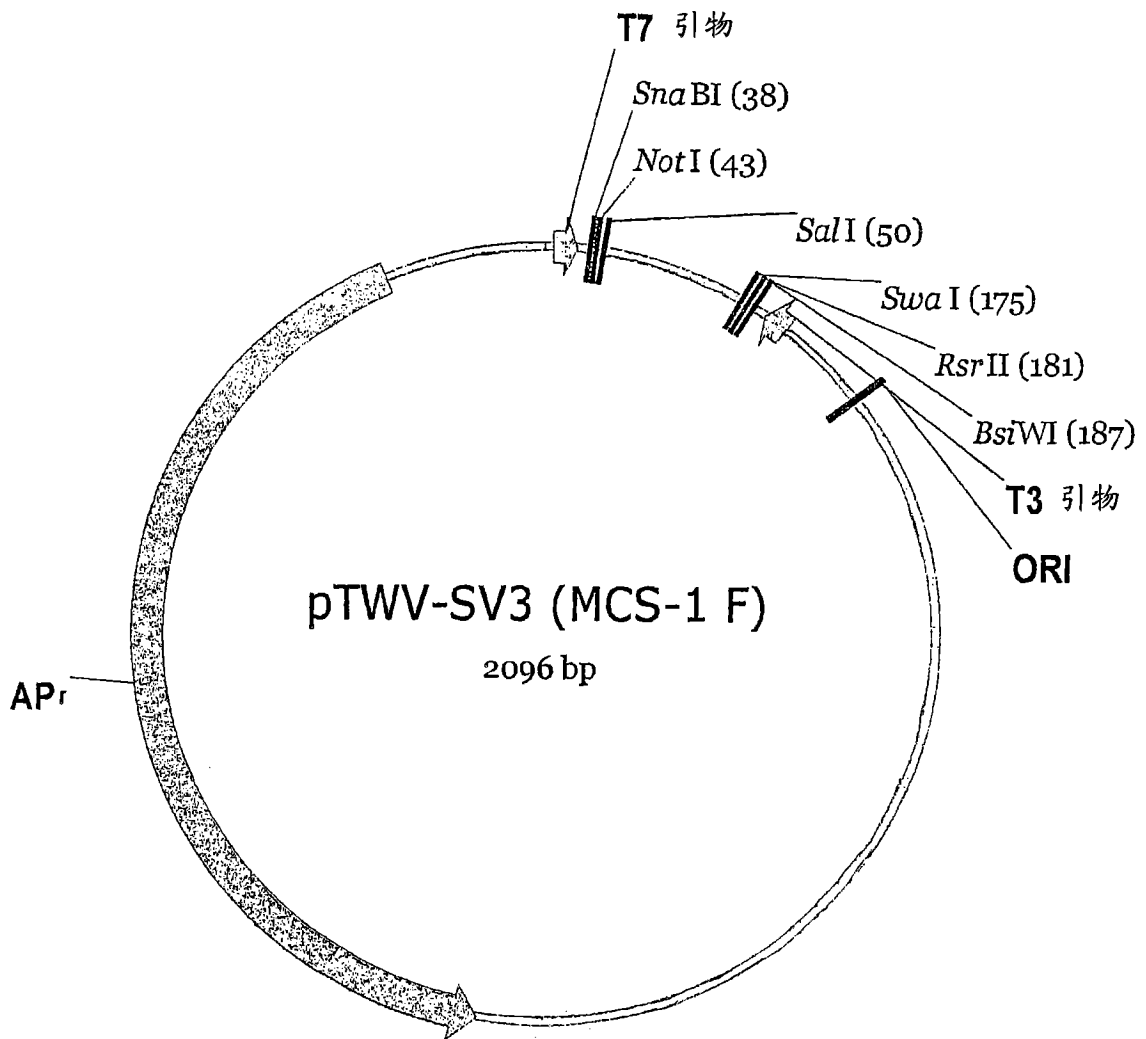


图 10

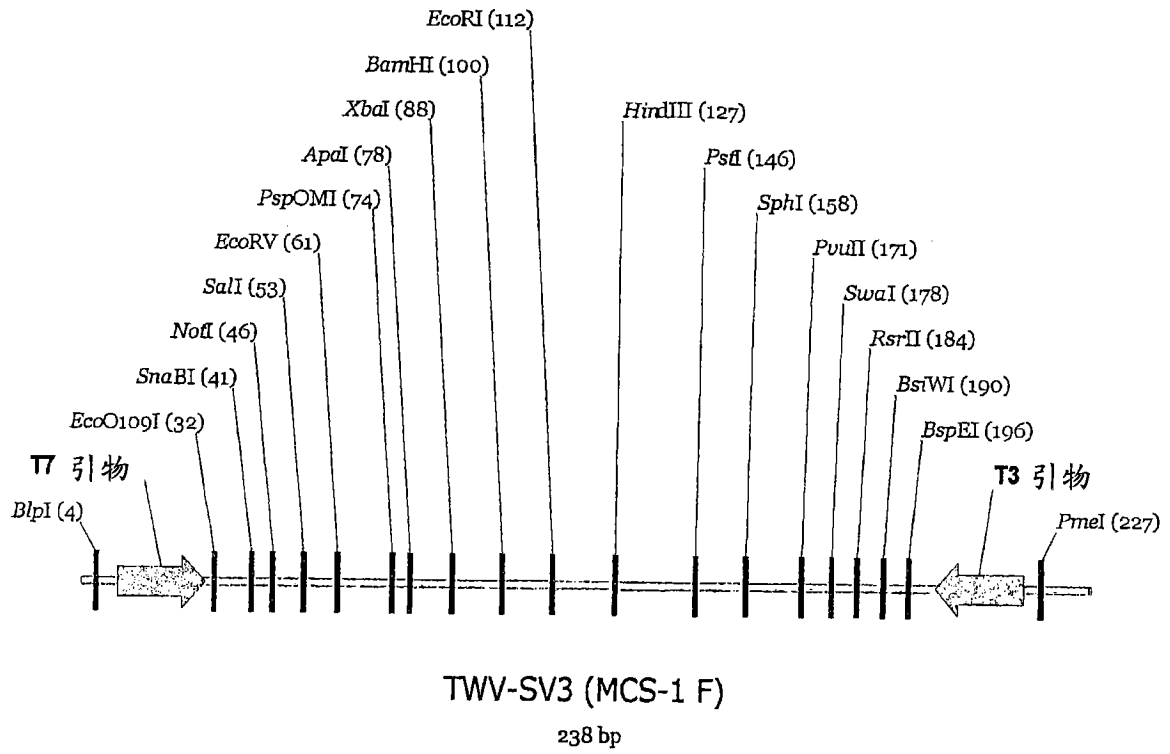


图 11