

[19] 中华人民共和国国家知识产权局

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

C12N 15/86

C12N 9/00 C12N 15/11

[12] 发明专利申请公开说明书

[21] 申请号 99805126.8

[43] 公开日 2001 年 5 月 30 日

[11] 公开号 CN 1297486A

[22] 申请日 1999.2.17 [21] 申请号 99805126.8  
[30] 优先权  
[32] 1998.2.17 [33] GB [31] 9803351.7  
[86] 国际申请 PCT/GB99/00325 1999.2.17  
[87] 国际公布 WO99/41397 英 1999.8.19  
[85] 进入国家阶段日期 2000.10.17  
[71] 申请人 牛津生物医学(英国)有限公司  
地址 英国牛津  
[72] 发明人 A·J·金斯曼 K·米特罗法诺斯  
N·金

[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 谭明胜

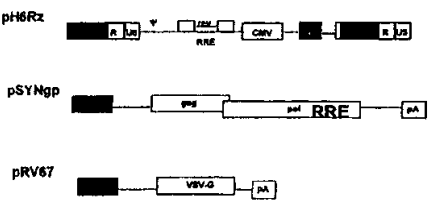
权利要求书 2 页 说明书 29 页 附图页数 8 页

[54] 发明名称 抗病毒载体

[57] 摘要

提供一种病毒载体生产系统,该系统包含:(i)一种包含至少一条第一核苷酸序列的病毒基因组,所述的第一核苷酸序列编码一种基因产物,该基因产物能结合第二核苷酸序列或其转录产物,并直接或间接地影响第二核苷酸序列或其转录产物的裂解,所述第二核苷酸序列编码一种对病毒颗粒装配必须的病毒多肽;(ii)一种编码病毒基因组装配成病毒颗粒所需的所述病毒多肽的第三核苷酸序列,所述的第三核苷酸序列拥有与第二核苷酸序列不同的核苷酸序列,结果所述的第三核苷酸序列或其转录产物抵抗由所述的基因产物指导的裂解。可以用该病毒载体生产系统生产用于治疗或预防病毒感染的病毒颗粒。

HIV 构建体



ISSN 1008-4274

## 权 利 要 求 书

1. 一种病毒载体系统，包括：

(i) 一种编码下列基因产物的第一核苷酸序列，所述的基因产物能结合第二核苷酸序列或其转录产物，并直接或间接地影响第二核苷酸序列或其转录产物的裂解，所述第二核苷酸序列编码一种对病毒颗粒的装配必须的病毒多肽；

(ii) 一种编码病毒颗粒装配所需的所述病毒多肽的第三核苷酸序列，所述的第三核苷酸序列拥有与第二核苷酸序列不同的核苷酸序列，使得所述的第三核苷酸序列或其转录产物抵抗由所述的基因产物指导的裂解。

2. 一种病毒载体生产系统，包括：

(i) 一种包含至少一条第一核苷酸序列的病毒基因组，所述的第一核苷酸序列编码一种基因产物，该基因产物能结合第二核苷酸序列或其转录产物，并直接或间接地影响第二核苷酸序列或其转录产物的裂解，所述的第二核苷酸序列编码一种对病毒颗粒的装配必须的病毒多肽；

(ii) 一种编码病毒基因组装配成病毒颗粒所需的所述病毒多肽的第三核苷酸序列，所述的第三核苷酸序列拥有与第二核苷酸序列不同的核苷酸序列，使得所述的第三核苷酸序列或其转录产物抵抗由所述的基因产物指导的裂解。

3. 一种根据权利要求 1 或 2 的系统，其中该基因产物选自核酶或反义核糖核酸。

4. 一种根据权利要求 1-3 之任一项的系统，其中该病毒载体是一种逆转录病毒载体。

5. 一种根据权利要求 4 的系统，其中该逆转录病毒载体是一种慢病毒载体。

6. 一种根据权利要求 5 的系统，其中该慢病毒载体是一种 HIV 载体。

7. 一种根据权利要求 4-6 之任一项的系统，其中病毒颗粒装配所需的该多肽选自 gag、pol 和 env 蛋白。

8. 一种根据权利要求 7 的系统，其中至少 gag 和 pol 蛋白来自一种慢病毒。

9. 一种根据权利要求 7 的系统，其中 env 蛋白来自一种慢病毒。

10. 一种根据权利要求 8 或 9 的系统，其中该慢病毒是 HIV。

11. 一种根据前述权利要求之任一项的系统，其中该第三核苷酸序列，由于在核苷酸序列中，由至少一个基因产物识别的裂解位点和/或针对至少一个基因产物的结合位点被移去而产生的一种或多种保守性改变导致对所述的基因产物所指导的裂解具有抗性。

12. 一种根据权利要求 1-10 之任一项的系统，其中该第三核苷酸序列适应于抵抗由所述的至少一种基因产物所指导的裂解。

13. 一种根据前述权利要求之任一项的系统，其中该第三核苷酸序列已被密码子优化，适于在生产细胞中表达。

14. 一种根据权利要求 13 的系统，其中该生产细胞是哺乳动物细胞。

15. 一种根据前述权利要求之任一项的系统，其中该系统包括多个由本文所定义的第一核苷酸序列和第三核苷酸序列。

16. 一种病毒颗粒，该病毒颗粒含有一种根据权利要求 2-15 之任一项所定义的病毒载体基因组和一种或多种根据权利要求 2-15 之任一项所定义的第三核苷酸序列。

17. 一种病毒颗粒，该病毒颗粒用一种根据权利要求 2-15 之任一项的病毒载体生产系统产生。

18. 一种产生病毒颗粒的方法，该方法包括导入到宿主细胞中的下列物质：(i) 一种根据权利要求 2-15 之任一项所定义的病毒基因组；(ii) 一种或多种根据权利要求 2-15 之任一项所定义的第三核苷酸序列；(iii) 不是由一种或多种第三核苷酸序列编码但编码其他必须的病毒包装组分的核苷酸序列。

19. 一种用权利要求 18 的方法产生的病毒颗粒。

20. 一种药物组合物，该组合物包括一种根据权利要求 16, 17 或 19 的病毒颗粒，与药学上可接受的载体或稀释剂。

21. 一种根据权利要求 1-16 之任一项的病毒系统，或一种根据权利要求 16, 17 或 19 的病毒颗粒用于治疗病毒感染。

22. 一种根据权利要求 1-16 之任一项的病毒系统，用于产生病毒颗粒的方法中。

## 说明书

## 抗病毒载体

## 发明领域

5 本发明涉及新的病毒载体，所述的病毒载体能把抗病毒的抑制性 RNA 分子传递到靶细胞中。

## 发明背景

基因治疗应用于治疗 AIDS 和 HIV 感染已经进行了广泛的讨论 (14)。提出的治疗基因的类型一般归入二种主要类型之一。在第一  
10 类型中，所述的基因编码蛋白产物，这些产物以许多可能的方式抑制病毒。上述蛋白的一个实例是 HIV Rev 蛋白的 RevM10 衍生物 (16)。RevM10 蛋白起反式显性 (transdominant) 负突变作用，所以在病毒中竞争性地抑制 Rev 的功能。象许多基于蛋白的策略，RevM10 蛋白是天然 HIV 蛋白的衍生物。尽管该蛋白提供抗 HIV 效应的基础，但也有  
15 严重的缺点。尤其是，这种类型的策略需要在缺乏病毒时极少或没有基因表达。反之，携带该基因的健康细胞成为宿主细胞毒 T 淋巴细胞 (CTL) 系统的靶物，CTL 系统识别外来蛋白 (17, 25)。治疗基因的第二种主要类型克服了这些 CTL 的问题。该治疗基因编码抑制性 RNA 分子；RNA 不是 CTL 识别的靶物。所述的 RNA 分子可以是反义 RNA (15,  
20 31)，核酶 (5) 或竞争性假目标 (1)。

核酶是有酶活性的 RNA 分子，它催化序列特异的 RNA 加工。在近几年的文献中，广泛描述了核酶的设计和结构 (3, 7, 31)。其中最有力系统是那些传递多靶核酶的种类，该多靶核酶能在多个位点裂解靶病毒 RNA (5, 21)。

25 近几年来，许多实验室开发了基于 HIV 的逆转录病毒载体系统 (2, 4, 18, 19, 22-24, 27, 32, 35, 39, 43)。在抗 HIV 基因治疗方面，相对于基于鼠的更常规载体如鼠白血病病毒 (MLV) 载体，这些载体有许多优点。首先，HIV 载体将精确地靶向那些易受 HIV 感染的细胞 (22, 23)。其次，基于 HIV 的载体能转导一般难以被鼠载体转导的细胞诸如巨噬细胞 (19, 20)。第三，抗 HIV 基因组能通过  
30 任何逃逸治疗策略的病毒 (HIV)，经 CD4+ 细胞群体增殖 (7)。这是因为该载体基因组具有能被病毒颗粒包装系统识别的包装信号。这些

各种特征使 HIV 载体在抗 HIV 基因治疗领域成为一个有力的工具。

多靶核酶和基于 HIV 的载体的组合作为一种治疗策略是吸引人的。不过，直到现在这种策略才有可能。载体颗粒的产生，发生在表达该颗粒的包装组分并包装该载体基因组的产生细胞中。因此，设计成摧毁病毒 RNA 的核酶也将在载体产生过程中，能打断基于 HIV 载体系统的组分的表达。本发明的目的是克服这个问题。

#### 本发明概述

因而本发明的一个目的是提供一种系统和方法，用于产生病毒颗粒，尤其是 HIV 颗粒，这些颗粒携带编码诸如核酶和/或反义 RNAs 的抑制性 RNA 分子的核苷酸构建体，这些分子在靶细胞中针对相应的病毒如 HIV，从而解决上面所提到的问题。所述的系统包括一种编码抑制性 RNA 分子的病毒基因组，和编码在生产细胞中包装病毒基因组所需要的组分的核苷酸构建体。然而，与以前的技术相反，尽管包装组分拥有基本上与相应的靶病毒组分相同的氨基酸序列，但是该抑制性 RNA 分子在生产细胞中不影响病毒颗粒的生产，因为在病毒系统中所用的包装组分的核苷酸序列已经被修饰，以避免抑制性 RNA 分子影响从构建体产生的 RNA 转录物的裂解或降解。上述的病毒颗粒可以用于治疗病毒感染，尤其是 HIV 的感染。

因此，本发明提供一种病毒载体系统，该系统包括：

(i) 一种编码下列基因产物的第一核苷酸序列，所述的基因产物能结合第二核苷酸序列或其转录产物，并直接或间接地影响第二核苷酸序列或其转录产物的裂解，所述第二核苷酸序列编码一种对病毒颗粒的装配必须的病毒多肽；

(ii) 一种编码病毒颗粒装配所需的所述病毒多肽的第三核苷酸序列，所述的第三核苷酸序列拥有与第二核苷酸序列不同的核苷酸序列，结果所述的第三核苷酸序列或其转录产物抵抗由所述的基因产物指导的裂解。

在另一方面，本发明提供一种病毒载体生产系统，该系统包括：

(i) 一种包含至少一条第一核苷酸序列的病毒基因组，所述的第一核苷酸序列编码一种基因产物，该基因产物能结合第二核苷酸序列或其转录产物，并直接或间接影响第二核苷酸序列或其转录产物的裂解，所述第二核苷酸序列编码一种对病毒颗粒的装配必须的病毒

多肽；

(ii) 一种编码病毒基因组装配进病毒颗粒所需的所述病毒多肽的第三核苷酸序列，所述的第三核苷酸序列拥有与第二核苷酸序列不同的核苷酸序列，结果所述的第三核苷酸序列或其转录产物抵抗由

5 所述的基因产物指导的裂解。

一般来说，该基因产物是一种选自核酶或反义核酸的 RNA 抑制性序列，优选的是核酶。

优选地是，所述的病毒载体是一种逆转录病毒载体，更优选的是一种慢病毒载体，如 HIV 载体。一般来说，第二核苷酸序列和第三核苷酸序列来自相同的病毒物种，更优选的是来自相同的病毒株。一般

10 来说，该病毒基因组来自相同的病毒物种，更优选的是来自相同的病毒株。

在逆转录病毒载体的情况下，病毒颗粒装配所需的多肽选自 gag、pol 和 env 蛋白。优选地是，至少 gag 和 pol 序列是慢病毒的序列，更优选的是 HIV 的序列。此外或另外，env 序列是一种慢病毒的序列，更优选的是 HIV 的序列。

15

在一优选的实施方案中，第三核苷酸序列抵抗由基因产物所指导的裂解，该是因为由至少一个基因产物识别的裂解位点和/或针对至少一个基因产物的结合位点被移去的核苷酸序列中，一种或多种保守性改变引起。例如，当基因产物是一种核酶时，所述的第三核苷酸序列适合于抵抗由该核酶引起的裂解。

20

优选地，第三核苷酸序列是优化适于在宿主细胞中表达的密码子。宿主细胞（该术语包括生产细胞和包装细胞）一般是哺乳动物细胞。

在一特别优选的实施方案中，(i) 病毒基因组是一种包含下列核苷酸序列的 HIV 基因组，所述的核苷酸序列编码抗 HIV 的核酶和/或抗 HIV 的反义序列，对抗在靶 HIV 中的 HIV 包装组分序列（如 gag.pol），以及 (ii) 用于产生包装 HIV 颗粒的病毒系统进一步包含编码与靶 HIV 中相同的包装组分（如 gag.pol 蛋白）的核苷酸构建体，其中，该核苷酸构建体的序列与靶 HIV 中发现的序列不同，所以

25

30 抗 HIV 的核酶和/或反义 HIV 序列，在生产细胞产生 HIV 颗粒的过程中，不影响 gag.pol 转录体的裂解或降解。

本发明也提供一种包含根据本发明病毒载体的病毒颗粒，和一种或多种由根据本发明的第三核苷酸序列编码的多肽。例如，本发明提供一种用本发明的病毒载体生产系统所产生的病毒颗粒。

在另一方面，本发明提供一种生产病毒颗粒的方法，该方法包括  
5 导入到宿主细胞中的下列物质：（i）一种根据本发明的病毒基因组载体；（ii）一种或多种根据本发明的第三核苷酸序列；和（iii）编码不是由一种或多种第三核苷酸序列编码的其他必须的病毒包装组分的核苷酸序列。

本发明进一步提供一种用本发明的方法所产生的病毒颗粒。

10 本发明也提供一种药物组合物，该组合物包括根据本发明的病毒颗粒与药学可接受的载体或稀释剂合用。

本发明的病毒系统或本发明的病毒颗粒可以用于治疗病毒感染，尤其是逆转录病毒感染，如包括 HIV 感染的慢病毒感染。因此，本发明提供一种治疗病毒感染的方法，该方法包括给患病毒感染的人  
15 或动物患者，施用有效量的本发明的病毒系统、病毒颗粒或药用组合物。

本发明特别涉及携带抗 HIV 核酶的基于 HIV 载体。然而，本发明可应用于任何其他的病毒，尤其是任何其他的慢病毒，用基因治疗来对付这些病毒可能是合乎需求的。本发明在此针对 HIV 进行说明，但  
20 是，不能认为这是把本发明的范围限制为基于 HIV 的抗 HIV 载体。  
本发明详述

术语“病毒载体”指的是一种核苷酸构建体，它包含能在宿主细胞中转录的一种病毒基因组，所述的基因组包含足够的病毒遗传信息，在包装组分的存在下允许包装病毒 RNA 基因组进入一种能感染靶  
25 细胞的病毒颗粒。靶细胞的感染包括逆转录并整合到适于具体病毒的靶细胞基因组中。一般来说，所用的病毒载体携带异源的编码序列（目的核苷酸），该编码序列能被载体传递到靶细胞中，例如编码核酶的第一核苷酸序列。所述的病毒载体不能进行独立复制，以在最终的靶细胞中产生感染性病毒颗粒。

30 可以认为，术语“病毒载体系统”意指一套含有各部分的药盒，当与病毒颗粒产生所必须的其他组分组合时，用该药盒可在宿主细胞中产生病毒颗粒。例如，第一核苷酸序列一般可存在于一种质粒载体

构建体中，该构建体适合于把第一核苷酸序列克隆到一种病毒基因组载体构建体中。当在药盒中和第三核苷酸序列（一般也存在于不同的质粒载体构建体中）组合时，含有第一核苷酸序列的质粒和含有第三核苷酸序列的质粒的组合结果包括了本发明的基本要素。然后，上述药盒可被技术人员用于产生合适的病毒载体基因组构建体，当和含有第三核苷酸序列的质粒或者编码其他病毒装配所需组分的核酸构建体一起转染到宿主细胞时，该表达载体基因组构建体可导致产生感染性病毒颗粒。

此外，第三核苷酸序列可稳定地存在于包装细胞系中，本药盒也包括该细胞系。

所述的药盒可以包括产生病毒颗粒所需的其他组分，如宿主细胞和编码病毒装配所需的基本病毒多肽的其他质粒。举例来说，该药盒可含有(i)一种含编码抗 HIV 核酶的第一核苷酸序列的质粒以及(ii)一种含编码修饰的 HIV gag. pol 构建体的第三核苷酸序列的质粒，所述的 HIV gag. pol 构建体不能被抗 HIV 的核酶裂解。然后，选择性组分是(a)一种具有合适的限制酶识别位点的 HIV 病毒基因组构建体，通过这些位点把第一核苷酸序列克隆到病毒基因组中；(b)一种编码 VSV-G env 蛋白的质粒。另外，编码病毒颗粒装配所需的病毒多肽的核苷酸序列也在药盒中作为含有所述核苷酸序列的包装细胞系形式提供，例如表达 VSV-G 的细胞系。

术语“病毒载体产生系统”指的是上面所述的病毒载体系统，其中，第一核苷酸序列已经被插入到合适的病毒载体基因组中。

一般来说，病毒载体是逆转录病毒载体，尤其是慢病毒载体如 HIV 载体。本发明的逆转录病毒载体可衍生于或者来源于任何合适的逆转录病毒。大量的不同逆转录病毒已经被鉴定出来。这些例子包括：鼠白血病病毒 (MLV)、人免疫缺陷病毒 (HIV)、猿猴免疫缺陷病毒、人 T 细胞白血病病毒 (HTLV)、马传染性贫血病毒 (EIAV)、小鼠乳头瘤病毒 (MMTV)、劳氏肉瘤病毒 (RSV)、Fujinami 肉瘤病毒 (FuSV)、莫洛尼鼠白血病病毒 (Mo-MLV)、FBR 鼠骨肉瘤病毒 (FBR MSV)、莫洛尼鼠肉瘤病毒 (Mo-MSV)、Abelson 鼠白血病病毒 (A-MLV)、鸟髓细胞瘤病毒-29 (MC29) 以及鸟成红细胞瘤病毒 (AEV)。逆转录病毒的详细表可在如下的书中找到，见 Coffin 等，1997，“逆转

录病毒”，冷泉港实验室出版社，编著者 JM Coffin, SM Hughes, HE Vamus, 758-763 页。

某些逆转录病毒的基因组结构的细节可以在本领域找到。举例来说，有关 HIV 和 Mo-MLV 的详细资料可从 NCBI 基因库获得（基因组入  
5 库号分别为 AF033819 和 AF033811）。

慢病毒组甚至能进一步分为“灵长类”和“非灵长类”二组。灵长类慢病毒的例子包括人免疫缺陷病毒（HIV，它是引起人自身免疫缺陷综合症（AIDS）的病原）和猿猴免疫缺陷病毒（SIV）。非灵长类慢病毒组包括“慢病毒”原型 visna/maedi 病毒（VMV）、和相关的羊关节炎-脑炎病毒（CAEV）、马传染性贫血病毒（EIAV）以及最近描述的猫免疫缺陷病毒（FIV）和牛免疫缺陷病毒（BIV）。  
10

逆转录病毒基因组的基本结构是一个 5' LTR 和一个 3' LTR（能使基因组包装的包装信号位于之间或其中），一个引物结合位点，能整合到宿主细胞基因组中的整合位点，以及编码包装组分的 gag、pol 和 env 基因—这些是病毒颗粒装配所需的多肽。更复杂的逆转录病毒有其他的特征，如在 HIV 中的 ren 和 RRE 序列，它们能把整合前病毒的 RNA 转录体有效地从细胞核运输到被感染靶细胞的细胞质中。  
15

在前病毒中，这些基因的两端侧翼是称为长末端重复（LTRs）的区域。所述的 LTRs 负责前病毒的整合和转录。LTRs 也用作增强子启动子的序列并能控制病毒基因的表达。逆转录病毒 RNAs 的颗粒化通过位于病毒基因组的 5' 末端的 psi 序列开始。  
20

LTRs 本身是相同的序列，这些序列可分为三个元件，称为 U3、R 和 U5。U3 来源于 RNA 3' 端特有的序列。R 来源于在 RNA 二个末端重复的序列而 U5 来源于 RNA 的 5' 末端特有的序列。三个元件的大小在不同的逆转录病毒中变化较大。  
25

在一缺陷逆转录病毒载体基因组中，gag、pol 和 env 可以缺乏或没有功能。位于 RNA 二个末端的所述 R 区是重复序列。U5 和 U3 分别代表 RNA 基因组的 5' 和 3' 末端的特有序列。

在一典型的用于基因治疗的逆转录病毒载体中，可从病毒中移去至少一个或多个复制所需的 gag、pol 和 env 蛋白编码区的部分。这样制得复制缺陷的逆转录病毒载体。移去的部分甚至可被诸如本发明的第一核苷酸序列等目的核苷酸序列（NOI）取代，以产生能把其基  
30

因整合到宿主基因组中的病毒，但是，其中修饰过的病毒基因组由于缺乏结构蛋白，其自身不能增殖。当在宿主基因组中整合时，出现 NOI 的表达—例如，结果表现出治疗和/或诊断效应。因此，把 NOI 转移到目的位点一般可通过下述方法获得，该方法把 NOI 整合到重组病毒载体中；把修饰病毒载体包装到病毒颗粒外壳上；然后使其转导到目的位点—如靶细胞或靶细胞群体中。

因此，用于本发明的最小逆转录病毒基因组可包括 (5') R-U5-1 个或多个第一核苷酸序列 -U3-R (3')。然而，用于在宿主细胞/包装细胞中产生逆转录病毒基因组的质粒载体，也包括转录调节控制序列，该调节控制序列操作性地与逆转录病毒基因组连接，以指导该基因组在宿主细胞/包装细胞中转录。这些调节序列可以是和转录逆转录病毒序列相连的天然序列，即 5' U3 区，或者它们是诸如另一种病毒启动子的异源启动子，例如 CMV 启动子。

某些逆转录病毒基因组为了有效地产生病毒，需要另外的序列。例如，在 HIV 的情况下，优选地包括 rev 和 RRE 序列。不过，对 rev 和 RRE 序列的需求能通过密码子优化而减少或取消。

一旦逆转录病毒载体基因组作为前病毒 DNA 被整合到其靶细胞的基因组中，必然表达核酶序列。在逆转录病毒中，启动子位于前病毒的 5' LTR U3 区。在逆转录病毒载体中，驱动治疗基因表达的启动子可以是在 5' U3 区的天然逆转录病毒启动子，或是被构建到载体中的替代启动子。该替代启动子基本上可取代逆转录病毒天然的 5' U3 启动子，或者，它可被整合到载体基因组之中如 LTRs 之间的不同位置。

因此，第一核苷酸序列也将操作性地与转录调节控制序列连接，以使第一核苷酸序列的转录在靶细胞中出现。该控制序列一般在哺乳动物细胞中具有活性。例如，所述的控制序列可以是一种病毒启动子，如天然病毒启动子或 CMV 启动子，或者它可以是哺乳动物启动子。特别优选地应用下列的启动子，该启动子优先地在该病毒最初感染的特殊细胞类型或组织类型中有活性。因此，在一实施方案中，可用组织特异调节的序列。驱动一个或多个第一核苷酸序列表达的调节控制序列可以是组成性的或调节的启动子。

例如，通过应用包装或辅助细胞系和重组载体结合的方法，可增殖复制缺陷的逆转录病毒载体，以制备用于下一步转导的逆转录病毒

载体的合适滴度。那就是说，以反式方式提供三个包装蛋白。

“包装细胞系”含有一个或多个逆转录病毒的 gag、pol 和 env 基因。该包装细胞系产生用于包装逆转录病毒 DNA 所需的蛋白，但由于缺乏 psi 区，它不能颗粒化。然而，当携带一个 NOI 和 psi 区的重组载体被导入到包装细胞系中时，所述的辅助蛋白能包装 psi 阳性的重组载体，以产生重组病毒贮存物。该病毒贮存物能用于转导细胞，以把 NOI 导入到靶细胞的基因组中。应用称为 psi 阳性的 psi 包装信号是优选的，psi 阳性含有横跨从拼接供体的上游至 gag 起始密码子下游的附加序列（Bender 等（46）），因为已经表明该包装信号增加病毒的滴度。

基因组缺乏形成病毒蛋白所需的全部基因的重组病毒仅能一次性转导，却不能增殖。这些仅能转导靶细胞一个循环的病毒载体被称为复制缺陷载体。所以，所述的 NOI 导入到宿主/靶细胞基因组中却不产生潜在的有害逆转录病毒。有效的包装系总结表见 Coffin 等，1997（同上）的文章中。

优选地使用下述的逆转录病毒包装细胞系，在这些细胞系中，gag、pol 和 env 病毒编码区在不同的表达质粒中携带，这些质粒独立地被转染进包装细胞系中。有时称为三质粒转染法（Soneoka 等（33））的这种策略，能降低产生有复制活性的病毒的潜力，因为产生野生型病毒需要三个重组事件。由于同源性大大地有利于重组，也用降低或消除载体和辅助体的基因组之间的同源性的方法，以降低产生有复制活性的辅助病毒的问题。

一种稳定地转染包装细胞系的替代方法是应用瞬时转染的细胞系。当正在开发载体时，使用瞬时转染法测定载体的产生水平具有优势。在这一方面，瞬时转染法避免了出现稳定的产生载体细胞系所需的较长时间，并且也可用于如果载体或逆转录病毒包装组分对细胞有毒性的情况。一般用于产生逆转录病毒载体的组分包括编码 gag/pol 蛋白的质粒，编码 env 蛋白的质粒和含有 NOI 的质粒。载体的生产涉及一种或多种这些组分瞬时转染到含有所需其他组分的细胞中。如果所述的载体编码毒性基因或干扰宿主细胞复制的基因，如细胞周期抑制蛋白或诱导凋亡的基因，出现稳定产生载体的细胞系就可能困难，但能用瞬时转染法，在细胞死亡之前产生载体。另外，应用瞬时转染

法开发出了许多细胞系，这些细胞产生的载体滴度水平可与从稳定产生载体的细胞系的水平相比拟（Pear 等（47））。

生产细胞/包装细胞可以是任何合适的细胞类型。最常用的是哺乳动物生产细胞，但是其他细胞如昆虫细胞也不排除在外。显然，生产细胞必须能有效地翻译 gag、pol 和 env 的 mRNA。许多合适的生产/包装细胞系在本领域众所周知。例如，本领域的技术人员通过把编码包装组分的核苷酸构建体稳定地导入细胞系中，也能制取合适的包装细胞系。

正如下面讨论的那样，如果逆转录病毒基因组编码一种能影响 gag、pol 和/或 env RNA 转录体裂解的抑制性 RNA 分子，所述的核苷酸序列以整合或质粒携带的方式存在于包装细胞系中，或者存在于瞬时转染的生产细胞中，编码 gag、pol 和/或 env 蛋白的这些核苷酸序列将被修饰，这样降低或防止与抑制性 RNA 分子的结合。用这种方法，抑制性 RNA 分子不能阻止在包装细胞系中病毒颗粒包装所需的组分的表达。

在实验和实际用途中，使用高滴度的病毒制品是十分合乎要求的。增加病毒滴度的技术包括使用上面讨论的 psi 阳性包装信号和一定浓度的病毒贮存物。此外，使用不同的包膜蛋白，如来自疱疹性口炎病毒的 G 蛋白，可提高滴度，使其浓度达到每毫升  $10^9$ （Cosset 等（48））。不过，一般选择包膜蛋白，这样病毒颗粒优先地感染被需要治疗的病毒感染的细胞。例如，如果一种 HIV 载体被用于治疗 HIV 感染，所用的 env 蛋白是 HIV env 蛋白。

按照本发明所用的合适的的第一核苷酸序列编码下述的基因产物，这些基因产物导致靶核苷酸序列的裂解和/或酶促降解，它们一般是核糖核酸。作为具体的实例，所指的是核酶和反义序列。

核酶是在特异的位点裂解 RNA 的 RNA 酶。可以这样构建核酶，使其特异地针对任何选择的含有核酶裂解位点的序列。因此，在转录病毒序列中选择识别位点而构建核酶。举例来说，由第一核苷酸序列编码的核酶，识别和裂解产生病毒颗粒所需的病毒基因组的必要元件，如包装组分。所以，对逆转录病毒基因组来说，上述必要元件包括 gag、pol 和 env 基因产物。一种能识别至少 gag、pol 和 env 基因序列之一，或更典型地识别从这些基因转录的 RNA 序列的合适的核酶，

能结合和裂解上述序列。作为合适的方法，这将降低或阻止 gag、pol 或 env 蛋白的产生，因而降低或阻止逆转录病毒颗粒的产生。

核酶以几种形式存在，包括锤头、发夹和肝炎  $\delta$  型抗基因组核酶。本发明优选地使用锤头核酶，部分原因是它们的体积相对较小，也因为它们的靶裂解位点的序列要求最小，以及它们的特性已得到全面的描述。在本发明研究中最常用的核酶是锤头核酶和发夹核酶。

每一个单独的核酶有一个基序，该基序识别在靶 RNA 中的识别位点并与之结合。该基序采取一个或多个“结合臂”的形式，一般是二个结合臂。在锤头核酶中的结合臂是侧翼序列螺旋 I 和螺旋 III，它们位于螺旋 II 的侧翼。这些序列的长度是可变的，一般每种为 6-10 个核苷酸，但可以更短或更长。侧翼序列的长度能影响裂解率。例如，已经发现降低侧翼序列中的核苷酸总数，从 20 降低到 12，能增加核酶裂解 HIV 序列的转换率，达到 10 倍（44）。在锤头核酶中，核酶螺旋 II 的催化基序，在一个称为裂解位点的位点中裂解靶 RNA。核酶是否裂解任何给定的 RNA，可通过存在或缺乏针对含合适的裂解位点的核酶的识别位点确定。

每一类型的核酶识别其自身的裂解位点。锤头核酶裂解位点拥有直接位于上游的核苷酸碱基三联体 GUX，其中，G 是鸟嘌呤，U 是尿嘧啶而 X 是任何的核苷酸碱基。发夹核酶拥有一个 BCUGNYR 的裂解位点，其中，B 是任何不是腺嘌呤的核苷酸碱基，N 是任何核苷酸，Y 是胞嘧啶或胸腺嘧啶而 R 是鸟嘌呤或腺嘌呤。由发夹核酶引起的裂解发生在裂解位点的 G 和 N 之间。

编码包装组分的核酸序列（第三核苷酸序列）可以抵抗所述的核酶或多种核酶，因为它们缺乏针对所述的核酶或多种核酶的任何裂解位点。这将阻止由所述的核酶或多种核酶引起的酶活性，因此不存在针对所述的核酶或多种核酶的有效识别位点。另外或此外，潜在的识别位点可在侧翼序列中改变，这些序列构成核酶与之结合的结合位点的一部分。这种改变消除核酶基序与识别位点的结合，或者足以使任何核酶-靶复合物的去稳定而降低结合能力，因此降低该核酶的特异性和催化活性。如果仅改变侧翼序列，它们优选地进行如下改变，结果在该改变的靶序列上的核酶催化活性可以被忽略或有效地消除。

优选地是，本发明使用了几种抗 HIV 的核酶系列（5，7，10，13，

21, 36, 38, 40)。这些核酶可以是任何的抗 HIV 的核酶，但是必须包括一种或多种裂解表达 gag、pol 或 env 所需的 RNA 的种类。优选地是，利用多个核酶，同时在多个位点能裂解原始逆转录病毒的 gag、pol 或 env 的 RNA。由于 HIV 以准物种群存在，并不是所有的核酶的靶序列包括在全部的 HIV 变体中。由该变异性所出现的问题能通过使用多靶核酶解决。多靶核酶包括成系列存在于单一载体中，并且当作为一条单一的 RNA 序列表达时，能独立起作用。一种含有二个或多个具有不同靶识别位点的核酶的单一 RNA 可以称为多靶核酶。已经表明：成系列地设置核酶可增强裂解作用。使用多个核酶并不局限于治疗 HIV 感染，而是可用于有关的其他病毒、逆转录病毒或其他病毒。

反义技术在本领域众所周知。一般认为反义序列通过各种机制抑制基因表达。据信反义序列起作用的一种机制是细胞蛋白 RNaseH 被招募到靶序列/反义构建体的异源双链体处，结果引起异源双链体的裂解和降解。因此反义构建体，与核酶的作用相反，据说能间接地引起靶序列的裂解/降解。因此根据本发明，第一核苷酸序列可以编码一种反义 RNA，该反义 RNA 与一种编码基本/包装组分的基因结合或与来自该基因转录的 RNA 结合，结果抑制该基因的表达，例如导致得到的异源双链体的 RNaseH 降解。对反义构建体来说，没有必要编码与下列基因互补的全长序列，所述的基因编码一种基本/包装组分——部分就足以起作用。技术人员可容易地确定如何设计合适的反义构建体。

相反，在宿主细胞/生产细胞/包装细胞中，编码病毒颗粒装配必需的病毒颗粒的基本/包装组分的核酸序列（第三核苷酸序列），抵抗由第一核苷酸序列编码的抑制性 RNA 分子。例如，在核酶的情况下，该抗性一般利用在序列中取消核酶识别位点的变异而实现。同时，保留针对基本/包装组分的氨基酸编码序列，这样由所述的序列编码的病毒组分仍然相同，或者至少十分相似，结果基本/包装组分的功能不被损害。

术语“病毒颗粒装配必需的病毒多肽”指的是一种一般由病毒基因组编码的被包装到病毒颗粒中的多肽，在缺乏该多肽时，病毒基因组就不能包装。例如，就逆转录病毒来说，上述多肽包括 gag、pol 和 env。术语“包装组分”和“基本组分”也包括在本定义之中。

在反义序列的情况下，第三核苷酸序列与编码靶病毒包装组分的反义序列的第二核苷酸序列不同，因为尽管反义序列能与第二核苷酸序列或其转录体结合，所述的反义序列不能有效地与第三核苷酸序列或从该序列转录的 RNA 结合。在第二和第三核苷酸序列之间的变化一般是保守性变化，虽然也可以忍受少数的氨基酸变化，只要基本/包装组分的功能不被明显地损害就可。

优选地，除了取消核酶识别位点外，针对病毒组分的编码序列的改变可改善在哺乳动物细胞或其他细胞中的密码子使用性，其中这些细胞作为产生逆转录病毒载体颗粒的生产细胞。在密码子使用方面的这种改善被称为“密码子优化”。包括 HIV 和其他慢病毒的许多病毒使用大量的稀有密码子，通过把这些密码子改变成相应的常规使用的哺乳动物密码子，能获得在哺乳动物生产细胞中包装组分的增加表达。针对哺乳动物细胞以及各种其他生物的密码子使用表，在本领域众所周知。

因而优选地，编码包装组分的序列是密码子优化的。更优选地是，该序列作为整体来说是密码子优化的。密码子优化后，可以发现，在野生型 gag、pol 和 env 序列中出现许多位点，它们能用作核酶识别位点并且不再存在于编码包装组分的序列中。在一替代但更不实用的策略中，编码包装组分的序列能通过靶向保守性改变而改变，只要使它们抵抗选出的能裂解野生型序列的核酶就行。

密码子优化的 HIV 包装组分的一个另外优点是能增加基因的表达。尤其是，它能使 gag、pol 的表达不依赖于 Rev，这样在基因组中就不需要包括 rev 和 RRE (11)。因而有可能得到不依赖 Rev 的载体。结果这种特点能在逆转录病毒载体中使用抗 rev 或 RRE 因子。

如上所述，用于逆转录病毒载体的包装组分包括 gag、pol 和 env 基因的表达产物。按照本发明，在包装系统中使用的 gag 和 pol 来源于该载体基因组以之为基础的靶逆转录病毒。因此，在 RNA 转录体形式中，gag 和 pol 一般可被存在于载体基因组中的核酶裂解。在包装系统中使用的 env 基因来源于不同的病毒，包括诸如 MLV 的其他逆转录病毒和诸如 VSV (棒状病毒) 的非逆转录病毒，在该情况下，为使其抵抗核酶裂解，不需要任何序列改变。此外，env 可来源于与 gag 和 pol 相同的逆转录病毒，在该情况下，任何针对核酶的识别位点需

要通过序列改变而消除。

产生包被蛋白不是逆转录病毒的原始包被的逆转录病毒载体的过程，被称为“假型化”（“pseudotyping”）。某些包被蛋白，如 MLV 包被蛋白和疱性口炎病毒 G（VSV-G）蛋白，非常有效地使逆转录病毒假型化。假型化能用于改变逆转录病毒的靶细胞范围。此外，为了维持需要治疗的被特异病毒感染的靶细胞的靶细胞特异性，该包被蛋白可与诸如 HIV 等靶病毒的包被蛋白相同。

其他治疗性编码序列可和第一核苷酸序列或各类序列一起存在。其他治疗性编码序列包括，但不局限于，编码下列分子的序列，这些分子是细胞因子、激素、抗体、免疫球蛋白融合蛋白、酶、免疫共刺激分子、反义 RNA、靶蛋白的反式显性阴性突变体、毒素、条件性毒素、抗原、单链抗体、肿瘤抑制蛋白和生长因子。当包括上述分子时，上述编码序列操作性与一合适的启动子连接，该启动子可以是驱动第一核苷酸序列表达的启动子，或不同的启动子或各类启动子。

因此本发明包含两种组分。第一种组分是一种基因组构建体，该构建体被病毒包装组分包装并携带一系列抗病毒的抑制性 RNA 分子，如抗 HIV 的核酶（5，7，10，13，21，36，38，40）。这些分子可以是任何抗 HIV 的核酶，但是，本发明的关键问题是某些分子裂解表达天然或野生型 HIV gag、pol 或 env 编码序列所需的 RNA。第二种组分是下述的包装系统，该包装系统含有一种表达 HIV gag、pol 的表达盒，和一种表达 HIV env 的表达盒或一种编码假型化包被蛋白的包被基因—该包装系统能抵抗抑制性 RNA 分子。

因此，本发明的病毒颗粒和病毒载体系统以及所用的产生载体的方法，可用于治疗或预防病毒感染，优选地治疗或预防逆转录病毒感染，尤其是慢病毒（具体来说 HIV）的感染。具体而言，本发明的病毒颗粒，一般用本发明的病毒载体系统产生，能用于把抑制性 RNA 分子传递到需要针对病毒感染治疗的人和动物中。

另外或此外，可用病毒产生系统转染离体从病人获得的细胞，然后回输到该病人体内。离体转染的病人细胞，在重新给病人施用之前，可制成一种药用组合物（见下面的描述）。

优选地是，所述的病毒颗粒可与药用上可接受的载体或稀释剂结合使用，以产生药用组合物。由此，本发明也提供一种用于治疗个体

的药用组合物，其中，该组合物包含治疗有效量的本发明的病毒颗粒，与药用上可接受的载体、稀释剂、赋形剂或佐剂一起施用。该药用组合物可用于人类或动物。

5 药用的载体、赋形剂或稀释剂可根据施用的预定途径和标准的药  
学实践选择。合适的载体和稀释剂包括等渗的盐溶液，例如磷酸盐缓冲液。该药用组合物可包括-除了载体、赋形剂或稀释剂之外的任何  
合适的结合剂、润滑剂、悬浮剂、包被剂、溶剂和其他载体药剂，这  
些药剂可辅助或增加靶位点的病毒进入（例如脂类传递系统）。

10 该药用组合物可制成制剂，适于胃肠道外、肌肉内、静脉内、颅内、皮下、眼内或透皮施用。

当条件合适时，该药组合物可通过下列任何一种或多种制剂方式施用：吸入剂，以栓剂或阴道栓剂的形式，用洗剂、液体剂、乳剂、药膏或粉末剂形式，通过皮肤垫在体表施用；口服施用可以是含诸如淀粉或乳糖的赋形剂的片剂形式，或单独或与赋形剂形成混合剂的胶囊或卵状囊形式，或者含有香味剂或增色剂的酏剂、液体剂或悬浮剂  
15 的形式；或它们在肠胃外，例如体腔内、静脉内、肌肉内或皮下注射使用。用于肠胃外施用时，该组合物最好以含有其他物质的无菌水溶液形式使用，例如含有足够的盐类或单糖以制成与血液等渗的溶液。用于口腔或舌下施用时，该组合物可以用常规的方法制成片剂或锭剂  
20 的形式施用。

所施用的病毒数量一般是  $10^3$ - $10^{10}$  pfu 的范围，优选地为  $10^5$ - $10^8$  pfu，更优选地为  $10^6$ - $10^7$  pfu。当使用注射方法时，一般施用 1-10  $\mu$ l 在药用上可接受的合适载体或稀释剂中的病毒。

25 当多核苷酸/载体作为裸核酸施用时，施用的核酸的数量一般为 1  $\mu$ g -10 mg 的范围，优选地为 100  $\mu$ g -1 mg。

如果第一核苷酸序列（或其他治疗用的序列）在可诱导的调节序列控制之下时，仅有必要在治疗期间诱导基因表达。一旦所述的病情被治愈，就可移去诱导剂而停止 NOI 的表达。这种方法显然具有临床优势。例如，上述体系可包含施用抗生素四环素，以通过其对 tet 阻遏物/VP16 融合蛋白的影响而活化基因表达。  
30

现在本发明进一步用实施例的方式描述，这些实施例的本意是用于帮助本领域的技术人员实施本发明，而无论如何也不能认为是限制

本发明的范围。实施例可参考附图。在附图中：

图 1 概括性显示插入到 4 个不同 HIV 载体中的核酶；

图 2 概括性显示如何用 PCR 制造一个合适的 3' LTR。

图 3 显示用于 HXB2 株的野生型 HIV gag, pol 的密码子使用表(保  
藏号: K03455)。

图 4 显示命名为 gag, pol-SYNgp 的密码子优化序列的密码子使用表。

图 5 显示称为 env-mn 的野生型 HIV env 的密码子使用表。

图 6 显示命名为 SYNgp160mn 的 HIV env 密码子优化序列的密码  
子使用表。

图 7 显示用于本发明的 3 个质粒构建体。

图 8 显示用于产生逆转录病毒载体颗粒的 2 个系统的原理。

现在本发明进一步用下列实施例描述，这些实施例仅用作阐述本发明，而不是限制本发明的范围。

## 15 实施例

### 实施例 1-基因组的构建

HIV gag. pol 序列是密码子优化的(图 4 和 SEQ I. D. NO. 1)，并用大约 40 个核苷酸的重叠寡聚物合成。该序列具有三个优点。首先，它允许基于 HIV 的载体携带核酶和其他的治疗因子。其次，密码子优化由于更高水平的基因表达而产生更高的载体滴度。第三，gag. pol  
20 的表达成为 rev 非依赖性，可允许应用抗 rev 或 RRE 因子。

在 gag. pol 中的保守序列参照 HIV 序列数据库鉴定，并用于设计核酶，所述的数据库位于 Los Alamos 国家实验室(<http://hiv-web.lanl.gov/>)。由于 HIV-1 亚型之间的变异性，所设计的核酶裂  
25 解属于 B 亚型的北美洲、拉丁美洲和加勒比、欧洲、日本和澳大利亚的主要亚型。所选择的位点相互参考合成的 gagpol 序列，以保证较低可能切割密码子优化的 gagpol mRNA。设计核酶分别在 5' 和 3' 末端具有 Xho I 和 Sal I 位点。这种设计可构建分隔并串联的核酶。

该核酶是下列普通结构的锤头结构 (25)：

|    |                |                        |            |
|----|----------------|------------------------|------------|
| 30 | 螺旋 I           | 螺旋 II                  | 螺旋 III     |
|    | 5' -NNNNNNNN ~ | CUGAUGAGGCCGAAAGGCCGAA | ~ NNNNNNNN |

在不降低催化转换率的前提下，该核酶的催化结构域（螺旋 II）能忍受一些变化。

具有基本 GUX（其中 X 是任何的核苷酸碱基）三联体，靶向 gag 和 pol 的裂解位点见如下：

|       |    |                         |
|-------|----|-------------------------|
| GAG 1 | 5' | UAGUAAGAAUGUAUAGCCCUAC  |
| GAG 2 | 5' | AACCCAGAUUGUAAGACUAUUU  |
| GAG 3 | 5' | UGUUUCAAUUGUGGCAAAGAAG  |
| GAG 4 | 5' | AAAAAGGGCUGUUGGAAAUGUG  |
| POL 1 | 5' | ACGACCCCTUCGUCACAAUAAAG |
| POL 2 | 5' | GGAAUUGGAGGUUUUAUCAAG   |
| POL 3 | 5' | AUAUUUUUCAGUUCUUAGAU    |
| POL 4 | 5' | UGGAUGAUUUGUAUGUAGGAUC  |
| POL 5 | 5' | CUUUGGAUGGGUUAUGAACUCC  |
| POL 6 | 5' | CAGCUGGACUGUCA AUGACAUA |
| POL 7 | 5' | AACUUUCUAUGUAGAUGGGGCA  |
| POL 8 | 5' | AAGGCCGCCUGUUGGUGGGCAG  |
| POL 9 | 5' | UAAGACAGCAGUACAAAUGGCA  |

5 所述的核酶被插入到 4 个不同的 HIV 载体中（pH4（10）、pH6、pH4.1 或 pH6.1）（图 1）。在 pH4 和 pH6 中，核酶的转录通过一个内部 HCMV 启动子驱动（9）。根据 pH4.1 和 pH6.1，该核酶从 5' LTR 表达。在 pH4 和 pH6（pH4.1 和 pH6.1）之间的主要差异在于生产质粒中的 3' LTR。pH4 和 pH4.1 在 3' LTR 中拥有 HIV U3。pH6 10 或 pH6.1 在 3' LTR 中拥有 HCMV。HCMV 启动子取代了大部分的 U3，以高组成水平地驱动表达，而 HIV-1 U3 仅在 Tat 的存在下支持高水平表达。

用 3 个 PCR 引物通过重组 PCR 产生 HCMV/HIV-1 杂交体 3' LTR（图 2）。用 RIB1 和 RIB2 引物以及 pH4（12）作为模板，进行第一轮 PCR，15 以扩增 HIV-1 HXB2 序列 8900-9123。经第二轮 PCR 从 pH4 扩增杂交体 5' LTR，制得 HIV-1 U3 的 5' 末端和 HCMV 启动子之间的连接序列。来自第一轮 PCR 反应的 PCR 产物和 RIB3，分别用作 5' 引物和 3' 引物。

RIB1: 5' -CAGCTGCTCGAGCAGCTGAAGCTTGCATGC-3'

RIB2: 5' -GTAAGTTATGTAACGGACGATATCTTGTCTTCTT-3'

RIB3: 5' -CGCATAGTCGACGGGCCCCGCGCCACTGCTAGAGATTTTC-3'

然后用 Sph I 和 Sal I 切下该 PCR 产物，并插入到 pH4 中，因而取代了 3' LTR。所得到的质粒命名为 pH6。为了构建 pH4.1 和 pH6.1，在 pH4 和 pH6 中的内部 HCMV 启动子 (Spe I -Xho I)，用 pBluescript II KS+(Stratagene) 的多克隆位点取代 (Spe I -Xho I)。

5 所述的核酶被插入到基因组载体骨架的 Xho I 位点上。以任何构象出现的任何核酶，用相似的方式使用。

### 实施例 2-包装体系的构建

包装系统能制成各种形式。在第一形式的包装系统中，HIV gag、pol 的组分与 HIV env 编码序列共表达。在该情况下，改变了 gag、pol 和 env 的编码序列，结果它们抵抗构建到基因组中的抗 HIV 的核酶。在改变密码子的使用性而获得抵抗性的同时，可选择密码子与最高水平表达的哺乳动物基因的密码子使用类型相匹配。这极大地增加表达水平 (28, 29)，因而增加了滴度。一种密码子优化的 HIV env 编码序列已被 Haas 等 (9) 描述。在本实施例中，使用一种修饰的密码子优化的 HIV env 序列 (SEQ I. D. No. 3)。相应的 env 表达质粒被命名为 pSYNgp160mn。该修饰序列含有未被 Haas 等使用的外加基序。该外加序列取自 MN 株的 HIV env 序列，并进行了密码子优化。只要使用的密码子与冗余 tRNAs 相应 (42)，任何相似修饰的核酸序列会起相似的作用，并且引起对基因组中的核酶的抗性。

20 在一个使用优化密码子的 gag、pol 的编码序列的实施例中，合成了重叠的寡核苷酸，然后连接在一起而产生合成的编码序列。野生型的序列 (基因库编号 K03455) 和合成的 (gagpol-SYNgp) gagpol 序列分别在 SEQ I. D. Nos1 和 2 中显示，而它们的密码子使用性分别在图 3 和图 4 显示。野生型 env 编码序列的序列 (基因库编号 M17449) 在 SEQ I. D. No 3 中给出，合成的密码子优化序列的序列在 SEQ I. D. No 4 中给出，而它们的密码子使用表分别在图 5 和图 6 给出。与 env 编码序列一样，可使用任何获得对各种核酶有抗性的 gag、pol 序列。所显示的合成序列被命名为 gag, pol-SYNgp，在其 5' 末端有一个 EcoR I 位点和在 3' 末端有一个 Not I 位点。它被插入到 pCIneo

(Promega) 中而产生质粒 pSYNgp。

在第二形式的包装系统中，一种合成的 gag, pol 表达盒与一种非 HIV 的包被编码序列共表达，该非 HIV 的包被编码序列产生一种假型化 HIV 的表面蛋白。例如，该蛋白可以是 VSV-G (20, 41)，双向 MLV env (6, 34) 或任何能掺入 HIV 颗粒中的其他蛋白 (37)。这也包括能把载体靶向到特异组织的分子。非 HIV 包被蛋白的编码序列不被各种核酶裂解，因此不需要序列修饰（虽然由于其他的原因，例如在哺乳动物细胞中密码子使用性的优化，某些序列的修饰是合乎需要的）。

### 10 实施例 3-产生载体颗粒

载体颗粒既能从与 Soneoka 等 (33) 描述相似的瞬时三质粒转染系统中产生，也可从用其他逆转录病毒载体 (20, 35, 39) 相似的生产细胞系中产生。这些原理在图 7 和 8 中阐明。例如，按照 Soneoka 等 (33) 描述的方法，用 pH6Rz、pSYNgp 和 pRV67 (VSV-G 表达质粒) 的三种质粒转染 293T 细胞 (图 8)，产生了命名为 H6Rz-VSV 的载体颗粒。这些方法把 H6Rz 基因组转导到诸如 C1866 或 Jurkat 的 CD4+ 细胞中，而产生多靶核酶。至此在这些细胞中的 HIV 复制被严格限制。

在以上说明中提到的全部出版物，在此引入仅供参考。在不偏离本发明的范围和精神的前提下，本发明描述的方法和系统的各种修饰和修改方法，对本领域的技术人员来说是显而易见的。尽管结合具体优选实施方案的形式描述了本发明，但应该明白，要求保护的本发明仅限制为上述的具体实施方案是不恰当的。事实上，应该认为：对分子生物学或相关领域的技术人员来说明显的，用于执行本发明的上述模式的各种修改方法，属于下列权利要求的范围之中。

## 参考文献

1. Bahner, I., K. Kearns, Q.L. Hao, E. M. Smogorzewska, 和 D. B. Kohn. 1996. 通过表达 RRE 假目标的逆转录病毒载体转导  
5 的人 CD34+造血祖细胞, 抑制在长期培养产生的骨髓单核细胞中人免  
疫缺陷 1 型病毒的复制. 病毒学杂志 70: 4352-60.
2. Blomer, U., L. Naldini, T. Kafri, D. Trono, I. M. Verma, F. H. Gage. 1997. 用慢病毒载体在成人神经元中高效  
并维持基因转移. 病毒学杂志 71: 6641-6649.
- 10 3. Breaker, R. R., 和 Joyce, G. F. 1994. 创造和改善核  
酶功能: 相对相互作用选择方法的合理设计. TIBTECH 12: 268-75.
4. Buchschacher, G. L., Jr., A. T. Panganiban. 1992.  
人免疫缺陷病毒载体用于诱导外源基因表达. 病毒学杂志 66:  
2731-2739.
- 15 5. Chen, C. J., A. C. Banerjee, G. G. Harmison, K.  
Haglund, 和 M. Schubert. 1992. 指导在多达 9 个的高度保守 HIV-1  
env RNA 区裂解的多靶核酶抑制 HIV-1 复制的潜在效力, 以对抗大多  
数目前测序的 HIV-1 分离物. 核酸研究 20: 4581-9.
- 20 6. Chesebro, B., K. Wehrly, W. Maury. 1990. 由鼠逆  
转录病毒假型化的人免疫缺陷病毒在人和小鼠细胞中的差异表达.  
病毒学杂志 64: 4553-7.
7. Couture, L. A., 和 Stinchcomb, D. T. 1996. 抗基因治  
疗: 应用核酶抑制基因功能. TIG 12: 510-5.
8. Dropulic, B., M. Hermankova, P. M. Pitha. 1996.  
25 一种条件性复制 HIV-1 载体干扰野生型 HIV-1 复制和传播. 美国国家  
科学院院报 93: 11103-8.
9. Foecking, M. K. 和 H. Hofstetter. 1986. 用于哺乳动物  
表达载体的强力和多样性增强子启动子单位. 基因 45: 101-105.
10. Gervaix, A., X. Li, G. Kraus, 和 F. Wong Staal.  
30 1997. 多基因抗病毒载体抑制多种人免疫缺陷 1 型病毒包被. 病毒学  
杂志 71: 3048-53.
11. Haas, J., E.-C. Park, 和 B. Seed. 1996. 在 HIV-1

包被糖蛋白表达中的密码使用性的限制。现代生物学 6: 315.

12. Kim, V. N., K. Mitrophanous, S. M. Kingsman, 和 K. A. J. 1998. 基于人免疫缺陷 1 型病毒的慢病毒载体的最低要求。病毒学杂志 72: 811-816.

5 13. Larsson, S., G. Hotchkiss, J. Su, T. Kebede, M. Andang, T. Nyholm, B. Johansson, A. Sonnerborg, A. Vahine, S. Britton, L. Ahrlund Richter. 1996. 一种位于 HIV-1 nef 开读框的新颖核酶靶位点。病毒学 219:161.

10 14. Lever, A. M. 1995. 用于 HIV 感染的基因治。英国医学学报 51: 149-66.

15 15. Liu, D., J. Donegan, G. Nuovo, D. Mitra, 和 J. Laurence. 1997. 用掺入到 U1 snRNA 中的多靶 HIV-1 反义序列处理的转化 CD4+ 单核细胞中稳定的人免疫缺陷 1 型病毒(HIV-1)的抗性。病毒学杂志 71: 4079-85.

16. Malim, M. H., S. Bohnlein, J. Hauber, 和 B. R. Cullen. 1989. Rev 作用的反式显性抑制子及 HIV-1 Rev 反式活化子衍生作用的功能分析。细胞 58: 205-14.

17. Miller, N., J. Whelan. 1997. 用于基因治疗的转录靶向和可调节的载体的进展。人类基因治疗 8: 803-15.

20 18. Naldini, L., U. Blomer, F. H. Gage, D. Trono, 和 I. M. Verma. 1996. 在用慢病毒载体注射的成年大鼠脑中转基因的有效转移、整合和维持长期表达。美国国家科学院院报 93: 11382-11388.

25 19. Naldini, L., U. Blomer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, 和 D. Trono. 1996. 通过慢病毒载体进行体内基因转移和稳定转导到非分裂的细胞[见评论]。科学 272: 263-7.

20. Ory, D. S., B. A. Neugeboren, 和 R. C. Mulligan. 1996. 一种用于产生高滴度的逆转录病毒/疱疹性口炎病毒 G 假型的人来源的稳定包装细胞系。美国国家科学院院报 93: 11400-6.

30 21. Paik, S. Y., A. Banerjee, C. J. Chen, Z. Ye, G. G. Harmison, 和 M. Schubert. 1997. 编码多靶核酶的缺陷性 HIV-1 前病毒抑制剪接和非剪接 HIV-1 mRNAs 的积累,降低病毒子代的感染性以

及保护细胞免受致病。人类基因治疗 8: 1115-24.

22. Poeschla, E., P. Corbeau, 和 F. Wong Staal. 1996. 用于抗 HIV 基因治疗的 HIV 载体的开发。美国国家科学院院报 93: 11395-9.

5 23. Poznansky, M., A. Lever, L. Bergeron, W. Haseltine, 和 J. Sodroski. 1991. 通过一种缺陷性人免疫缺陷 1 型病毒载体把基因转移到人淋巴细胞中。病毒学杂志 65: 532-6.

24. Ramezani, A, 和 S. Joshi. 1996. 在 HIV-1 RNA 中 5 个高度保守的靶位点对体外和体内锤头核酶介导的裂解敏感性的比较分析。反义核酸药物开发 6: 229-35.

25. Riddell, S. R., M. Elliott, D. A. Lewinsohn, M. J. Gilbert, L. Wilson, S. A. Manley, S. D. Lupton, R. W. Overell, T. C., Reynolds, L. Corey, 和 P. D. Greenberg. 1996. 在 HIV 感染的病人中基因修饰的 HIV 特异的细胞毒 T 淋巴细胞的 T 细胞介导的排斥作用 (见评论)。自然医学 2: 216-23.

26. Ruffner, D. E., S. C. Dahm, 和 O. C. Uhlenbeck. 1989. 锤头 RNA 自裂解结构域的研究。基因 82: 31-41.

27. Sarver, N., E. M. Cantin, P. S. Chang, J. A. Zaia, P. A. Ladne, D. A. Stephens, 和 J. J. Rossi. 1990. 作为潜在的抗 HIV-1 治疗剂的核酶。科学 247:1222.

28. Schneider, R., M. Campbell, G. Nasioulas, B. K. Felber, 和 G. N. Pavlakis. 1997. 人免疫缺陷 1 型病毒抑制元件的失活允许 Rev 非依赖的 Gag 和 Gag/蛋白酶的表达和颗粒形成。病毒学杂志 71: 4892-903.

29. Schwartz, S., M. Campbell, G. Nasioulas, J. Harrison, B. K. Felber, 和 G. N. Pavlakis. 1992. 在人免疫缺陷 1 型病毒中抑制性序列的突变失活导致 Rev 非依赖的 gag 基因的表达。病毒学杂志 66: 7176-82.

30. Scott, W. G., 和 Klug, A. 1996. 核酶: 在 RNA 催化中的结构和机制。TIBS. 21: 220-4.

31. Sczakiel, G., 和 M. Pawlita. 1991. 在人 T 细胞中稳定表达反义 RNA 抑制人免疫缺陷 1 型病毒的复制。病毒学杂志 65: 468-

72.

32. Shimada, T., H. Fujii, H. Mitsuya, A. W. Nienhuis. 1991. 通过重组人免疫缺陷病毒逆转录病毒载体靶向和高效率地把基因转移到 CD4<sup>+</sup> 细胞中。临床研究杂志 88: 1043-47.

5 33. Soneoka, Y., P. M. Cannon, E. E. Ramsdale, J. C. Griffiths, G. Romano, S. M. Kingsman, 和 A. J. Kingsman. 1995. 一种用于产生高滴度逆转录病毒载体的瞬时三质粒表达系统。核酸研究 23: 628-33.

34. Spector, D. H., E. Wade, D. A. Wright, V. Koval, C. Clark, D. Jaquish, S. A. Spector. 1990. 具有扩展的细胞和物种嗜性的人免疫缺陷病毒假型。病毒学杂志 64: 2298-2308.

35. Srinivasakumar, N., N. Chazal, C. Helga Maria, S. Prasad, M. L. Hammarskjold, D. Rekosh. 1997. 病毒调节蛋白的表达对由在稳定包装细胞系中产生的人免疫缺陷 1 型病毒载体的基因转移的影响。病毒学杂志 71: 5841-8.

36. Sun, L. Q., L. Wang, W. L. Gerlach, 和 G. Symonds. 1995. 通过针对 tat RNA 的核酶靶序列特异地抑制 HIV-1 的复制。核酸研究 23: 2909-13.

37. Valsesia Wittmann, S., A. Drynda, G. Deleage, M. Aumailley, J. M. Heard, O. Danos, G. Verdier, 和 F. L. Cosset. 1994. 在鸟类逆转录病毒包被蛋白的结合结构域的修饰重新指导逆转录病毒载体的宿主范围。病毒学杂志 68: 4609-19.

38. Yamada, O., G. Kraus, M. C. Leavitt, M. Yu, 和 F. Wong Staal. 1994. 在人 T 细胞中抗 HIV-1 发卡核酶的活性和裂解位点的特异性。病毒学 205: 121-6.

39. Yu, H., A. B. Rabson, M. Kaul, Y. Ron, 和 J. P. Dougherty. 1996. 可诱导的人免疫缺陷 1 型病毒包装细胞系。病毒学杂志 70: 4530-37.

40. Zhou, C., I. Bahner, J. J. Rossi, 和 D. B. Kohn. 1996. 通过逆转录病毒载体表达的锤头核酶抑制 HIV-1 的复制: RNA 水平和病毒抑制的比较。反义核酸药物开发 6: 17-24.

41. Zhu, Z. H., S. S. Chen, 和 A. S. Huang. 1990. 混杂

有人免疫缺陷病毒和疱疹性口炎病毒或单纯疱疹病毒的表型。 获得性免疫缺陷综合症杂志 3: 215-9.

42. Zolotukhin, S., M. Potter, W. W. Hauswirth, J. Guy, 和 N. Muzyczka. 1996. 一种适于在哺乳动物细胞中高水平表达的“人源化”绿色荧光蛋白 cDNA. 病毒学杂志 70: 4646-54.
43. Zufferey, R., D. Nagy, R. J. Mandel. L. Maldini. 和 D. Trono. 1997. 多点减毒的慢病毒载体获得在体内的有效基因转移. 自然生物技术 15: 871-875.
44. Goodchild, J., V. Kohli. 1991. 裂解来自人免疫缺陷病毒的 RNA 序列的核酶: 侧翼序列对裂解率的影响. 生物物理生物化学学报 2 月 1 日 284(2): 386-391.
45. Hertel, Klemens J., Alessio Peracchi, Olke C. Uhlenbeck, 和 Daniel Herschlag. 1997. RNA 酶利用内部结合能催化. 美国国家科学院院报 94, 8497-8502, 8 月.
46. Bender 等, 1987, 病毒学杂志 61: 1639-1646
47. Pear 等, 1993, 美国国家科学院院报 90: 8392-8396
48. Cosset 等, 1995, 病毒学杂志 69: 7430-7436

## 说明书的序列表部分

## SEQ.ID.NO.1-HXB2株的野生型gagpol序列(登记号K03455)

ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG GGAGAATTAG ATCGATGGGA AAAAATTCGG 60  
 TTAAGGCCAG GGGGAAAGAA AAAATATAAA TTA AACATA TAGTATGGGC AAGCAGGGAG 120  
 CTAGAACGAT TCGCAGTTAA TCCTGGCCTG TTAGAACAT CAGAAGGCTG TAGACAAATA 180  
 CTGGGACAGC TACAACCATC CCTTCAGACA GGATCAGAAG AACTTAGATC ATTATATAAT 240  
 ACAGTAGCAA CCCTCTATTG TGTGCATCAA AGGATAGAGA TAAAAGACAC CAAGGAAGCT 300  
 TTAGACAAGA TAGAGGAAGA GCAAAACAAA AGTAAGAAAA AAGCACAGCA AGCAGCAGCT 360  
 GACACAGGAC ACAGCAATCA GGTCAGCCAA AATTACCCTA TAGTGCAGAA CATCCAGGGG 420  
 CAAATGGTAC ATCAGGCCAT ATCACCTAGA ACTTTAAATG CATGGGTAA AGTAGTAGAA 480  
 GAGAAGGCTT TCAGCCCAGA AGTGATACCC ATGTTTTTCAG CATTATCAGA AGGAGCCACC 540  
 CCACAAGATT TAAACACCAT GCTAAACACA GTGGGGGGAC ATCAAGCAGC CATGCAAATG 600  
 TTAAGAGAGA CCATCAATGA GGAAGCTGCA GAATGGGATA GAGTGCATCC AGTGCATGCA 660  
 GGGCCTATTG CACCAGGCCA GATGAGAGAA CCAAGGGGAA GTGACATAGC AGGAACTACT 720  
 AGTACCCTTC AGGAACAAAT AGGATGGATG ACAAATAATC CACCTATCCC AGTAGGAGAA 780  
 ATTTATAAAA GATGGATAAT CCTGGGATTA AATAAAATAG TAAGAATGTA TAGCCCTACC 840  
 AGCATTCTGG ACATAAGACA AGGACCAAAG GAACCCTTTA GAGACTATGT AGACCGGTTT 900  
 TATAAAACTC TAAGAGCCGA GCAAGCTTCA CAGGAGGTAA AAAATTGGAT GACAGAAACC 960  
 TTGTTGGTCC AAAATGCGAA CCCAGATTGT AAGACTATTT TAAAAGCATT GGGACCAGCG 1020  
 GCTACACTAG AAGAAATGAT GACAGCATGT CAGGGAGTAG GAGGACCCGG CCATAAGGCA 1080  
 AGAGTTTTGG CTGAAGCAAT GAGCCAAGTA ACAAATTCAG CTACCATAAT GATGCAGAGA 1140  
 GGCAATTTTA GGAACCAAAG AAAGATTGTT AAGTGTTCAT ATTGTGGCAA AGAAGGGCAC 1200  
 ACAGCCAGAA ATTGCAGGGC CCCTAGGAAA AAGGGCTGTT GGAAATGTGG AAAGGAAGGA 1260  
 CACCAAATGA AAGATTGTAC TGAGAGACAG GCTAATTTTT TAGGGAAGAT CTGGCCTTCC 1320  
 TACAAGGGAA GGCCAGGGAA TTTTCTTCAG AGCAGACCAG AGCCAACAGC CCCACCAGAA 1380  
 GAGAGCTTCA GGTCTGGGGT AGAGACAACA ACTCCCCCTC AGAAGCAGGA GCCGATAGAC 1440  
 AAGGAACTGT ATCCTTTAAC TTCCCTCAGG TCACTCTTTG GCAACGACCC CTCGTCACAA 1500  
 TAAAGATAGG GGGGCAACTA AAGGAAGCTC TATTAGATAC AGGAGCAGAT GATACAGTAT 1560  
 TAGAAGAAAT GAGTTTGCCA GGAAGATGGA AACCAAAAAT GATAGGGGGA ATTGGAGGTT 1620  
 TTATCAAAGT AAGACAGTAT GATCAGATAC TCATAGAAAT CTGTGGACAT AAAGCTATAG 1680  
 GTACAGTATT AGTAGGACCT ACACCTGTCA ACATAATTGG AAGAAATCTG TTGACTCAGA 1740  
 TTGGTTGCAC TTTAAATTTT CCCATTAGCC CTATTGAGAC TGTACCAGTA AAATTAAAGC 1800  
 CAGGAATGGA TGGCCCAAAA GTTAAACAAT GGCCATTGAC AGAAGAAAAA ATAAAGCAT 1860  
 TAGTAGAAAT TTGTACAGAG ATGGAAAAGG AAGGGAAAAT TTCAAAAAT GGCCTGAAA 1920  
 ATCCATACAA TACTCCAGTA TTTGCCATAA AGAAAAAAGA CAGTACTAAA TGGAGAAAAT 1980  
 TAGTAGATTT CAGAGAACTT AATAAGAGAA CTCAAGACTT CTGGGAAGTT CAATTAGGAA 2040  
 TACCACATCC CGCAGGGTTA AAAAAGAAAA AATCAGTAAC AGTACTGGAT GTGGGTGATG 2100  
 CATATTTTTT AGTTCCCTTA GATGAAGACT TCAGGAAGTA TACTGCATTT ACCATACCTA 2160  
 GTATAAACAA TGAGACACCA GGGATTAGAT ATCAGTACAA TGTGCTTCCA CAGGGATGGA 2220  
 AAGGATCACC AGCAATATTC CAAAGTAGCA TGACAAAAAT CTTAGAGCCT TTAGAAAAAC 2280  
 AAAATCCAGA CATAGTTATC TATCAATACA TGGATGATTT GTATGTAGGA TCTGACTTAG 2340  
 AAATAGGGCA GCATAGAACA AAAATAGAGG AGCTGAGACA ACATCTGTTG AGGTGGGGAC 2400  
 TTACCACACC AGACAAAAAA CATCAGAAAG AACCTCCATT CTTTGGATG GGTTATGAAC 2460  
 TCCATCCTGA TAAATGGACA GTACAGCCTA TAGTGCTGCC AGAAAAAGAC AGCTGGACTG 2520  
 TCAATGACAT ACAGAAGTTA GTGGGGAAAT TGAATTGGGC AAGTCAGATT TACCCAGGGA 2580

```

TTAAAGTAAG GCAATTATGT AAATCCTTA GAGGAACCAA AGCACTAACA GAAGTAATAC 2640
CACTAACAGA AGAAGCAGAG CTAGAACTGG CAGAAAACAG AGAGATTCTA AAAGAACCAG 2700
TACATGGAGT GTATTATGAC CCATCAAAAG ACTTAATAGC AGAAATACAG AAGCAGGGGC 2760
AAGGCCAATG GACATATCAA ATTTATCAAG AGCCATTTAA AAATCTGAAA ACAGGAAAAAT 2820
ATGCAAGAAT GAGGGGTGCC CACACTAATG ATGTAAAACA ATTAACAGAG GCAGTGCAA 2880
AAATAACCAC AGAAAGCATA GTAATATGGG GAAAGACTCC TAAATTTAAA CTGCCCATAC 2940
AAAAGGAAAC ATGGGAAACA TGGTGGACAG AGTATTGGCA AGCCACCTGG ATCCTGAGT 3000
GGGAGTTTGT TAATACCCCT CCCTTAGTGA AATTATGGTA CCAGTTAGAG AAAGAACCCA 3060
TAGTAGGAGC AGAAACCTTC TATGTAGATG GGGCAGCTAA CAGGGAGACT AAATTAGGAA 3120
AAGCAGGATA TGTTACTAAT AGAGGAAGAC AAAAAGTTGT CACCCTAACT GACACAACAA 3180
ATCAGAAGAC TGAGTTACAA GCAATTTATC TAGCTTTGCA GGATTCGGGA TTAGAAGTAA 3240
ACATAGTAAC AGACTCACAA TATGCATTAG GAATCATTCA AGCACAACCA GATCAAAGTG 3300
AATCAGAGTT AGTCAATCAA ATAATAGAGC AGTTAATAAA AAAGGAAAAG GTCTATCTGG 3360
CATGGGTACC AGCACACAAA GGAATTGGAG GAAATGAACA AGTAGATAAA TTAGTCAGTG 3420
CTGGAATCAG GAAAGTACTA TTTTATAGTG GAATAGATAA GGCCCAAGAT GAACATGAGA 3480
AATATCACAG TAATTGGAGA GCAATGGCTA GTGATTTTAA CCTGCCACCT GTAGTAGCAA 3540
AAGAAATAGT AGCCAGCTGT GATAAATGTC AGCTAAAAGG AGAAGCCATG CATGGACAAG 3600
TAGACTGTAG TCCAGGAATA TGGCAACTAG ATTGTACACA TTTAGAAGGA AAAGTTATCC 3660
TGGTAGCAGT TCATGTAGCC AGTGGATATA TAGAAGCAGA AGTTATTCCA GCAGAAACAG 3720
GGCAGGAAAC AGCATATTTT CTTTTAAAT TAGCAGGAAG ATGGCCAGTA AAAACAATAC 3780
ATACTGACAA TGGCAGCAAT TTCACCGGTG CTACGGTTAG GGCCGCTGT TGGTGGCGG 3840
GAATCAAGCA GGAATTTGGA ATTCCCTACA ATCCCCAAAG TCAAGGAGTA GTAGAATCTA 3900
TGAATAAGA ATTAAGAAA ATTATAGGAC AGGTAAGAGA TCAGGCTGAA CATCTTAAGA 3960
CAGCAGTACA AATGGCAGTA TTCATCCACA ATTTAAAAG AAAAGGGGGG ATTGGGGGGT 4020
ACAGTGCAGG GGAAAGAATA GTAGACATAA TAGCAACAGA CATACAAAT AAAGAATTAC 4080
AAAAACAAAT TACAAAAATT CAAAATTTTC GGGTTTATTA CAGGGACAGC AGAAATTCAC 4140
TTTGGAAAGG ACCAGCAAAG CTCCTCTGGA AAGGTGAAGG GGCAGTAGTA ATACAAGATA 4200
ATAGTGACAT AAAAGTAGTG CCAAGAAGAA AAGCAAAGAT CATTAGGGAT TATGGAAAC 4260
AGATGGCAGG TGATGATTGT GTGGCAAGTA GACAGGATGA GGATTAG 4307

```

# SEQ.ID.NO.2-gagpol-SYNgp-密码子优化的gagpol序列

```

ATGGGCGCCC GCGCCAGCGT GCTGTGCGGC GGCGAGCTGG ACCGCTGGGA GAAGATCCGC 60
CTGCGCCCCG GCGGCAAAAA GAAGTACAAG CTGAAGCACA TCGTGTGGGC CAGCCGCGAA 120
CTGGAGCGCT TCGCCGTGAA CCCCAGGCTC CTGGAGACCA GCGAGGGGTG CCGCCAGATC 180
CTCGECCAAC TGCAGCCCAG CCTGCAAACC GGCAGCGAGG AGCTGCGCAG CCTGTACAAC 240
ACCGTGGCCA CGCTGTACTG CGTCCACCAG CGCATCGAAA TCAAGGATAC GAAAGAGGCC 300
CTGGATAAAA TCGAAGAGGA ACAGAATAAG AGCAAAAAGA AGGCCAACA GGCCGCCGCG 360
GACACCGGAC ACAGCAACCA GGTGAGCCAG AACTACCCA TCGTGCAGAA CATCCAGGGG 420
CAGATGGTGC ACCAGGCCAT CTCCCCCGC ACGCTGAACG CCTGGGTGAA GGTGGTGGAA 480
GAGAAGGCTT TTAGCCCGGA GGTGATACCC ATGTTCTCAG CCCTGTCAGA GGGAGCCACC 540
CCCCAAGATC TGAACACCAT GCTCAACACA GTGGGGGGAC ACCAGCCGC CATGCAGATG 600
CTGAAGGAGA CCATCAATGA GGAGGCTGCC GAATGGGATC GTGTGCATCC GGTGCACGCA 660
GGGCCCATCG CACCGGGCCA GATGCGTGAG CCACGGGGCT CAGACATCGC CGGAACGACT 720
AGTACCCTTC AGGAACAGAT CGGCTGGATG ACCAACAACC CACCCATCCC GGTGGGAGAA 780
ATCTACAAAC GCTGGATCAT CCTGGGCTG AACAAGATCG TGCGCATGTA TAGCCCTACC 840
AGCATCCTGG ACATCCGCCA AGGCCGAAG GAACCTTTC GCGACTACGT GGACCGGTTC 900

```

TACAAAACGC TCCGCGCCGA GCAGGCTAGC CAGGAGGTGA AGAACTGGAT GACCGAAACC 960  
CTGCTGGTCC AGAACGCGAA CCCGGACTGC AAGACGATCC TGAAGGCCCT GGGCCCAGCG 1020  
GCTACCCTAG AGGAAATGAT GACCGCTGT CAGGGAGTGG GCGGACCCGG CCACAAGGCA 1080  
CGCGTCCTGG CTGAGGCCAT GAGCCAGGTG ACCAACTCCG CTACCATCAT GATGCAGCGC 1140  
GGCAACTTTC GGAACCAACG CAAGATCGTC AAGTGCTTCA ACTGTGGCAA AGAAGGGCAC 1200  
ACAGCCCGCA ACTGCAGGGC CCCTAGGAAA AAGGGCTGCT GGAAATGCGG CAAGGAAGGC 1260  
CACCAGATGA AAGACTGTAC TGAGAGACAG GCTAATTTTT TAGGGAAGAT CTGGCCTTCC 1320  
TACAAGGGAA GGCCAGGGAA TTTTCTTCAG AGCAGACCAG AGCCAACAGC CCCACCAGAA 1380  
GAGAGCTTCA GGTCTGGGGT AGAGACAACA ACTCCCCCTC AGAAGCAGGA GCCGATAGAC 1440  
AAGGAACTGT ATCCTTTAAC TTCCCTCAGA TCACTCTTTG GCAACGACCC CTCGTCACAA 1500  
TAAAGATAGG GGGGCAGCTC AAGGAGGCTC TCCTGGACAC CGGAGCAGAC GACACCGTGC 1560  
TGGAGGAGAT GTCGTTGCCA GGCCGCTGGA AGCCGAAGAT GATCGGGGGA ATCGGCGGTT 1620  
TCATCAAGGT GCGCCAGTAT GACCAGATCC TCATCGAAAT CTGCGGCCAC AAGGCTATCG 1680  
GTACCGTGCT GGTGGGCCCC ACACCCGTCA ACATCATCGG ACGCAACCTG TTGACGCAGA 1740  
TCGGTTGCAC GCTGAACTTC CCCATTAGCC CTATCGAGAC GGTACCGGTG AAGCTGAAGC 1800  
CCGGGATGGA CGGCCCCAAG GTCAAGCAAT GGCCATTGAC AGAGGAGAAG ATCAAGGCAC 1860  
TGGTGGAGAT TTGCACAGAG ATGGAAAAGG AAGGGAAAAT CTCCAAGATT GGGCCTGAGA 1920  
ACCCGTACAA CAGGCCGGTG TTCGCAATCA AGAAGAAGGA CTCGACGAAA TGGCGCAAGC 1980  
TGGTGGACTT CCGCGAGCTG AACAAAGCGCA CGCAAGACTT CTGGGAGGTT CAGCTGGGCA 2040  
TCCCGCACCC CGCAGGGCTG AAGAAGAAGA AATCCGTGAC CGTACTGGAT GTGGGTGATG 2100  
CCTACTTCTC CGTTCCCCTG GACGAAGACT TCAGGAAGTA CACTGCCTTC ACAATCCCTT 2160  
CGATCAACAA CGAGACACCG GGGATTGAT ATCAGTACAA CGTGCTGCCC CAGGGCTGGA 2220  
AAGGCTCTCC CGCAATCTTC CAGAGTAGCA TGACCAAAAT CCTGGAGCCT TTCCGCAAAC 2280  
AGAACCCCGA CATCGTCATC TATCAGTACA TGGATGACTT GTACGTGGGC TCTGATCTAG 2340  
AGATAGGGCA GCACCGCACC AAGATCGAGG AGCTGCGCCA GCACCTGTTG AGGTGGGGAC 2400  
TGACCACACC CGACAAGAAG CACCAGAAGG AGCCTCCCTT CCTCTGGATG GGTTACGAGC 2460  
TGCACCCTGA CAAATGGACC GTGCAGCCTA TCGTGCTGCC AGAGAAAGAC AGCTGGACTG 2520  
TCAACGACAT ACAGAAGCTG GTGGGGAAGT TGAAGTGGG CAGTCAGATT TACCCAGGGA 2580  
TTAAGGTGAG GCAGCTGTGC AAACCTCTCC GCGGAACCAA GGCACCTACA GAGGTGATCC 2640  
CCCTAACCGA GGAGGCCGAG CTCGAACTGG CAGAAAACCG AGAGATCCTA AAGGAGCCCG 2700  
TGCACGGCGT GTACTATGAC CCCTCCAAGG ACCTGATCGC CGAGATCCAG AAGCAGGGGC 2760  
AAGGCCAGTG GACCTATCAG ATTTACCAGG AGCCCTTCAA GAACCTGAAG ACCGGCAAGT 2820  
ACGCCCGGAT GAGGGGTGCC CACACTAACG ACGTCAAGCA GCTGACCGAG GCCGTGCAGA 2880  
AGATCACCAC CGAAAGCATC GTGATCTGGG GAAAGACTCC TAAGTTCAAG CTGCCATCC 2940  
AGAAGGAAAC CTGGGAAACC TGGTGGACAG AGTATTGSCA GGCCACCTGG ATTCTGAGT 3000  
GGGAGTTCGT CAACACCCCT CCCCTGGTGA AGCTGTGGTA CCAGCTGGAG AAGGAGCCCA 3060  
TAGTGGGCGC CGAAACCTTC TACGTGGATG GGGCCGCTAA CAGGGAGACT AAGCTGGGCA 3120  
AAGCCGGATA CGTACTAAC CGGGGCAGAC AGAAGGTTGT CACCCTCACT GACACCACCA 3180  
ACCAGAAGAC TGAGCTGCAG GCCATTTACC TCGCTTTGCA GGAAGTGGG CTGGAGGTGA 3240  
ACATCGTGAC AGACTCTCAG TATGCCCTGG GCATCATTCA AGCCAGCCA GACCAGAGTG 3300  
AGTCCGAGCT GGTCAATCAG ATCATCGAGC AGCTGATCAA GAAGGAAAAG GTCTATCTGG 3360  
CCTGGGTACC CGCCACAAA GGCATTGGCG GCAATGAGCA GGTGACAAAG CTGGTCTCGG 3420  
CTGGCATCAG GAAGGTGCTA TTCCTGGATG GCATCGACAA GGCCAGGAC GAGCAGGAGA 3480  
AATACCACAG CAACTGGCGG GCCATGGCTA GCGACTTCAA CCTGCCCCCT GTGGTGGCCA 3540  
AAGAGATCGT GGCCAGCTGT GACAAGTGTG AGCTCAAGGG CGAAGCCATG CATGGCCAGG 3600  
TGGACTGTAG CCCCGGCATC TGGCAACTCG ATTGCACCCA TCTGGAGGGC AAGGTTATCC 3660  
TGGTAGCCGT CCATGTGGCC AGTGGCTACA TCGAGGCCGA GGTCATTCCC GCCGAAACAG 3720  
GGCAGGAGAC AGCCTACTTC CTCCTGAAGC TGGCAGGCCG GTGGCCAGTG AAGACCATCC 3780

ATACTGACAA TGGCAGCAAT TTCACCAGTG CTACGGTTAA GGCCGCCTGC TGGTGGGCGG 3840  
 GAATCAAGCA GGAGTTCGGG ATCCCTACA ATCCCAGAG TCAGGGCGTC GTCGAGTCTA 3900  
 TGAATAAGGA GTTAAAGAAG ATTATCGGCC AGGTCAGAGA TCAGGCTGAG CATCTCAAGA 3960  
 CCGCGGTCCA AATGGCGGTA TTCATCCACA ATTTCAAGCG GAAGGGGGG ATTGGGGGGT 4020  
 ACAGTGC GGGG GGAGCGGATC GTGGACATCA TCGCGACCGA CATCCAGACT AAGGAGCTGC 4080  
 AAAAGCAGAT TACCAAGATT CAGAATTTCC GGGTCTACTA CAGGGACAGC AGAAATCCCC 4140  
 TCTGGAAAGG CCCAGCGAAG CTCCTCTGGA AGGGTGAGGG GGCAGTAGTG ATCCAGGATA 4200  
 ATAGCGACAT CAAGGTGGTG CCCAGAAGAA AGGCGAAGAT CATTAGGGAT TATGGCAAAC 4260  
 AGATGGCGGG TGATGATTGC GTGGCGAGCA GACAGGATGA GGATTAG 4307

SEQ.ID.NO.3-来自HIV-1 MN的包被基因 (基因库登记号M17449)

ATGAGAGTGA AGGGGATCAG GAGGAATTAT CAGCACTGGT GGGGATGGGG CACGATGCTC 60  
 CTTGGGTTAT TAATGATCTG TAGTGCTACA GAAAAATTGT GGGTCACAGT CTATTATGGG 120  
 GTACCTGTGT GGAAAGAAGC AACCACCACT CTATTTTGTG CATCAGATGC TAAAGCATAT 180  
 GATACAGAGG TACATAATGT TTGGGCCACA CAAGCCTGTG TACCCACAGA CCCCAACCCA 240  
 CAAGAAGTAG AATTGGTAAA TGTGACAGAA AATTTTAACA TGTGGAAAAA TAACATGGTA 300  
 GAACAGATGC ATGAGGATAT AATCAGTTTA TGGGATCAAA GCCTAAAGCC ATGTGTAAAA 360  
 TTAACCCAC TCTGTGTTAC TTAAATTGC ACTGATTGA GGAATACTAC TAATACCAAT 420  
 AATAGTACTG CTAATAACAA TAGTAATAGC GAGGGAACAA TAAAGGGAGG AGAAATGAAA 480  
 AACTGCTCTT TCAATATCAC CACAAGCATA AGAGATAAGA TGCAGAAAAG ATATGCACTT 540  
 CTTTATAAAC TTGATATAGT ATCAATAGAT AATGATAGTA CCAGCTATAG GTTGATAAGT 600  
 TGTAATACCT CAGTCATTAC ACAAGCTTGT CCAAAGATAT CCTTTGAGCC AATCCCATA 660  
 CACTATTGTG CCGCGGCTGG TTTTGCGATT CTAATATGTA ACGATAAAAA GTTCAGTGGA 720  
 AAAGGATCAT GTAAAAATGT CAGCACAGTA CAATGTACAC ATGGAATTAG GCCAGTAGTA 780  
 TCAACTCAAC TGCTGTTAAA TGGCAGTCTA GCAGAAGAAG AGGTAGTAAT TAGATCTGAG 840  
 AATTTCACTG ATAATGCTAA AACCATCATA GTACATCTGA ATGAATCTGT ACAAATTAAT 900  
 TGTACAAGAC CCAACTACAA TAAAAGAAAA AGGATACATA TAGGACCAGG GAGAGCATTT 960  
 TATACAACAA AAAATATAAT AGGAACTATA AGACAAGCAC ATTGTAACAT TAGTAGAGCA 1020  
 AAATGGAATG ACACTTTAAG ACAGATAGTT AGCAAATTAA AAGAACAATT TAAGAATAAA 1080  
 ACAATAGTCT TTAATCAATC CTCAGGAGGG GACCCAGAAA TTGTAATGCA CAGTTTTAAT 1140  
 TGTGGAGGGG AATTTTTCTA CTGTAATACA TCACCACTGT TTAATAGTAC TTGGAATGGT 1200  
 AATAACTTT GGAATAATAC TACAGGGTCA AATAACAATA TCACACTTCA ATGCAAAATA 1260  
 AAACAAATTA TAAACATGTG GCAGGAAGTA GGAAAAGCAA TGTATGCCCC TCCCATTGAA 1320  
 GGACAAATTA GATGTTTCATC AAATATTACA GGGCTACTAT TAACAAGAGA TGGTGGTAAG 1380  
 GACACGGACA CGAACGACAC CGAGATCTTC AGACCTGGAG GAGGAGATAT GAGGGACAAT 1440  
 TGGAGAAGTG AATTATATAA ATATAAAGTA GTAACAATTG AACCATTAGG AGTAGCACCC 1500  
 ACCAAGGCAA AGAGAAGAGT GGTGCAGAGA GAAAAAAGAG CAGCGATAGG AGCTCTGTTC 1560  
 CTTGGGTTCT TAGGAGCAGC AGGAAGCACT ATGGGCGCAG CGTCAGTGAC GCTGACGGTA 1620  
 CAGGCCAGAC TATTATTGTC TGGTATAGTG CAACAGCAGA ACAATTTGCT GAGGGCCATT 1680  
 GAGGCGCAAC AGCATATGTT GCAACTCACA GTCTGGGGCA TCAAGCAGCT CCAGGCAAGA 1740  
 GTCCTGGCTG TGGAAAGATA CCTAAAGGAT CAACAGCTCC TGGGGTTTTG GGGTTGCTCT 1800  
 GGAAACTCA TTTGCACCAC TACTGTGCCT TGGAATGCTA GTTGGAGTAA TAAATCTCTG 1860  
 GATGATATTT GGAATAACAT GACCTGGATG CAGTGGGAAA GAGAAATTGA CAATTACACA 1920  
 AGCTTAATAT ACTCATTACT AGAAAAATCG CAAACCCAAC AAGAAAAGAA TGAACAAGAA 1980  
 TTATTGGAAT TGGATAAATG GGCAAGTTTG TGGAATTGGT TTGACATAAC AAATTGGCTG 2040  
 TGGTATATAA AAATATTCAT AATGATAGTA GGAGGCTTGG TAGGTTTAAG AATAGTTTTT 2100

GCTGTACTTT CTATAGTGAA TAGAGTTAGG CAGGGATACT CACCATTGTC GTTGCAGACC 2160  
 CGCCCCCAG TTCCGAGGGG ACCCGACAGG CCCGAAGGAA TCGAAGAAGA AGGTGGAGAG 2220  
 AGAGACAGAG ACACATCCGG TCGATTAGTG CATGGATTCT TAGCAATTAT CTGGGTGAC 2280  
 CTGCGGAGCC TGTCCTCTT CAGCTACCAC CACAGAGACT TACTCTTGAT TGCAGCGAGG 2340  
 ATTGTGGAAC TTCTGGGACG CAGGGGGTGG GAAGTCCTCA AATATTGGTG GAATCTCCTA 2400  
 CAGTATTGGA GTCAGGAAGT AAAGAGTAGT GCTGTTAGCT TGCTTAATGC CACAGCTATA 2460  
 GCAGTAGCTG AGGGGACAGA TAGGGTTATA GAAGTACTGC AAAGAGCTGG TAGAGCTATT 2520  
 CTCCACATAC CTACAAGAAT AAGACAGGGC TTGGAAAGGG CTTTGCTATA A 2571

#### SE Q.ID.NO.4-SYNgp-160mn-密码子优化的env序列

ATGAGGGTGA AGGGGATCCG CCGCAACTAC CAGCACTGGT GGGGCTGGGG CACGATGCTC 60  
 CTGGGGCTGC TGATGATCTG CAGCGCCACC GAGAAGCTGT GGGTGACCGT GTACTACGGC 120  
 GTGCGCGTGT GGAAGGAGGC CACCACCACC CTGTTCTGCG CCAGCGACGC CAAGGCGTAC 180  
 GACACCGAGG TGCACAACGT GTGGGCCACC CAGGCGTGCG TGCCCACCGA CCCCAACCCC 240  
 CAGGAGGTGG AGCTCGTGAA CGTGACCGAG AACTTCAACA TGTGAAGAA CAACATGGTG 300  
 GAGCAGATGC ATGAGGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG 360  
 CTGACCCCCC TGTGCGTGAC CCTGAACTGC ACCGACCTGA GGAACACCAC CAACACCAAC 420  
 AACAGCACCG CCAACAACAA CAGCAACAGC GAGGGCACCA TCAAGGGCGG CGAGATGAAG 480  
 AACTGCAGCT TCAACATCAC CACCAGCATC CGCGACAAGA TGCAGAAGGA GTACGCCCTG 540  
 CTGTACAAGC TGGATATCGT GAGCATCGAC AACGACAGCA CCAGCTACCG CCTGATCTCC 600  
 TGCAACACCA GCGTGATCAC CCAGGCCTGC CCCAAGATCA GCTTCGAGCC CATCCCCATC 660  
 CACTACTGCG CCCCCGCCGG CTTCGCCATC CTGAAGTGCA ACGACAAGAA GTTCAGCGGC 720  
 AAGGGCAGCT GCAAGAACGT GAGCACCCTG CAGTGCACCC ACGGCATCCG GCCGGTGGTG 780  
 AGCACCCAGC TCCTGCTGAA CGGCAGCCTG GCCGAGGAGG AGGTGGTGAT CCGCAGCGAG 840  
 AACTTCACCG ACAACGCCAA GACCATCATC GTGCACCTGA ATGAGAGCGT GCAGATCAAC 900  
 TGCACGCGTC CCAACTACAA CAAGCGCAAG CGCATCCACA TCGGCCCCGG GCGCGCCTTC 960  
 TACACCACCA AGAACATCAT CGGCACCATC CGCCAGGCC ACTGCAACAT CTCTAGAGCC 1020  
 AAGTGAACG ACACCCTGCG CCAGATCGTG AGCAAGCTGA AGGAGCAGTT CAAGAACAAG 1080  
 ACCATCGTGT TCAACCAGAG CAGCGGCGGC GACCCCGAGA TCGTGATGCA CAGCTTCAAC 1140  
 TGCGGCGGGC AATTCTTCTA CTGCAACACC AGCCCCCTGT TCAACAGCAC CTGGAACGGC 1200  
 AACAACACCT GGAACAACAC CACCGGCAGC AACAACAATA TTACCCTCCA GTGCAAGATC 1260  
 AAGCAGATCA TCAACATGTG GCAGGAGGTG GGCAAGGCCA TGTACGCCCC CCCCATCGAG 1320  
 GGCCAGATCC GGTGCAGCAG CAACATCACC GGTCTGCTGC TGACCCGCGA CGGCGGCAAG 1380  
 GACACCGACA CCAACGACAC CGAAATCTTC CGCCCCGGCG GCGGCGACAT GCGCGACAAC 1440  
 TGGAGATCTG AGCTGTACAA GTACAAGGTG GTGACGATCG AGCCCCCTGGG CGTGCCCCC 1500  
 ACCAAGGCCA AGCGCCGCGT GGTGCAGCGC GAGAAGCGGG CCGCCATCGG CGCCCTGTTT 1560  
 CTGGGCTTCC TGGGGGCGGC GGGCAGCACC ATGGGGGCGG CCAGCGTGAC CCTGACCGTG 1620  
 CAGGCCCCGC TGCTCTGAG CGGCATCGTG CAGCAGCAGA ACAACCTCCT CCGCGCCATC 1680  
 GAGGCCCAGC AGCATATGCT CCAGCTCACC GTGTGGGGCA TCAAGCAGCT CCAGGCCCCG 1740  
 GTGCTGGCCG TGGAGCGCTA CCTGAAGGAC CAGCAGCTCC TGGGCTTCTG GGGCTGCTCC 1800  
 GGCAAGCTGA TCTGCACCAC CACGGTACCC TGGAAACGCT CCTGGAGCAA CAAGAGCCTG 1860  
 GACGACATCT GGAACAACAT GACCTGGATG CAGTGGGAGC GCGAGATCGA TAACTACACC 1920  
 AGCCTGATCT ACAGCCTGCT GGAGAAGAGC CAGACCCAGC AGGAGAAGAA CGAGCAGGAG 1980  
 CTGCTGGAGC TGGACAAGTG GGCAGCCTG TGGAACTGGT TCGACATCAC CAACTGGCTG 2040  
 TGGTACATCA AAATCTTCAT CATGATTGTG GCGGCCTGG TGGGCTCCG CATCGTGTTC 2100  
 GCCGTGCTGA GCATCGTGAA CCGCGTGCGC CAGGGCTACA GCCCCCTGAG CCTCCAGACC 2160

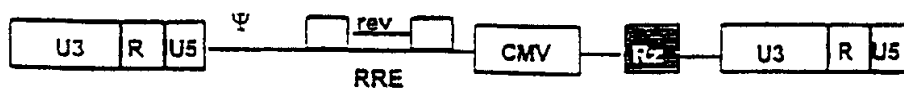
00.10.17

```
CGGCCCCCGG TGCCGCGCGG GCCCGACCGC CCCGAGGGCA TCGAGGAGGA GGGCGGCGAG 2220
CGCGACCGCG ACACCAGCGG CAGGCTCGTG CACGGCTTCC TGGCGATCAT CTGGGTCGAC 2280
CTCCGCAGCC TGTTCTGTT CAGCTACCAC CACCGCGACC TGCTGCTGAT CGCCGCCCCG 2340
ATCGTGGAAC TCCTAGGCCG CCGCGGCTGG GAGGTGCTGA AGTACTGGTG GAACCTCCTC 2400
CAGTATTGGA GCCAGGAGCT GAAGTCCAGC GCCGTGAGCC TGCTGAACGC CACCGCCATC 2460
GCCGTGGCCG AGGGCACCGA CCGCGTGATC GAGGTGCTCC AGAGGGCCCG GAGGGCGATC 2520
CTGCACATCC CCACCCGCAT CCGCCAGGGG CTCGAGAGGG CGCTGCTGTA A 2571
```

## 说明书附图

pH4Rz

质粒

整合的  
载体载体基因组  
RNA/核酶  
核酶

pH4.1Rz

质粒

整合的  
载体载体基因组  
RNA/核酶

pH6Rz

质粒

整合的  
载体载体基因组  
RNA/核酶  
核酶

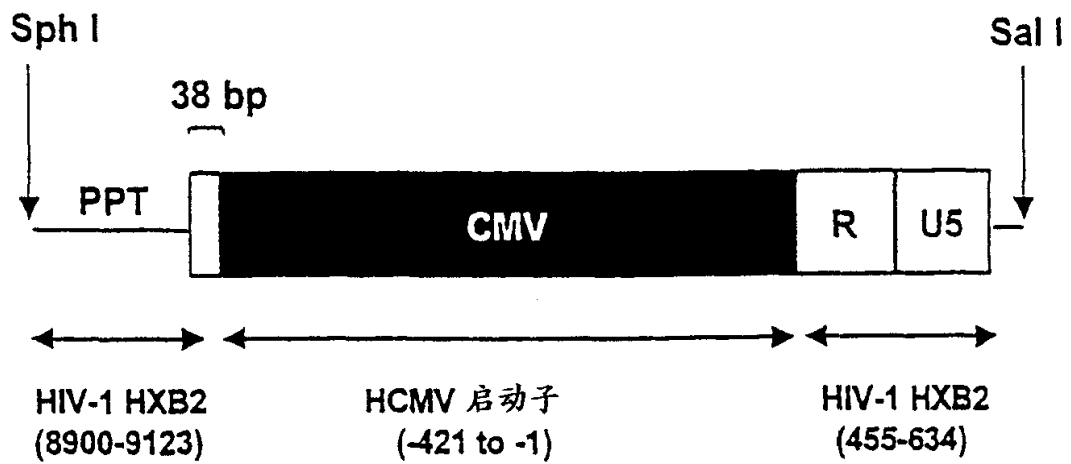
pH6.1Rz

质粒

整合的  
载体载体基因组  
RNA/核酶

图 1

A



B

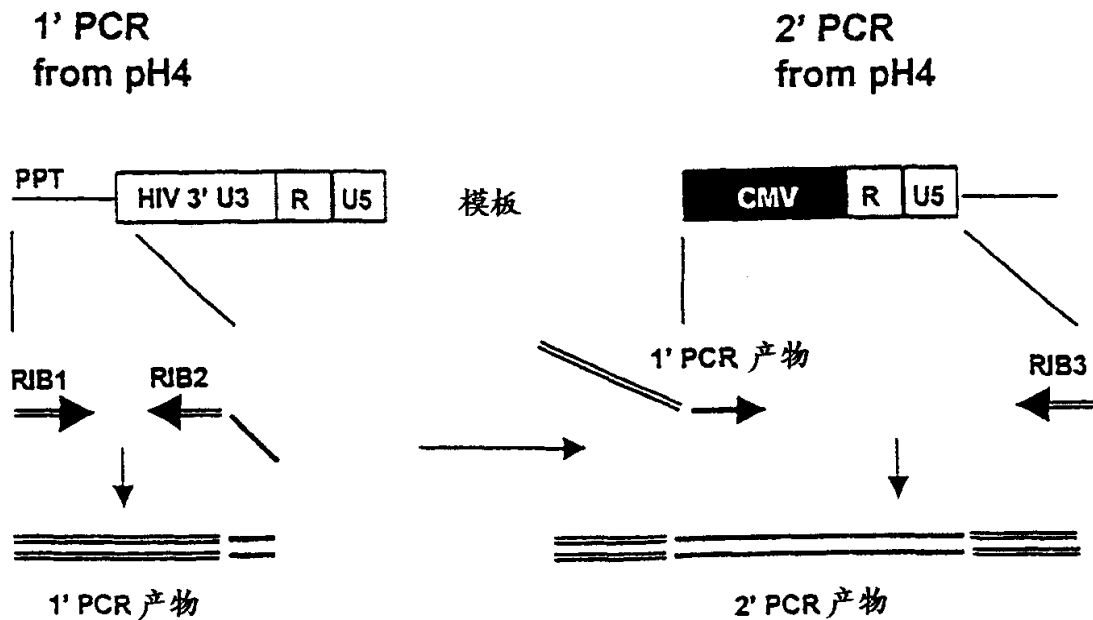


图 2

gagpol-HXB2 -> 密码子使用性

DNA 序列 4308 b.p. ATGGGTGCGAGA ... GATGAGGATTAG 线性

1436 密码子

MW : 161929 Dalton CAI(S.c.) : 0.083 CAI(E.c.) : 0.151

|           |    |           |    |           |    |           |    |
|-----------|----|-----------|----|-----------|----|-----------|----|
| TTT phe F | 21 | TCT ser S | 3  | TAT tyr Y | 30 | TGT cys C | 18 |
| TTC phe F | 14 | TCC ser S | 3  | TAC tyr Y | 9  | TGC cys C | 2  |
| TTA leu L | 46 | TCA ser S | 19 | TAA OCH Z | -  | TGA OPA Z | -  |
| TTG leu L | 11 | TCG ser S | 1  | TAG AMB Z | 1  | TGG trp W | 37 |
| CTT leu L | 13 | CCT pro P | 21 | CAT his H | 20 | CGT arg R | -  |
| CTC leu L | 7  | CCC pro P | 14 | CAC his H | 7  | CGC arg R | -  |
| CTA leu L | 17 | CCA pro P | 41 | CAA gln Q | 56 | CGA arg R | 3  |
| CTG leu L | 16 | CCG pro P | -  | CAG gln Q | 39 | CGG arg R | 3  |
| ATT ile I | 30 | ACT thr T | 24 | AAT asn N | 42 | AGT ser S | 18 |
| ATC ile I | 14 | ACC thr T | 20 | AAC asn N | 16 | AGC ser S | 16 |
| ATA ile I | 56 | ACA thr T | 43 | AAA lys K | 88 | AGA arg R | 45 |
| ATG met M | 29 | ACG thr T | 1  | AAG lys K | 34 | AGG arg R | 18 |
| GTT val V | 15 | GCT ala A | 17 | GAT asp D | 37 | GGT gly G | 11 |
| GTC val V | 11 | GCC ala A | 19 | GAC asp D | 26 | GGC gly G | 10 |
| GTA val V | 55 | GCA ala A | 55 | GAA glu E | 75 | GGA gly G | 61 |
| GTG val V | 15 | GCG ala A | 5  | GAG glu E | 32 | GGG gly G | 26 |

图 3

gagpol-SYNgp [1 to 4308] -> 密码子使用性

DNA 序列 4308 b.p. ATGGGCGCCCGC ... GATGAGGATTAG 线性

1436 密码子

MW : 161929 Dalton CAI(S.c.) : 0.080 CAI(E.c.) : 0.296

|           |    |           |    |           |    |           |    |
|-----------|----|-----------|----|-----------|----|-----------|----|
| TTT phe F | 5  | TCT ser S | 5  | TAT tyr Y | 10 | TGT cys C | 6  |
| TTC phe F | 30 | TCC ser S | 11 | TAC tyr Y | 29 | TGC cys C | 14 |
| TTA leu L | 2  | TCA ser S | 4  | TAA OCH Z | -  | TGA OPA Z | -  |
| TTG leu L | 7  | TCG ser S | 6  | TAG AMB Z | 1  | TGG trp W | 37 |
| CTT leu L | 3  | CCT pro P | 14 | CAT his H | 6  | CGT arg R | 2  |
| CTC leu L | 22 | CCC pro P | 39 | CAC his H | 21 | CGC arg R | 34 |
| CTA leu L | 6  | CCA pro P | 10 | CAA gln Q | 14 | CGA arg R | 3  |
| CTG leu L | 70 | CCG pro P | 13 | CAG gln Q | 81 | CGG arg R | 10 |
| ATT ile I | 17 | ACT thr T | 11 | AAT asn N | 13 | AGT ser S | 7  |
| ATC ile I | 79 | ACC thr T | 48 | AAC asn N | 45 | AGC ser S | 27 |
| ATA ile I | 4  | ACA thr T | 13 | AAA lys K | 25 | AGA arg R | 7  |
| ATG met M | 29 | ACG thr T | 16 | AAG lys K | 97 | AGG arg R | 13 |
| GTT val V | 5  | GCT ala A | 15 | GAT asp D | 19 | GGT gly G | 10 |
| GTC val V | 27 | GCC ala A | 56 | GAC asp D | 44 | GCC gly G | 54 |
| GTA val V | 6  | GCA ala A | 13 | GAA glu E | 29 | GGA gly G | 16 |
| GTG val V | 58 | GCG ala A | 12 | GAG glu E | 78 | GGG gly G | 28 |

图 4

env-mn [1 to 2571] -> 密码子使用性

DNA 序列 2571 b.p. ATGAGAGTGAAG ... GCTTTGCTATAA 线性

857 个密码子

MW : 97078 Dalton CAI(S.c.) : 0.083 CAI(E.c.) : 0.140

|           |    |           |    |           |    |           |    |
|-----------|----|-----------|----|-----------|----|-----------|----|
| TTT phe F | 13 | TCT ser S | 7  | TAT tyr Y | 15 | TGT cys C | 16 |
| TTC phe F | 11 | TCC ser S | 3  | TAC tyr Y | 7  | TGC cys C | 5  |
| TTA leu L | 20 | TCA ser S | 13 | TAA OCH Z | 1  | TGA OPA Z | -  |
| TTG leu L | 17 | TCG ser S | 2  | TAG AMB Z | -  | TGG trp W | 30 |
| CTT leu L | 9  | CCT pro P | 5  | CAT his H | 8  | CGT arg R | -  |
| CTC leu L | 11 | CCC pro P | 9  | CAC his H | 6  | CGC arg R | 2  |
| CTA leu L | 12 | CCA pro P | 12 | CAA gln Q | 22 | CGA arg R | 1  |
| CTG leu L | 15 | CCG pro P | 2  | CAG gln Q | 19 | CGG arg R | 1  |
| ATT ile I | 21 | ACT thr T | 16 | AAT asn N | 50 | AGT ser S | 18 |
| ATC ile I | 10 | ACC thr T | 14 | AAC asn N | 13 | AGC ser S | 11 |
| ATA ile I | 32 | ACA thr T | 28 | AAA lys K | 32 | AGA arg R | 30 |
| ATG met M | 17 | ACG thr T | 5  | AAG lys K | 14 | AGG arg R | 15 |
| GTT val V | 8  | GCT ala A | 16 | GAT asp D | 18 | GGT gly G | 10 |
| GTC val V | 9  | GCC ala A | 7  | GAC asp D | 14 | GGC gly G | 6  |
| GTA val V | 26 | GCA ala A | 20 | GAA glu E | 36 | GGA gly G | 28 |
| GTG val V | 12 | GCG ala A | 5  | GAG glu E | 10 | GGG gly G | 12 |

图 5

SYNgp160mn -&gt; 密码子使用性

DNA 序列 2571 b.p. ATGAGGGTGAAG ... GCGCTGCTGTAA 线性

857 个密码子

MW : 97078 Dalton CAI(S.c.) : 0.074 CAI(E.c.) : 0.419

|           |    |           |    |           |    |           |    |
|-----------|----|-----------|----|-----------|----|-----------|----|
| TTT phe F | -  | TCT ser S | 2  | TAT tyr Y | 1  | TGT cys C | -  |
| TTC phe F | 24 | TCC ser S | 4  | TAC tyr Y | 21 | TGC cys C | 21 |
| TTA leu L | -  | TCA ser S | -  | TAA och Z | 1  | TGA opa Z | -  |
| TTG leu L | -  | TCG ser S | -  | TAG amb Z | -  | TGG trp W | 30 |
| CTT leu L | -  | CCT pro P | -  | CAT his H | 2  | CGT arg R | 1  |
| CTC leu L | 20 | CCC pro P | 26 | CAC his H | 12 | CGC arg R | 36 |
| CTA leu L | 1  | CCA pro P | -  | CAA gln Q | -  | CGA arg R | -  |
| CTG leu L | 63 | CCG pro P | 2  | CAG gln Q | 41 | CGG arg R | 4  |
| ATT ile I | 2  | ACT thr T | -  | AAT asn N | 2  | AGT ser S | -  |
| ATC ile I | 61 | ACC thr T | 59 | AAC asn N | 61 | AGC ser S | 48 |
| ATA ile I | -  | ACA thr T | -  | AAA lys K | 1  | AGA arg R | 2  |
| ATG met M | 17 | ACG thr T | 4  | AAG lys K | 45 | AGG arg R | 6  |
| GTT val V | -  | GCT ala A | -  | GAT asp D | 2  | GGT gly G | 1  |
| GTC val V | 1  | GCC ala A | 40 | GAC asp D | 30 | GGC gly G | 47 |
| GTA val V | 1  | GCA ala A | -  | GAA glu E | 3  | GGA gly G | -  |
| GTG val V | 53 | GCG ala A | 8  | GAG glu E | 43 | GGG gly G | 8  |

图 6

# HIV 构建体

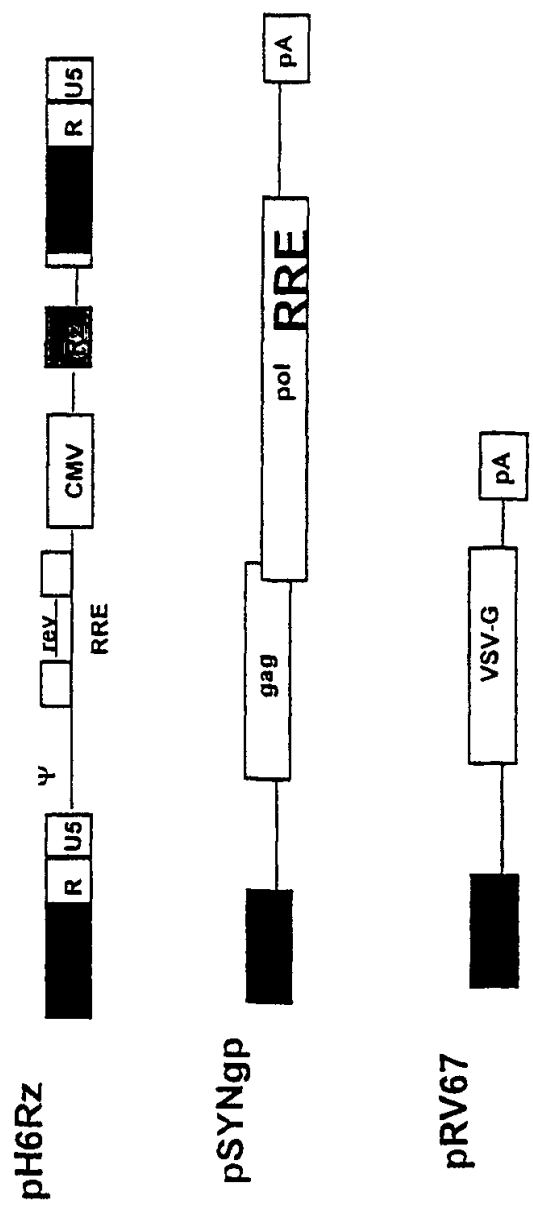
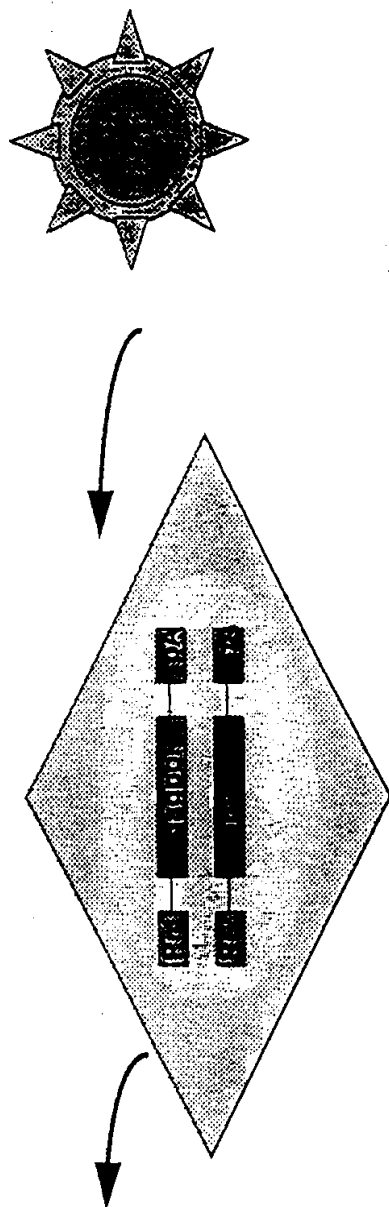


图 7

# Hit 载体体系

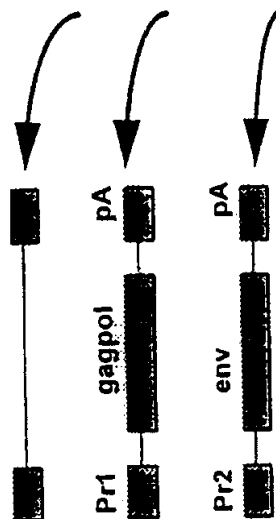
辅助包装细胞系

载体基因组构建体



## 三质粒共转染 (HIT)

载体基因组构建体



e.g. 293T, COS

图 8