

[19] 中华人民共和国国家知识产权局



# [12] 发明专利说明书

专利号 ZL 02823087.6

[51] Int. Cl.

*C12N 15/48 (2006.01)*

*C12N 15/09 (2006.01)*

*C12N 15/49 (2006.01)*

*A61P 31/18 (2006.01)*

*C12N 15/869 (2006.01)*

*A61K 39/12 (2006.01)*

[45] 授权公告日 2008 年 6 月 4 日

[11] 授权公告号 CN 100392085C

[22] 申请日 2002.9.18 [21] 申请号 02823087.6

[30] 优先权

[32] 2001.9.20 [33] GB [31] PCT/GB01/04207

[32] 2001.12.11 [33] GB [31] 0129604.5

[32] 2002.3.19 [33] GB [31] 0206462.4

[86] 国际申请 PCT/EP2002/010592 2002.9.18

[87] 国际公布 WO2003/025003 英 2003.3.27

[85] 进入国家阶段日期 2004.5.20

[73] 专利权人 葛兰素集团有限公司

地址 英国梅得塞克斯

[72] 发明人 A·贝顿 P·F·埃尔特尔

G·W·古 A·利尔

J·P·蒂特 C·A·范维利

[56] 参考文献

WO0136614A2 2001.5.25

WO0039304A2 2000.7.6

WO9839463A2 1998.9.11

WO0039302A2 2000.7.6

WO0147955A2 2001.7.5

A rev - independent human immunodeficiency virus type 1 (HIV - 1) - based vector that exploits a codon - optimized HIV - 1 gag - pro gene. Kotsopoulou E et al. Journal of Virology, Vol. 74 No. 10. 2000

审查员 杨振宇

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 刘 玥 徐雁漪

权利要求书 2 页 说明书 42 页 附图 50 页

[54] 发明名称

疫苗

[57] 摘要

本发明提供的核苷酸序列, 编码了 HIV - 1 gag 蛋白或其含有 gag 表位的片段以及第二 HIV 抗原, 或编码所述第二 HIV 抗原表位的片段, 并与异源启动子可操作地连接。优选的核苷酸序列进一步编码 nef 或其片段和 RT 或其片段。

1. 一种核苷酸序列，编码 HIV-1 gag 蛋白或其含有 gag 表位的片段，HIV-1 Nef 蛋白或其含有 nef 表位的片段，以及 RT 蛋白或含有 RT 表位的片段，并与异源启动子可操作地连接。

2. 如权利要求 1 所述的核苷酸序列，其中所述 gag 蛋白包含 p17。

3. 如权利要求 2 所述的核苷酸序列，其中所述 gag 蛋白还包含 p24。

4. 如权利要求 1-3 中任何一项所述的核苷酸序列，其中所述核苷酸序列中的 gag 编码部分的密码子被优化成与 RSCU 值为 0.5 的高度表达的人类基因的密码子利用率相似。

5. 如权利要求 1-3 中任何一项所述的核苷酸序列，其中所述核苷酸序列中的 RT 或其片段编码部分的密码子被优化成与高度表达的人类基因相似。

6. 如权利要求 1-3 中任何一项所述的核苷酸序列，其中所述序列的顺序是 RT gag Nef 或 RT Nef gag。

7. 一种核苷酸序列，其中所述序列选自：

- Gag p17、p24；RT；Nef 截短物
- 密码子优化的 Gag p17、p24；RT；Nef 截短物
- 密码子优化的 Gag p17、p24；密码子优化的 RT；Nef 截短物
- 密码子优化的 RT；密码子优化的 Gag p17、p24；Nef 截短物
- 密码子优化的 RT；Nef 截短物；密码子优化的 gag p17、p24。

8. 如权利要求 1-3 和 7 中任何一项所述的核苷酸序列，其中所述异源启动子是 HCMV IE 基因的启动子。

9. 如权利要求 8 所述的核苷酸序列，其中所述启动子的 5'端含有外显子 1。

10. 如权利要求 1-3 和 7 中任何一项所述的核苷酸序列，其中所述 RT 编码一个突变以灭活任何反转录酶的活性。

11. 如权利要求 10 所述的核苷酸序列，其中所述 RT 中的突变是由色氨酸 229 替代赖氨酸。

12. 一种含有如权利要求 1-11 中任何一项所述的核苷酸序列的

载体。

13. 一种如权利要求 12 所述的载体，其中所述载体是病毒载体。

14. 一种如权利要求 13 所述的病毒载体，其中所述载体是复制缺陷型腺病毒。

15. 一种如权利要求 12 所述的载体，其中所述载体是双链 DNA 质粒。

16. 一种由如权利要求 1-11 中任何一项所述的核苷酸序列编码的蛋白质。

17. 一种药物组合物，其含有如权利要求 1-11 中任何一项所述的核苷酸序列或权利要求 12、13、14 或 15 中所述的载体以及药用的赋形剂、稀释剂、载体或佐剂。

18. 一种如权利要求 17 所述的药物组合物，其中所述药物组合物适合于肌肉内或皮内给予。

19. 一种如权利要求 17 或 18 所述的药物组合物，其中所述载体是金小珠。

20. 一种皮内给予装置，其特征在于，所述装置包含权利要求 17、18 或 19 所述的药物组合物。

21. 如权利要求 1-11 中任何一项所述的核苷酸序列，如权利要求 12-15 中所述的载体，或如权利要求 17-19 中任何一项所述的组合物在制备用于治疗疾病的药物中的用途。

22. 一种生产如权利要求 1-11 中任何一项所述的核苷酸的方法，其包括将编码 HIV-1 gag 蛋白或其片段以及第二 HIV-1 Nef 蛋白或其片段的核苷酸序列与异源启动子序列可操作地连接。

## 疫苗

## 发明领域

本发明涉及核酸构建体、含有这种构建体的宿主细胞以及它们在核酸疫苗中的应用。本发明进一步涉及含有这种构建体的疫苗制剂以及这种制剂在药物中的应用。本发明特别涉及用于 HIV 感染的预防 and 治疗的 DNA 疫苗，更特别的是以颗粒介导的释放方式施用疫苗。

## 发明背景

HIV-1 是获得性免疫缺陷综合症 (AIDS) 的主要原因，这种疾病被认为是全球主要的健康问题之一。尽管全世界为了生产相关疫苗进行了广泛的研究，但是迄今为止这些努力尚未获得成功。

已经描述了 HIV-1 的非包膜蛋白，包括例如内部结构蛋白，如 gag 和 pol 基因的产物，及其它非结构蛋白如 Rev、Nef、Vif 和 Tat (Green 等, New England J. Med, 324, 5, 308 以及下列等等 (1991) 和 Bryant 等, (Pizzo 编辑), Pediatr. Infect. Dis. J., 11, 5, 390 以及下列等等 (1992))。

Gag 基因的全长 RNA 翻译产生多蛋白前体，随后该前体被酶切成为 3-5 个衣壳蛋白；基质蛋白、衣壳蛋白和核酸结合蛋白以及蛋白酶 (1. Fundamental Virology, Fields BN, Knipe DM 和 Howley M 1996; 2. Fields Virology, vol 2 1996)。

所述 gag 基因产生 55-千道尔顿 (kD) 的 Gag 前体蛋白，也称为 p55，它是从未剪接的病毒 mRNA 表达产生的。在翻译过程中，p55 的 N 末端被十四酰化，使其与细胞膜的胞质相缔合。膜缔合的 Gag 多蛋白使两个拷贝的病毒基因组 RNA 连同其它病毒和细胞蛋白聚集而引起病毒颗粒在感染细胞的表面产生芽接。芽接后，在病毒成熟过程中 p55 被病毒编码的蛋白酶 (pol 基因的产物) 酶切成为被称做 MA (基质[p17])、CA (衣壳[p24])、NC (核壳[p9]) 和 p6. (4) 的四个较小蛋白质。

除了 3 个主要的 Gag 蛋白之外，所有的 Gag 前体含有一些其它的区域，它们被酶切成各种大小的肽，保留于病毒粒子中。这些蛋白质具有不同的作用，例如 p2 蛋白被认为在调节蛋白酶的活性方面具有作用，并有助于校正蛋白水解过程的时间。

MA 多蛋白衍生自 p55 的十四酰化 N 末端。大多数 MA 分子保持与病毒粒子的脂双分子层的内表面连接，以稳定病毒颗粒。在病毒粒子较深层的内部聚集其它的 MA，而成为携带病毒 DNA 进入细胞核的复合体的一部分。(5) 这些 MA 分子促进病毒基因组的核转移，因为 MA 上的亲核信号可被细胞核输入系统所识别。这种情况使 HIV 可以感染未分裂的细胞，这是反转录病毒的一个不寻常的特性。

p24 (CA) 蛋白形成病毒颗粒的锥形核。已证实亲环蛋白 A 与 p55 的 p24 区域相互作用，而掺合到 HIV 颗粒中。因为环孢菌素可以阻止这种相互作用而抑制病毒的复制，所以 Gag 和亲环蛋白 A 之间的相互作用是必需的。

Gag 的 NC 区域负责特异性识别所谓的 HIV 包装信号。所述包装信号由位于病毒 RNA 的 5' 末端附近的 4 个茎环结构组成，足以介导异源 RNA 整合到 HIV-1 病毒粒子中。NC 通过由两个锌指基序介导的相互作用而与包装信号相结合。NC 还促进逆转录进行。

p6 多肽区域介导了 p55 Gag 和辅助蛋白 Vpr 之间的相互作用，使 Vpr 整合到组装好的病毒粒子中。p6 区域还含有一个所谓的晚期结构域，这是芽接病毒粒子从感染细胞有效释放所需要的。

Pol 基因编码包含病毒早期感染时所需要的具有各别活性的两种蛋白质，即将病毒 DNA 整合进入细胞 DNA 所需的 RT 和整合酶蛋白。Pol 的最初产物被病毒粒子蛋白酶切割，产生含有 DNA 合成所必需的活性 (RNA 和 DNA 定向的 DNA 聚合酶，核酶 H) 的氨基端 RT 肽以及羧基端整合酶蛋白。

HIV RT 是全长 RT (p66) 和缺乏羧基末端的 RNA 酶整合酶结构域的酶切产物 (p51) 的异源二聚体。

RT 是反转录酶病毒基因组所编码的保守性最高的蛋白质之一。RT 的两个主要活性是 DNA Pol 和核酶 H。RT 的 DNA Pol 活性可交替使用 RNA 和 DNA 为模板，并且如所有已知的 DNA 聚合酶一样，不能

起始 DNA 的从头合成,而需要一个预先存在的分子作为引物(RNA)。

当 DNA 合成发生时,在所有的 RT 蛋白中所固有的 RNA 酶 H 的活性在除去 RNA 基因组复制早期中起着重要作用。它选择性地从所有 RNA-DNA 杂合分子中降解 RNA。聚合酶和核酶 H 在结构上是分离的、具有非重叠结构域,其中 Pol 覆盖了 Pol 的三分之二的氨基。

p66 催化亚基折叠成 5 个不同的亚结构域。这些亚结构域的氨基末端 23 包含具有 RT 活性的部分,羧基末端是 RNA 酶 H 结构域。

感染宿主细胞后,逆转录病毒 RNA 基因组通过感染颗粒中的反转录酶而被复制成线性双链 DNA。整合酶(参见 Skalka AM '99 Adv in Virus Res 52 271-273)识别和修饰病毒 DNA 的末端,并伴随病毒 DNA 进入宿主染色体位点以进行催化整合。宿主 DNA 中有许多位点可成为整合的靶标位点。尽管整合酶在体外足以催化整合,但在体内它不是唯一与病毒 DNA 结合的蛋白质,从感染的细胞中分离的大蛋白-病毒 DNA 复合体被称为预整合复合体。这促进子代病毒基因组获得宿主细胞的基因。

整合酶由 3 个不同的结构域组成,即 N 末端结构域、催化核心和 C 末端结构域。催化核心结构域中包含了多核苷酸转移化学反应的所有要素。

已知 Nef 蛋白可引起 CD4、HIV 受体从细胞表面的去除,但对于这种功能的生物学重要性还存在争议。此外, Nef 与 T 细胞的信号途径相互作用并诱导成激活状态,这反过来可促进更有效的基因表达。一些 HIV 分离物在该区域存在突变,导致不编码有功能的蛋白而严重影响它们在体内的复制和致病过程。

DNA 疫苗通常包括插入了强启动子的细菌质粒载体和感兴趣的基因所组成,所述基因编码抗原肽和聚腺苷酸化/转录终止序列。感兴趣的基因可编码全蛋白或仅编码与想要预防的病原体、肿瘤或其它病因有关的抗原肽序列。在施用给宿主之前,质粒可在细菌例如大肠杆菌中生长,然后分离并在适当介质中制备质粒,所用介质取决于施用途径。施用后,质粒为宿主细胞所摄取,并在其中产生编码的肽。为了防止质粒在哺乳动物宿主中复制和在有关动物中对染色体 DNA 的整合,优选地,质粒载体被制备成不具有在真核细胞中发挥功能的

复制起点。

相对于传统的疫苗接种技术，DNA 疫苗接种具有许多优越性。首先，由于 DNA 序列编码的蛋白质在宿主中合成，可预测蛋白质的结构和构象将与疾病有关的天然蛋白质相似。通过产生可以识别保守蛋白质的抗原决定簇的细胞毒素 T 淋巴细胞反应，DNA 疫苗接种还可能抵抗不同病毒株系。而且，由于质粒为宿主细胞所摄取并在其中生产抗原蛋白，可激发长期的免疫反应。DNA 疫苗接种还可结合各种不同的免疫原而可能制成单一制剂，因而促进对许多种疾病的同时免疫。

有关 DNA 疫苗接种的有用的背景资料描述于 Donnelly 等 “DNA vaccines” *Ann. Rev Immunol.* 1997 15: 617-648，本文引用该文公开的全部内容作为参考。

#### 发明概述

本发明提供用于预防和治疗 HIV 感染和 AIDS 的核酸疫苗的新构建体。

相应地，作为第一个方面，本发明提供了一种核酸分子，其含有编码 HIV gag 蛋白或其片段的核苷酸序列和与其连接的编码了其它 HIV 抗原或其片段的核苷酸序列，这些核苷酸序列与异源启动子可操作地连接。所述核苷酸序列的片段编码 HIV 表位，并典型地编码至少 8 个氨基酸的肽。核苷酸序列优选地是 DNA 序列，并优选地包含于没有复制起始的质粒中。这种核酸分子与药用的赋形剂、载体、稀释剂或佐剂一起配制，生产适合于治疗和/或预防 HIV 感染和 AIDS 的药物组合物。

在一个优选的实施方案中，DNA 序列被配制在适合于以颗粒作为介质传递药物的惰性颗粒(particle)或小珠(bead)的表面。优选地，小珠是金的。

在本发明一个优选的实施方案中，提供含有 gag 蛋白高度表达的密码的 DNA 序列，所述序列已被优化成与哺乳动物细胞中基因所使用的密码子相似。具体地说，所述 gag 蛋白被优化成类似于高度表达的人类基因。

DNA 密码分别由 4 个字母 (A、T、C 和 G) 表示, 并使用它们拼写成代表生物体基因编码蛋白质的氨基酸的三个字母所表示的“密码子”。沿着 DNA 分子的顺序排列的密码子被翻译成由那些基因编码的蛋白质中氨基酸的序列。密码子具有高度简并性, 其中 61 个密码子编码 20 个常见氨基酸(即蛋白质氨基酸)和 3 个密码子代表“终止”信号。因此, 大多数氨基酸由不止一个密码子所编码—事实上, 有些氨基酸是由四个或更多不同的密码子所编码。

不止一个密码子编码了同一个特定的氨基酸, 人们已观察到生物体中密码子的使用模式是高度不随机的。不同物种, 在密码子选择上显示了不同的偏好性, 而且在同一物种中, 在以高水平和低水平表达的基因之间的密码子使用模式可以非常不同。这种偏好性在病毒、植物、细菌和哺乳动物细胞中是不同的, 并且某些物种比其他物种显示出更高的非随机密码子选择的偏好性。例如, 人和其他哺乳动物比某些细菌或病毒的偏好性要弱。由于这些原因, 在大肠杆菌中表达哺乳动物基因或在哺乳动物细胞中表达外源或重组基因时, 对于有效表达而言密码子分布是非常有可能不适当。在异源 DNA 序列中存在的密码子簇或者丰富的密码子在将要用于表达的宿主中很少观察到, 则往往预示着在该宿主中对该异源基因的表达水平将很低。

在本发明的一个实施方案中, 提供了一个编码氨基酸序列的 gag 多核苷酸序列, 其中多核苷酸序列的密码子使用模式与高度表达的哺乳动物基因的密码子使用模式相似。优选地, 多核苷酸序列是 DNA 序列。理想地, 多核苷酸序列的密码子使用模式是典型的高度表达的人类基因。

在本发明的多核苷酸中, 密码子使用模式从人免疫缺陷病毒的密码子典型模式改变为更接近靶标生物体的密码偏好性的使用模式, 例如哺乳动物, 特别是人。“密码子使用系数(codon usage coefficient)”衡量给定的多核苷酸序列的密码子模式与靶标物种的密码子模式相似的程度。许多物种的高度表达基因的密码子出现频率可从已有文献资料中获得(参见, 例如, Nakamura 等, *Nucleic Acids Research* 1996, 24: 214-215)。将 61 个密码子中每一个密码子的出现频率(用所选择基因类别中每 1000 个密码子中出现的数量来表示)对 20 个常见氨基酸



中每一个进行标准化，将每一个氨基酸最频繁使用的密码子的值设定为 1，使用较少的密码子的频率范围将在 0 和 1 之间。这样，对于靶标物种的高度表达基因，61 个中的每一个密码子被确定为值 1 或更低。为了计算某一特定多核苷酸的密码子相对于该物种的高度表达基因的使用系数，记录该特定多核苷酸的各个密码子的测量值，并且计算出所有这些值的几何平均值（这些值的自然对数之和除以密码子的总数，并取反对数）。该系数的值在 0 和 1 之间，并且密码子的系数越高，多核苷酸中对该密码子的使用更为频繁。如果多核苷酸序列的某一密码子使用系数为 1，对于靶标物种的高度表达基因而言，所有密码子均为“最频繁出现”的密码子。

依据本发明，多核苷酸的密码子利用模式优先地排除在靶标生物体的高度表达基因中 RSCU 值小于 0.2 的密码子。可选地，该密码子利用模式排除编码特定氨基酸少于 10% 的密码子。如果所述氨基酸的所有密码子利用频率相同，那么相对同义密码子利用率（RSCU）值是密码子的观测数量除以预测数量。对于高度表达的人类基因，本发明的多核苷酸一般具有大于 0.3，优选大于 0.4，最优选大于 0.5 的密码子使用系数（或 RSCU）。人的密码子利用率表还可在 Genebank 中查到。

经比较，高度表达的β肌动蛋白基因的 RSCU 值为 0.747。人(homo sapiens)的密码子利用率表如下：

#### 密码子利用率表 1:

人(homo sapiens) [gbpri]: 27143 CDS's (12816923 个密码子)

范围 (fields): [三联体][频率: 每 1000 个]([数量])

|                  |                  |                  |                  |
|------------------|------------------|------------------|------------------|
| UUU 17.0(217684) | UCU 14.8(189419) | UAU 12.1(155645) | UGU 10.0(127719) |
| UUC 20.5(262753) | UCC 17.5(224470) | UAC 15.8(202481) | UGC 12.3(157257) |
| UUA 7.3( 93924)  | UCA 11.9(152074) | UAA 0.7( 9195)   | UGA 1.3( 16025)  |
| UUG 12.5(159611) | UCG 4.5( 57572)  | UAG 0.5( 6789)   | UGG 12.9(165930) |
| CUU 12.8(163707) | CCU 17.3(222146) | CAU 10.5(134186) | CGU 4.6( 59454)  |
| CUC 19.3(247391) | CCC 20.0(256235) | CAC 14.9(190928) | CGC 10.8(137865) |
| CUA 7.0( 89078)  | CCA 16.7(214583) | CAA 12.0(153590) | CGA 6.3( 80709)  |
| CUG 39.7(509096) | CCG 7.0( 89619)  | CAG 34.5(441727) | CGG 11.6(148666) |
| AUU 15.8(202844) | ACU 12.9(165392) | AAU 17.0(218508) | AGU 12.0(154442) |
| AUC 21.6(277066) | ACC 19.3(247805) | AAC 19.8(253475) | AGC 19.3(247583) |
| AUA 7.2( 92133)  | ACA 14.9(191518) | AAA 24.0(308123) | AGA 11.5(147264) |
| AUG 22.3(285776) | ACG 6.3( 80369)  | AAG 32.6(418141) | AGG 11.3(145276) |

GUU 10.9(139611) GCU 18.5(236639) GAU 22.4(286742) GGU 10.8(138606)  
 GUC 14.6(187333) GCC 28.3(362086) GAC 26.1(334158) GGC 22.7(290904)  
 GUA 7.0( 89644) GCA 15.9(203310) GAA 29.1(373151) GGA 16.4(210643)  
 GUG 28.8(369006) GCG 7.5( 96455) GAG 40.2(515485) GGG 16.4(209907)

编码 GC 52.51% 第一个字母 GC 56.04% 第二个字母 GC 42.35%  
 第三个字母 GC 59.13%。

密码子利用率表 2 (优选的):

人 (高度表达的) 基因 1/24/91 的密码子利用率 (human-high.cod)

| 氨基酸 | 密码子 | 数量      | /1000 | 分数(fraction) |
|-----|-----|---------|-------|--------------|
| Gly | GGG | 905.00  | 18.76 | 0.24         |
| Gly | GGA | 525.00  | 10.88 | 0.14         |
| Gly | GGT | 441.00  | 9.14  | 0.12         |
| Gly | GGC | 1867.00 | 38.70 | 0.50         |
| Glu | GAG | 2420.00 | 50.16 | 0.75         |
| Glu | GAA | 792.00  | 16.42 | 0.25         |
| Asp | GAT | 592.00  | 12.27 | 0.25         |
| Asp | GAC | 1821.00 | 37.75 | 0.75         |
| Val | GTG | 1866.00 | 38.68 | 0.64         |
| Val | GTA | 134.00  | 2.78  | 0.05         |
| Val | GTT | 198.00  | 4.10  | 0.07         |
| Val | GTC | 728.00  | 15.09 | 0.25         |
| Ala | GCG | 652.00  | 13.51 | 0.17         |
| Ala | GCA | 488.00  | 10.12 | 0.13         |
| Ala | GCT | 654.00  | 13.56 | 0.17         |
| Ala | GCC | 2057.00 | 42.64 | 0.53         |
| Arg | AGG | 512.00  | 10.61 | 0.18         |
| Arg | AGA | 298.00  | 6.18  | 0.10         |
| Ser | AGT | 354.00  | 7.34  | 0.10         |
| Ser | AGC | 1171.00 | 24.27 | 0.34         |
| Lys | AAG | 2117.00 | 43.88 | 0.82         |
| Lys | AAA | 471.00  | 9.76  | 0.18         |
| Asn | AAT | 314.00  | 6.51  | 0.22         |
| Asn | AAC | 1120.00 | 23.22 | 0.78         |
| Met | ATG | 1077.00 | 22.32 | 1.00         |
| Ile | ATA | 88.00   | 1.82  | 0.05         |
| Ile | ATT | 315.00  | 6.53  | 0.18         |
| Ile | ATC | 1369.00 | 28.38 | 0.77         |
| Thr | ACG | 405.00  | 8.40  | 0.15         |

|     |     |         |       |      |
|-----|-----|---------|-------|------|
| Thr | ACA | 373.00  | 7.73  | 0.14 |
| Thr | ACT | 358.00  | 7.42  | 0.14 |
| Thr | ACC | 1502.00 | 31.13 | 0.57 |
| Trp | TGG | 652.00  | 13.51 | 1.00 |
| End | TGA | 109.00  | 2.26  | 0.55 |
| Cys | TGT | 325.00  | 6.74  | 0.32 |
| Cys | TGC | 706.00  | 14.63 | 0.68 |
| End | TAG | 42.00   | 0.87  | 0.21 |
| End | TAA | 46.00   | 0.95  | 0.23 |
| Tyr | TAT | 360.00  | 7.46  | 0.26 |
| Tyr | TAC | 1042.00 | 21.60 | 0.74 |
| Leu | TTG | 313.00  | 6.49  | 0.06 |
| Leu | TTA | 76.00   | 1.58  | 0.02 |
| Phe | TTT | 336.00  | 6.96  | 0.20 |
| Phe | TTC | 1377.00 | 28.54 | 0.80 |
| Ser | TCG | 325.00  | 6.74  | 0.09 |
| Ser | TCA | 165.00  | 3.42  | 0.05 |
| Ser | TCT | 450.00  | 9.33  | 0.13 |
| Ser | TCC | 958.00  | 19.86 | 0.28 |
| Arg | CGG | 611.00  | 12.67 | 0.21 |
| Arg | CGA | 183.00  | 3.79  | 0.06 |
| Arg | CGT | 210.00  | 4.35  | 0.07 |
| Arg | CGC | 1086.00 | 22.51 | 0.37 |
| Gln | CAG | 2020.00 | 41.87 | 0.88 |
| Gln | CAA | 283.00  | 5.87  | 0.12 |
| His | CAT | 234.00  | 4.85  | 0.21 |
| His | CAC | 870.00  | 18.03 | 0.79 |
| Leu | CTG | 2884.00 | 59.78 | 0.58 |
| Leu | CTA | 166.00  | 3.44  | 0.03 |
| Leu | CTT | 238.00  | 4.93  | 0.05 |
| Leu | CTC | 1276.00 | 26.45 | 0.26 |
| Pro | CCG | 482.00  | 9.99  | 0.17 |
| Pro | CCA | 456.00  | 9.45  | 0.16 |
| Pro | CCT | 568.00  | 11.77 | 0.19 |
| Pro | CCC | 1410.00 | 29.23 | 0.48 |

本发明的另一个方面提供了一种表达载体，含有并能够调控本发明第一个方面所述多核苷酸序列的表达，特别是 gag 多核苷酸序列的密码子使用模式是典型的高度表达的哺乳动物基因，优选地是高度表达的人类基因。所述载体可适合于在细菌、昆虫或哺乳动物细胞、特别是人细胞中的表达异源 DNA。在一个实施方案中，所述表达载体是 p7313（见图 1）。

在第三个实施方案中，提供了与编码 NEF 及其片段、或 HIV 反转录酶 (RT) 或其片段的 DNA 序列融合的异源启动子控制下的 gag 基因。所述基因的 gag 部分可以在融合体的 N 或 C 端部分。

在一个优选的实施方案中，gag 基因不编码 gag p6 肽。优选地，NEF 基因是截短的即除去了编码 N 末端区域的序列，也即除去 30-85，优选地 60-85，典型的约 81，优选地 N 末端的 65 个氨基酸。

在另一个实施例中，RT 基因也被优化成与高度表达的人类基因相似。所述 RT 优选地编码一个突变以基本上失活了任何反转录酶的活性。一个优选的失活突变包括用 W 色氨酸 229 代替 K (赖氨酸)。

根据本发明的另一个方面，宿主细胞含有本发明所述的多核苷酸序列，或者本发明所提供的表达载体。所述宿主细胞可以是细菌，例如大肠杆菌；哺乳动物，例如人；或者是昆虫细胞。含有本发明提供的载体的哺乳动物细胞可以是培养细胞，可体外转染，或者可通过给予哺乳动物载体而发生体内转染。

本发明进一步提供了含有本发明的多核苷酸序列的药物组合物。优选地，所述组合物含有 DNA 载体。在优选的实施方案中，所述组合物含有大量颗粒，优选地是金颗粒，为含有编码本发明的多核苷酸序列的载体的 DNA 所包被。优选地，编码 HIV gag 氨基酸序列的多核苷酸序列的密码使用模式是典型的高度表达的哺乳动物基因，特别是人的基因。在可选的实施方案中，所述组合物含有药用的赋形剂以及根据本发明第二个方面的 DNA 载体。所述组合物还可包括佐剂。

因而，本发明的一个实施方案是本发明的载体与免疫刺激剂一起利用。优选地，免疫刺激剂与本发明的核酸载体同时给药，并且在优选的实施方案中它们一起配制。这种免疫刺激剂包括，但这里决非全面列举，也不排除其他的免疫刺激剂：合成咪唑并喹啉 (imidazoquinoline) 例如咪喹莫特 (imiquimod) (p14) [S-26308, R-837], (Harrison 等, “单独使用咪喹莫特 或与糖苷配基蛋白质疫苗结合使用减少 HSV 疾病的复发”, Vaccine 19: 1820-1826, (2001)); 和 resiquimod [S-28463, R-848] (Vasilakos 等, “免疫反应修饰基因 R-848 的佐剂活性: 与 CpG ODN’ 的比较”, 细胞免疫学 (Cellular immunology) 204: 64-74 (2000)), 在抗原递呈细胞和 T-细胞表

面组成型表达的羧基和胺的 Schiff 碱基, 例如 tucaresol (Rhodes, J. 等, “通过共刺激 Schiff-碱基-形成药物免疫系统的治疗效果增强”, Nature 377: 7175 (1995)), 细胞因子、趋化因子和共-刺激分子的蛋白质或肽, 包括促炎细胞因子例如 GM-CSF、IL-1 $\alpha$ 、IL-1 $\beta$ 、TGF- $\alpha$ 和 TGF- $\beta$ , Th1 诱导剂例如干扰素 $\gamma$ 、IL-2、IL-12、IL-15 和 IL-18, Th2 诱导剂例如 IL-4、IL-5、IL-6、IL-10 和 IL-13 和其它趋化因子以及共-刺激基因例如 MCP-1、MIP-1 $\alpha$ 、MIP-1 $\beta$ 、RANTES、TCA-3、CD80、CD86 和 CD40L, 以及其它免疫刺激靶向配基例如 CTLA-4 和 L-选择蛋白、细胞程序死亡刺激蛋白和肽例如 Fas, (49)、合成脂基佐剂, 例如 vaxfectin, (Reyes 等, “Vaxfectin 增强抗原特异性抗体的效价并保持对质粒 DNA 免疫的 Th1 型的免疫反应”, Vaccine 19: 3778-3786) 角鲨烯、 $\alpha$ -生育酚、聚山梨醇酯 80、DOPC 和胆固醇、内毒素、[LPS] (Beutler, B, “内毒素, Toll-样受体 4 和先天免疫性的传入支”, 现代微生物学观点 (Current Opinion in Microbiology), 3: 23-30 (2000)); CpG 寡-和二-核苷酸 (Sato, Y. 等, “有效皮内基因免疫所必需的免疫刺激 DNA 序列”, Science 273 (5273): 352-354 (1996)。Hemmi, H 等, “一种 Toll-样受体识别细菌 DNA”, Nature 408: 740-745, (2000)), 以及其它引发 Toll 受体产生 Th1-诱导细胞因子的潜在的配基, 例如合成的分枝杆菌脂蛋白、分枝杆菌蛋白 p19、肽聚糖磷壁质和类脂 A。

激发主要的 Th1-型反应的某些优选的佐剂包括, 例如, 类脂 A 的衍生物, 例如单磷酸类脂 A, 或优选地 3-去-O-酰化单磷酸类脂 A。MAL<sup>®</sup>佐剂可从 Corixa Corporation 购得 (Seattle, WA; 见例如, 美国专利 4,436,727、4,877,611、4,866,034 和 4,912,094)。包含 CpG 的寡核苷酸 (其中 CpG 二核苷酸未甲基化) 也诱导主要的 Th1 反应。这种寡核苷酸是众所周知的, 并描述于例如, WO 96/02555、WO 99/33488 和美国专利 6,008,200 和 5,856,462。免疫刺激 DNA 序列也在例如 Sato 等, Science 273: 352, 1996 中有描述。另一种优选的佐剂包括皂苷, 例如 Quil A, 或其衍生物, 包括 QS21 和 QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA)、七叶皂苷、毛地黄皂苷、或 Gypsophila 或奎葵籽皂苷 (Chenopodium quinoa saponins)。

还提供了根据本发明的多核苷酸或根据本发明的载体在治疗或预防 HIV 感染的应用。

本发明还提供了治疗 HIV 或预防 HIV 感染、及任何与之有关的症状或疾病的方法，包括给予有效量的本发明多核苷酸、载体或药物组合物。药物组合物可采取一或多个单独剂量方式给予，例如用相同 DNA 质粒的重复剂量，或以一种“初次免疫-加强免疫”（prime-boost）的治疗接种方式。在某些情况下，“初次”接种可通过颗粒介导 DNA 释放本发明的多核苷酸，优选地整合到质粒衍生载体中，而“加强免疫”通过给予含有同样的多核苷酸序列的重组病毒载体进行，或者用佐剂中的蛋白质加强。相反地，初次免疫可使用病毒载体或典型地在佐剂中配制蛋白质的蛋白质制剂，而加强免疫使用本发明的 DNA 疫苗。可采用多次剂量的初次免疫和/或加强免疫。

涉及 gag、nef 或 RT 蛋白的片段的实施方案也得到了考虑。例如，本发明的多核苷酸可编码 HIV gag、nef 或 RT 蛋白的片段。一种多核苷酸编码至少 8 个氨基酸的片段，例如长度为 8-10 个氨基酸或高达 20、50、60、70、80、100、150 或 200 个氨基酸，只要所编码的寡肽或多肽显示了 HIV 抗原性，均认为属于本发明的范围。特别地，但并不是唯一的，本发明的这一方面包含多核苷酸编码完全 HIV 蛋白序列的片段，其代表着 HIV 蛋白质的一个或多个分离(discrete)表位时的情况。这种片段可以含优化密码子，以使该片段就具有与高度表达的哺乳动物基因相似的密码子使用模式。

根据本发明的优选构建体包括：

1. p17、p24，与截短的 NEF 融合（缺乏编码末端氨基酸 1-65 的核苷酸）
2. p17、p24、RT 截短的 NEF（缺乏编码末端氨基酸 1-65 的核苷酸）
3. p17、p24（优化的 gag）截短的 NEF（缺乏编码末端氨基酸 1-65 的核苷酸）
4. p17、p24（优化的 gag）RT（优化的）截短的 NEF（缺乏编码末端氨基酸 1-85 的核苷酸）
5. p17、p24、RT（优化的）截短的 NEF（缺乏编码末端氨基酸 1-65

的核苷酸)

6. 截短的 NEF- (缺乏核苷酸 1-65) 与优化的 p17、p24 gag 融合.

7. 本发明特别优选的构建体包括三重融合 RT-NEF-Gag, 尤其是 RT-Gag-Nef;

8. 优化的 RT、截短的 NEF 和优化的 p17、p24 (gag) (RNG) 和

9. 优化的 RT、优化的 p17、24 (gag)、Nef 截短物 (缺乏氨基酸 1-65) RGN

优选地, HIV 构建体衍生自 HIV 分化体(Clade)B 或分化体 C, 尤其是分化体 B.

如上所述, 本发明包括含有本发明核苷酸序列的表达载体. 这种表达载体可用分子生物学领域的常规技术进行构建, 并可能例如涉及利用质粒 DNA 和适当的起始子、启动子、增强子、和其它元件例如多聚腺苷酸化信号, 它可能是必需的并位于正确的位置和方向, 以能够使蛋白质表达. 其它适合的载体是本领域熟练的技术人员已知的. 关于这一点的其它的实施例我们参考 Sambrook 等, 分子克隆: 实验室手册 (Molecular Cloning: a Laboratory Manual), 第二版, CSH Laboratory Press. (1989)。

优选地, 本发明多核苷酸, 或在本发明中用于载体的多核苷酸, 与能够在宿主细胞中表达编码序列的控制序列可操作地连接, 即该载体是一种表达载体. 术语“可操作地连接”是指所描述的成分毗邻并以允许它们以其预定的方式发挥功能的关系存在. 调控序列如启动子与编码序列“可操作地连接”, 意指以这样的方式连接, 即在与调控序列相容的条件下使编码序列的成功表达.

载体可以是, 例如, 有复制起点的质粒、人工染色体 (例如 BAC、PAC、YAC)、病毒或噬菌体载体, 任选地含有用于多核苷酸表达的启动子和任选地含有该启动子的调节子. 该载体可包括一种或多种可选择的标记基因, 例如在细菌质粒中可包含氨苄青霉素或卡那霉素抗性基因, 或用于真菌载体的抗性基因. 载体可在体外使用, 例如用于 DNA 或 RNA 的生产或用于转染或转化宿主细胞, 例如哺乳动物宿主细胞, 如用于生产由载体编码的蛋白质. 该载体也可适应于体内使用,

例如用于 DNA 疫苗接种或基因治疗中。

选择的启动子和其它表达调节信号与用于表达的宿主细胞相兼容。例如，哺乳动物启动子包括可诱导对重金属例如镉的反应的金属硫蛋白启动子，以及 $\beta$ -肌动蛋白启动子。病毒启动子例如 SV40 大 T 抗原启动子，人细胞肥大病毒 (CMV) 立即早期 (IE) 启动子、劳氏肉瘤病毒 LTR 启动子、腺病毒启动子或 HPV 启动子，特别是还可使用 HPV 上游调节区 (URR)。所有这些启动子在本领域中已被详细描述，并可容易地获得。

一种优选的启动子元件是 CMV 立即早期启动子，它缺乏内含子 A 但包括外显子 1。启动子元件可以是最小的启动子元件或增强的启动子，增强的启动子是优选的。因此，这里提供包含了在 HCMV IE 早期启动子控制下的含本发明多核苷酸的载体。

适合的病毒载体例如包括单纯疱疹病毒载体、牛痘或 $\alpha$ -病毒载体和逆转录病毒，包括慢病毒、腺病毒和腺伴随病毒。利用这些病毒的基因转移技术是本领域技术人员所熟知的。例如逆转录病毒载体可用于将本发明多核苷酸稳定地整合到宿主基因组，尽管这种重组不是优选的。相反，复制-缺陷型腺病毒载体保持游离状态，并因此允许瞬时表达。可使用能在昆虫细胞（例如杆状病毒载体）、人细胞、酵母或细菌中表达的载体，以产生大量的由本发明多核苷酸编码的 HIV 蛋白，例如可用做亚单位疫苗或用于免疫测定。

根据本发明多核苷酸可用于表达生产编码的蛋白质，其表达可在体外、体内进行或来自体内。因此，核苷酸可涉及重组蛋白质合成，例如增加产量，或实际上可凭其自身的特性用做治疗剂，用于 DNA 疫苗接种技术。本发明的多核苷酸用于体外或来自体内、细胞中生产编码的蛋白质，例如在细胞培养物中进行，则这些细胞被修饰以包含被表达的多核苷酸。这种细胞包括瞬时的、或优选稳定的哺乳动物细胞系。关于插入编码本发明多肽的载体而修饰的细胞例如包括哺乳动物 HEK293T、CHO、HeLa、293 和 COS 细胞。优选地，所选择的细胞系不仅仅是稳定的，而且能对多肽进行成熟糖基化和在细胞表面表达。表达可通过转化卵母细胞来达到。可在转基因非-人动物，优选地是小鼠细胞中表达本发明多核苷酸以获得多肽。将本发明多核苷酸



表达成多肽的转基因非-人动物包含于本发明的范围中。

本发明进一步提供接种哺乳动物受检者的方法，包括给予有效量的这种疫苗或疫苗组合物。最优选地，用于 DNA 疫苗、疫苗组合物和免疫疗法的表达载体为质粒载体。

DNA 疫苗可以“裸 DNA”的形式给药，例如配制成液体用注射器或高压喷射器进行给药，或将 DNA 与脂质体或刺激性转染增强子一起配制，或通过颗粒介导 DNA 释放 (PMDD) 给药。所有这些释放系统为本领域所熟知的。例如通过病毒载体释放系统可将载体导入哺乳动物。

本发明的组合物可通过许多途径来给药释放，例如通过肌肉内、皮下、腹膜内、静脉内或粘膜。

在一个优选的实施方案中，组合物通过皮内给药释放。具体地，组合物通过基因枪给予，尤其是粒子轰击进行给药，它包括在粒子（例如金颗粒）上包被载体，然后该粒子在高压下进入表皮；例如，其方法如 Haynes 等，J Biotechnology 44: 37-42 (1996) 中所描述。

在一个说明性的实施例，气体-推动颗粒加速可利用例如由 Powderject Pharmaceuticals PLC (Oxford, UK) 和 Powderject Vaccines Inc. (Madison, WI) 生产的设备来达到，其中某些例子描述于美国专利 5,846,796、6,010,478、5,865,796、5,584,807 和欧洲专利 0500799 中。这种方法提供了无针给药方法，其中显微颗粒的干粉制剂，例如多核苷酸，在由手动仪器产生的氦气喷射器中被加速到高速，将颗粒推进感兴趣的靶组织，典型的是皮肤。

优选的颗粒是 0.4-4.0 $\mu$ m 的金颗粒，更优选地直径为 0.6-2.0 $\mu$ m，并且 DNA 合成物包被在这些颗粒上，然后装入药筒或弹夹而放入“基因枪”中。

在一个相关的实施方案中，可用于本发明组合物的气体-推动无针注射的其它设备和方法包括由 Bioject, Inc. (Portland, OR) 提供的设备，其中的具体例子描述于美国专利 4,790,824、5,064,413、5,312,335、5,383,851、5,399,163、5,520,639 和 5,993,412 中。

含有编码抗原肽的核苷酸序列的载体将以预防或治疗有效量进行给药。对于颗粒介导释药，给予的量通常为每剂核苷酸为 1 微微克

(picogram)到1毫克,优选地为1微微克到10微克,而对于其它途径,剂量为100纳克到1毫克,优选地为10微克到1毫克。精确剂量很大程度上取决于被免疫的受检者的体重和给药途径。

对于含有编码抗原肽的核苷酸序列的免疫原成分可以一次给予或重复多次给予,例如1-7次,优选地1-4次,间隔为1天到18个月。然而这种治疗方式将很大程度上根据有关受检者的大小(size)、给予核苷酸序列的数量、给予的途径以及其它为熟练兽医或医师所熟知的因素而进行调整。受检者可接受一种或多种其它抗HIV逆转录病毒药物作为它们整个治疗方式中的一部分。此外,核酸免疫原可与佐剂一起给予。

类似地,本文列举的佐剂成分可通过各种不同的途径给予,例如通过口、鼻、肺、肌肉、皮下、皮内或体表给予。优选地,佐剂成分通过皮内或体表途径给予。最优选通过体表给予。佐剂的给予可在核苷酸序列给予前约14天到给予后约14天之间进行,优选地在核苷酸序列给予前约1天到给予后约3天之间。在一个实施方案中,佐剂成分与核苷酸序列的给予基本上同时进行。“基本上同时”是指佐剂成分优选地在核苷酸序列给予的同一时间给予,否则,至少在核苷酸序列给予的前后几个小时内进行给予。在最优选的治疗方案中,佐剂成分与核苷酸序列的给予基本上同时进行。显而易见,根据上述变量的类型,该方案可随需要而改变。优选地佐剂是1H-咪唑并[4,5c]喹啉-4-胺衍生物例如咪喹莫特。典型地,咪喹莫特为体表的膏状制剂,并根据上述方案进行给予。

同样,衍生物给予的剂量也依据这些变量进行调整,但是可在例如,从约0.1mg/kg到约100mg/kg之间,这里“每千克(/kg)”指相关哺乳动物的体重。1H-咪唑并[4,5-c]喹啉-4-胺衍生物的给予优选地随每次核苷酸序列给予后或加强给予而重复。最优选地,给予剂量为约1mg/kg到50mg/kg。在如本文所描述的“初次-加强”方案的情况下,咪喹莫特或其它1H-咪唑并[4,5-c]喹啉-4-胺衍生物可随初次免疫、或加强免疫给予、或在初次和加强免疫中都进行给予。

虽然,佐剂成分可能含有以原始化学状态给予唯一的1H-咪唑并[4,5-c]喹啉-4-胺衍生物,以药物制剂的形式给予是优选的。即佐剂

成分优选地含有与一种或多种药用载体结合的 1H-咪唑并[4,5-c] 喹啉-4-胺衍生物，以及任选的其它治疗成分。载体必须是“可接受的”，是指与制剂中其它成分相兼容，并对其接受者无害。制剂特性将根据给予途径自然地进行调整，并可通过制药领域熟知的方法来制备。所有方法都包括将 1H-咪唑并[4,5-c] 喹啉-4-胺衍生物与适当的载体结合的步骤。通常，通过均匀和紧密地将衍生物与液体载体或良好分离的固体载体或二者结合制备制剂，然后如果需要，将产品制成预期的剂型。适合于口服的本发明制剂可制成分离的单位，例如制成每一个均含有预定量活性成分的胶囊、扁囊剂或片剂；粉剂或颗粒；溶液或存在于水成液或非水成液中的悬浮液；或为水包油性液体乳液或油包水性乳状液。活性成分还可制成大丸剂(bolus)、药糖剂或软膏。

可任选地与一种或多种助剂压缩或塑形(moulding)制备成片剂。压缩的片剂可通过在一个合适的机器中以自由流动形式如粉剂或颗粒的活性成分压缩制备而成，任选地与粘合剂、润滑剂、惰性稀释剂、润滑剂、表面活性剂或分散剂混合。塑形片剂可通过在适当的机器中将惰性液体稀释剂润湿的粉状化合物的混合物塑模制成。

片剂可任选地包被或刻划，并可配制成活性成分缓慢或控制释放的片剂。

用于通过例如肌内、腹膜内或皮下途径注射的制剂包括水性和非水性无菌注射溶液，可含有抗氧化剂、缓冲剂、抑菌剂和与预定的接受者的血液等渗的溶质；而水性和非水性无菌悬浮液可包括悬浮剂和增稠剂。该制剂可制成单剂量或多剂量的容器中，例如密封的安瓿和管瓶，并可保存于仅需要添加无菌液体载体的冻干条件下，例如在使用前立即添加注射用水。可从前面描述的无菌粉剂、颗粒和片剂类型制备临时的注射液和悬浮液。适合于通过口或鼻进行肺部给予的制剂被制成理想地具有 0.5 到 7 微米直径的含有活性成分的颗粒，被释放到接受者的树状支气管。这种制剂可以精细研磨的粉剂形式存在，可方便地制成可穿透(piercable)的胶囊，合适地如明胶，以用于吸入设备，或者可选地，作为一种含有活性成分、一种适当的液体推进剂和任选地其它成分如表面活性剂和/或一种固体稀释剂的自我推进制剂。也可采用活性成分以溶液或悬浮液液滴的形式配制成自我推进制

剂。这种自我推进制剂类似于本领域已知的那些制剂，并可通过成熟方法来制备。手工操作或者用具有所需喷雾特征的自动功能阀来制备，优选地，该阀是具有能测量传递固定体积的类型，例如每次操作的固定体积为 50-100 $\mu$ L。

另一种可能性，佐剂成分可以液体的形式用于喷雾器或雾化器，其中利用加速气流或超声搅拌产生用于吸入的微滴薄雾。

适合于经鼻给予的制剂通常包括与上述用于肺部给予的制剂相似的制剂，对于这种制剂优选地具有约 10-200 微米的颗粒直径，使之能够保持于鼻腔内。相应地，可通过使用颗粒大小适当的粉剂或选择适当的阀来获得。其它适当的制剂包括具有颗粒直径为约 20-500 微米的粗粉，用于从与鼻靠近的容器中迅速吸入鼻内而经鼻通道给予，以及在水溶液或油性溶液中含有约 0.2-5% w/w 的活性成分的鼻滴剂。在本发明的一个实施方案中，含有编码抗原肽的核苷酸序列的载体可以与 1H-咪唑并[4,5-c] 喹啉-4-胺衍生物采用相同的制剂中给予。因此，在这个实施方案中，免疫原性成分和佐剂成分并存于相同的制剂中。

在一个实施方案中，佐剂成分以适合于基因枪给予的形式制备，并经该途径与核苷酸序列基本上同时施用。为了配制适合于以这种方式使用的制剂，对于 1H-咪唑并[4,5-c] 喹啉-4-胺衍生物可能需要冻干并附着在例如适合于基因枪给予的金小珠上。

在另一可选实施方案中，佐剂成分可作为干粉经高压气体推进器给予。

即使没有配制在一起，对于佐剂成分可能在核苷酸序列给予位点或者其基本上相同的给予位点进行给予是合适的。

药物制备的其它细节可在 Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pennsylvania (1985) 中找到，其公开的全部内容被本文引入作为参考。

用于将裸多核苷酸或载体引入患者的合适技术也包括以适当的方式进行局部给予。核酸可经皮肤局部给予，或粘膜表面，例如经鼻内、口腔、阴道内或直肠内给予。裸多核苷酸或载体可与药用的赋形剂例如磷酸缓冲盐 (PBS) 一起存在。通过使用易化剂例如丁呱卡而进一

步促进 DNA 的摄取, 该易化剂或者单独使用或者包含于 DNA 制剂中。其它将核酸直接给予接受者的方法包括超声、电刺激、电穿孔和显微接种, 在美国 5,697,901 中对此有描述。

可通过一些已知的转染技术增强对核酸构建体的吸收, 例如包括使用转染试剂的技术。这些试剂的例子包括阳离子试剂, 例如, 磷酸钙和 DEAE-葡聚糖和脂转染试剂 (lipofectant), 例如 lipofectam 和 transfectam。给予核酸的剂量可以进行调整。

本发明的核酸序列还可通过基因疗法中使用的特异性释放载体给予。基因治疗方法例如由 Verme 等在 Nature 1997, 389: 239-242 中描述。病毒和非病毒载体系统均可使用。基于病毒的系统包括基于逆转录病毒、慢病毒、腺病毒、腺伴随病毒、疱疹病毒、Canarypox 和痘苗-病毒系统。基于非病毒的系统包括核酸的直接给予、微球包裹技术 (聚 (丙交酯-共-乙交酯)) 和基于脂质体的系统。病毒和非病毒释放系统可结合使用, 其中在初始接种后进行强化注射是理想的, 例如使用非病毒载体例如质粒进行最初的 “初次” DNA 接种, 随后使用病毒载体或基于非病毒的系统进行一次或多次 “加强” 接种。同样地, 本发明设计了具有本发明多核苷酸的初次免疫导入系统, 随后用存在于佐剂中的蛋白质进行加强免疫, 反之亦然。

本发明的核酸序列还可通过转化细胞给予。这种细胞包括从被试者收集的细胞。在体外将本发明的裸多核苷酸或载体导入这种细胞, 随后将该转化的细胞返回到被试者体内。本发明的多核苷酸可通过同源重组整合进入细胞中已存在的核酸中。如果需要, 转化细胞可在体外生长, 并且获得的一个或多个细胞可用于本发明。可通过已知的外科或显微外科技术 (例如移植、显微注射等) 在患者适当的部位给予这种细胞。

本发明的药物组合物可包括如上所述的佐剂化合物, 或其它可用于提高由 DNA 编码的蛋白质诱导的免疫反应的物质。这些物质可由 DNA 编码, 或者单独存在或与抗原进行了融合, 或者作为非 DNA 元件包含在制剂中。可包含于本发明制剂中的佐剂型物质的例子包括泛素、溶酶体伴随膜蛋白 (LAMP)、乙肝病毒核心抗原、FLT3-配体 (一种在消化性抗原呈递细胞发生中重要的细胞因子, 具体是树枝状

细胞)以及其它细胞因子例如 IFN- $\gamma$ 和 GM-CSF。其它优选的佐剂包括咪喹莫特、Resimquimod 和 Tucarasol。咪喹莫特是特别优选的。

在本发明一个优选的实施方案中,提供了如本文所描述的核酸分子在治疗或预防 HIV 感染中的用途。核酸分子优选地与咪喹莫特配合给予。咪喹莫特优选局部给予,而核酸分子优选通过颗粒介导释放的方式给予。

因此,本发明提供了治疗患有 HIV 或易于遭受 HIV 感染的被检者的方法,包括给予本发明描述的核酸分子和咪喹莫特。

现在本发明将参考如下实施例作进一步的描述:

### 实施例

实施例 1: 优化 p55 gag (p17、p24、p13) 以与高度表达的人类基因的密码子利用率相似

#### 感兴趣的基因

用于哺乳动物细胞中表达的编码 HIV-1 分化体 B 菌株 HXB2 的 p55gag 抗原的优化合成基因 (GenBank 记录号 K03455), 利用 PCR 扩增的重叠寡核苷酸进行装配。

优化涉及改变病毒基因的密码子使用模式以获得与在高度表达的人类基因中发现的密码子频率接近的密码子频率。利用被称做 Syngene 的统计学 Visual Basic 程序确定密码子 (由 R.S. Hale 和 G. Thompson 编写的 Calcgene 的最新版本, 蛋白质表达和纯化, 12 卷, 185-188 页, 1998)。

#### 克隆:

1528bp 的 gag PCR 产物通过凝胶纯化, 用限制性核酸内切酶 Not I 和 Bam HI 酶切, 并连接到 NotI/BamHI 酶切的载体 WRG7077 中。即将基因置于 CMV 启动子/内含子 A 和牛生长素多腺苷酸化信号之间。

对克隆进行测序并检查存在的错误。没有一个克隆是 100%正确的。因此, 通过重叠 PCR 将两个克隆中正确序列的区域结合, 其重叠 PCR 利用优化的寡核苷酸的适当组合以得到密码子优化的全长 gag 基因。随后发现最终克隆中含有一个核苷酸缺失, 这导致移码和翻译提前终止。通过切除含有不正确序列的基因区域, 并将另一个克隆中

含有正确序列的相应区域克隆进来，而修复了上述缺失。这样得到密码子被优化 p55 gag 的最终克隆：Gagoptpr2。（见图 2）。

#### 实施例 2：制备 p17/p24 截短的 Nef 融合基因

##### 感兴趣的基因

通过 PCR 从质粒 pHXB?Pr (B.Maschera, E Furfine 和 E.D.Blair 1995 J.Virol 69 5431-5436) 扩增衍生自 HIV-1 分化体 B 菌株 HXB2 的 p55gag 基因的 p17 和 p24 部分。从该质粒中扩增来自 HXB2 nef 基因的 3' 末端的 pHXB?Pr. 426bp。由于 HXB2 nef 基因含有提前的终止密码子，因此用两个重叠 PCR 修复密码子（TGA[终止]到 TGG[Trp]）。

p17/p24 接头和 trNEF 接头的 PCR 产物通过一个 PCR 反应而接合形成 p17p24trNEF 融合基因（图 3）（反义）。

1542bp 的产物通过凝胶纯化，用限制性核酸内切酶 NotI 和 BamHI 酶切，并克隆进载体 WRG7077 的 NotI BamHI 位点。即将基因置于 CMV 启动子/内含子 A 和牛生长素多腺苷酸化信号之间。

#### 实施例 3：制备 Gag p17/24opt/trNef1（'Gagopt/Nef'）融合基因

##### 感兴趣的基因

以质粒 pGagOPTpr2 为模板进行 PCR 扩增，获得密码子优化的衍生自 HIV-1 分化体 B 菌株 HXB2 的 p55gag 基因的 p17 和 p24 部分。通过 PCR 从质粒 7077trNef20 扩增具有修复的提前终止密码子（TGA[终止]到 TGG[Trp]）的截短的 HXB2 Nef 基因。这两个 PCR 产物被设计成具有重叠的末端，因而这两个基因在第二轮 PCR 中发生接合。

1544bp 的产物通过凝胶纯化，用限制性核酸内切酶 NotI 和 BamHI 酶切，并克隆（见图）进载体 WRG7077 的 NotI BamHI 位点。即将基因置于 CMV 启动子/内含子和牛生长素多腺苷酸化信号之间。

#### 实施例 4：质粒：p7077-RT3 克隆#A

##### 感兴趣的基因

编码 HIV-1 分化体 B 菌株 HXB2 的 pol 基因的 RT 部分，被优化

用于在哺乳动物细胞中表达的合成基因是通过 PCR 扩增的重叠寡核苷酸进行装配。克隆的序列与 HXB2 参考序列 (GeneBank 记录号 K03455) 的位置 2550-4222 相应。为了确保表达, 在克隆序列的 5' 末端具有两个附加的不存在于原来基因中的密码子—AUG GGC (Met Gly)。

优化涉及改变病毒基因的密码子使用模式, 以得到与在高度表达的人类基因中发现的密码子频率相接近的密码子频率, 但不包括很少使用的密码子。使用被称做 Syngene 的统计学 Visual Basic 程序确定密码子 (由 R.S. Hale 和 G. Thompson 编写的 Calcgene 的最新版本, 蛋白质表达和纯化, 12 卷, 185-188 页, 1998)。

通过两个中间克隆, 即 #16 和 #21 而构建成最终克隆。

克隆:

1.7kb 的 PCR 产物通过凝胶纯化, 用 NotI 和 BamHI 酶切并 PCR 净化, 然后与 NotI/BamHI 酶切的 pWRG7077 连接。这将基因置于 CMV 启动子和牛生长素多腺苷酸化信号之间。对克隆进行测序。没有任何克隆是 100% 正确的, 但通过用来自克隆 #21 的正确的 KpnI-BamHI 片段取代克隆 #16 中含有 3 个错误的 403bp 的 KpnI-BamHI 片段, 这样克隆 #16 被修复。通过测序证实最终克隆。(见图 5)

#### 实施例 5: 优化的 RT

##### 感兴趣的基因

编码 HIV-1 分化体 B 菌株 HXB2 的 pol 基因的 RT 部分, 被优化用于在哺乳动物细胞中表达的合成基因是从质粒 p7077-RT3 切割获得, 即 1697bp 的 NotI/BamHI 片段, 通过凝胶纯化, 并克隆进入 p7313-ie (衍生自 pspC31) 的 NotI 和 BamHI 位点, 而将基因置于 Iowa 长 HCWV 启动子+外显子 1 的下游, 和兔球蛋白多腺苷酸化信号的上游。(R7004 p27) (图 6)

#### 实施例 6

质粒: 7077trNef20

##### 感兴趣的基因



插入片段含有来自 HIV-1 分化体 B 菌株 HXB2 的 Nef 基因部分。从基因的 5' 末端去除 195bp, 即除去 Nef 的最初 65 个氨基酸的密码子。另外, 如前面描述的质粒 p17/24trNEF1 的构建中, 将已公开的 HXB2 nef 序列中的提前终止密码子进行了修复 (TAG 到 TGG[Trp])。截短的 nef 序列是从质粒 p17/24trNef1 扩增获得。克隆序列与 HXB2 参考序列 (GenBank 记录号 K03455) 的位置 8992-9417 相应。为了确保表达, 在克隆序列 5' 末端具有附加的在原来基因中不存在的密码子—AUG (Met)。

引物:

StrNef (有义) ATAAGAATGCGGCCGCCATGGTGGGTTTTCCAGTCACACCTT  
[SEQ ID NO: 1]  
AstrNef (反义)  
CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEQ ID NO: 2]

PCR: 94℃ 2 分钟, 然后按如下参数进行 25 个循环: 94℃ 30 秒, 50℃ 30 秒, 72℃ 2 分钟, 最后 72℃ 5 分钟。

克隆:

455bp 的 RT PCR 产物通过凝胶纯化, 用限制性核酸内切酶 NotI 和 Bam HI 酶切, 并连接到 NotI/BamHI 酶切的载体 WRG7077 上。即将基因置于 CMV 启动子/内含子 A 和牛生长素多腺苷酸化信号之间。

## 实施例 7

质粒: 7077RT 8

感兴趣的基因

pol 基因的 RT 部分衍生自 HIV-1 分化体 B 菌株 HXB2。通过 PCR 从质粒 p7077Pol14 扩增获得。

克隆序列与 HXB2 参考序列 (GeneBank 记录号 K03455) 的位置 2550-4234 相应。为了确保表达, 克隆序列在 5' 末端具有两个附加的不存在于原来基因中的密码子—AUG GGC (Met Gly)。

引物:

SRT (有义) ATAAGAATGCGGCCGCCATGGGCCCCATTAGCCCTATTGAGACT

[SEQ ID NO: 3]

ASRT (反义)

CGCGGATCCTTAATCTAAAAATAGTACTTTCCTGATT [SEQ ID NO: 4]

PCR: 94℃2 分钟, 然后按如下参数进行 25 个循环: 94℃30 秒, 50℃30 秒, 72℃4 分钟, 最后 72℃5 分钟。

克隆:

1720bp 的 RT PCR 产物通过凝胶纯化, 用限制性核酸内切酶 NotI 和 Bam HI 酶切, 并连接到 NotI/BamHI 酶切的载体 WRG7077 上。即将基因置于 CMV 启动子/内含子 A 和牛生长素多腺苷酸化信号之间。

### 实施例 8

p17/24opt/RT/trNef13 ( 'Gagopt/RT/Nef' )

该构建体含有引起由 R 到 H 氨基酸改变的 PCR 产物。

感兴趣的基因

衍生自 HIV-1 分化体 B 菌株 HXB2 的优化 p55gag 基因的密码子的 p17/p24 部分从质粒 pGagOPTpr2 经 PCR 扩增获得。RT 编码序列从质粒 7077RT 8 经 PCR 扩增获得。具有修复的提前终止密码子 (TGA[终止]到 TGG[Trp]) 的截短的 HXB2 Nef 基因经 PCR 从质粒 7077trNef20 扩增得到。三个 PCR 产物被设计为具有重叠末端, 因此这三个基因在第二轮 PCR 中发生连接。

引物:

( P17/24 )

Sp17p24opt (有义)

ATAAGAATGCGGCCGCCATGGGTGCCCCGAGCTTCGGT [SEQ ID NO: 5]

ASp17p24optRT 接头 (反义)

TGGGGCCCATCAACACTCTGGCTTTGTGTC [SEQ ID NO: 6]

PCR: 94℃1 分钟, 然后按如下参数进行 20 个循环: 94℃30 秒, 50℃30 秒, 72℃2 分钟, 最后 72℃4 分钟。

1114bp 的 p17/24opt 产物通过凝胶纯化。

( RT )

Sp17p24optRT 接头 (有义)

CAGAGTGTGGATGGGCCCATTTAGCCCTAT [SEQ ID NO: 7]

ASRTtrNef 接头 (反义)

AACCCACCATATCTAAAAATAGTACTTTCC [SEQ ID NO: 8]

PCR: 如上文

1711bp 的 RT PCR 产物通过凝胶纯化。

(5'截短的 nef)

SRTtrNef 接头 (有义)

CTATTTTGTAGATATGGTGGGTTTCCAGTCAC [SEQ ID NO: 9]

AstrNef (反义)

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEQ ID NO: 10]

PCR 如上文

448bp 的产物通过凝胶纯化。

然后 Sp17/24opt 和 AstrNef 为引物进行第二轮 PCR 将上述三个 PCR 产物接合在一起。

PCR: 94℃ 1 分钟, 然后按如下参数进行 30 个循环: 94℃ 30 秒, 50℃ 30 秒, 72℃ 4 分钟, 最后 72℃ 4 分钟。

3253bp 的产物通过凝胶纯化, 用限制性核酸内切酶 NotI 和 Bam HI 酶切, 并克隆到载体 WRG7077 的 NotI BamHI 位点。即将基因置于 CMV 启动子/内含子 A 和牛生长素多腺苷酸化信号之间。

#### 实施例 9

质粒: pGRN#16 (p17/p24opt corr/RT/trNef.)

感兴趣的基因

由于在 p24 中氨基酸 270 附近存在不利的密码子簇, 由 p17/24opt /RT/trNef13 ('Gagopt/RT/Nef') 产生的多蛋白被表达为约 30kDa 的截短产物。这些密码子被 PCR 接合诱变产生的优选密码子所取代。以 p17/24opt/RT/trNef13 作模板, Sp17/p24opt 和 GTR-A 为引物扩增

Gag 5'端的部分以产生突变体，和以 GTR-S 和 Asp17/p24optRT 接头为引物扩增 Gag 3'端的部分以产生突变体。利用 Sp17/p24opt 和 Asp17/p24optRT 接头引物将包含了改变的密码子的重叠产物和凝胶纯化的产物连接在一起。所得到的产物用 NotI 和 AgeI 酶切，并插入用相同酶切割的 p17 / 24opt/ RT/trNef 13 中，获得 pGRN。对克隆#16 进行验证并用于进一步的研究。

#### 引物:

##### 5' PCR:

Sp17p24opt (有义)

ATAAGAATGCGGCCCGCCATGGGTGCCCCGAGCTTCGGT [SEQ ID NO: 11]

GTR-A (反义)

GCGCACGATCTTGTTTCAGGCCAGGATGATCCACCGTTTATAGATTTCTCC [SEQ ID NO: 12]

##### 3' PCR

有义: GTR-S (有义)

ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCC [SEQ ID NO: 13]

Asp17p24optRT 接头 (反义)

TGGGGCCCATCAACACTCTGGCTTTGTGTC [SEQ ID NO: 14]

用于各个产物和连接，使用 PWO DNA 聚合酶 (Roche) 的 PCR 条件是:

95℃ 1 分钟，然后按下述参数进行 20 个循环: 95℃ 30 秒，55℃ 30 秒，72℃ 180 秒，最后 72℃ 120 秒，在 4℃ 保持。

1114bp 的产物通过凝胶纯化，并用 NotI 和 AgeI 酶切产生 6647bp 的片段，经凝胶纯化后连接到 NotI/AgeI 酶切的、凝胶纯化的 p17/24opt/RT/trNef13 中，获得 pGRN#16。

#### 实施例 10:

质粒: p73i-GRN2 克隆#19

(p17/p24 (opt) /RT (opt) trNef) -修复的  
感兴趣的基因

来自于 HIV -1 分化体 B 菌株 HXB2 的密码子被优化的 gag 的 p17/p24 部分、密码子优化的 RT 和截短的 Nef 基因位于 Iowa 长 HCMV 启动子+外显子 1 的下游和兔  $\beta$ -球蛋白多腺苷酸化信号的上游之间。

包含衍生自质粒 p17/24trNef1 的 trNef 基因的质粒包含一个 PCR 错误，导致从 nef 末端的 19 个氨基酸的 R 到 H 的氨基酸改变。这通过 PCR 诱变得以校正，校正的 nef PCR 与 p7077-RT3 的密码子优化 RT 进行连接，并用 ApaI 和 BamHI 酶切连接的片段，并克隆到用 ApaI/BamHI 酶切的 p73i-GRN 中。

引物:

利用引物进行 PCR 以从 p7077-RT3 获得 coRT:

(聚合酶=全部为 PWO (Roche) )

有义: U1

GAATTTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTG

AAACCCGGGAT [SEQ ID NO: 15]

AScoRT-Nef

GGTGTGACTGGAAAACCCACCATCAGCACCTTTCTAATCCCCGC [SEQ ID NO:

16]

循环: 95°C (30 秒), 然后如下参数进行 25 个循环: 95°C (30 秒), 55°C (30 秒), 72°C (180 秒), 然后 72°C (120 秒), 最后保持在 4°C。

1.7kp 的 PCR 产物通过凝胶纯化。

利用下述引物进行 PCR 以从 p17/24trNef1 获得 5'Nef:

有义: S-Nef

ATGGTGGGTTTTCCAGTCACACC [SEQ ID NO: 17]

反义 : ASNef-G:

GATGAAATGCTAGGCGGCTGTCAAACCTC [SEQ ID NO: 18]

循环: 95°C (30 秒), 然后按如下参数进行 15 个循环: 95°C (30

秒), 55℃ (30 秒), 72℃ (60 秒), 然后 72℃ (120 秒), 最后保持在 4℃。

利用如下引物进行 PCR 以从 p17/24trNef1 获得 3'Nef:

有义: SNEF-G

GAGGTTTGACAGCCGCCTAGCATTTTCATC [SEQ ID NO: 19]

反义:

AstrNef (反义)

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEQ ID NO: 20]

循环: 95℃ (30 秒), 然后按如下参数进行 15 个循环: 95℃ (30 秒), 55℃ (30 秒), 72℃ (60 秒), 然后 72℃ (120 秒), 最后保持在 4℃。

PCR 产物通过凝胶纯化。使用 5' (S-Nef) 和 3' (AstrNef) 引物将两个最初 Nef 产物连接。

循环: 95℃ (30 秒), 然后按如下参数 15 个循环: 95℃ (30 秒), 55℃ (30 秒), 72℃ (60 秒), 然后 72℃ (180 秒), 最后保持在 4℃。

PCR 产物被 PCR 净化, 并用 U1 和 AstrNef 引物将其与 RT 产物连接:

循环: 95℃ (30 秒), 然后按如下参数进行 20 个循环: 95℃ (30 秒), 55℃ (30 秒), 72℃ (180 秒), 然后 72℃ (180 秒), 最后保持在 4℃。

2.1kb 的产物通过凝胶纯化, 并用 ApaI 和 BamHI 酶切。质粒 p73I-GRN 也用 ApaI 和 BamHI 酶切、凝胶纯化并与 ApaI-Bam RT3trNef 连接产生 p17/p24 (opt) /RT (opt) trNef 基因。

## 实施例 11

p73i-GN2 克隆 #2 (p17/p24opt/trNef) -修复的  
感兴趣的基因:

来自于 HIV -1 分化体 B 菌株 HXB2 的密码子优化的 gag 的 p17/p24 部分、截短的 Nef 基因位于 Iowa 长 HCMV 启动子+外显子 1

的下游和兔 $\beta$ -球蛋白多腺苷酸化信号的上游之间。

包含衍生自质粒 p17/24trNef1 的 trNef 基因的质粒含有一个 PCR 错误，导致从 nef 末端的 19 个氨基酸 R 到 H 的氨基酸改变。这通过 PCR 诱变得以校正，校正的片段用 BglII 和 BamHI 酶切，并克隆进 BglII/BamHI 酶切的 p73I-GN 中。（图 12）重建了 Iowa 长 HCMV 启动子+外显子 1 的下游以及兔 $\beta$ -球蛋白多腺苷酸化信号的上游之间经校正的 p17/p24opt/trNef 所产生的融合基因。

利用下述引物以从 p17/24trNef1 获得 5' Nef:

聚合酶=全部为 PWO (Roche)

有义: S-Nef

ATGGTGGGTTTTCCAGTCACACC [SEQ ID NO: 21]

反义: ASNef-G:

GATGAAATGCTAGGCGGCTGTCAAACCTC [SEQ ID NO: 22]

循环: 95°C (30 秒)，然后按如下述参数进行 15 个循环: 95°C (30 秒)，55°C (30 秒)，72°C (60 秒)，然后 72°C (120 秒)，最后保持在 4°C。

利用下述引物以从 p17/24trNef1 获得 3' Nef:

有义: SNEF-G

GAGGTTTGACAGCCGCTAGCATTTCATC [SEQ ID NO: 23]

反义: AstrNef

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEQ ID NO: 24]

循环: 95°C (30 秒)，然后按如下参数进行 15 个循环: 95°C (30 秒)，55°C (30 秒)，72°C (60 秒)，然后 72°C (120 秒)，最后保持在 4°C。

PCR 产物通过凝胶纯化，并用 5' (S-Nef) 和 3' (AstrNef) 引物进行连接。

循环: 95°C (30 秒)，然后按如下参数进行 15 个循环: 95°C (30 秒)，55°C (30 秒)，72°C (60 秒)，然后 72°C (180 秒)，最后

保持在 4℃。

PCR 产物经 PCR 净化，用 BglII/BamHI 酶切，得到的 367bp 片段经凝胶纯化后，克隆到 BglII/BamHI 酶切、凝胶纯化的 p73i-GN 中。

#### 实施例 12

质粒：p73I-RT w229k （失活的 RT）

感兴趣的基因：

获得 Iowa 长 HCMV 启动子+外显子 1 的下游和兔 β-球蛋白多腺苷酸化信号的上游之间的失活 RT 基因。

考虑到在治疗性疫苗中使用活性 HIV RT 的种类，需要将该基因失活。这可通过 PCR 诱变而将 RT（衍生自 P73I-GRN2）第 229 位氨基酸从 Trp 改变为 Lys（R7271 p1-28）来达到。

引物

PCR 5' RT + 突变用引物：

（聚合酶=全部为 PWO（Roche））

有义：RT3-u:1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTG

AAACCCGGGAT [SEQ ID NO: 25]

反义：AScoRT-Trp229Lys

GGAGCTCGTAGCCCATCTTCAGGAATGGCGGCTCCTTCT [SEQ ID NO: 26]

循环：

1 × [ 94℃ ( 30 秒 ) ]

15 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 60 秒 ) ]

1 × [ 72℃ ( 180 秒 ) ]

PCR 凝胶纯化



**PCR 3' RT + 突变用引物:**

反义 : RT3- 1:1

GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCAGCTTGTCGACCTG  
CTCGTTGCCGC [SEQ ID NO: 27]

有义 : ScoRT-Trp229Lys

CCTGAAGATGGGCTACGAGCTCCATG [SEQ ID NO: 28]

**循环:**

1 × [ 94℃ ( 30 秒 ) ]

15 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 60 秒 ) ]

1 × [ 72℃ ( 180 秒 ) ]

**PCR 凝胶纯化**

PCR 产物经凝胶纯化, 用 5' (RT3-U1) 和 3' (RT3-L1) 引物将 RT 的 5'和 3'末端连接。

**循环:**

1 × [ 94℃ ( 30 秒 ) ]

15 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 120 秒 ) ]

1 × [ 72℃ ( 180 秒 ) ]

PCR 产物经凝胶纯化, 并利用 NotI 和 BamHI 限制性位点克隆进入 p7313ie, 产生 p73I-RT w229k。(见图 13)

**实施例 13**

质粒: p73i-Tgrn (#3)

**感兴趣的基因:**

来自 HIV -1 分化体 B 菌株 HXB2 的密码子被优化的 gag 的 p17/p24 部分、密码子优化 RT 和截短的 Nef 基因位于 Iowa 长 HCMV 启动子+外显子 1 的下游和兔β-球蛋白多腺苷酸化信号的上游之间。

含有活性形式 RT 的三重融合的构建体对于用于人类的调节目可能不被接受, 因此通过将 p73i-RT w229k 的 NheI 和 ApaI 酶切片段插入到 NheI/ApaI 酶切的 p73i-GRN2#19 中而获得失活的 RT(图 14)。这导致在 RT 的位置 229 上发生 W 改变为 K。

## 实施例 14

## p73I-Tnrg (#16)

感兴趣的基因:

来自 HIV -1 分化体 B 菌株 HXB2 的截短的 Nef、密码子优化的失活 RT 和密码子优化的 gag 基因的 p17/p24 部分位于 Iowa 长度 HCMV 启动子+外显子 1 的下游以及兔 $\beta$ -球蛋白多腺苷酸化信号的上游之间。

由 p73i-Tgrn 编码的多蛋白中的基因顺序通过 PCR 和 PCR 结合重排以产生 p73I-Tnrg (图 15)。在基因的 PCR 结合以形成单一的多蛋白之前,每个基因都经 PCR 扩增并经凝胶纯化。该产物经凝胶纯化、用 NotI/BamHI 消化后连接到 NotI/BamHI 酶切的 p7313ie 中。

引物:

trNef PCR

S-Nef (Not I)

CATTAGAGCGGCCGCGATGGTGGGTTTTCCAC [SEQ ID NO: 29]

AS-Nef-coRT 接头

GATGGGACTGATGGGGCCCATGCAGTTCTTGAAGTACTCCGG [SEQ ID NO:  
30]

RTw229k PCR

S-coRT

ATGGGCCCCATCAGTCCCATCGAG [SEQ ID NO: 31]

AS-coRT-p17p24 接头

CAGTACCGAAGCTCGGGCACCCATCAGCACCTTTCTAATCCCCGC [SEQ ID NO:  
32]

p17p24opt PCR**S-p17p24opt**

ATGGGTGCCCCGAGCTTCGGTACTG [SEQ ID NO: 33]

**AS-p17p24opt (BamHI)**

GATGGGGGATCCTCACAACACTCTGGCTTTGTGTCC [SEQ ID NO: 34]

用 VENT DNA 聚合酶 (NEB) 的对于各个产物和连接的 PCR 条件如下:

1 × [ 94℃ ( 30 秒 ) ]

25 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 120 秒 [p17p24 或 RT]  
或 60 秒 [trNef] ) ]

1 × [ 72℃ ( 240 秒 ) ]

PCR 产物经凝胶纯化, 并用引物 S-trNef (NotI) 和 AS- p17p24opt  
(BamHI) 进行 PCR 结合。

1 × [ 94℃ ( 30 秒 ) ]

25 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 210 秒 ) ]

1 × [ 72℃ ( 240 秒 ) ]

3000bp 的产物经凝胶纯化, 并用 NotI 和 BamHI 酶切后经 PCR  
净化, 连接到 NotI/BamHI 消化的、凝胶纯化的 p7313ie 中以产生  
p73i-Tnrg。

**实施例 15:****1. 质粒: P73i-Tngr (#3)****感兴趣的基因:**

来自 HIV -1 分化体 B 菌株 HXB2 的截短的 Nef、密码子优化的  
gag 的 p17/p24 部分、密码子优化的失活 RT 基因位于 Iowa 长 HCMV  
启动子+外显子 1 的下游的兔β-球蛋白多腺苷酸化信号的上游之间获  
得的。

p73i-Tgrn 编码的多蛋白中的基因顺序通过 PCR 重排以产生 p73I-

Tngr (图 16)。密码子优化的 p17/p24 和 RT 产生单一产物, 并经 PCR 接合到扩增的 trNef 上。该产物经凝胶纯化、NotI/BamHI 消化并连接到 NotI/BamHI 酶切的 p7313ie 中。

引物:

P17/p24-RT 3'PCR

Sp17p24opt (有义)

ATGGGTGCCCCGAGCTTCGGTACTG [SEQ ID NO: 35]

RT3 1:1 (反义)

GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCAGCTTGTCGACCTG

CTCGTTGCCGC [SEQ ID NO: 36]

TrNef 5'PCR

S-Nef (NotI)

CATTAGAGCGGCCGCGATGGTGGGTTTTCCAC [SEQ ID NO: 37]

AS-Nef-p17p24

CAGTACCGAAGCTCGGGCACCCATGCAGTTCTTGAACTACTCCGG [SEQ ID NO: 38]

用 VENT DNA 聚合酶 (NEB) 的对于各个产物和连接的 PCR 条件如下:

1 × [ 94°C ( 30 秒 ) ]

25 × [ 94°C ( 30 秒 ) / 55°C ( 30 秒 ) / 72°C ( 180 秒 [p17p24+RT] 或 60 秒 [trNef]或 210 秒[连接] ) ]

1 × [ 72°C ( 240 秒 ) ]

3000bp 的产物经凝胶纯化, 并用 NotI 和 BamHI 酶切, 经 PCR 净化后, 连接到 NotI/BamHI 消化的、凝胶纯化的 p7313ie 中以产生 p73i-Tngr。

实施例 16:

质粒: p73I-Trgn (#6)

感兴趣的基因:

来自 HIV -1 分化体 B 菌株 HXB2 密码子优化的失活 RT、密码子优化的 gagp17/p24 部分和截短的 Nef 基因位于 Iowa 长 HCMV 启动子+外显子 1 的下游和兔 $\beta$ -球蛋白多腺苷酸化信号的上游之间。

构建体内基因的顺序通过 PCR 分别扩增来自质粒 p73I-GN2 和 p73I-RTw229k 的 p17p24-trNef 和 RTw229k 达到。进行 PCT 连接反应, 其产物经凝胶纯化并用 NotI/BamHI 酶切, 然后与 NotI/BamHI 消化的 p73I3ie 进行连接。序列测定揭示 p17p24 没有被完全优化, 故用 AgeI/MunI 从编码区切去 700bp 的片段, 被而用来自含有正确编码序列的 p73i-Tgrn#3 的 MunI/Age 片段进行取代。(见图 17)

引物:

#### p17p24-trNef PCR

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [SEQ ID NO: 39]

AstrNef ( BamHI )

#### RTw229k

RT3-U:1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTG  
AAACCCGGGAT [SEQ ID NO: 40]

AS-coRT-p17p24opt 接头

CAGTACCGAAGCTCGGGCACCCATCAGCACCTTTCTAATCCCCGC [SEQ ID NO:  
41]

用 VENT DNA 聚合酶 (NEB) 的对于各个产物和连接的 PCR 条件如下:

1 × [ 94°C ( 30 秒 ) ]

25 × [ 94°C ( 30 秒 ) / 55°C ( 30 秒 ) / 72°C ( 120 秒 ( PCR ) 或 180  
秒 ( 连接 ) ]

1 × [72℃ (240 秒)]

PCR 接合产生的 3000bp 产物经凝胶纯化, 并用 NotI 和 BamHI 酶切, 经 PCR 净化后, 连接到 NotI/BamHI 消化的、凝胶纯化的 p7313ie 中以产生 p73i-Tngr。序列分析显示从 p73I-GN2 获得的 p17p24 序列密码子没有完全被优化, 这被携带到新的质粒中。从 MunI 和 AgeI 酶切的 p73i-Tngr 中切掉 700bp 的片段, 用来自 p73i-Tgrn 的 700bp 的 MunI/AgeI 消化产物进行连接而取代它, 产生构建体 p73I-Tngr#6。

实施例 17:

质粒: p73i-Trng (#11)

感兴趣的基因:

Iowa 长度 HCMV 启动子+外显子 1 下游, 以及兔β-球蛋白多腺苷酸化信号上游的、从 HIV-1 分化体 B 菌株 HXB2 获得的密码子优化的失活 RT、截短的 Nef 和密码子优化的 gag 基因的 p17/p24 部分。

构建体内基因的顺序通过 PCR 扩增来自 p73i-Tgrn 的 RT-trNef 和 p17p24 基因而获得。进行两个 DNA 片段的 PCT 连接反应, 其 3kb 的产物经凝胶纯化, 并用 NotI/BamHI 酶切, 然后与 NotI/BamHI 消化的 p7313ie 进行连接, 产生 p73I Trng (#11)

引物:

RTw229k-trNef

RT3-u:1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTG  
AAACCCGGGAT [SEQ ID NO: 42]

AS-Nef-p17p24opt 接头

CAGTACCGAAGCTCGGGCACCCATGCAGTTCTTGAACTACTCCGG [SEQ ID NO:  
43]

p17p24

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [SEQ ID NO: 44]

### AS-p17p24opt (BamHI)

GATGGGGGATCCTCACAACACTCTGGCTTTGTGTCC [SEQ ID NO: 45]

用 VENT DNA 聚合酶 (NEB) 的对于各个产物和连接的 PCR 条件如下:

1 × [ 94℃ ( 30 秒 ) ]

25 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 120 秒 ( 基因的 PCR )  
或 180 秒 ( 连接 ) ]

1 × [ 72℃ ( 240 秒 ) ]

PCR 接合产生的 3000bp 产物经凝胶纯化, 用 NotI 和 BamHI 酶切, 经 PCR 净化后, 连接到 NotI/BamHI 消化的、凝胶纯化的 p7313ie 中, 以产生 p73i-Tngr.

### 实施例 18:

p73i-Tngr (#f1)

感兴趣的基因:

Iowa 长度 HCMV 启动子+外显子 1 下游以及兔β-球蛋白多腺苷酸化信号上游的、从 HIV -1 分化体 B 菌株 HXB2 获得的密码子优化的 gag 的 p17/p24 部分、截短的 Nef 和密码子优化的失活 RT 基因。

构建体内基因的顺序通过 PCR 分别扩增来自质粒 p73I-GN2 和 p73I-RTw229k 的 p17p24-trNef 和 RTw229k 来获得。其对 PCR 产物进行 PCT 连接反应, 经凝胶纯化, 并用 NotI/BamHI 酶切后, 与 NotI/BamHI 消化的 p7313ie 进行连接。在序列 (p17p24 和 RT) 中发现两个序列错误, 随后这种错误通过利用多蛋白中的限制性位点用正确的基因部分进行取代而得到修复。(见图 19)

引物:

p17p24-trNef PCR

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [SEQ ID NO: 46]

### AS-Nef-coRT 接头

GATGGGACTGATGGGGCCCATGCAGTTCTTGAACTACTCCGG [SEQ ID NO:  
47]

### RTw229k

#### S-coRT

ATGGGCCCCATCAGTCCCATCGAG [SEQ ID NO: 48]

### RT3-1:1

GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCAGCTTGTGACCTG  
CTCGTTGCCGC [SEQ ID NO: 49]

用 VENT DNA 聚合酶 (NEB) 的对于各个产物和连接的 PCR 条件如下:

1 × [ 94℃ ( 30 秒 ) ]

25 × [ 94℃ ( 30 秒 ) / 55℃ ( 30 秒 ) / 72℃ ( 120 秒 ( PCR ) 或 180  
秒 ( 连接 ) ]

1 × [ 72℃ ( 240 秒 ) ]

3000bp 的产物经凝胶纯化, 用 NotI 和 BamHI 酶切, 经 PCR 净化后, 连接到 NotI/BamHI 消化的、凝胶纯化的 p7313ie 中, 以产生 p73i-Tngr. 序列测定揭示 p17p24 的密码子不是完全优化的, 进而用 AgeI/MunI 将编码区中的一个 700bp 的片段切掉, 并用来自含有正确编码序列的 p73i-Tgrn#3 的 MunI/AgeI 片段取代。该多蛋白还含有一个点突变 (G2609A), 导致多蛋白的 RT 部分中产生 Thr 改变成 Ala 的氨基酸取代。该突变通过 ApaI/BamHI 消化构建体并经 PCR 净化而除去突变的序列, 并用来自 p73i-Tngr 的 RT 的 ApaI/BamHI 消化部分连接而取代该序列而得到校正。

### 实施例 19:

制备用于“基因枪”DNA 药筒(cartridges)的质粒包被“金浆液”



质粒 DNA (大约  $1\mu\text{g}/\mu\text{l}$ )，例如  $100\mu\text{g}$ ，和  $2\mu\text{m}$  的金颗粒，例如  $50\text{mg}$  (PowderJect)，悬浮于例如  $100\mu\text{l}$  的  $0.05\text{M}$  的亚精胺中 (Sigma)。通过加入例如  $100\mu\text{l}$   $1\text{M}$  的  $\text{CaCl}_2$  (American Pharmaceutical Partners, Inc., USA) 使 DNA 沉淀到金颗粒上。该 DNA/金颗粒复合物在室温下保持 10 分钟，用无水乙醇洗涤 3 次，例如  $3 \times 1\text{ml}$ ，(事先在分子筛 3A (BDH) 上干燥)。样品重悬浮于含有  $0.05\text{mg}/\text{ml}$  的聚乙烯吡咯烷酮 (PVP, Sigma) 的无水乙醇中，并将其分成三等份于  $1.5\text{ml}$  的显微离心管中 (Eppendorf)。该等分试样分别用于 (a) “金浆液”的分析，(b) 洗出液—从 (a) 洗提出的质粒的分析和 (c) 制备用于“基因枪”的包被乙烯-四氟乙烯共聚物 (Tefzel) 药筒的金颗粒/质粒 (见下文的实施例 3)。为了制备样品 (a) 和 (b)，含有质粒 DNA/乙醇中的“金浆液”/PVP 的离心管在 Eppendorf 5418 显微离心机中高速旋转 2 分钟，除去上清液，“金浆液”在室温下干燥 10 分钟。样品 (a) 在 TE pH 8.0 中重悬浮到质粒 DNA 为  $0.5\text{--}1.0\mu\text{g}/\mu\text{l}$ ，假定大约 50% 被包被。对于洗出液，样品 (b) 在 TE pH 8.0 中重悬浮到质粒 DNA 为  $0.5\text{--}1.0\mu\text{g}/\mu\text{l}$ ，并在  $37^\circ\text{C}$  温育 30 分钟，激烈摇动，然后在 Eppendorf 5418 显微离心机中高速旋转 2 分钟，其上清液、洗出液被除去后，储藏在  $-20^\circ\text{C}$ 。洗提的精确 DNA 浓度用 GenequantII (Pharmacia Biotech) 通过分光光度法定量测定。

#### 实施例 20

##### 用于 DNA 免疫的药筒的制备

如以前所描述的制备用于 Accell 基因转移装置的药筒 (Eisenbraun 等, DNA 和细胞生物学 (DNA and Cell Biology), 1993 vol 12 No 9, 791-797 页; Pertner 等)。简而言之，质粒 DNA 被包被到  $2\mu\text{m}$  的金颗粒上 (DeGussa Corp., South Plainfield, N.J., USA) 并装载到 Tefzel 管中，该管随后被切成  $1.27\text{cm}$  长以用做药筒，并在  $4^\circ\text{C}$  下干燥储存直至使用。在典型的疫苗接种中，各个药筒含有  $0.5\text{mg}$  被包被的金颗粒，总量为  $0.5\mu\text{g}$  DNA/药筒。

#### 实施例 21:

### 利用基因枪进行 DNA 疫苗接种后对 HIV 抗原的免疫反应

用位于两个载体中的核酸所编码的抗原接种小鼠 ( $n=3/\text{组}$ )。P7077 利用包含内含子 A 和外显子 I 的 HCMV IE 启动子 (fcmv 启动子)。P73I 产生相同的抗原, 但包含的是缺乏内含子 A 而包括外显子 1 的 HCMV IE 启动子 (icmv 启动子)。

质粒被给予到 F1 (C3H  $\times$  Balb/c) 小鼠的腹部皮肤切开的靶位点。在第 0 天, 对小鼠以  $2 \times 0.5 \mu\text{g}$  DNA 进行最初免疫, 在第 35 天用  $2 \times 0.5 \mu\text{g}$  DNA 进行加强免疫, 在第 40 天用 IFN- $\gamma$  Elispot 检测细胞反应。

|             |   |                                |
|-------------|---|--------------------------------|
| · P73I      | - | 空载体                            |
| · P7077     | - | 空载体                            |
| · P7077 GRN | - | (fCMV 启动子) Gag, RT, Nef        |
| · P73I GRN  | - | (iCMV 启动子) Gag, RT, Nef        |
| · P73I GR3N | - | (CMV 启动子) 优化的 Gag, 优化的 RT, Nef |
| · P7077 GN  | - | (fCMV 启动子) Gag, Nef            |
| · P73I GN   | - | (iCMV 启动子) Gag, Nef            |

### 细胞毒素 T 细胞反应

细胞毒素 T 细胞反应通过对 5 天后收集的脾细胞用 CD8 $^{+}$  T 细胞-限制性 IFN- $\gamma$  ELISPOT 进行分析来评估。经颈部离位杀死小鼠, 将脾收集到冰冷的 PBS 中。将梳理出脾细胞放入磷酸缓冲盐溶液中 (PBS), 随后溶解红血细胞 (在含有 155mM  $\text{NH}_4\text{Cl}$ 、10mM  $\text{KHCO}_3$ 、0.1mM EDTA 的缓冲液中 1 分钟)。用 PBS 洗涤两次除去颗粒物质后, 单个细胞悬浮液按等分放入事先用俘获的 IFN- $\gamma$  抗体包被的 ELISPOT 平板中, 并用 CD8-限制性同源肽 (Gag、Nef 或 RT) 刺激。过夜培养后, 加入抗-小鼠 IFN- $\gamma$ -生物素标记抗体 (PharMingen) 继而加入链霉抗生物素-偶联的碱性磷酸酶即可观察到产生 IFN- $\gamma$  的细胞, 用图象分析进行定量测定。

该实验结果显示于图 20、21 和 22 中。

## 实施例 22:

### 疫苗构建体的免疫原性

#### 1. 细胞实验

细胞免疫反应包括细胞毒素 CD8 细胞和辅助 CD4 细胞。检测特异性 CD8 和 CD4 细胞的灵敏方法是 ELIspot 分析, 它可用于定量测定能够分泌干扰素- $\gamma$  或 IL-2 的细胞。ELIspot 分析基于俘获的从各个细胞分泌的细胞因子。简而言之, 特化的微量滴定板用抗-细胞因子抗体包被。从免疫动物分离的脾细胞与含有已知表位 (CD8) 或蛋白 (CD4) 的特异性肽一起培养过夜。如果细胞经刺激释放细胞因子, 它们将在滴定板表面上各个产生细胞的位置周围与抗体结合。在细胞被溶解和洗涤平板后, 细胞因子仍然保持与包被抗体结合。该分析利用生物素/抗生物素蛋白扩增系统、以与 ELISA 分析相似的方式进行。斑点的数量与细胞因子产生细胞的数量相等。

对 6 个构建体均记录了可对如下 K2<sup>d</sup>-限制性小鼠表位: Gag (AMQMLKETI)、Nef (MTYKAAVDL) 和 RT (YYDPSKDLI) 发生反应的 CD8 细胞数量, 以及可对 Gag 和 RT 蛋白发生反应的 CD4 细胞数量。这些实验的结果经统计学分析, 且根据构建体的免疫原性进行分级。结果显示于图 23 中。

#### 2. 体液实验

从两个实验中进行加强免疫后 7 和 14 天, 收集血液样品用于抗体分析。分离血清并冰冻储存直到可以用特异性 ELISA 分析测定抗体的滴定度。检测了所有样品对 Gag、Nef 和 RT 的抗体。简而言之, ELISA 平板用相关蛋白包被。洗去多余的蛋白质后, 将稀释的血清样品加入到加样孔中进行培育。洗去血清样品, 加入与适当的 tag 偶联的抗小鼠抗血清。培养平板并在平板读出器上读取数据。结果显示于图 24 中。

#### 3. 抗体数据

在四个实验中测定了所有六个构建体的抗体滴定度。构建体 p73i-GNR 始终对 Gag 不产生抗体反应, 对 Nef 产生有限的抗体反应。

其原因不清楚，由于分离自同一小鼠的脾细胞中观察到 T-细胞反应，表明 Gag 蛋白是体内表达的。

产生 Gag 特异性抗体的等级如下：

RNG>GRN>NRG>RGN>NGR>GNR

#### 分析细胞免疫数据

该目的是在 3 个免疫学实验中根据点记数资料对 6 个构建体进行分级。按下述三类反应进行评估：

在第 7 天（初次免疫始后 7 天）CD8 对 Gag、Nef 和 RT 发生反应，

第 35 天（加强免疫后 7 天）CD4 对 Gag 和 RT 发生反应，

第 35 天（加强免疫后 7 天）CD8 对 Gag、Nef 和 RT 发生反应。

每个反应（例如 CD8 对 Gag 的反应）使用 SAS 第 8 版中的线性混合效果模型进行模拟。不管具体的抗原（Gag、Nef 或 RT）存在与否，也不管 IL-2 存在与否，该模型包括构建体的固定效果。另外，对于 CD8 的反应，其数据可从各个小鼠获得，受检者作为随机效果被包含于模型中。所述模型包括相互作用以考虑到构建体在抗原（存在/缺乏）和 IL-2（存在/缺乏）之间的各个组合中的不同效果。

通过模型，在抗原（存在/缺乏）和 IL-2（存在/缺乏）的各个组合中，分别评估了各个构建体和 p7313（对照组）之间的校正平均反应的差异，及 p 值以显示差异是否具有统计显著性。基于抗原存在而 IL-2 缺乏时的差异和 p 值，对构建体进行分级，即具有最大差异的构建体的值为 6，其次为 5 等，但是对于其差异在 5%的水平上差异不显著的任何构建体均为 0。

模型假定——残差随常数方差正态分布，使用图解方法和灵敏度分析进行评估，其中首先对反应取对数，接着取平方根而将反应转换模型值。构建体的等级对与模型假定的背离不敏感。

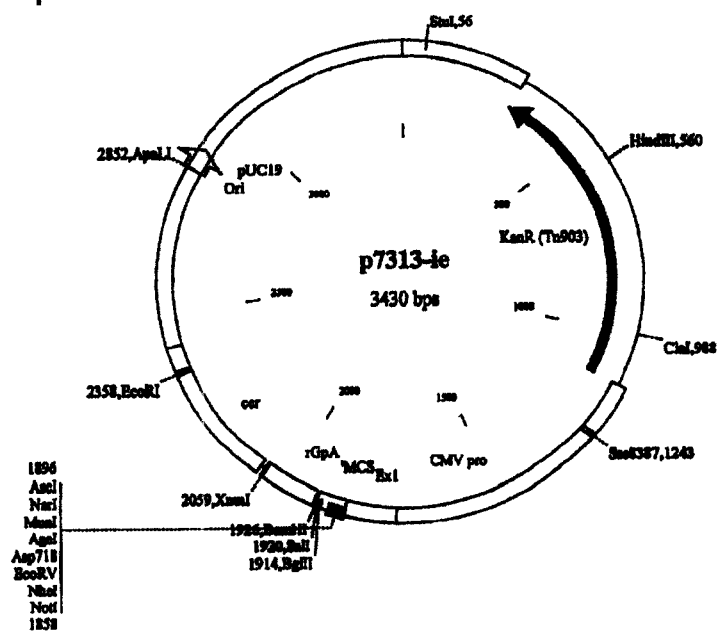
分别计算了各个实验中的各个反应的等级，对于所有 3 个实验中计算了 3 类反应的等级之和。下表显示了 3 个实验总的等级之和。

结合 3 个免疫学实验，3 类反应中各个反应的构建体的等级之和。

| 构建体 | 第 7 天（初次免疫后 7 天） | 第 35 天（加强后 7 天） |           |
|-----|------------------|-----------------|-----------|
|     | CD8              | CD4             | CD8       |
| GRN | 5                | 18              | 3         |
| GNR | 17               | 24              | 28        |
| RGN | <b>28</b>        | 23              | 33        |
| RNG | 25               | <b>27</b>       | <b>37</b> |
| NRG | 25               | 19              | 0         |
| NGR | 4                | 14              | 10        |

在第 35 天时，对于两类反应而言 RNG 具有最高的等级。而第二高的等级是第 7 天的 RGN。在第 35 天，对于两组反应而言 RGN 也都显示了比较高的等级。

**图 1**



5

10

15

20

图 2

pGagOptrrpr2 中 p55 gag 插入序列

```

5  ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
   AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
   AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
   CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
   TCCTTGTATAACACAGTGGCTACCCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
10  GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
   GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
   ATTGTCCAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
   AATGCCCTGGGTGAAGGTTGTGGAAGAGAAAGGCCTTTTCTCCTGAGGTTATCCCCATG
   TTCTCCGCTTTGAGTGAGGGGGCCACTCCCTCAGGACCTCAATACAATGCTTAATACC
15  GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
   GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
   CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
   GGATGGATGACCAACAATCCTCCCATCCCAAGTTGGAGAAATCTATAAACGGTGGATC
   ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
20  AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
   CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
   CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
   CTGGAAGAGATGATGACCGCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
   GTGTTGGCCGAAGCCATGAGCCAGGTGACGAACTCCGCAACCATCATGATGCAGAGA
25  GGGAACTTCCGCAATCAGCGGAAGATCGTGAAGTGTTTCAATTGCGGCAAGGAGGGT
   CATACCGCCCGCAACTGTGCGGCCCTTAGGAAGAAAGGGTGTTGGAAAGTGCGGCAAG
   GAGGGACACCAGATGAAAGACTGTACAGAACGACAGGCCAATTTTCTTGAAAGATT
   TGGCCGAGCTACAAGGGGAGACCTGGTAATTTCTTGCAAAGCAGGCCCGAGCCACC
   GCCCCCCTGAGGAATCCTTCAGGTCCGGAGTGGAGACCACAACGCCTCCCCAAAAA
30  CAGGAACCAATCGACAAGGAGCTGTACCCTTTAACTTCTCTGCGTTCTCTCTTTGGC
   AACGACCCGTCGTCTCAATAA [SEQ ID NO: 50]

```

```

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCROI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIBIKDTKEA
35  LDKIEEEONK SKKKAQAAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
   TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAMQM
   LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
   TNNPPIPVGE IYKRWIILGL NKIVRMYSP SILDIRQGP EPFRDYVDRF
   YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTIKALGPA ATLEEMMTAC
40  QGVGGPGHKA RVLAEAMSQV TNSATIMMOR GNFRNQKIV KCFNCGKEGH
   TARNCRAPRK KGCWKCGKEG HQMKDCTERQ ANFLGKIWPS YKGRPGNFLO
   SRPEPTAPPE ESFRSGVETT TPPQKQEPID KELYPLTSLR SLFGNDPSSQ
   *

```

[SEQ ID NO: 51]

图 3

p17/24trNEF1 中 p17/24trNEF 插入序列

```

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGAGAATTAGATCGATGGGAAAAAATT
5 CGGTTAAGGCCAGGGGAAAGAAAAATATAAATTAAACATATAGTATGGGCAAGC
AGGGAGCTAGAACGATTTCGAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGT
AGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAAGCTTAGA
TCATTATATAATACAGTAGCAACCCCTCTATTGTGTGCATCAAAGGATAGAGATAAAA
GACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAA
10 GCACAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCT
ATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAAGCTTTA
AATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTTCAGCCAGAAAGTGATACCCATG
TTTTTCAGCATTATCAGAAGGAGCCACCCACAAGATTTAAACACCATGCTAAACACA
GTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCT
15 GCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATG
AGAGAACCAAGGGGAAGTGACATAGCAGGAAGTACTAGTACCCTTCAGGAACAAATA
GGATGGATGACAAATAATCCACCTATCCAGTAGGAGAAATTTATAAAGATGGATA
ATCCTGGGATTAAATAAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATA
AGACAAGGACCAAAAGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAACTCTA
20 AGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTC
CAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACA
CTAGAAGAAATGATGACAGCATGTTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGA
GTTTTTGGTGGGTTTTCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAG
GCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATT
25 CACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTAC
TTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTT
GGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAA
GGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAG
AGAGAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCCGA
30 GAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [SEQ ID NO: 52]

```

```

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCROI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIRIKDTKEA
LDKIEEBQNK SKKKAQQAQ DTGHSNQVBSQ NYPIVQNIQG QMVHQAI SPR
35 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNIMLNT VGGHQAAMQM
LKETINBEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSP T SILDIRQGP EPFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQNPDC KTLKALGPA ATLEEMMTAC
QGVGGPGHKA RVLVGFPVTP QVPLRPMTYK AAVDLSHFLK EKGGLLEGLIH
40 SQRQDILD L WIYHTQGYFP DWQNYTPGPG VRYPLTFGWC YKLVPVEPDK
VEEANKGENT SLLHPVSLHG MDDPEREVLE WRFDShLAFH HVARELHPEY
FKNC*
[SEQ ID NO: 53]

```



图 4

p17/24opt/trNef1 中 p17/24opt/trNef 插入序列

```

5  ATGGGTGCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
   AGGCTGCGCCCGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
   AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
   CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
   TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
10  GATACCAAGGAGGCCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
   GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
   ATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACGCTC
   AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCATG
   TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
15  GTGGGCGGCCATCAGGCCGCCATGCAATGTTGAAGGAGACTATCAACGAGGAGGCA
   GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
   CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
   GGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATC
   ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
20  AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
   CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
   CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
   CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
   GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
25  AAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTA
   ATTCACCTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
   TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC
   TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
   AAAGGAGAGAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCG
30  GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCC
   CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [SEQ ID NO: 54]

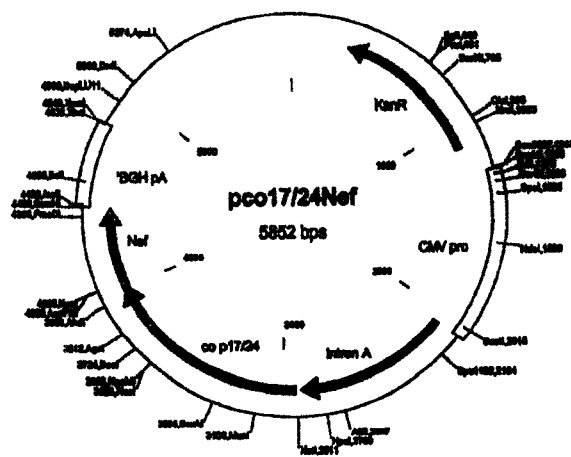
```

```

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCRQI LGQLQPSLOT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
35  LDKIEEEQNK SKKKAQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
   TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAAMQM
   LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
   TNPPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPB EPFRDYVDRF
   YKTLRAEQAS QEVKNWMTET LLVQNPANPD KTIKALGPA ATLEEMMTAC
40  QGVGGPGHKA RVLVGFPPVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
   HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYPPLTFW CYKLVPVEPD
   KVEEANKGEN TSLHHPVSLH GMDDPEREVL EWRFDShLAF HHVARELHPE
   YFKNC*
   [SEQ ID NO: 55]

```

45



5

10

15

20

25

30

图 5

p7077-RT3 的RT插入序列:

```

5  ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
   GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
   ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
   TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
   GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
10  ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
   GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
   ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC
   CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
   CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
15  GTGGGCTCTGACCTGGAAATCGGGCAGCATCGACGAAGATTGAGGAGCTGAGGCAG
   CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
   TTCCTGTGGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTC
   CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
   AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
20  CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCTCACGGAGGAAGCCGAGCTCGAG
   CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCC
   TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
   ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
   GCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
25  TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
   TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
   GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
   GCGCGGAGACATTCTACGTGGACGGCGCGGCCAACC CGGAAACAAAGCTCGGGAAG
   GCCGGGTACGTACCAACCGGGGCCGCCAGAAGGTCGTACCCCTGACCGACACCACC
30  AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
   GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTTCAGGCCAGCCGGAC
   CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
   GTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTCGAC
   AAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGTAA [SEQ ID NO: 56]
35  MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKKBKIS
   KIGPENPYNT PVFAIKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
   GLKKKKSVTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
   LPQGWKGSFA IFQSSMTKIL EPFRKQNPDI VIYQYMDDL YVGSLEIGQH
40  RTKIEBLRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
   KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEB
   AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
   LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
   ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
45  ETKLGKAGYV TNRGRQKVV LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
   SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLAWVPA HKGIGGNEQV

```



图 6

p731-RT3 中编码插入序列:

```

5  ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
   GGCCCCAAGGTCAAGCAGTGGCCACTCACCAGGAGAGAAGATCAAGGCCCTGGTGGAG
   ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
   TACAACACCCCGTGTGTTGCCATCAAGAAGAAGGACAGCACCAAGTGCGCGAAGCTG
   GTGGATTTCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
10  ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
   GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
   ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCTCTCCC
   CAGGGCTGGAAGGGCTCTCCCGCCATTTTCAGAGCTCCATGACCAAGATCCTGGAG
   CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
15  GTGGGCTCTGACCTGGAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
   CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
   TTCCTGTGGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTC
   CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
   AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
20  CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
   CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCC
   TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
   ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
   GCCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
25  TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
   TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCAGTGGGAGTTC
   GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
   GCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAA
   GGCCGGGTACGTACCAACCGGGGCCCGCCAGAAGGTGCTCACCTGACCGACACCAC
30  CAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGA
   GGTGAACATCGTGACGACAGCCAGTACGCGCTGGGCATTATTACGGCCCAGCCGGA
   CCAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAA
   GGTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTCCA
   CAAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGTAA [SEQ ID NO: 58]
35
   MGPISPIETV SVKCLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
   KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
   GLKKKKSIVTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
   LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDLY VGSLEIGQH
40  RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
   KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
   AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
   LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
   ETWWTEYWQA TWIPEWBFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
45  ETKLGKAGYV TNRGRQKVVV LDTDTNQKTE LQAIYLALQD SGLEVNIVTD
   SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLAWVPA HKGIGGNEQV

```

DKLVSAGIRK VL\*  
[SEQ ID NO: 59]

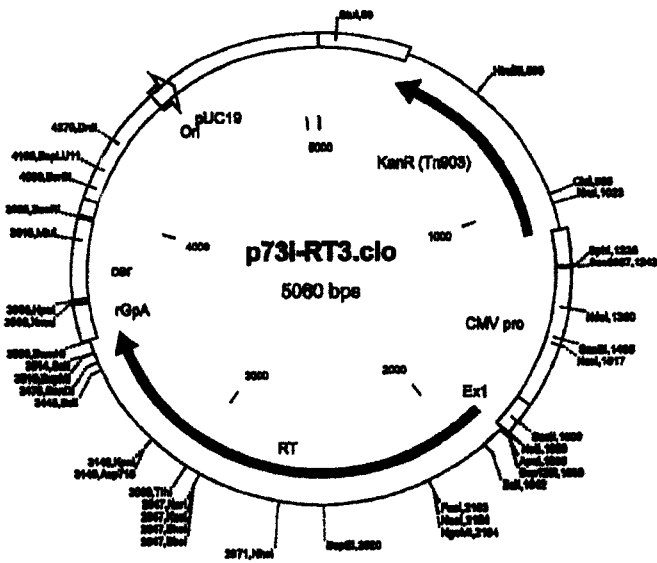


图 7

## 7077trNef20 中 Nef 插入序列

5 ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCA  
GCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCAC  
TCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC  
CCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGA  
TGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAGGA  
GAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGA  
10 GAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTCATCACGTGGCCCGAGAG  
CTGCATCCGGAGTACTTCAAGAACTGCTGA [SEQ ID NO: 60]

MVGFPVTPQV PLRPMTYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDWI  
YHTQGYFPDW QNYTPGPGVR YPLTFGWCYK LVPVEPDKVE EANKGENTSL  
15 LHPVSLHGMD DPEREVLEWR FDSVLAFHHV ARELHPEYFK NC\*  
[SEQ ID NO: 61]

20

25

30

图 8

7077RT8 中RT插入序列

```

ATGGGCCCCATTAGCCCTATTGAGACTGTGTCAGTAAAATTAAAGCCAGGAATGGAT
GGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAAGCATTAGTAGAA
ATTTGTACAGAGATGGAAGGAAGGGAATTTCAAAATTGGGCCTGAAAATCCA
5 TACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAATTA
GTAGATTTTACAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGA
ATACCACATCCCGCAGGGTTAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGT
GATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCATTTACC
ATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCA
10 CAGGGATGGAAAGGATCACCAGCAATATTCAGAGTAGCATGACAAAAATCTTAGAG
CCTTTTAGAAAACAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTAT
GTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAA
CATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAACATCAGAAAGAACCTCCA
TTCTTTTGGATGGGTTATGAACTCCATCTGATAAATGGACAGTACAGCCTATAGTG
15 CTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGGAAATTG
AATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAACTCCTT
AGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAA
CTGGCAGAAAAACAGAGAGATTCTAAAAGAACCACTACATGGAGTGTATTATGACCCA
TCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAA
20 ATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGGAAAAATATGCAAGAATGAGGGGT
GCCACACTAATGATGTAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAA
AGCATAGTAATATGGGGAAAGACTCCTAAATTTAACTGCCCATACAAAAGGAAACA
TGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTT
GTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCTAGTA
25 GGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAA
GCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACA
AATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAA
GTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAAACCAGAT
CAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAAG
30 GTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGAT
AAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTATAGATTAA
[SEQ ID NO: 62]

```

```

MGPISPIETV SVKLPQMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
35 KIGPENPYNT PVFAIKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
GLKKKKSQTV LDVGDYFVS PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
LPQGWKGSFA IFQSSMTKIL EPFRKQNPDI VIYQYMDDLY VGSLEIGQH
RTKIEELRQH LLRWGLTTPD KKHQKEPFL WMGYELHPDK WTVQPIVLPE
KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
40 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGGQWT YQIYQEPFKN
LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
ETWWTIEYQA TWIPEWFEVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
ETKLKGAGYV TNRGRQKVV LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
SQYALGIIQA QPDQSESELV NQIEQLIKK EKVYLAVPA HKGIGGNEQV
45 DKLVSAIRK VLFLD*

```



[SEQ ID NO: 63]

5

10

15

20

25

图 9

p17/24opt/RT/trNef13 中 p17/24opt/RT/trNef 插入序列

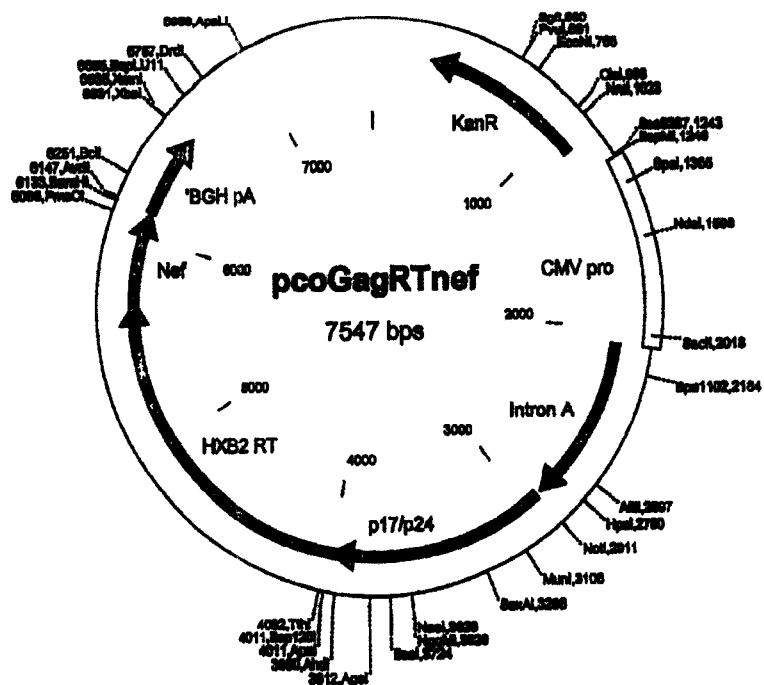
```

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
5  AGGCTGCGCCCGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
   AGGGAGCTTGAAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
   CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
   TCCTTGTATAACACAGTGGCTACCCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
   GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
10  GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
   ATTGTCCAAAACATTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
   AATGCCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
   TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
   GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
15  GCCGAGTGGGACAGAGTGCAATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
   CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
   GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
   ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
   AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTATATAAGACCCTG
20  CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
   CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
   CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
   GTGTTGATGGGCCCCATTAGCCCTATTGAGACTGTGTGAGTAAAAATTAAAGCCAGGA
   ATGGATGGCCCAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAA
25  GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAA
   AATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGA
   AAATTAGTAGATTTTCAAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAA
   TTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGAT
   GTGGGTGATGCATATTTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCA
30  TTTACCATACCTAGTATAAACAAATGAGACACCAGGGATTAGATATCAGTACAATGTG
   CTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATC
   TTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGAT
   TTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTG
   AGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAA
35  CCTCCATTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCT
   ATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAACTTAGTGGGG
   AAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAA
   CTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAGCAGAG
   CTAGAACTGGCAGAAAAAGAGAGATTCTAAAAGAACCAGTACATGGAGTGATTAT
40  GACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACA
   TATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAAATATGCAAGAATG
   AGGGGTGCCCACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACC
   ACAGAAAGCATAGTAATATGGGAAAGACTCCTAAATTTAACTGCCCATACAAAAG
   GAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGG

```

GAGTTTGTTAATACCCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCC  
 ATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTA  
 GGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCAACCTAACTGAC  
 ACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTGCGGA  
 5 TTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAA  
 CCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAG  
 GAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA  
 GTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTTAGATATGGTGGGT  
 TTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT  
 10 CTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGA  
 AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGG  
 CAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTAC  
 AAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACC  
 AGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTA  
 15 GAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCG  
 GAGTACTTCAAGAACTGCTGA [SEQ ID NO: 64]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL  
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIBIKDTKEA  
 20 LDKIEEBQNK SKKKAQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAISPR  
 TLNAWVKVVE EKAPSPVIP MFSALSEGAT PQDLNMLNT VGGHQAAMQM  
 LKETINEEAA EWDVRVPHVA GPIAPQMRE PRGSDIAGTT STLQEQIGWM  
 TNNPPPIVGE IYKRWIILGL NKIVRMYSP T SILDIRQGPB EPFRDYVDRF  
 YKTLRAEQAS QEVKNWMTET LLVQANPDC KTLKALGPA ATLEHMTAC  
 25 QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV  
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ  
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI  
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVYQYMD  
 DLYVGSLEI QHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH  
 30 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG  
 TKALTEVIPL TEEAELELAE NREILKEPVH GVVYDPSKDL IAEIQKQGGG  
 QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWVK  
 TPKFKLPIQK ETWETWNTY WQATWIPEWE FVNTPLVLK WYQLEKEPIV  
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA  
 35 LQDSGLEVNI VTDSQYALGI IQAQPQSES ELVNQIIEQL IKKEKVYLAW  
 VPAHKGIGGN EQVDKLVSA IRKVLFLDMV GFPVTPQVPL RPMTYKAAVD  
 LSHFLKEKGG LEGLIHSQRR QDILDWYH TQGYFPDWON YTPGPGVRYP  
 LTFGWYCYLV PVEPDVVEA NKGENTSLLH PVS LHGMDDP EREVLEWRFD  
 SHLAFHHVAR ELHPEYFKNC \*  
 40 [SEQ ID NO: 65]



5

10

15

图 10

WRG7077 中 p17/p24opt(cor)/RT/trNef 编码插入序列

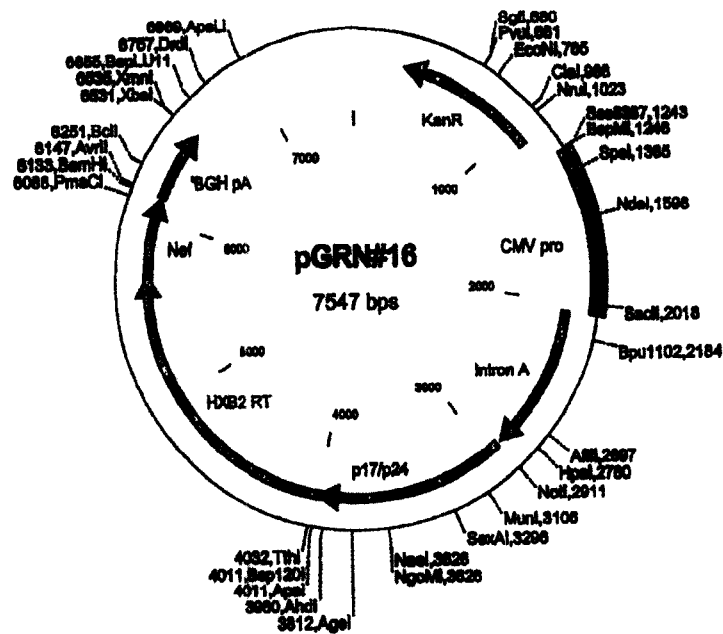
```

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
5 AGGCTGCGCCCCGGGAGGCCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGCCTCG
AGGGAGCTTGAACGGTTTGGCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
TCCTTGTATAACACAGTGGCTACCCCTCTACTGCGTACACCAGAGGATCGAGATTAG
GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
10 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
ATTGTCCAAACATTTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACGCTC
AATGCCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
GTGGGCGGCCATCAGGCCGCCATGCAAAATGTTGAAGGAGACTATCAACGAGGAGGCA
15 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGGCCAATCGCGCCCGGACAGATG
CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAATC
GGATGGATGACCAACAATCCTCCCATCCCAAGTTGGAGAAATCTATAAACGGTGGATC
ATCCTGGGCCTGAACAAGATCGTCCCATGTACTCTCCGACATCCATCCTTGACATT
AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
20 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
CAGAACGCTAACCCCGACTGCAAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
GTGTTGATGGGCCCCATTAGCCCTATTGAGACTGTGTGAGTAAATTAAGCCAGGA
ATGGATGGCCCCAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAGCATTAA
25 GTAGAAATTTGTACAGAGATGAAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAA
AATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGA
AAATTAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAA
TTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGAT
GTGGGTGATGCATATTTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCA
30 TTTACCATACCTAGTATAAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTG
CTTCCACAGGGATGGAAAGGATCACCAGCAATATTCAAAGTAGCATGACAAAAATC
TTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGAT
TTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTG
AGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAA
35 CCTCCATTCCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCT
ATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAAGTTAGTGGGG
AAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAA
CTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAG
CTAGAAGTGGCAGAAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTAT
40 GACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGCCCAATGGACA
TATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAAATATGCAAGAATG
AGGGGTGCCCACTAATGATGTAAACAATTAACAGAGGCAGTGCAAAAATAACC
ACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCCATACAAAG
GAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGTAGTGG
45 GAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCC

```

ATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTA  
 GGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCAACCTAACTGAC  
 ACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGA  
 TTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAA  
 5 CCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAG  
 GAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA  
 GTAGATAAATTAGTCAGTGTCTGGAATCAGGAAAGTACTATTTTTAGATATGGTGGGT  
 TTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT  
 CTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGA  
 10 AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGG  
 CAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGTCTAC  
 AAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACC  
 AGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCGAGAGAGAAAGTGTTA  
 GAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCG  
 15 GAGTACTTCAAGAACTGCTGA [SEQ ID NO: 66]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL  
 LETSEGCRQI LGQLQPSLOT GSEELRSLYN TVATLYCVHQ RIRIKDTKEA  
 20 LDKIEEBQNK SKKKAQQA AA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR  
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAAQMOM  
 LKETINEEAA BWD RVHPVHA GPIAPGQMR PRGSDIAGTT STLQEBQIGWM  
 TNNPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPB EPFRDYVDRF  
 YKTLRAEQAS QEVKNWMTET LLVQANPDC KTLKALGPA ATLEEMMTAC  
 25 QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV  
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DPRELNKRTQ  
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY PSVPLDEDFR KYTAFTIPSI  
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVIIYQYMD  
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH  
 30 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG  
 TKALTEVIPL TEEAELELAE NREILKEPVH GVVYDPSKDL IAEIQKQGGQ  
 QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWVK  
 TPKFKLPIQK ETWETWWTBY WQATWIPEWE FVNTPLVLK WYQLEKEPIV  
 GAETFFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA  
 35 LQDSGLEVNI VTDSQYALGI IQAQPDSSES ELVNQIIEQL IKKEKVYLAN  
 VPAHKGIGGN EQVDKLVSA IRKVLFLDMV GPFVTPQVPL RPTTYKAAVD  
 LSHFLKEKGG LEGLIHSQRR QDILDWIIYH TQGYFPDWQN YTPGPGVRYF  
 LTFGWICYLV PVEPKVBEA NKGENTSLLH PVS LHGMDDP EREVLEWRFD  
 SHLAFHHVAR ELHPEYFKNC \*  
 40 [SEQ ID NO: 67]



5

10

15

**图 11**

**p73l-GRN2 中 p17/p24(opt)/RT(opt)trNef 插入序列**

```
ATGGGTGCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGA
5  GAAAATTAGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCAT
  ATCGTGTGGGCTCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCT
  GCTGGAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCAGCCAT
  CCCTCCAGACCGGGAGTGAAGAGCTGAGGTCCTTGTATAACACAGTGGCT
  ACCCTCTACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGC
10  CTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAGGCCAG
  CAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCC
  TATTGTCCAAACATTTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCC
  GGACGCTCAATGCCTGGGTGAAGGTTGTCGAAGAGAAGGCCTTTTCTCCT
  GAGGTTATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGAC
15  CTCAATACAATGCTTAATACCGTGGGCGGCCATCAGGCCGCCATGCAAT
  GTTGAAGGAGACTATCAACGAGGAGGCAGCCGAGTGGGACAGAGTGCAT
  CCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAGCCTCGCGG
  CTCTGACATTGCCGGCACCACTCTACACTGCAAGAGCAAATCGGATGGA
  TGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
20  ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTT
  GACATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTT
  TTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGA
  TGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAAACAATC
  TTGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTG
25  TCAGGGAGTAGGCGGACCCGGACACAAAGCCAGAGTGTTGATGGGCCCC
  ATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGG
  CCCCAGGTC AAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTG
  GTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCG
  GGCCTGAGAACCCATACAACACCCCCGTGTTTGCCATCAAGAAGAAGGAC
30  AGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGAC
  CCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCATCCGGCCGGCCTGA
  AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGC
  GTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATCT
```



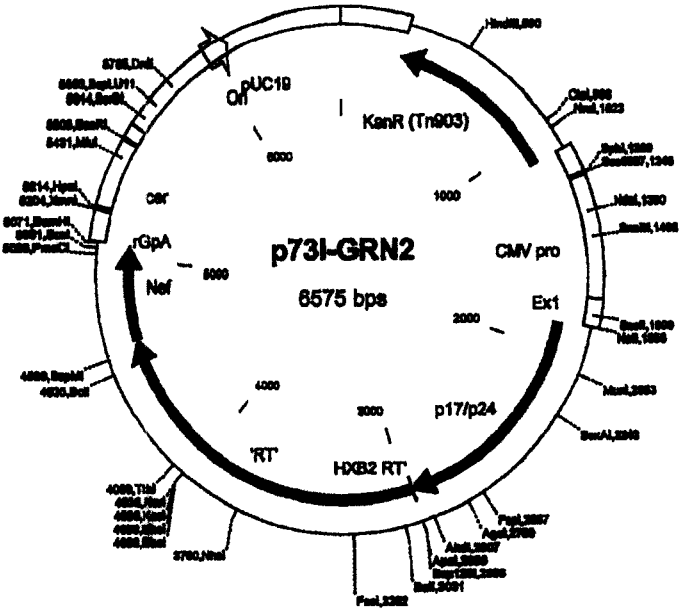
ATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCCCA  
GGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCC  
TGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATG  
GACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAA  
5 GATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGG  
ACAAGAAGCATCAGAAGGAGCCGCCATTCTGTGGATGGGCTACGAGCTC  
CATCCCGACAAGTGGACCGTGCAGCCTATCGTCCTCCCCGAGAAGGACAG  
CTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTA  
GCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGC  
10 GGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCT  
CGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGT  
ACTATGACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAG  
GGGCAGTGGACATAACCAGATTTACCAGGAGCCTTTCAAGAACTCAAGAC  
CGGCAAGTACGCCCCGATGAGGGGCGCCACACCAACGATGTCAAGCAG  
15 CTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGG  
GAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGT  
GGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTTCGTG  
AATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCAT  
CGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACA  
20 AAGCTCGGGAAGGCCGGGTACGTCACCAACCGGGGCCGCGCAGAAGGTG  
TCACCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCCATCTAT  
CTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCA  
GTACGCGCTGGGCATTATTCAGGCCAGCCGGACCAAGTCCGAGAGCGAAC  
TGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTC  
25 GCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTGACAA  
GCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTCA  
CACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTT  
AGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATCACTCCCA  
AAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACT  
30 TCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTG  
ACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGA  
AGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGC

ATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAG  
 CC~~C~~CCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCA  
 AGAACTGCTGA [SEQ ID NO: 68]

```

5  MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
   LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
   LDKIEEBQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
   TLNAWVKVVE EKAPSPVIP MFSALSEGAT PQDLNIMLNT VGGHQAMQM
   LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
10 TNNPPIPVGE IYKRWII LGL NKIVRMYSPT SILDIRQGP EPFRDYVDRF
   YKTLRAEQAS QEVKNWMTET LLVQANANPDC KTIKALGPA ATLEEMMTAC
   QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV
   EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
   DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
15 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVIIQYMD
   DLYVGSDBEI GQHRTKIBEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
   PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
   TKALTEVIPL TEEAELELAE NREILKEPVH GVEYDPSKDL IAEIQKQGQG
   QWTYQIYQEP FKNLKTGKYA RMGAHTNDV KQLTEAVQKI TTESIWIWVK
20 TPKFKLPIQK ETWETWWTEY WQATWIPWE FVNTPLVLKL WYQLEKEPIV
   GAETFFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDITNQ KTELQAIYLA
   LQDSGLEVNI VTDSQYALGI IQAQPQSES ELVNQIIEQL IKKEKVYLAW
   VPAHKGIGGN EQVDKLVSAG IRKVL MVGFP VTPQVPLRPM TYKAAVDLSH
   FLKEKGGLG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF
25 GWCYKLVPE PDKVEEANKG ENTSLHPVS LHGMDDPERE VLEWRFDSRL
   AFHHVARELH PEYFKNC*
   [SEQ ID NO: 69]

```



5

10

15

图 12

p731-GN2 中 p17/p24opt/trNef 插入序列

```

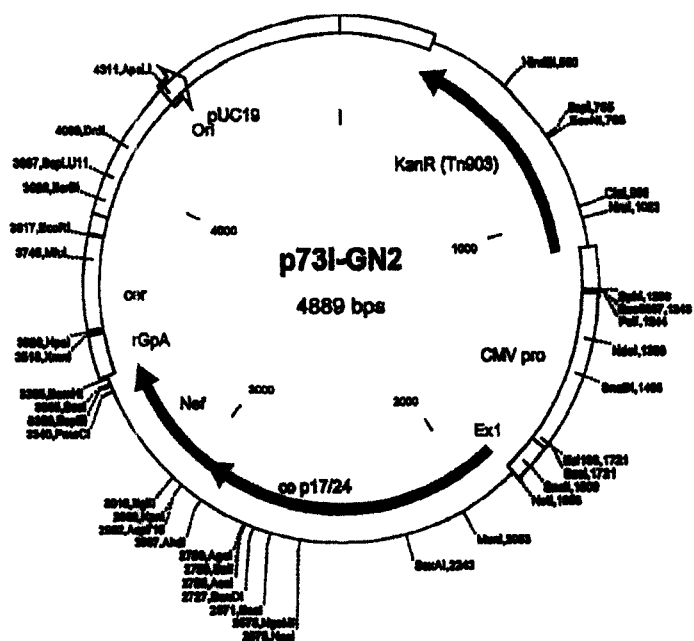
ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
5 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
AGGGAGCTTGAAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
TCCTTGTATAACACAGTGGCTACCTCTACTGCGTACACCAGAGGATCGAGATTAAG
GATACCAAGGAGGCCTTGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
10 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
ATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACGCTC
AATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
GTGGGCGGCCATCAGGCCGCCATGCAATGTTGAAGGAGACTATCAACGAGGAGGCA
15 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAAATC
GGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATC
ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCCTG
20 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
AAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTA
25 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC
TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
AAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCG
GAGAGAGAAGTGTAGAGTGGAGGTTTGACAGCCGCTAGCATTTCATCAGTGGCC
30 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [SEQ ID NO: 70]

```

```

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
35 LDKIEEBQNK SKKKAQQA AA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLN TMLNT VGGHQAA MQM
LKETINEEAA EWDRVHPVHA GPIAPGQ MRE PRGSDIAGTT STLQEBQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGP K EPFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQNaNPDC K TILKALGPA ATLEEMMTAC
40 QGVGGPGHKA RVL MVGFPVT PQVPLRPMTY KAAVDLSHFL KEKGGLGLI
HSQRQDILD LWIYHTQGYF PDWQNYTPGP G VRYPLTFGW CYKLVPVEPD
KVEEANKGEN TSL LHPVSLH GMDDPEREVL EWRFD SRLAF HHVARELHPE
YFKNC* [SEQ ID NO: 71]

```



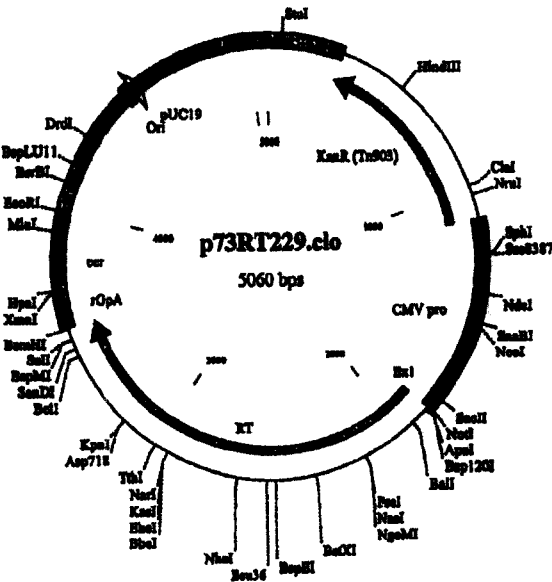
5

10

**图 13****RT w229k 序列**

```
ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGGCCC
5 CAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCG
AGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCATACAACACCCCGTG
TTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAA
TAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCATCCGGCCGGCCTGAAGA
AGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAG
10 GACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACCCCTGGCATCAG
ATATCAGTACAACGTCCTCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTCAGAGCTCCA
TGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAACCCGATATCGTCATCTACCAGTACATG
GACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCT
GAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGC
15 CATTCTCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTCCTC
CCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGC
TAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGGGCACCAGG
CCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCGAGCTCGAGCTGGCTGAGAACCGGGAG
ATCCTGAAGGAGCCCGTGACGGCGTGACTATGACCCCTCCAAGGACCTGATCGCCGAAAT
20 CCAGAAGCAGGGCCAGGGGCAGTGGACATAACGATTTACCAGGAGCCTTTCAAGAACCTCA
AGACCGGCAAGTACGCCCGCATGAGGGCGGCCACACCAACGATGTCAAGCAGCTGACCGAG
GCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCT
GCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTTC
CCGAGTGGGAGTTCTGTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAG
25 CCCATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACC CGGAAACAAAGCTCGG
GAAGGCCGGGTACGTCAACCAACCGGGGCCGCCAGAAGGTCGTACCCTGACCGACACCACCA
ACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAGGTGAAC
ATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTTCAGGCCAGCCGACCAAGTCCGAGAG
CGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGG
```

TCCCGGCCCATTAAGGGCATTGGCGGCAACGAGCAGGTCGACAAGCTGGTGAGTGCCTGGGATT  
AGAAAGGTGCTGTAA [SEQ ID NO: 72]



5

10

15





ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT  
AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCCGGTTTATAAGACCCTG  
CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA  
CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC  
5 CTGGAAGAGATGATGACCGCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA  
GTGTTGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGG  
ATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTG  
GTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAG  
AACCATAACAACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGC  
10 AAGCTGGTGGATTTCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAG  
CTGGGCATCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGAC  
GTGGGCGAGCCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCC  
TTTACCATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTC  
CTCCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTCAGAGCTCCATGACCAAGATC  
15 CTGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGAC  
CTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTG  
AGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAG  
CCGCCATTCTGaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCT  
ATCGTCCTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGC  
20 AAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAG  
CTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAG  
CTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGGTGACTAT  
GACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCACTGGACA  
TACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATG  
25 AGGGGCGCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCAGC  
ACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAG  
GAGACCTGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCAGTGG  
GAGTTCGTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCC  
ATCGTGGGCGCGGAGACATTTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTC  
30 GGGAAGGCCGGGTACGTACCAACCGGGGCGCCAGAAGGTGCTCACCTGACCGAC  
ACCACCAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGC  
CTGGAGGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTAGGCCCAG  
CCGGACCACTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAA  
GAGAAGGTCTACCTCGCTGGGTCCCGGCCCATAGGGCATTTGGCGGCAACGAGCAG  
35 GTCGACAAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTC  
ACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCAC  
TTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACCTCCCAAAGAGCAAGAT  
ATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTAC  
ACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTA  
40 CCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTA  
CACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGG  
TTTGACAGCCGCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTC  
AAGAAGTCTGA [SEQ ID NO: 73]

45

```

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCROI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
LDKIEEBQNK SKKKAQAAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
5 TLNANVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAA MQM
LKETINERAA EWDRVHPVHA GPIAPGOMRE PRGSDIAGTT STLQEQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPB EPFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTILKALGPA ATLEEMMTAC
QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV
10 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVIQYMD
DLYVGSDEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
PDKNTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
15 TKALTEVIPL TEEAELELAE NREILKEPVH GVIYDPSKDL IAEIQKQGGQ
QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIVINGK
TPKFKLPIQK ETWETWNTY WQATWIPEWE FVNTPLLVKL WYQLEKEPIV
GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
LQDSGLEVNI VTDSQYALGI IQAQPQSES ELVNQIIEQL IKKEKVYLAW
20 VPAHKGIGGN EQVDKLVSA G IRKVL MVGFP VTPQVPLRPM TYKAAVDLSH
FLKEKGGLEG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF
GWCYKLVPE PDKVEEANKG ENTSLHHPVS LHGMDDPERE VLEWRFD SRL
AFHHVARELH PEYFKNC*
[SEQ ID NO: 74]
25

```

30

35

图 15

Tnrg 的序列:

```
ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCA
5 GCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCAC
TCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC
CCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGA
TGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGA
GAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGA
10 GAAGTGTTAGAGTGGAGGTTTGACAGCCGCCCTAGCATTTTCATCACGTGGCCCCGAGAG
CTGCATCCGGAGTACTTCAAGAACTGCATGGGCCCCATCAGTCCCATCGAGACCGTG
CCGGTGAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACC
GAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAG
ATCAGCAAGATCGGGCCTGAGAACCCATACAACACCCCCGTGTTTGCCATCAAGAAG
15 AAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACC
CAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCCATCCGGCCGGCCTGAAGAAGAAG
AAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAG
GACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACCCCTGGC
ATCAGATATCAGTACAACGTCCTCCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTC
20 CAGAGCTCCATGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAACCCCCGATATCGTC
ATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCAT
CGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCG
GACAAGAAGCATCAGAAGGAGCCGCCATTCCTgaaGATGGGCTACGAGCTCCATCCC
GACAAGTGGACCGTGAGCCTATCGTCCTCCCCGAGAAGGACAGCTGGACCGTGAAC
25 GACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCCCGGATC
AAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATT
CCCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAG
CCCGTGACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAG
CAGGGCCAGGGCAGTGGACATACCAGATTTACCAGGAGCCCTTCAAGAACCTCAAG
30 ACCGGCAAGTACGCCCGCATGAGGGGCGCCACACCAACGATGTCAAGCAGCTGACC
GAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAG
TTCAAGCTGCCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATATTGGCAG
GCCACCTGGATTCCCGAGTGGGAGTTTCGTGAATACACCTCCTCTGGTGAAGCTGTGG
TACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCG
35 GCCAACCGCGAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGCCCGCCAG
AAGGTCGTACCCCTGACCGACACCAACCAAGACGAGCTGCAGGCCATCTAT
CTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAGTACGCG
CTGGGCATTATTACAGGCCAGCCGACCAAGTCCGAGAGCGAACTGGTGAACCAGATT
ATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCCATAAG
40 GGCATTGGCGGCAACGAGCAGGTGACAAAGCTGGTGAAGTGGGGGATTAGAAAGGTG
CTGATGGGTGCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAA
ATTAGGCTGCGCCCGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCC
TCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGA
TGTGCCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTG
45 AGGTCTTGTATAACACAGTGGCTACCTCTACTGCGTACACCAGAGGATCGAGATT
```

AAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAG  
 AAGGCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCAGAGAACTAT  
 CCTATTGTCCAAAACATTGAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACG  
 CTTCAATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCC  
 5 ATGTTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAAT  
 ACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAG  
 GCAGCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAG  
 ATGCGGGAGCCTCGCGGCTCTGACATTGCGGGCACCACCTCTACACTGCAAGAGCAA  
 ATCGGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGG  
 10 ATCATCCTGGGCCTGAACAAGATCGTGCATGTACTCTCCGACATCCATCCTTGAC  
 ATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACC  
 CTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTG  
 GTACAGAACGCTAACCCCGACTGCAAAAACAATCTTGAAGGCACTAGGCCCGGCTGCC  
 ACCCTGGAAGAGATGATGACCGCCTGTGAGGAGTAGGCGGACCCGGACACAAAGCC  
 15 AGAGTGTGTGA [SEQ ID NO: 75]

MVGFVTPQV PLRPMTYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDWI  
 YHTQGYFPDW QNYTPGPGVR YPLTFGWCYK LVFVEPDKVE EANKGENTSL  
 20 LHPVSLHGMD DPEREVLEWR FDSRLAFHHV ARELHPEYFK NCMGPISPIE  
 TVSVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY  
 NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKS  
 TVLDVGDAYF SVPLDDEFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS  
 PAIFQSSMTK ILEPFRQNP DIVIYQYMDL LYVGSdleig QHRTKIEELR  
 25 QHLLRWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND  
 IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTEVIPLT EEAELELAEN  
 REILKEPVHG VYYDPSKDLI AEIQKQGGGQ WTYQIYQEPF KNLKTGKYAR  
 MRGAHTNDVK QLTHAVQKIT TESIVINGKT PKFKLPIQKE TWETWWTEYW  
 QATWIPEWEF VNTPLVLKLW YOLEKEPIVG AETFYVDGAA NRETCLGKAG  
 30 YVTNRGRQKV VTLTDITNOK TELQAIYLAL QDSGLEVNIV TDSQYALGII  
 QAQPDQSESE LVNQIIEQLI KKEKVYLAWV PAHKGIGGNE QVDKLVSAGI  
 RKVLMGARAS VLSGGELDRW EKIRLRPGGK KKYKLKHIVW ASRELERFAV  
 NPGLLETSEG CRQILGQLQP SLQTGSEELR SLYNTVATLY CVHQRIKID  
 TKEALDKIEE EQNKSKKKAQ QAAADTGHSN QVSQNYPIVQ NIQQQMVHQA  
 35 ISPRTLNAWV KVVEEKAFSP EVIPMFSALS EGATPQDLNT MLNTVGGHQA  
 AMQMLKETIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI AGTTSTLQEQ  
 IGWMTNNPPI PVGEIYKRWI ILGLNKIVRM YSPTSILDIR QGPKEPFRDY  
 VDRFYKTLRA EQASQEVKNW MTETLLVQNA NPDCKTILKA LGPAATLEEM



图 16

p7313le 中的 Tngr 插入序列:

```

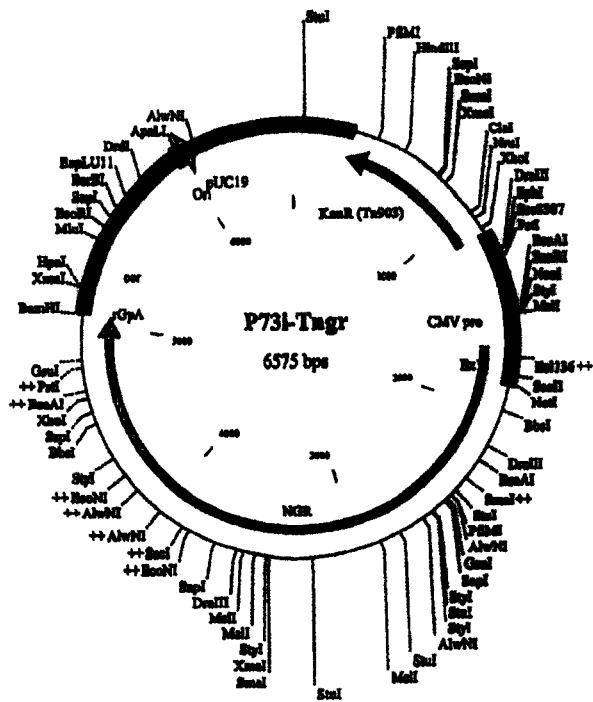
ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGT
5 AGATCTTAGCCACTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACCTCCCAAAGAA
GACAAGATATCCTTGATCTGTGGATCTACACACACAAGGCTACTTCCCTGATTGGCAGAAC
TACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACC
AGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTAGAGTGGAGGTTTGACAGCCGC
10 CTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCATGGGTGC
CCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATTAGGCTGCGCCCGG
GAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTT
GCCGTGAACCCAGGCTGCTGGAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCA
GCCATCCCTCCAGACCGGAGTGAAGAGCTGAGGTCTTGTATAACACAGTGGCTACCCCTCT
15 ACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGCCCTTGGACAAAATTGAGGAG
GAGCAAAACAAGAGCAAGAAGAAGGCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCA
GGTATCACAGAACTATCCTATTGTCCAAACATTGAGGCCAGATGGTTTCATCAGGCCATCA
GCCCCCGGACGCTCAATGCCTGGGTGAAGGTTGTGAAGAGAAGGCCTTTTCTCCTGAGGTT
ATCCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAA
20 TACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCAG
CCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGACAGATGCGGGAG
CCTCGCGGCTCTGACATTGCCGACACCTCTACACTGCAAGAGCAAATCGGATGGATGAC
CAACAATCCTCCCATCCAGTTGGAGAAATCTATAACCGTGGATCATCCTGGGCCTGAACA
AGATCGTGCGCATGTACTCTCGACATCCATCCTTGACATTAGACAGGGACCCAAAGAGCCT
25 TTTAGGGATTACGTCGACCGGTTTTTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGT
CAAAAAGTGGATGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAAACAATCT
TGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGC
GGACCCGGACACAAAGCCAGAGTGTGATGGGCCCATCAGTCCCATCGAGACCGTGCCGGT
GAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGA
30 TCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGG
CCTGAGAACCATACAACACCCCGTGTGTCATCAAGAAGAAGGACAGCACCAAGTGGCG
CAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGG
GCATCCCCCATCCGGCCGGCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGAC
GCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATC
35 TATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCCCAGGGCTGGAAGG
GCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAAC
CCCGATATCGTCATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGG
GCAGCATCGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTC
CGGACAAGAAGCATCAGAAGGAGCCGCCATTCTGaaGATGGGCTACGAGCTCCATCCCGAC
40 AAGTGGACCGTGACGCTATCGTCTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCA
GAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCCGGATCAAGGTGCGCCAGC
TCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATTCCCTCACGGAGGAAGCC
GAGCTCGAGCTGGCTGAGAACCAGGAGATCCTGAAGGAGCCCGTGACGCGGTGTACTATGA
CCCCCTCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAGA
45 TTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGCGCCAC

```

ACCAACGATGTCAAGCAGCTGACCGAGGCCCTCCAGAAGATCAAGACCGAGTCCATCGTGAT  
 CTGGGGGAAGACACCCAAGTTCAAGCTGCCCTATCCAGAAGGAGACCTGGGGAGACGTGGTGA  
 CCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTCGTGAATACACCTCCTCTGGTG  
 AAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGACGG  
 5 CGCGGCCAACC CGAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGCCGCCAGA  
 AGGTCGTCACCCTGACCGACACCACCAACCAGAAGACCGAGCTGCAGGCCATCTATCTCGCT  
 CTCCAGGACTCGGCCTGGAGGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTAT  
 TCAGGCCCAGCCGACCACTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCA  
 AGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCCCATAGGGCATTGGCGGCAACGAGCAG  
 10 GTCGACAAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGTAA [SEQ ID NO: 77]

MVGFPPVTPQV PLRPMYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDWI  
 YHTQGYFPDW QNYTPGPGVR YPLTFGWYK LVPVEPDKE EANKGENTSL  
 15 LHPVSLHGM DPREFVLEWR FDSRLAFHHV ARELHPFYFK NCMGARASVL  
 SGGELDRWEK IRLRPGGKKK YKLKHIVWAS RELERFAVNP GLLETSEGR  
 QILGQLQPSL QTGSEELRSL YNTVATLYCV HQRIBIKDTK EALDKIEEQ  
 NKSKKKAQQA AADTGHSNQV SQNYPIVQNI QGQMVHQAIS PRTLNAWVKV  
 VEEKAFSPFV IPMFSALSEG ATPQDLNML NTVGGHQAAM QMLKETINEE  
 20 AAEDWRVHPV HAGPIAPGOM REPRGSDIAG TTSTLQEQIG WMTNNPPIPV  
 GEIYKRWIIL GLNKIVRMYS PTSILDIRQG PKEPFRDYVD RPYKTLRARQ  
 ASQEVKNWMT ETLVQNANP DCKTILKALG PAATLEEMMT ACQGVGGPGH  
 KARVLMGPIS PIETVSVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK  
 EGKISKIGPE NPYNTPVFAI KKKDSTKNRK LVDFRELNKR TQDFWEVQLG  
 25 IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR  
 YQYNVLPQGW KGSPAIFQSS MTKILEPFRK QNPDIYIYQY MDDLIVGSDL  
 EIGQHRTKIE ELRQHLLRWG LTTDPKKHQK EPPFLWMGYE LHPDKWTVQP  
 IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTALTEVI  
 PLTEEALELE AENREILKEP VHGVYDPSK DLIAEIQKQG QGQWYQIYQ  
 30 EPFKNLKTGK YARMGAHTN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI  
 QKETWETWWT EYWQATWIPE WEFVNTPLV KLWYQLEKEP IVGAETFYVD  
 GAANRETKLG KAGYVTNRGR QKVVTLDTT NQKTELQAIY LALQDSGLEV  
 NIVTDSQYAL GIIQAQPDQS ESELVNQIIE QLIKKEKVYL ANVPAHKGIG  
 GNEQVDKLVS AGIRKVL\*  
 35 [SEQ ID NO: 78]

40



5

10

15



图 17

## Trgn #6 插入序列

```

5  ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
   GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
   ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
   TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
   GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
10  ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
   GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
   ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAAAGTCCTCCCC
   CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
   CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
15  GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
   CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
   TTCCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTC
   CTCCCCGAGAAGGACAGCTGGACCGTGAAACGACATCCAGAAGCTGGTGGGCAAGCTC
   AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
20  CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
   CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCC
   TCCAAGGACCTGATCGCCGAATCCAGAAGCAGGGCCAGGGGCAGTGGACATAACAG
   ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
   GCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
25  TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
   TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
   GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
   GCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAAG
   GCCGGGTACGTACCAACCGGGGCCGCCAGAAGGTCGTCAACCTGACCGACACCACC
30  AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
   GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTTCAGGCCAGCCGGAC
   CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
   GTCTACCTCGCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTCGAC
   AAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGGTGCCCGAGCTTCGGTACTG
35  TCTGGTGGAGAGCTGGACAGATGGGAGAAAAATTAGGCTGCGCCCGGGAGGCAAAAAG
   AAATACAAGCTCAAGCATATCGTGTGGCCTCGAGGGAGCTTGAACGGTTTGCCGTG
   AACCAGGCCTGCTGGAAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCAG
   CCATCCCTCCAGACCGGGAGTGAAGAGCTGAGGTCTTGTATAACACAGTGGCTACC
   CTCTACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGCCCTTGGACAAA
40  ATTGAGGAGGAGCAAAACAAGAGCAAGAAGAGGCCAGCAGGCAGCTGCTGACACT
   GGCATAGCAACCAGGTATCACAGAACTATCCTATTGTCCAAAACATTTCAGGGCCAG
   ATGGTTCATCAGGCCATCAGCCCCGACGCTCAATGCCTGGGTGAAGGTTGTGAA
   GAGAAGGCCCTTTCTCCTGAGGTATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCC
   ACTCCTCAGGACCTCAATACAATGCTTAATACCGTGGGCGGCCATCAGGCCGCCATG

```

CAAATGTTGAAGGAGACTATCAACGAGGAGGCAGCCGAGTGGGACAGAGTGCATCCC  
 GTCCACGCTGGCCCAATCGCGCCCGACAGATGCGGGAGCCTCGCGGCTCTGACATT  
 GCCGGCACCACCTCTACACTGCAAGAGCAAATCGGATGGATGACCAACAATCCTCCC  
 ATCCCAGTTGGAGAAATCTATAACGGTGGATCATCCTGGGCCTGAACAAGATCGTG  
 5 CGCATGTACTCTCCGACATCCATCCTTGACATTAGACAGGGACCCAAAGAGCCTTTT  
 AGGGATTACGTGACCCGTTTTTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAG  
 GTCAAAACTGGATGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAA  
 ACAATCTTGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTGT  
 CAGGGAGTAGGCCGACCCGGACACAAAGCCAGAGTGTGATGGTGGGTTTTCCAGTC  
 10 ACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCAC  
 TTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAGAAGACAAGAT  
 ATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTAC  
 ACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGTCTACAAGCTAGTA  
 CCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTA  
 15 CACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGG  
 TTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTC  
 AAGAACTGCTGA [SEQ ID NO: 79]

20 MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TMEKEGKIS  
 KIGPENPYNT PVFAIKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA  
 GLKKKKS MTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV  
 LPQGWKQSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDLY VGSLEIGQH  
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLP  
 25 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE  
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGGQWT YQIYQEPFKN  
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW  
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR  
 ETKLGKAGYV TNRGRQKVV LTDTTNQKTE LQAIYLALQD SGLEVNIVTD  
 30 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLAWVPA HKGIGGNEQV  
 DKLVSAIRK VLMGARASVL SGGEIDRWEK IRLRPGGKKK YKLKHIVWAS  
 RELERFAVNP GLLETSEGCR QILGQLQPSL QTGSEELRSL YNTVATLYCV  
 HORIEIKDTK EALDKIEEEQ NKSKKKAQQA AADTGHSNQV SQNYPIVQNI  
 QGQMVHQAI S PRTLNAWVKV VEEKAFSPKV IPMFSALSEG ATPQDLNMTL  
 35 NTVGGHQAM QMLKETINEE AAEWDRVHPV HAGPIAPGQM REPRGSDIAG  
 TTSTLQEQIG WMTNNPPIPV GEIYKRWIL GLNKIVRMYS PTSILDIRQG  
 PKEPFRDYVD RFYKTLRAEQ ASQEVKNWMT ETLLVQANP DCKTILKALG  
 PAATLEEMMT ACQGVGGPGH KARVLMVGFP VTPQVPLRPM TYKAAVDLSH  
 FLKEKGGLG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTF GPGVRYPLTF

GWCYKLVPE PDKVEANKG ENTSLHPVS LHGMDDPERE VLEWRFDSRL  
AFHHVARELH PEYFKNC \* [SEQ ID NO: 80]

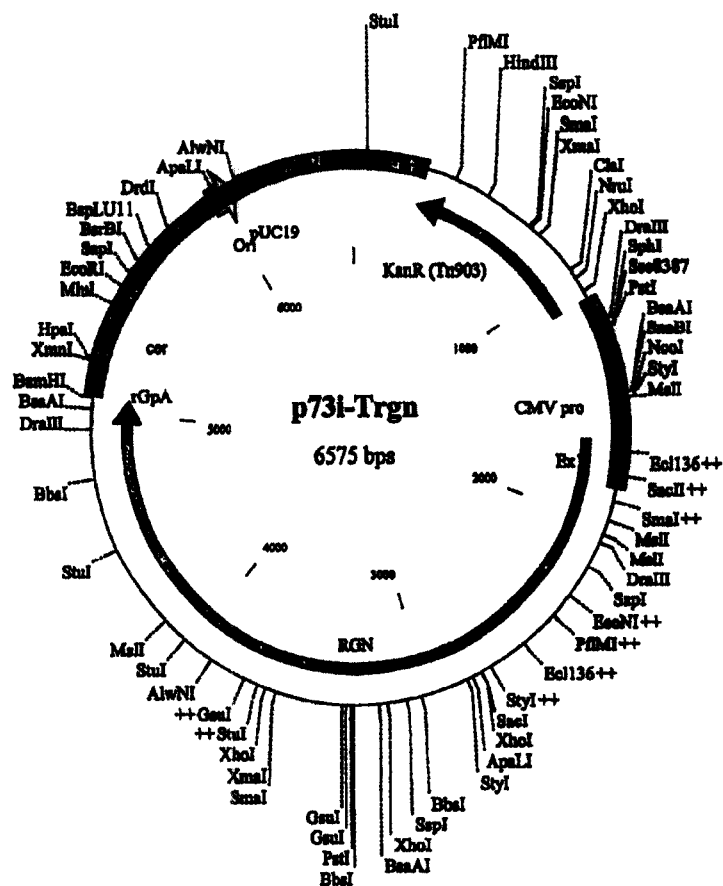


图 18

## Trng #11 插入序列

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC  
5 GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG  
ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCGGAGAACCCA  
TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG  
GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC  
ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC  
10 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC  
ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC  
CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG  
CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC  
GTGGGCTCTGACCTGGAATCCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG  
15 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA  
TTCTTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTC  
CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC  
AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG  
CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCTCACGGAGGAAGCCGAGCTCGAG  
20 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACCGCGTGTAATGACCCC  
TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCGAGTGGACATACCAG  
ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCCGATGAGGGGC  
GCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG  
TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC  
25 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC  
GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG  
GGCGCGGAGACATTTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAAG  
GCCGGGTACGTACCAACCGGGCCGCCAGAAGGTGCTCACCCTGACCGACACCAAC  
AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG  
30 GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTCAGGCCAGCCGGAC  
CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG  
GTCTACCTCGCTGGGTCCCGGCCATAAGGGCATTTGGCGGCAACGAGCAGGTTCGAC  
AAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTCAACCT  
CAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTTA  
35 AAAGAAAAGGGGGGACTGGAAGGGCTAATTCACCTCCCAAAGAAGACAAGATATCCTT  
GATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCA  
GGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTT  
GAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCCT  
GTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGAC  
40 AGCCGCCTAGCATTTTCATCACGTGGCCCCGAGAGCTGCATCCGGAGTACTTCAAGAAC  
TGCTGAATGGGTGCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAG  
AAAATTAGGCTGCGCCCGGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGG  
GCCTCAGAGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAG  
GGATGTGCGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAG  
45 CTGAGGTCTTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAG

ATTAAGGATACCAAGGAGGCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAG  
 AAGAAGGCCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAAC  
 TATCCTATTGTCCAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGG  
 ACGCTCAATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATC  
 5 CCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCCT  
 AATACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAG  
 GAGGCAGCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCCGA  
 CAGATGCGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAG  
 CAAATCGGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGG  
 10 TGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTT  
 GACATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCCGGTTTATAAG  
 ACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACCGAGACACTC  
 CTGGTACAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCT  
 GCCACCCTGGAAGAGATGATGACCGCTGTGAGGAGTAGGCGGACCCGGACACAAA  
 15 GCCAGAGTGTG [SEQ ID NO: 81]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKKGKIS  
 KIGPENFYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA  
 GLKKKKSVTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV  
 20 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDL YVGSDLSEIGQH  
 RTKIEBLRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE  
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEB  
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGGQWY YQIYQEPFKN  
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVWNGKTPK PKLPIQKETW  
 25 ETWWTIFYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR  
 ETKLKAGYV TNRGRQKVVT LTDTTNQKTE LQAIYLALQD SGLEVNIVTD  
 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLANVPA HKGIGGNEBQV  
 DKLVSAGIRK VLMVGFPVTP QVPLRPMTYK AAVDLSEFLK EKGGLLEGLIH  
 SQRQDILD LWIYHTQGYFP DWQNYTPGPG VRYPLTFGWC YKLVPVEPDK  
 30 VEEANKGENT SLLHPVSLHG MDDPEREVLE WRFDSRLAFH HVARELHPEY  
 FKNCMGARAS VLSGGELDRW EKIRLRPGGK KKYKLKHIVW ASRELERFAV  
 NPGLLETSEG CRQILGQLQP SLQTSSEELR SLYNTVATLY CVHQRIEIKD  
 TKEALDKIEE EQNKSKKKAQ QAAADTGHEN QVSONYPIVQ NIQGMVHQA  
 ISPRTLNAWV KVVEEKAFSP EVIPMFSALE EGATPQDLNT MLNTVGGHQA  
 35 AMQMLKETIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI AGTTSTLQEQ  
 IGWMTNNPPI PVGEIYKRWI ILGLNKIVRM YSPTSILDIR QGPKEPFRDY  
 VDRFYKTLRA EQASQEVKNW MTETLLVQNA NPDCKTILKA LGPAATLEEM  
 MTACQGVGGP GHKARVL [SEQ ID NO: 82]

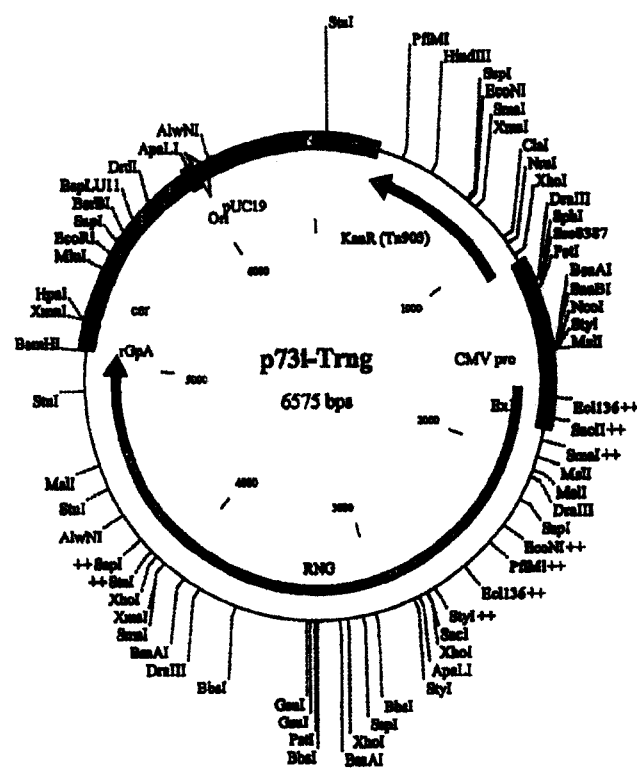


图 19

TgnR (F1) 序列:

```
ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
5 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCCTCG
AGGGAGCTTGAACGGTTTGGCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAAG
10 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
ATTGTCCAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGACGCTC
AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
15 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACTCTACACTGCAAGAGCAAATC
GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
ATCCTGGGCCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT
AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
20 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
AAGGCAGCTGTAGATCTTAGCCACTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTA
25 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC
TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
AAAGGAGAGAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCG
GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTTCATCAGTGGCC
30 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCATGGGCCCCATCAGTCCCATCGAG
ACCGTGCCCGTGAAGCTGAAACCCGGGATGACGCGCCCCAAGGTCAAGCAGTGGCCA
CTCACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAG
GGCAAGATCAGCAAGATCGGGCCTGAGAACCATAACAACCCCCCGTGTGTGCCATC
AAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAG
35 CGGACCCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCCATCCGGCCGGCTGAAG
AAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTG
GACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACC
CCTGGCATCAGATATCAGTACAACGTCCTCCCCCAGGGCTGGAAGGGCTCTCCCGCC
ATTTTCCAGAGCTCCATGACCAAGATCCTGGAGCCGTTTTCGGAAGCAGAACCCCGAT
40 ATCGTCATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGGG
CAGCATCGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACC
ACTCCGGACAAGAAGCATCAGAAGGAGCCGCCATTCTGaaGATGGGCTACGAGCTC
CATCCCGACAAGTGGACCGTGCAGCCTATCGTCCTCCCCGAGAAGGACAGCTGGACC
GTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCC
45 GGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAG
```

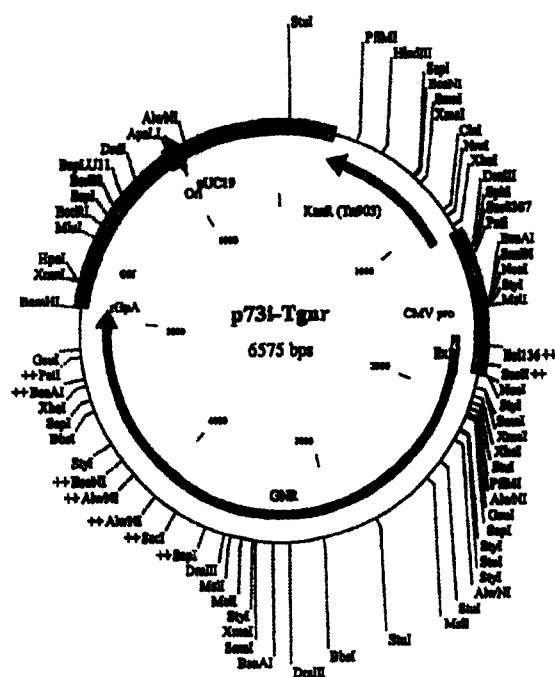
GTGATTCCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAGATCCTG  
 AAGGAGCCCGTGACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAATC  
 CAGAAGCAGGGCCAGGGCCAGTGGACATACCAGATTTACCAGGAGCCTTTCAAGAAC  
 CTCAAGACCGGCAAGTACGCCCGCATGAGGGGCGCCACACCAACGATGTCAAGCAG  
 5 CTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACA  
 CCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATAT  
 TGGCAGGCCACCTGGATTCCCGAGTGGGAGTTCTGTGAATACACCTCCTCTGGTGAAG  
 CTGTGGTACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGAC  
 GGCGCGGCCAACCGCGAAACAAAGCTCGGGAAGGCCGGGTACGTCACCAACCGGGGC  
 10 CGCCAGAAGGTCGTCAACCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCC  
 ATCTATCTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAG  
 TACGCGCTGGGCATTATTCAGGCCCAGCCGACAGTCCGAGAGCGAACTGGTGAAC  
 CAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCC  
 CATAAGGGCATTGGCGGCAACGAGCAGGTGACAAGCTGGTGAGTGCGGGGATTAGA  
 15 AAGGTGCTGTAA [SEQ ID NO: 83]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGI  
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA  
 LDKIBEEQNK SKKKAQQA AAA DTGHSNQVSQ NYPIVQNIQG QMVHQAISPR  
 20 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAAMQM  
 LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM  
 TNNPPIPVGE IYKRWIILGL NKIVRMYSP SILDIRQGPB EPFRDYVDRF  
 YKTLRAEQAS QEVKNWMTET LLVQANANPDC KTIKALGPA ATLEEMMTAC  
 QGVGGPGHKA RVLVGVFPVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI  
 25 HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYPITFGW CYKLVPVEPD  
 KVEEANKGEN TSLHHPVSLH GMDDPEREVL EWRFDSEILAF HHVARELHPE  
 YFKNCMPIS PIETVSVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK  
 EGKISKIGPE NPYNTFVFAI KKKDSTKWRK LVDFRELNKR TQDFWEVQLG  
 IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR  
 30 YQYNVLPQGW KGSPAIFQSS MTKILEPFRK QNPDIYIYQY MDDLIVGSDL  
 EIGQHRTKIE ELRQHLLRWG LTPDKKHQK EPPFLWMGYE LHPDKWTVQP  
 IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTALTEVI  
 PLTEEAEL EL AENREILKEP VHGVYDPSK DLIABIQKQG QGQWYQIYQ  
 EPFKNLKTGK YARMRGATN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI  
 35 QKETWETWWT EYWQATWIPE WEFVNTPLV KLWYQLEKEP IVGAETFYVD  
 GAANRETKLG KAGYVTNRGR QKVVTLTDTT NQKTELQAIY LALQDSGLEV  
 NIVTDSQYAL GIIQAQPDQS ESELVNIIE QLIKKEKVYL AWVPAHKGIG  
 GNEQVDKLVS AGIRKVL\*  
 [SEQ ID NO: 84]

40

45





5

10

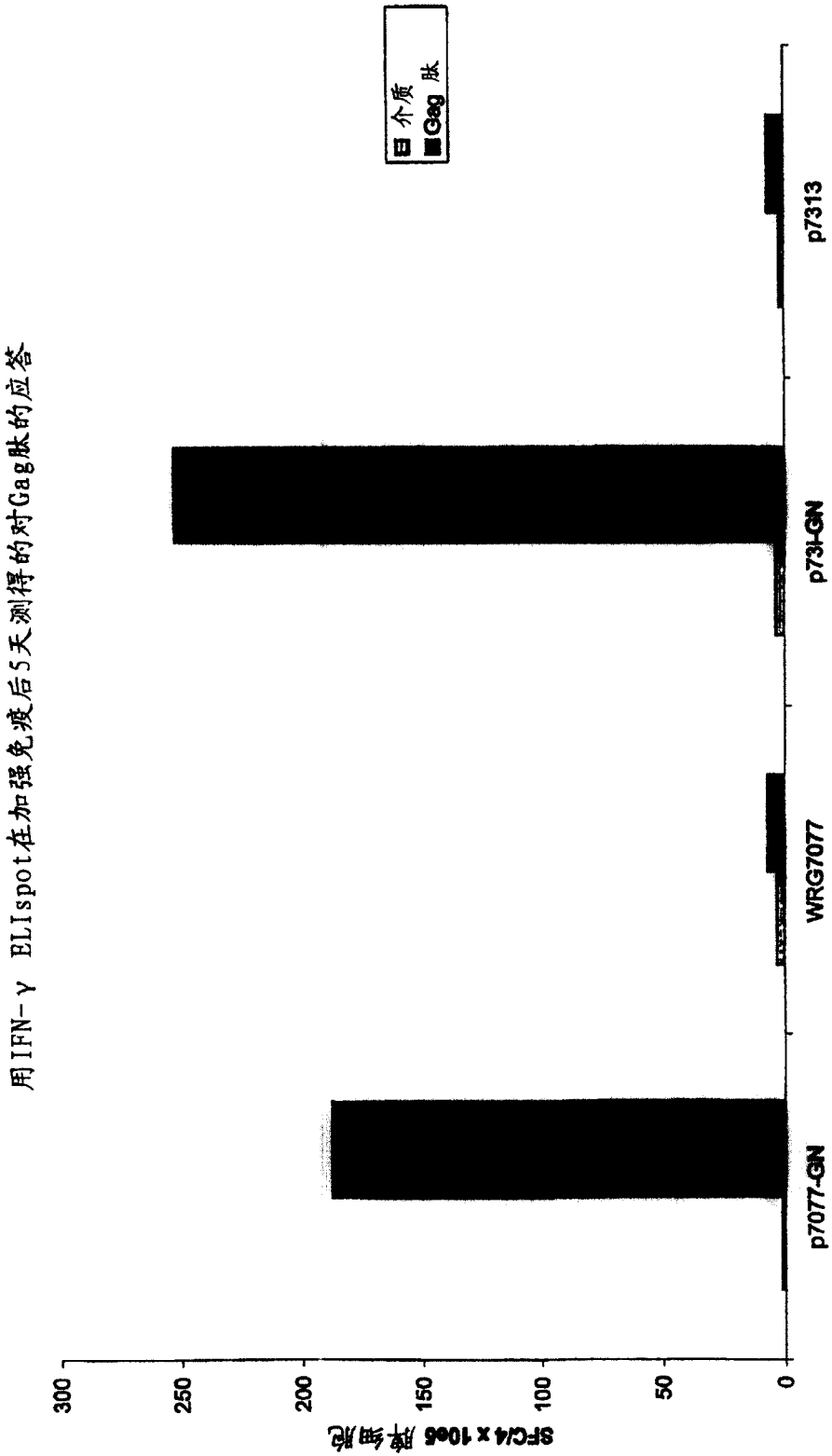


图 20

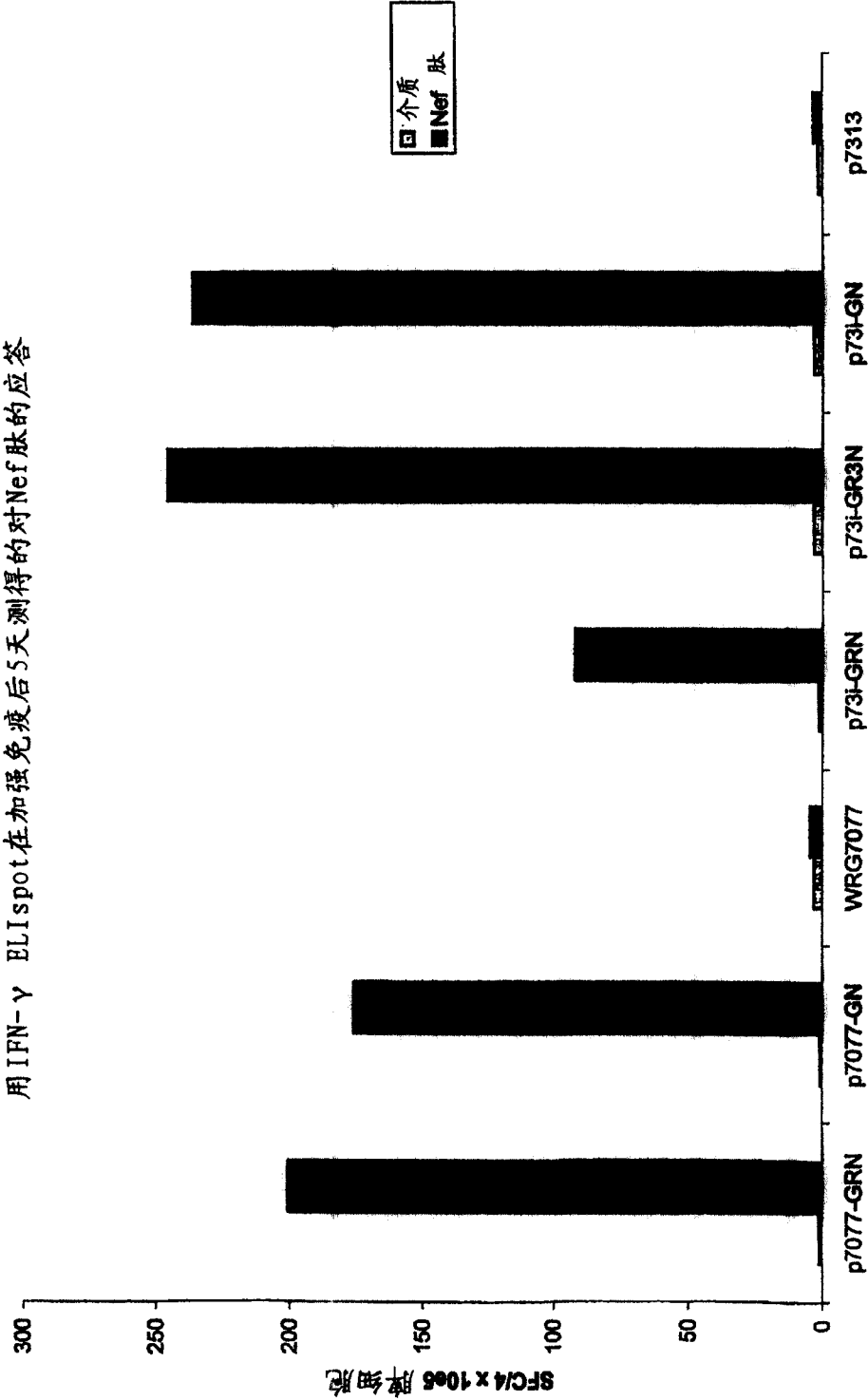


图 21

PG4653

用IFN- $\gamma$  ELIspot在加强免疫后5天测得的对Rt肽的应答

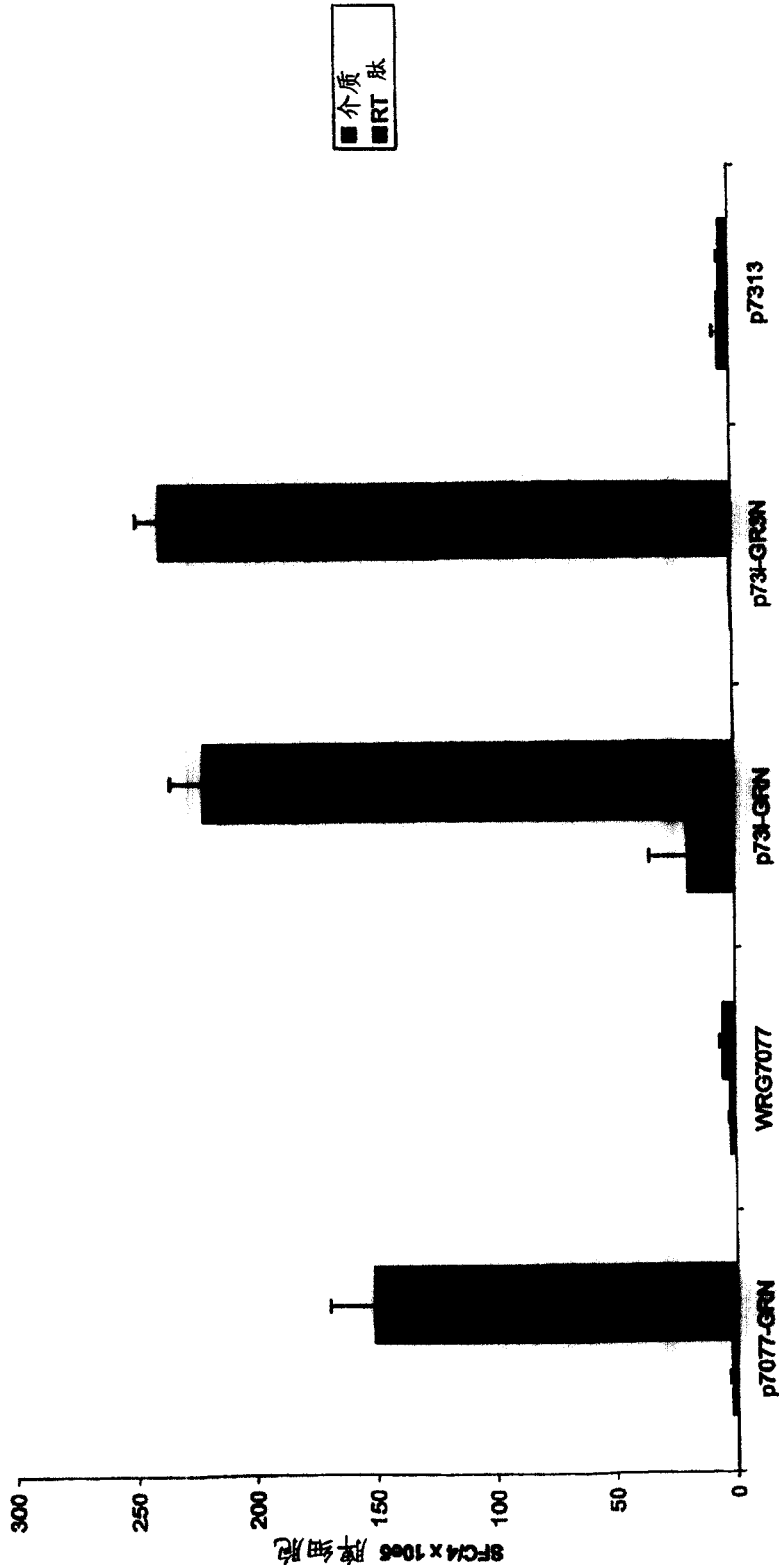
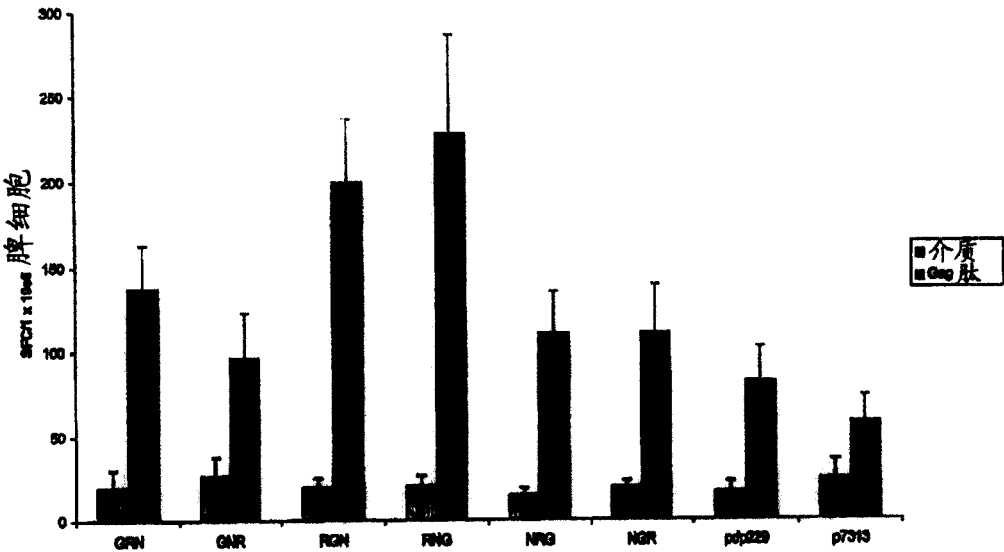


图 22

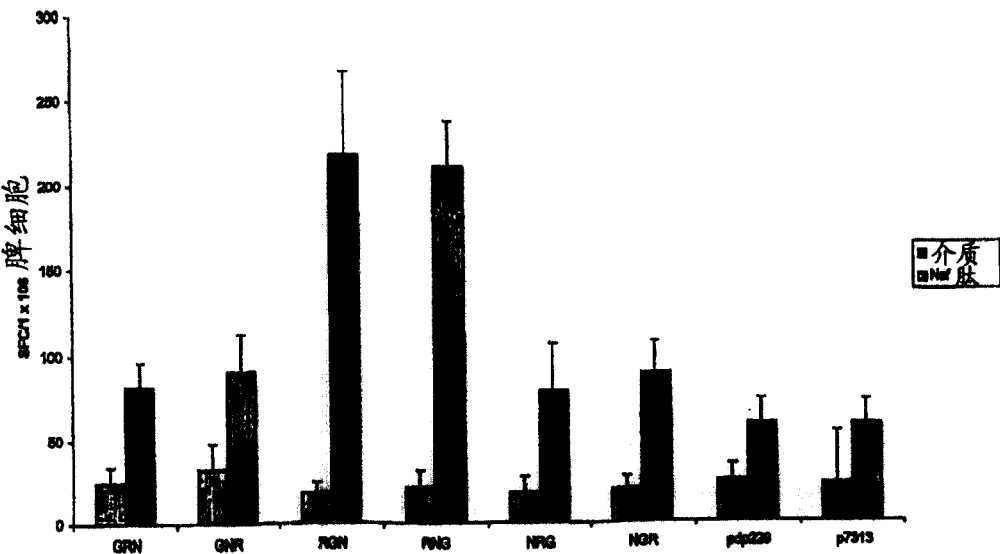
图 23

在用重建构建体免疫之后测得的小鼠模型中CD8免疫应答的实例。  
5 构建体包括Gag、RT和Nef作为三重融合。G是Gag，R是RT，N是Nef

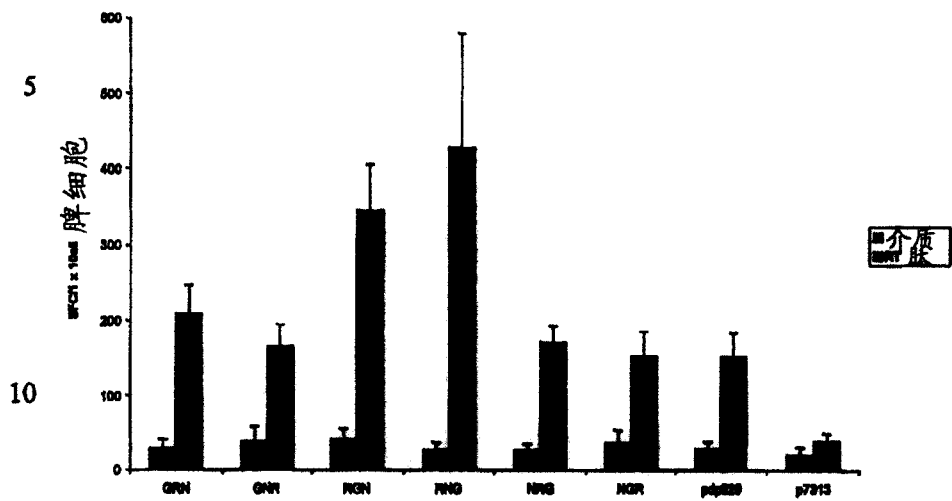
用IFN- $\gamma$  ELISpot在第35天（加强免疫后7天）测得的对Gag肽的应答



用IFN- $\gamma$  ELISpot在第35天（加强免疫后7天）测得的对Nef肽的应答



用IFN- $\gamma$  ELIspot在第35天（加强免疫后7天）测得的对RT肽的应答



用ELISA在第35天（加强免疫后7天）测得的Gag抗体

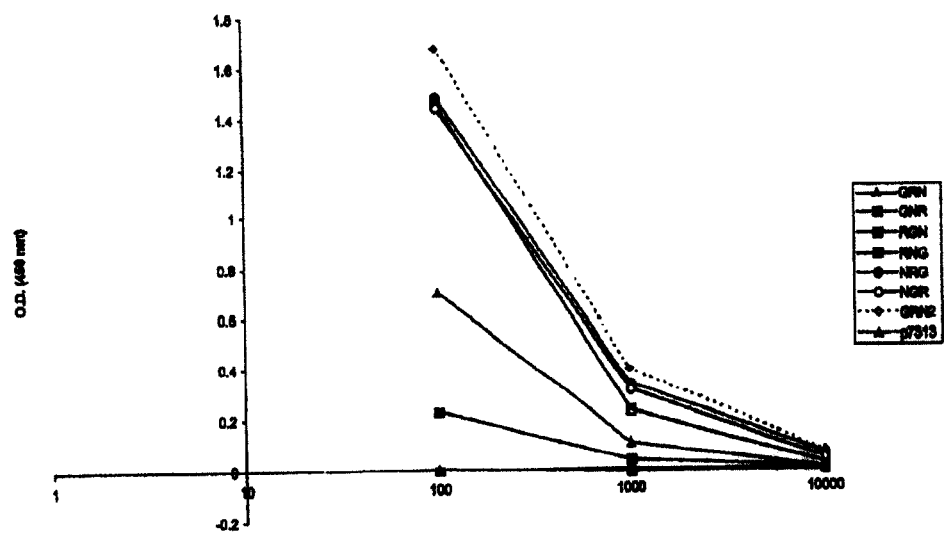


图 24