(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105624193 B
(45)授权公告日 2021.07.13

(21)申请号 201510624987.9

(22)申请日 2009.09.26

(65)同一申请的已公布的文献号
申请公布号 CN 105624193 A

(43)申请公布日 2016.06.01

(30)优先权数据
61/100,666 2008.09.26 US
61/120,618 2008.12.08 US
61/186,823 2009.06.13 US

(62)分案原申请数据
200980147515.1 2009.09.26

(73)专利权人 索元生物医药(美国)有限公司
地址 美国加利福尼亚州

(72)发明人 哈利·E·格鲁贝尔
道格拉斯·乔利 奥马尔·佩雷斯
克里斯托弗·R·洛格

(74)专利代理机构 北京市万慧达律师事务所 11111
代理人 谢敏楠

(51)Int.Cl.
C12N 15/867(2006.01)
C12N 15/48(2006.01)
A61K 35/768(2015.01)
A61P 35/00(2006.01)

(56)对比文件
US 6410313 B1,2002.06.25
Chien-Kuo Tai等.Single-Shot, Multicycle Suicide Gene Therapy by Replication-Competent Retrovirus Vectors Achieves Long-Term Survival Benefit in Experimental Glioma.《Molecular Therapy》.2005,第12卷(第5期),第843页左栏第45行-第844页左栏第11行和图2.
Chien-Kuo Tai等.Single-Shot, Multicycle Suicide Gene Therapy by Replication-Competent Retrovirus Vectors Achieves Long-Term Survival Benefit in Experimental Glioma.《Molecular Therapy》.2005,第12卷(第5期),第843页左栏第45行-第844页左栏第11行和图2.
刘扬 等.逆转录病毒载体介导shRNA对人胚胎肾细胞中p53 基因的沉默作用.《吉林大学学报(医学版)》.2006,第32卷(第2期),228-231页.

审查员 黄磊

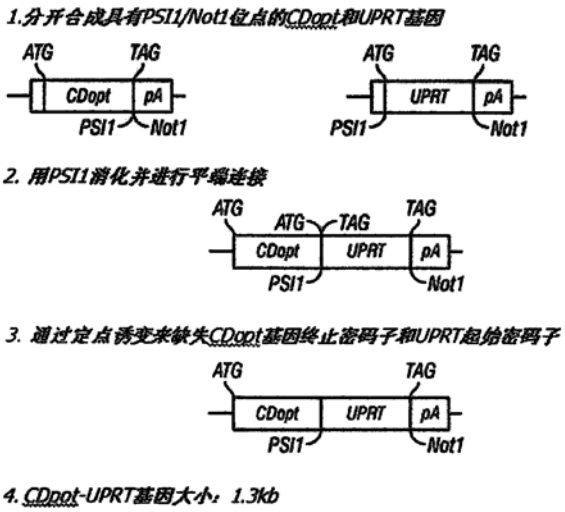
权利要求书2页 说明书67页
序列表81页 附图87页

(54)发明名称

重组载体

(57)摘要

本公开书提供经修饰的胞嘧啶脱氨酶(CD)。本公开书进一步涉及表达或包含所述经修饰CD的细胞和载体,和使用所述经修饰CD治疗疾病和失调的方法。



CN 105624193 B

1.一种重组可复制逆转录病毒,其包含:

逆转录病毒GAG蛋白;

逆转录病毒POL蛋白;

逆转录病毒包膜;

逆转录病毒多聚核苷酸,其包含:在所述逆转录病毒多聚核苷酸序列的3'端的长末端重复(LTR)序列、在所述逆转录病毒多聚核苷酸的5'端的启动子序列、gag核酸域、pol核酸域和env核酸域,所述启动子适合在哺乳动物细胞中表达;

包含可操作地连接于异源多聚核苷酸的内部核糖体进入位点的IRES盒和包含位于RNAi多聚核苷酸上游并可操作地连接于该RNAi多聚核苷酸的polIII启动子的polIII盒,其中所述IRES盒和polIII盒位于所述3'LTR的5'和编码所述逆转录病毒包膜的所述env核酸域的3';和

在靶细胞中逆转录、包装和整合所必需的顺式作用序列,

其中所述env核酸域由SEQ ID NO:19或22的核苷酸编号6359到核苷酸8323的序列组成,和

所述3'LTR来源于γ逆转录病毒,所述3'LTR包括U3-R-U5域,并且所述3'LTR由如SEQ ID NO:19或22的核苷酸9405到9998所述的序列组成。

2.根据权利要求1所述的逆转录病毒,其中所述逆转录病毒多聚核苷酸序列从由选自以下各项组成的组的病毒被重组工程化:鼠白血病病毒(MLV)、莫洛尼鼠白血病病毒(MoMLV)、猫白血病病毒(FeLV)、狒狒内源性逆转录病毒(BEV)、猪内源性病毒(PERV)、源于猫的逆转录病毒RD114、松鼠猴逆转录病毒、异嗜性小鼠白血病病毒相关病毒(XMRV)、禽网状内皮组织增生病毒(REV)或长臂猿白血病病毒(GALV)。

3.根据权利要求1所述的逆转录病毒,其中所述逆转录病毒包膜是双嗜性MLV包膜。

4.根据权利要求1所述的逆转录病毒,其中所述逆转录病毒是γ逆转录病毒。

5.根据权利要求1所述的逆转录病毒,其中所述靶细胞是具有细胞增殖失调的细胞。

6.根据权利要求1所述的逆转录病毒,其中所述靶细胞是赘生性细胞。

7.根据权利要求5所述的逆转录病毒,其中所述细胞增殖失调选自由肺癌、结肠直肠癌、乳腺癌、前列腺癌、尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌、类风湿性关节炎和其它自身免疫疾病组成的组。

8.根据权利要求1所述的逆转录病毒,其中在所述逆转录病毒多聚核苷酸的5'端的所述启动子序列与生长调节基因关联。

9.根据权利要求1所述的逆转录病毒,其中在所述逆转录病毒多聚核苷酸的5'端的所述启动子序列包括组织特异性启动子序列。

10.根据权利要求9所述的逆转录病毒,其中所述组织特异性启动子序列包含至少一个雄激素反应元件(ARE)。

11.根据权利要求1所述的逆转录病毒,其中所述启动子包括具有如SEQ ID NO:19、20或22的核苷酸1到核苷酸582所述的序列的CMV启动子,且可包含对一个或多个核酸碱基的修饰,所述启动子能引导和起始转录。

12.根据权利要求1所述的逆转录病毒,其中所述启动子包括CMV-R-U5域多聚核苷酸。

13.根据权利要求12所述的逆转录病毒,其中所述CMV-R-U5域包括与MLV R-U5区域连

接的来自人类巨细胞病毒的立即早期启动子。

14.根据权利要求13所述的逆转录病毒，其中所述CMV-R-U5域多聚核苷酸包括如SEQ ID NO:19、20或22的核苷酸1到核苷酸1202所述的序列，或与如SEQ ID NO:19、20或22所述的序列具有至少95％同一性的序列，其中所述多聚核苷酸促进与其可操作地连接的核酸分子的转录。

15.根据权利要求1所述的逆转录病毒，其中所述gag多聚核苷酸来源于γ逆转录病毒。

16.根据权利要求15所述的逆转录病毒，其中所述gag核酸域包括SEQ ID NO:19或22的核苷酸编号1203到核苷酸2819的序列，或与其具有至少95％、98％、99％或99.8％同一性的序列。

17.根据权利要求1所述的逆转录病毒，其中所述多聚核苷酸的pol域来源于γ逆转录病毒。

18.根据权利要求17所述的逆转录病毒，其中所述pol域包括SEQ ID NO:19或22的核苷酸编号2820到核苷酸6358的序列，或与其具有至少95％、98％、99％或99.9％同一性的序列。

19.根据权利要求1所述的逆转录病毒，其中所述polIII启动子是H1或U6启动子。

20.根据权利要求19所述的逆转录病毒，其中所述polIII启动子是H1启动子。

21.根据权利要求1所述的逆转录病毒，其中所述异源核酸序列包括抑制性多聚核苷酸。

22.根据权利要求21所述的逆转录病毒，其中所述抑制性多聚核苷酸包括miRNA、RNAi或siRNA序列。

23.一种用于产生权利要求1的逆转录病毒的重组逆转录病毒多聚核苷酸基因组。

24.权利要求1的逆转录病毒与抗癌剂或化学治疗剂在制备用于治疗细胞增殖失调的药物的用途，其中所述药物被配制为在允许所述逆转录病毒感染具有细胞增殖失调的细胞的条件下，对患有所述细胞增殖失调的受试对象施用所述逆转录病毒，和使所述受试对象与抗癌剂或化学治疗剂接触。

25.根据权利要求24所述的用途，其中所述抗癌剂选自由贝伐单抗、派加他尼、兰尼单抗、索拉非尼(sorafenib)、舒尼替尼(sunitinib)、AE-941、VEGF Trap、帕唑帕尼(pazopanib)、凡德他尼(vandetanib)、瓦他拉尼(vatalanib)、西地尼布(cediranib)、芬维A胺(fenretinide)、角鲨胺(squalamine)、INGN-241、口服四硫钼酸盐、四硫钼酸盐、Panzem NCD、2-甲氧基雌二醇、AEE-788、AG-013958、贝伐西尼钠、AMG-706、阿西替尼(axitinib)、BIBF-1120、CDP-791、CP-547632、PI-88、SU-14813、SU-6668、XL-647、XL-999、IMC-1121B、ABT-869、BAY-57-9352、BAY-73-4506、BMS-582664、CEP-7055、CHIR-265、CT-322、CX-3542、E-7080、ENMD-1198、OSI-930、PTC-299、Sirna-027、TKI-258、Veglin、XL-184和ZK-304709组成的组。

26.根据权利要求24所述的用途，其中所述逆转录病毒以每克脑重量10^3到10^7TU施用。

27.根据权利要求26所述的用途，其中所述逆转录病毒以每克脑重量10^4到10^6TU施用。

重组载体

[0001] 本申请是申请日为2009年9月26日、申请人为托卡根公司、发明名 称为“重组载体”的中国专利申请200980147515.1的分案申请。

[0002] 相关申请的交叉参考

[0003] 本申请主张2008年9月26日提交的美国临时申请第61/100,666号、200 年12月8日提交的美国临时申请第61/120,618号和2009年6月13日提交的美 国临时申请第61/186,823号的优先权,其揭示内容以引用方式并入本文中。

技术领域

[0004] 本公开书涉及用于治疗细胞增殖的可复制逆转录病毒载体。本公开书 另外涉及使用所述可复制逆转录病毒载体递送和表达异源核酸。

背景技术

[0005] 对细胞和受试对象递送基因和异源核酸的有效方法一直是科学研发 和疾病和失调的可能疗法的研究目标。

发明内容

[0006] 本公开书提供一种重组可复制逆转录病毒(RCR),其包含:逆转录 病毒GAG蛋白;逆转录病毒POL蛋白;逆转录病毒包膜;逆转录病毒 多聚核苷酸,其包含位于逆转录病毒多聚核苷酸序列3’端的长末端重复 (LTR)序列、位于逆转录病毒多聚核苷酸5’端的启动子序列、gag核酸 域、pol核酸域和env核酸域,所述启动子适合在哺乳动物细胞中表达; 盒,其包含可操作地连接于异源多聚核苷酸的内部核糖体进入位点 (IRES),其中所述盒位于3’LTR的5’和编码逆转录病毒包膜的env核 酸域的3’;和靶细胞中的逆转录、包装和整合所必需的顺式作用序列, 其中与pACE载体(SEQ ID NO:21)相比,RCR在6次传代后保持较高 的复制能力。在一个实施方案中,逆转录病毒多聚核苷酸序列来源于鼠白 血病病毒(MLV)、莫洛尼鼠白血病病毒(MoMLV)、猫白血病病毒(FeLV) 或长臂猿白血病病毒(GALV)。在另一个实施方案中,MLV是双嗜性 MLV。在另一个实施方案中,逆转录病毒是致癌逆转录病毒或γ逆转录 病毒。在另一个实施方案中,靶细胞是具有细胞增殖失调的细胞。细胞增 殖失调可选自由(但不限于)肺癌、结肠直肠癌、乳腺癌、前列腺癌、尿 道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌、类风湿 性关节炎和其它自身免疫疾病组成的组。在一个实施方案中,启动子包括 具有如SEQ ID NO:19、20或22的核苷酸1到约核苷酸582所述的序列的CMV启动子,且可包含对一个或多个核酸碱基的修饰,所述启动子能 引导和引发转录。在另一个实施方案中,启动子包括如SEQ ID NO:19、20或22的核苷酸1到约核苷酸582所述的序列。在另一个实施方案中, 启动子包括CMV-R-U5域多聚核苷酸。在一个实施方案中,CMV-R-U5 域包括与MLV R-U5区域连接的来自人类巨细胞病毒的立即早期启动子。 在另一个实施方案中,CMV-R-U5域多聚核苷酸包括如SEQ ID NO:19、20或22的约核苷酸1到约核苷酸1202所述的序列,或与如SEQ ID NO:19、20或22所述的序列具有至少95%同一性的序列,

其中所述多聚核苷酸促 进与其可操作地连接的核酸分子的转录。在另一个实施方案中，多聚核苷 酸的gag和pol来源于致癌逆转录病毒或γ逆转录病毒。gag核酸域可包 括SEQ ID NO:19或22的约核苷酸编号1203到约核苷酸2819的序列，或与其具有至少95％、98％、99％或99.8％同一性的序列。pol域可包括 SEQ ID NO:19或22的约核苷酸编号2820到约核苷酸6358的序列，或 与其具有至少95％、98％、99％或99.9％同一性的序列。在一个实施方案中，env域编码双嗜性env蛋白。env域可包含SEQ ID NO:19或22的约 核苷酸编号6359到约核苷酸8323的序列，或与其具有至少95％、98％、 99％或99.8％同一性的序列。载体的IRES域可以是任何IRES，但在一个 实施方案中，IRES来源于脑心肌炎病毒。在另一个实施方案中，IRES包 括SEQ ID NO:19或22的约核苷酸编号8327到约核苷酸8876的序列，或与其具有至少95％、98％或99％同一性的序列。

[0007] 载体可包含任何数量的不同的异源多聚核苷酸。举例来说，异源多聚 核苷酸可包含细胞因子、siRNA、miRNA或RNAi分子、靶向序列、结 合域、细胞毒性基因、单链抗体或其任何组合。当异源多聚核苷酸是用于 例如siRNA、miRNA或RNAi等非翻译RNA时，不需要IRES，但也可 包括IRES以用于另一种被翻译基因RNA，且可使用任何种类的逆转录 病毒。在另一个实施方案中，异源多聚核苷酸包括具有如SEQ ID NO:3、 5、11、13、15或17所述的序列的多聚核苷酸。在另一个实施方案中， 异源序列编码包括如SEQ ID NO:4所述的序列的多肽。异源核酸是经人 类密码子优化的且编码如SEQ ID NO:4所述的多肽。在另一个实施方案中，异源核酸包括如SEQ ID NO:19或22的约核苷酸编号8877到约9353 所述的序列。在一个实施方案中，3'LTR来源于致癌逆转录病毒或γ逆 转录病毒。在另一个实施方案中，3'LTR包括U3-R-U5域。在另一个实 施方案中，3'LTR包括如SEQ ID NO:19的约核苷酸9405到约9998所述 的序列，或与其具有至少95％、98％或99.5％同一性的序列。

[0008] 本公开书提供一种包括如SEQ ID NO:19、20或22所述的序列的多 聚核苷酸。

[0009] 本公开书提供一种分离的多聚核苷酸，其从5'到3'包含：人类巨细 胞病毒的立即早期启动子与MLV R-U5区的CMV-R-U5融合体；逆转录 酶的引物结合位点PBS；5'剪接位点；ψ包装信号；MLV群特异性抗原 的gag编码序列；MLV聚合酶多聚蛋白的pol编码序列；3'剪接位点； MLV品系4070A的包膜蛋白的4070A env编码序列；脑心肌炎病毒的内 部核糖体进入位点(IRES)；经修饰的胞嘧啶脱氨酶编码序列；聚嘌呤 束(polypurine tract)；和U3-R-U5MLV长末端重复序列。

[0010] 本公开书提供一种治疗患有细胞增殖失调的受试对象的方法，其包含 使受试对象与编码具有胞嘧啶脱氨酶活性的本公开书的多肽的多聚核苷 酸在允许多聚核苷酸表达的条件下接触，和使受试对象与5-氟胞嘧啶接 触。

[0011] 本公开书还提供一种治疗受试对象的细胞增殖失调的方法，其包含使 受试对象与本发明的逆转录病毒接触，其中异源核酸序列编码抑制赘生性 细胞增殖的治疗性蛋白。在一个实施方案中，逆转录病毒包含编码具有如 SEQ ID NO:4、12、14、16或18所述的序列的多肽的多聚核苷酸。

[0012] 本公开书的一个或多个实施方案的细节阐述于附图和以下具体实施 方式中。根据具体实施方式和附图并根据权利要求书，其它特征、目标和 优势可显而易见。

附图说明

[0013] 图1A-C显示野生型酵母胞嘧啶脱氨酶(SEQ ID NO:2)和本公开书 的胞嘧啶脱氨酶(SEQ ID NO:4)和本公开书的其它序列的比对。

[0014] 图2显示细胞杀死数据的曲线图,其显示经修饰的载体与原始野生 型CD相比更为有效。所述曲线图还显示新的经修饰骨架(T5.0007)在 细胞杀死方面比老的骨架(pACE-CD)更为有效。还显示将各种载体构 建体和其名称按目录分类的表格。

[0015] 图3A-D显示(a)本公开书的重组逆转录病毒载体的图解;(b)本 公开书的多聚核苷酸的质粒图谱;(c和d)本公开书的多聚核苷酸的序列 (SEQ ID NO:19)。

[0016] 图3E显示修饰pACE-GFPemd质粒的先前骨架的3种方式。

[0017] 图3F-1和图3F-2显示pAC3-yCD2(aka T5.0002)的结构。

[0018] 图3F-3至图3F-5显示pAC3-yCD2的连续序列。

[0019] 图3F-6至图3F-18显示pAC3-yCD2(唯一消化物序列)。

[0020] 图4显示在感染的U-87细胞中观察到的yCD2蛋白的含量高于野生 型yCD蛋白。

[0021] 图5显示如使用PCR扩增所评定,本公开书的载体在12个病毒传代 周期后保持遗传稳定。所述图还显示本公开书的载体与载体pACE-CD (Kasahara等人)相比在更长传代后更为稳定。具体来说,pAC3-CD比 pACE-CD更稳定,证明经改变的骨架使载体更稳定。此外,pACE-yCD1 (T5.0001)和-yCD2(T5-0002)都比pAC-yCD稳定得多,证明对转基 因的编码序列所作的微小和沉默改变对稳定性具有很大的影响,从而产生 优良性质。T5.0003最不稳定且T5.0004和T5.0005似乎与在小鼠肿瘤中 显示功效的pACE-CD(W.Wang等人Hum Gene Ther 14:1172003;Tai 等人Mol Ther 12:8422005)大致相等。

[0022] 图6显示(A)细胞杀死分析;和(B)用不同载体感染的细胞的胞 嘧啶脱氨酶比活性。(A)显示被感染U87细胞暴露于渐增水平的5-FC 时本公开书的胞嘧啶脱氨酶和载体比原始pACE-CD至少一样或更好地杀 死被感染细胞。(B)显示本公开书的CD比活性(T5.0007、T5.0001和 T5.0002)均高于pACE-CD(T5.0000),次序为 T5.0000＜T5.0007＜T5.0001＜T5.0002。

[0023] 图7显示用本公开书的CD载体体内处理且从用4个5-FC周期处理 的小鼠外植的U-87肿瘤对药物仍然敏感。

[0024] 图8显示小鼠脑癌模型的Kaplan-Meyer生存分析中的给药信息和治 疗效果。

[0025] 图9显示同系小鼠模型的Kaplan-Meyer生存分析中的给药信息和治 疗效果。

[0026] 图10A-D显示用于产生包含具有CD、OPRT和UPRT活性的多肽的 本公开书的各种实施方案的方案。

[0027] 图11:A是含有人类初级前体miR-128-1、人类初级前体miR-128-2 和与人类H1启动子连接的人类前体miR-128的多聚核苷酸序列的MLV 逆转录病毒载体pAC3骨架(分别命名为pAC3-miR128-1、pAC3-miR-128-2和pAC3-H1-shRNAmiR128)的图解载体图谱。B是含有与人类H1启动子连接的人类前体miR-128的多聚核苷酸序列的MLV 逆转录病毒载体pAC3-yCD2骨架(命名为pAC3-yCD2-H1-shRNAmiR128) 的图解载体图谱。

[0028] 图12:A显示由qPCR分析的含miR-128载体(pAC3-miR-128-1、pAC3-miR-128-2和pAC3-H1-shRNAmiR128)在人类纤维肉瘤细胞 HT1080中的复制动力学的比较。所述图是通过将从qPCR获得的C(t)值 的倒数与病毒复制期间的各个时间点作图来产生。B显示由qPCR

分析的 含miR-128载体(pAC3-miR-128-1、pAC3-miR-128-2和 pAC3-H1-shRNAmiR128)在人类神经胶质瘤细胞U87-MG中的复制动力 学的比较。所述图是通过将从qPCR获得的C(t)值的倒数与病毒复制期间 的各个时间点作图来产生。

[0029] 图13显示从用含miR-128载体转导的细胞中的成熟miR-128表达的 相对定量。

[0030] 图14显示从用含miR-128载体转导的细胞中的Bmi-1基因表达的相 对定量。

[0031] 图15是含有单拷贝142-3pT靶序列(命名为pAC3-emd-142-3pT)和 4个串联重复的142-3pT(命名为pAC3-emd-142-3pT4X)的MLV逆转录 病毒载体pAC3-emd的图解载体图谱。

[0032] 图16是含有142-3pT靶序列的单拷贝(命名为pAC3-yCD2-142-3pT) 和142-3pT的4个串联重复序列(命名为pAC3-yCD2-142-3pT4X)的MLV 逆转录病毒载体pAC3-yCD2的图解载体图谱。

[0033] 图17:A显示由qPCR分析的含142-3pT载体(pAC3-emd-142-3pT、 pAC3-emd-142-3pT4X、pAC3-yCD2-142-3pT和pAC3-yCD2-142-3pT4X) 和其母体载体(pAC3-emd和pAC3-yCD2)在人类纤维肉瘤细胞HT1080 中的复制动力学的比较。所述图是通过将从qPCR获得的C(t)值的倒数与 病毒复制期间的各个时间点作图来产生。B显示通过在载体传播期间的各个时间点的GFP表达的流式细胞分析所分析的,含GFP载体(pAC3-emd、 pAC3-emd-142-3pT和pAC3-emd-142-3pT4X)在人类纤维肉瘤细胞 HT1080中的复制动力学的比较。

[0034] 图18:A显示由qPCR分析的含142-3pT载体(pAC3-emd-142-3pT、 pAC3-emd-142-3pT4X、pAC3-yCD2-142-3pT和pAC3-yCD2-142-3pT4X) 和其母体载体(pAC3-emd和pAC3-yCD2)在人类神经胶质瘤细胞 U87-MG中的复制动力学的比较。所述图是通过将从qPCR获得的C(t)值 的倒数与病毒复制期间的各个时间点作图来产生。B显示通过在载体传播 期间的各个时间点的GFP表达的流式细胞分析所分析的,含GFP载体 (pAC3-emd、pAC3-emd-142-3pT和pAC3-emd-142-3pT4X)在人类神经 胶质瘤细胞U87-MG中的复制动力学的比较。

[0035] 图19显示通过在载体传播期间的各个时间点的GFP表达的流式细胞 分析所分析的含GFP载体(pAC3-emd)在小鼠和人类造血细胞中的复制 动力学。

[0036] 图20:A显示通过在载体传播期间的各个时间点的GFP表达的流式 细胞分析所分析的含GFP载体(pAC3-emd、pAC3-emd-142-3pT和 pAC3-emd-142-3pT4X)在小鼠T淋巴细胞EL4中的复制动力学的比较。 B显示通过在载体传播期间的各个时间点的GFP表达的流式细胞分析所 分析,含GFP载体(pAC3-emd、pAC3-emd-142-3pT和 pAC3-emd-142-3pT4X)在人类T淋巴细胞SUP-T1中的复制动力学的比 较。C显示通过在载体传播期间的各个时间点的GFP表达的流式细胞分 析所分析,含GFP载体(pAC3-emd、pAC3-emd-142-3pT和 pAC3-emd-142-3pT4X)在人类单核细胞U937中的复制动力学的比较。

[0037] 图21A和图21B是在肿瘤内CED输注Toca 511和钆期间从病患狗 获得的MRI影像的静止帧。应注意影像左侧的大肿瘤挤压脑的两侧并使 中线结构向右位移。白色区域是钆-Toca 511输注物。图21B显示在肿瘤 中安置两个导管。

[0038] 图22是pAC3骨架中分别编码人类IFN-γ(hIFNg)和小鼠IFN-γ (mIFNg)的MLV逆转录病毒载体的图解载体图谱。

[0039] 图23显示来自用pAC3-mIFNg载体感染的人类纤维肉瘤细胞系 HT1080中RNA水平下的mIFN-γ表达。通过RT-PCR来检测表达。

[0040] 图24显示用pAC3-hIFNg载体感染的人类纤维肉瘤细胞系HT1080 所分泌的hIFN-

γ蛋白的表达。

[0041] 图25显示用pAC3-mIFNg载体感染的人类纤维肉瘤细胞系HT1080 所分泌的mIFN-γ蛋白的表达。

[0042] 图26显示在裸小鼠模型中肿瘤内或静脉内递送AC3-GFP载体后，U87细胞中的GFP表达的流式细胞分析。通过流式细胞术测量细胞的GFP 阳性百分比。从稚裸小鼠脑分离的细胞、来自组织培养物的U87细胞或 以1的感染复数用AC3-GFP(V)体外转导的U87细胞用作对照。实施例 27(静脉内注射GFP载体)。

[0043] 图27显示另外对来自实施例27(静脉内注射GFP载体)的第1、3 和5组进行直方图分析，来测量分离的U87细胞中GFP信号的分布。GFP 表达是来自从在载体处理后14天后的小鼠脑分离的U87肿瘤细胞。

[0044] 图28是pAC3骨架中的编码人类IL-2的MLV逆转录病毒载体的图 解载体图谱。

[0045] 各个图式中的类似参考符号表示类似元件。

具体实施方式

[0046] 除非上下文明确指示，否则如本文和随附权利要求书中所用，单数形 式“一”和“所述”包括复数个指代物。因此，举例来说，提及“一细胞”时包 括多个所述细胞，且提及“所述试剂”时包括提及一种或一种以上所属领域 技术人员已知的试剂，诸如此类。

[0047] 除非另外说明，否则“或”也意味着“和/或”。类似地，“包含”和“包括” 可互换使用且不打算具有限制性。

[0048] 应进一步了解，当各个实施方案的描述使用术语“包含”时，所属领域 技术人员将认为在一些特定情况下，可替代性地使用措辞“基本上由..... 组成”或“由.....组成”来描述实施方案。

[0049] 除非另外定义，否则本文所用的所有科学技术术语的含义与本公开书 所属领域的一般技术人员通常所了解的含义相同。虽然与本文所述的方法 和材料类似或相等的方法和材料也可用于所公开的方法和组合物的实施 中，但本文描述示例性的方法、装置和材料。

[0050] 描述适用于本文的分子生物学技术(包括载体、启动子的使用和许多 其它相关课题)的一般文章包括Berger和Kimmel，Guide to Molecular Cloning Techniques，Methods in Enzymology第152卷，(Academic Press, Inc.，San Diego，Calif.)(″Berger″)；Sambrook等人，Molecular Cloning--A Laboratory Manual，第2版，第1-3卷，Cold Spring Harbor Laboratory，Cold Spring Harbor，N.Y.，1989(″Sambrook″)以及Current Protocols in Molecular Biology，F.M.Ausubel等人编，Current Protocols，a joint venture between Greene Publishing Associates，Inc.and John Wiley&Sons，Inc.，(1999年增 补)(″Ausubel″)。足以在体外扩增方法(包括聚合物链反应(PCR)、连 接酶链反应(LCR)、Qβ-复制酶扩增和其它RNA聚合酶介导的技术(例 如NASBA))中指导所属领域技术人员例如用于产生本公开书的同源核 酸的方案的实例见于Berger、Sambrook和Ausubel，以及Mullis等人(1987) 美国专利第4,683,202号；Innis等人编，(1990)PCR Protocols:A Guide to Methods and Applications(Academic Press Inc.San Diego，Calif.)(″Innis″)； Arnheim&Levinson(1990年10月1日)C&EN 36-47；The Journal OfNIH Research

(1991)3:81-94;Kwoh等人(1989)Proc.Natl.Acad.Sci.USA 86: 1173;Guatelli等人(1990)Proc.Nat′l.Acad.Sci.USA 87:1874;Lomell 等人(1989)J.Clin.Chem 35:1826;Landegren等人(1988)Science 241: 1077-1080;Van Brunt(1990)Biotechnology 8:291-294;Wu和Wallace (1989)Gene 4:560;Barringer等人(1990)Gene 89:117;和Sooknanan和Malek(1995)Biotechnology 13:563-564。Wallace等人,美国专利第 5,426,039号描述用于克隆体外扩增的核酸的改良方法。Cheng等人(1994) Nature 369:684-685和其中所引用的参考文献总结了通过PCR扩增大核 酸的改良方法,其中产生高达40kb的PCR扩增子。所属领域技术人员 将了解,使用逆转录酶和聚合酶基本上可将任何RNA转化为适合于限制 性酶切、PCR扩增和测序的双链DNA。参看例如Ausubel、Sambrook和 Berger(所有见上文)。

[0051] 本文讨论的发表内容仅涉及其在本申请的提交日期之前的公开内容。本文中的任何内容不应被理解为承认发明人没有权利因为先前公开书而 早于所述公开书。

[0052] 本公开书提供适用于细胞或受试对象的基因或蛋白质递送的方法和 组合物。所述方法和组合物可用于治疗受试对象的各种疾病和失调,包括 癌症和其它细胞增殖疾病和失调。本公开书提供用于基因递送的可复制逆 转录病毒载体。

[0053] 术语“载体”、“载体构建体”和“表达载体”意思是可将DNA或RNA序 列(例如外来基因)引入宿主细胞中,以便转化宿主并促进所引入的序列 的表达(例如转录和翻译)的运载体。载体典型地包含可传播病原体的 DNA,其中通过限制酶技术插入编码蛋白质的外来DNA。常见类型的载 体是“质粒”,其一般是双链DNA的自包含式分子,容易接受其它(外来)DNA且容易引入合适宿主细胞中。已描述了可在多种真核和原核宿主中 复制和/或表达的许多载体,包括质粒和真菌载体。非限制性实例包括pKK 质粒(Clonetech)、pUC质粒、pET质粒(Novagen,Inc.,Madison,Wis.、 pRSET或pREP质粒(Invitrogen,San Diego,Calif.)或pMAL质粒(New England Biolabs,Beverly,Mass.)和许多适当的宿主细胞,其中使用本文公开或引用的方法或者相关领域技术人员已知的方法。重组克隆载体通常 包括一种或一种以上用于克隆或表达的复制系统、一种或一种以上用于宿 主筛选的标志物(例如抗生素抗性)和一种或一种以上表达盒。

[0054] 术语“表达”意味着允许或导致基因或DNA序列的信息变得明白,例 如通过激活对应基因或DNA序列的转录和翻译中所涉及的细胞功能来产 生蛋白质。细胞表达DNA序列,从而形成“表达产物”,例如蛋白质。表 达产物自身(例如所得蛋白质)也可被称为由细胞“表达”。举例来说,当 多聚核苷酸或多肽是在外来或天然启动子的控制下在外来宿主细胞中表 达或产生时,或在外来启动子的控制下在天然宿主细胞中表达或产生时, 所述多聚核苷酸或多肽是重组表达。

[0055] 本公开书提供可复制性病毒载体,其含有编码例如胞嘧啶脱氨酶或其 突变体、miRNA或siRNA、细胞因子、抗体结合域等的异源多聚核苷酸, 所述异源多聚核苷酸可被递送到细胞或受试对象中。病毒载体可以是腺病 毒载体、麻疹载体、疱疹载体、逆转录病毒载体(包括慢病毒载体)、弹 状病毒载体(rhabdoviral vector)(例如水泡性口炎病毒载体)、呼肠孤病 毒载体(reovirus vector)、西尼加谷病毒载体(Seneca Valley Virus vector)、痘病毒载体(poxvirus vector)(包括动物痘或牛痘递送型载体)、细小病 毒载体(parvovirus vector)(包括AAV载体)、α病毒载体或所属领域技 术人员已知的其它病毒载体(也参看例如Concepts in Genetic Medicine, Boro Dropulic和Barrie Carter编,

Wiley,2008,Hoboken,NJ.;The Development of Human Gene Therapy,Theodore Friedmann编,Cold Springs Harbor Laboratory Press,Cold springs Harbor,New York,1999;Gene and Cell Therapy,Nancy Smyth Templeton编,Marcel Dekker Inc.,New York, New York,2000以及GeneTherapy:Therapeutic Mechanism and Strategies,Nancy Smyth Templetone和Danilo D Lasic编,Marcel Dekker,Inc.,New York,NewYork,2000;其揭示内容以引用的方式并入本文中)。

[0056] 在一个实施方案中,病毒载体可以是仅能感染复制中的哺乳动物细胞 的可复制逆转录病毒载体。在一个实施方案中,可复制逆转录病毒载体在 编码例如胞嘧啶脱氨酶、miRNA、siRNA、细胞因子、受体、抗体等的异 源多聚核苷酸的5′包含内部核糖体进入位点(IRES)。当异源多聚核苷酸 编码例如siRNA、miRNA或RNAi等非翻译RNA时,不需要IRES,但也可包括IRES以用于另一种被翻译基因,且可使用任何种类的逆转录病 毒(见下文)。在一个实施方案中,多聚核苷酸位于逆转录病毒载体的 ENV多聚核苷酸的3’。

[0057] 在其它实施方案中,提供用本公开书的可复制逆转录病毒载体转染的 宿主细胞。宿主细胞包括真核细胞,例如酵母细胞、昆虫细胞或动物细胞。 宿主细胞也可包括原核细胞,例如细菌细胞。

[0058] 还提供用本文提供的载体(例如可复制逆转录病毒载体)转导(转化 或转染)的经工程改造的宿主细胞。经工程改造的宿主细胞可以在经适当 改良以用于激活启动子、筛选转化株或扩增编码多聚核苷酸的常规营养培 养基中培养。培养条件(例如温度、pH等)是如先前筛选宿主细胞表达 中所用,且对于所属领域技术人员和根据本文引用的参考文献是显而易见 的,所述参考文献包括例如Sambrook、Ausubel和Berger,以及例如 Freshney(1994)Culture of Animal Cells:A Manual of Basic Technique,第3 版(Wiley-Liss,New york)和其中引用的参考文献。

[0059] 适当表达宿主的实例包括:细菌细胞,例如大肠杆菌(E.coli)、枯 草杆菌(B.subtilis)、链霉菌(Streptomyces)和鼠伤寒沙门氏菌(Salmonella typhimurium);真菌细胞,例如酿酒酵母(Saccharomyces cerevisiae)、毕赤酵母(Pichia pastoris)和粗糙链孢霉(Neurospora crassa);昆虫细 胞,例如果蝇(Drosophila)和草地贪夜蛾(Spodoptera frugiperda);哺 乳动物细胞,例如CHO、COS、BHK、HEK 293或Bowes黑色素瘤;或 植物细胞或外植体等。典型地可使用人类细胞或细胞系,然而,可能需要 将本公开书的载体和多聚核苷酸克隆到非人类宿主细胞中以用于测序、扩 增和克隆。

[0060] 本公开书还提供相对于先前逆转录病毒载体具有提高的稳定性的可 复制逆转录病毒载体。在感染和复制期间所述提高的稳定性对于治疗细胞 增殖失调很重要。可复制逆转录病毒提供转导效率、转基因稳定性和靶选 择性的组合。组合物和方法提供插入物稳定性且维持转基因的转录活性和 所编码多肽的翻译活力。

[0061] 本公开书提供经修饰的逆转录病毒载体。经修饰的逆转录病毒载体可 从逆转录病毒家族的成员衍生得到。逆转录病毒家族由三组组成:泡沫病 毒,例如人泡沫病毒(HFV);慢病毒,例如绵羊髓鞘脱落病毒;和致癌 病毒(但不是此组内的所有病毒都是致癌的)。术语“慢病毒”以其常规含 义使用,用于描述含有逆转录酶的病毒属。慢病毒包括“免疫缺陷病毒”, 其包括人免疫缺陷病毒(HIV)1型和2型(HIV-1和HIV-2)和猴免疫 缺陷病毒(SIV)。基于在病毒成熟期间于电子显微镜下观察的粒子形态, 致癌病毒在历史上被进一

步再分为A、B、C和D组。A型粒子代表在被 感染细胞的细胞质中看到的B型和D型病毒的不成熟粒子。这些粒子不 具有感染性。通过胞质内A型粒子的包膜作用,B型粒子作为成熟病毒粒子从质膜发芽。所述粒子在膜上具有75nm的环形核,长的糖蛋白从 其中突出。在发芽后,B型粒子含有偏心定位的电子致密核心。原型B 型病毒是小鼠乳腺瘤病毒(MMTV)。在被C型病毒感染的细胞中没有 观察到胞质内粒子。相反,成熟粒子经由月牙C式缩合直接从细胞表面 生长出,然后自我闭合并由质膜封闭起来。可见到包膜糖蛋白突出物以及 均匀的电子致密核心。出芽可从表面质膜开始或直接进入胞内液泡。C型 病毒被研究地最为广泛,且包括许多禽和鼠白血病病毒(MLV)。因为 牛白血病病毒(BLV)和人T细胞白血病病毒I和II型(HTLV-I/II)从 细胞表面的出芽形态学,它们被类似地归类为C型粒子。然而,它们还 具有规则的六边形形态和比原型C型病毒(例如鼠白血病病毒(MLV)) 更为复杂的基因组结构。D型粒子与B型粒子的相似之处在于它们都在 被感染的细胞质中显示环状结构。D型粒子从细胞表面开始出芽,但病毒 粒子包含短的表面糖蛋白突出物。电子致密核心在粒子内也是偏心定位。梅森费舍猴病毒(MPMV)是原型D型病毒。

[0062] 逆转录病毒的分类存在多种方式,但近十年来命名法已趋于标准化 (参看ICTVdB-The Universal Virus Database,v 4 on the World Wide Web (www)at ncbi.nlm.nih.gov/ICTVdb/ICTVdB/and the text book“Retroviruses” Coffin,Hughs和Varmus编,Cold Spring Harbor Press 1997,其公开内容以 引用的方式并入本文中)。在一个实施方案中,可复制逆转录病毒载体可 包含正向逆转录病毒(Orthoretrovirus)或更典型地包含γ逆转录病毒载体。

[0063] 逆转录病毒可根据其复制遗传材料的方式来定义。RNA在复制期间 转化为DNA。在细胞感染之后,通过被称为逆转录的分子方法从病毒粒 子中携带的两个RNA分子产生双链DNA分子。所述DNA形式作为前病 毒共价地整合到宿主细胞基因组中,在细胞和/或病毒因子的帮助下从所 述DNA表达病毒RNA。所表达的病毒RNA包装到粒子中且作为感染性 病毒粒子释放。

[0064] 逆转录病毒粒子由两个相同的RNA分子构成。每个野生型基因组具 有正义单链RNA分子,其在5’端被加帽并在3’尾端被聚腺苷酸化。二倍 体病毒粒子含有与gag蛋白复合的两条RNA链、病毒酶(pol基因产物) 和gag蛋白“核心”结构内的宿主tRNA分子。从宿主细胞膜衍生且含有病 毒包膜(env)蛋白的脂质双层包围和保护着此衣壳。env蛋白与病毒的细胞受体结合,且粒子典型地通过受体介导的内吞作用和/或膜融合而进 入宿主细胞。

[0065] 在外部包膜脱落后,病毒RNA通过逆转录被拷贝到DNA中。逆转 录是由pol区编码的逆转录酶催化且使用包装到病毒粒子中的宿主细胞 tRNA作为DNA合成引物。以此方式,RNA基因组被转化为更为复杂的 DNA基因组。

[0066] 逆转录所产生的双链线性DNA可能必须或可能不是必须在核内成环 形。前病毒现在在每一端具有两个相同的重复序列,所述重复序列被称为 长末端重复序列(LTR)。这两个LTR序列的末端产生被pol产物(催 化整合的整合酶蛋白)识别的位点,使得前病毒总是与距离LTR末端两 个碱基对(bp)的宿主DNA结合。在两个LTR的末端看到细胞序列的 复制,暗示可换位基因元件的整合模式。认为整合基本上是在靶细胞基因 组内随机发生。然而,通过修饰长末端重复序列,可以控制逆转录病毒基 因组的整合。

[0067] 所整合的病毒DNA的转录、RNA剪接和翻译是由宿主细胞蛋白介 导。产生以各种方

式剪接的转录物。在人类逆转录病毒HIV-1/2和 HTLV-I/II的情况下，也使用病毒蛋白来调节基因表达。细胞因子和病毒 因子之间的相互作用是病毒潜伏期和表达病毒基因的时间序列的控制中 的一个因素。

[0068] 逆转录病毒可水平地和垂直地传播。逆转录病毒的有效感染传播需要 在具有可特异性识别病毒包膜蛋白的受体的靶细胞上表达，但病毒也可低 效地利用不依赖受体的非特异性进入途径。此外，靶细胞类型必须能够支 持在病毒结合和穿透之后复制循环的所有阶段。当病毒基因组整合入宿主 的生殖细胞系中时，发生垂直传播。前病毒然后从一代传代到另一代，就 好像它是细胞基因一样。因此，产生了内源前病毒，其通常潜伏在体内，但在宿主暴露于适当媒介物时可被激活。

[0069] 如上所述，整合的DNA中间物被称为前病毒。先前基因疗法或基因 递送系统使用需要在适当辅助病毒存在下或在含有允许衣壳化而不会同 时产生污染性辅助病毒的适当序列的细胞系中，转录前病毒并组装到感染 性病毒中的方法和逆转录病毒。如下所述，产生本公开书的重组逆转录病 毒不需要辅助病毒，这是因为在基因组中提供用于衣壳化的序列，从而提 供可复制逆转录病毒载体以用于基因递送或疗法。

[0070] 现有的可复制逆转录病毒载体也容易在水平或垂直传播期间和复制 期间从被感染细胞或宿主细胞丢失。这可能是部分因为存在包括重复序列 或降低聚合酶效率的额外核苷酸序列。

[0071] 本公开书的逆转录病毒基因组和前病毒DNA具有至少三个基因：gag、pol和env，这些基因可侧接一个或两个长末端(LTR)重复序列，或在 前病毒中侧接两个长末端重复序列(LTR)和含有顺式作用序列(例如 psi)的序列。gag基因编码内部结构(基质、衣壳和核衣壳)蛋白；pol 基因编码RNA指导的DNA聚合酶(逆转录酶)、蛋白酶和整合酶；且 env基因编码病毒包膜糖蛋白。5’和/或3’LTR用于促进病毒粒子RNA的 转录和聚腺苷酸化。LTR含有病毒复制所必需的所有其它顺式作用序列。 慢病毒具有其它基因，包括vif、vpr、tat、rev、vpu、nef和vpx(在HIV-1、 HIV-2和/或SIV中)。

[0072] 与5’LTR相邻的序列是基因组(tRNA引物结合位点)的逆转录和 病毒RNA有效衣壳化入粒子(Psi位点)中所必需的。如果衣壳化(或 逆转录病毒RNA包装到感染性病毒粒子中)所必需的序列从病毒基因组 丢失，那么将导致顺式缺陷(cis defect)，其防止基因组病毒RNA的衣 壳化。这类经修饰的载体是先前基因递送系统(意即缺少病毒粒子衣壳化 所必需的元件的系统)中典型使用的载体。

[0073] 在第一实施方案中，本公开书提供一种能感染非分裂细胞、分裂细胞 或具有细胞增殖失调的细胞的重组逆转录病毒。本公开书的重组可复制逆 转录病毒包含编码病毒GAG、病毒POL、病毒ENV的多聚核苷酸序列， 在封装于病毒粒子内的内部核糖体进入位点(IRES)之后的异源多聚核 苷酸。

[0074] 短语“非分裂”细胞是指未经历有丝分裂的细胞。非分裂细胞可以在细 胞周期的任何点(例如G_0/G_1、$G_{1/S}$、$G_{2/M}$)被阻断，只要细胞没有积极地 分裂。例如，在间接体内感染中，可通过所属领域技术人员所用的标准技 术(包括照射、阿非迪霉素(aphidocolin)处理、血清饥饿和接触抑制) 处理分裂细胞以阻断细胞分裂。然而，应了解间接体内感染通常在不阻断 细胞的情况下进行，这是因为许多细胞已被抑制(例如干细胞)。举例来 说，不论使用何种机制阻断细胞分裂或细胞在细胞周期中哪一点被阻断， 重组慢病毒载体能感染任何

非分裂细胞。体内预先存在的非分裂细胞的实 例包括神经元、肌肉、肝、皮肤、心脏、肺和骨髓细胞和其衍生物。对于 分裂细胞，可使用致癌逆转录病毒载体。

[0075] “分裂的”细胞意思是经历活跃的有丝分裂或减数分裂的细胞。所述分 裂细胞包括干细胞、皮肤细胞(例如成纤维细胞和角质形成细胞)、配子 和所属领域中已知的其它分裂细胞。术语分裂的细胞特别关注且涵盖具有 细胞增殖失调的细胞，例如赘生性细胞。术语“细胞增殖失调”是指特征为 细胞数目异常的状况。所述状况可包括肥大(细胞连续繁殖，导致细胞群 体在组织内过度生长)和不足(组织内的细胞缺乏或不足)细胞生长，或 细胞过量流入或迁移入身体某区域。细胞群体不一定是被转化、致瘤或恶 性细胞，也可包括正常细胞。细胞增殖失调包括与结缔组织的过度生长有 关的失调，例如各种纤维化失调，包括硬皮病、关节炎和肝硬化。细胞增 殖失调包括赘生失调，例如头颈癌。头颈癌包括例如口癌、食道癌、咽喉 癌、喉癌、甲状腺癌、舌癌、唇癌、唾液腺癌、鼻癌、副鼻窦癌、鼻咽癌、上鼻道癌和鼻窦癌、嗅神经母细胞瘤、鳞状细胞癌、恶性黑色素瘤、鼻腔 鼻窦未分化癌(SNUC)、脑肿瘤(包括胶质母细胞瘤)或血液肿瘤。还 包括区域淋巴结癌，所述区域淋巴结包括颈淋巴结、喉前淋巴结、肺食管 旁淋巴结和下颌下淋巴结(Harrison′s Principles of Internal Medicine (Isselbacher等人编，McGraw-Hill，Inc.，第13版，第1850-1853页，1994)。其它癌症类型包括(但不限于)肺癌、结肠直肠癌、乳腺癌、前列腺癌、尿道癌、子宫癌、淋巴瘤、口腔癌、胰腺癌、白血病、黑色素瘤、胃癌、皮肤癌和卵巢癌。细胞增殖疾病还包括类风湿性关节炎(O’Dell NEJM 350:25912004)和特征通常为免疫系统细胞不当增殖的其它自身免疫失 调(Mackay等人NEJM 345:3402001)。

[0076] 异源核酸序列与IRES可操作地连接。如本文所用，术语“异源”核酸 序列或转基因是指(i)在野生型逆转录病毒中通常不存在的序列，(ii) 来源于外来物种的序列，或(iii)如果来自相同物种，那么可从其原始形 式经实质改变。或者，在细胞中通常不表达的未改变的核酸序列是异源核 酸序列。

[0077] 取决于本公开书的逆转录病毒载体的预期用途，可向逆转录病毒载体 中插入任何数目的异源多聚核苷酸或核酸序列。举例来说，在体外研究中 可使用常用的标志基因或报告基因，包括抗生素抗性和荧光分子(例如 GFP)。编码任何期望多肽序列的其它多聚核苷酸序列也可插入本公开书 的载体中。当寻求异源核酸序列的体内递送时，可使用治疗性序列和非治 疗性序列两者。举例来说，异源序列可编码涉及与细胞增殖失调相关的特 定基因或与疾病或失调有关的其它基因相关的的治疗性分子，包括反义分 子(miRNA、siRNA)或核酶。异源序列可以是自杀基因(例如HSV-tk 或PNP或胞嘧啶脱氨酶；经修饰或未经修饰)、生长因子或治疗性蛋白 (例如因子IX、IL2等)。适用于本公开书的其它治疗性蛋白在所属领域 中也容易鉴别。

[0078] 在一个实施方案中，载体内的异源多聚核苷酸包含已经经优化用于在 人类细胞中表达的胞嘧啶脱氨酶。在另一个实施方案中，胞嘧啶脱氨酶包 含已经人类密码子优化的序列，所述序列包含与野生型胞嘧啶脱氨酶相比 增加胞嘧啶脱氨酶稳定性(例如降解减少或热稳定性提高)的突变。在另 一个实施方案中，异源多聚核苷酸编码融合构建体，所述构建体包含与编 码具有UPRT或OPRT活性的多肽的多聚核苷酸可操作地连接的胞嘧啶 脱氨酶(经人类密码子优化或未经优化，突变或未突变)。在另一个实施 方案中，异源多聚核苷酸包括本公开书的CD多聚核苷酸(例如SEQ ID NO:3、5、11、13、15或17)。

[0079] 在另一个实施方案中,可复制逆转录病毒载体可包含编码包括胞嘧啶 脱氨酶(如本文所述)的多肽的异源多聚核苷酸,且许多载体可进一步包 含包括作为病毒启动子的初级转录物的一部分或连接于启动子的多聚核 苷酸的miRNA或siRNA分子,所述启动子可以是细胞型或组织特异型。

[0080] 微小RNA(miRNA)是小的非编码RNA。其位于编码或非编码基因 的内含子、非编码基因的外显子或基因间区域内。miRNA基因由RNA 聚合酶II转录,产生被称为初级前体miRNA(pri-miRNA)的前体多聚 核苷酸。核中的pri-miRNA由核糖核酸酶Drosha加工,产生miRNA前 体(pre-miRNA),其形成短发夹结构。随后,pre-miRNA通过输出蛋白 (Exportin)5被输送到细胞质,且被称为切酶(Dicer)的另一种核糖核 酸酶进一步加工,产生活性的成熟miRNA。

[0081] 成熟miRNA的长度约为21个核苷酸。通过与被靶向基因的mRNA 的3'未翻译区域结合,并通过抑制蛋白质翻译或降解mRNA来抑制蛋白 表达,所述成熟miRNA发挥其功能。miRNA涉及包括发育、细胞增殖、分化和癌症发展的生物过程。miRNA制谱(profiling)研究表明一些 miRNA表达是组织特异性的或在某些组织中富集。举例来说,miR-142-3p、miR-181和miR-223表达已被证明在人类和小鼠的造血组织中富集 (Baskerville等人,2005 RNA 11,241-247;Chen等人,2004 Science 303, 83-86)。

[0082] 已经观察到一些miRNA在若干肿瘤中受到上调(致癌miRNA)或 下调(阻遏miRNA)(Spizzo等人,2009Cell 137,586e1)。举例来说, miR-21在胶质母细胞瘤、乳腺癌、肺癌、前列腺癌、结肠癌、胃癌、食 道癌和宫颈癌、子宫平滑肌肉瘤、DLBCL、头颈癌中过度表达。对比之 下,let-7的成员已被报导在胶质母细胞瘤、肺癌、乳腺癌、胃癌、卵巢 癌、前列腺癌和结肠癌中受到下调。癌症中的miRNA表达的内稳态的重 新建立是抑制或逆转癌症发展的必要机制。

[0083] 由于miRNA在癌症中调节的重要功能,开发潜在治疗方法的焦点已经 集中在致癌miRNA的反义介导性抑制(antigomer)。然而,miRNA替换可 能代表同样有效的策略。在这种方法中,治疗上最有用的miRNA是在肿瘤 中表达水平很低,但在正常组织中表达水平很高且因此可被耐受的miRNA。

[0084] 在癌症中被下调的miRNA可用作抗癌剂。实例包括mir-128-1、let-7、 miR-26、miR-124和miR-137(Esquela-Kerscher等人,2008 Cell Cycle 7, 759-764;Kumar等人,2008 Proc Natl Acad Sci USA 105,3903-3908; Kota等人,2009 Cell137,1005-1017;Silber等人,2008 BMC Medicine 6:14 1-17)。已报导了miR-128表达在中枢神经系统中水平很高,且已观察到其 在胶质母细胞瘤中受到下调(Sempere等人,2004Genome Biology 5:R13-5-11;Godlewski等人,2008 Cancer Res 68:(22)9125-9130)。miR-128 是由两种不同的基因miR-128-1和miR-128-2编码。这两种基因被加工成相 同的成熟序列。已报导Bmi-1和E2F3a是miR-128的直接靶标(Godlewski等 人,2008 Cancer Res 68:(22)9125-9130;Zhang等人,2009 J.Mol Med 87:43-51)。此外,已观察到Bmi-1表达在各种人类癌症中被上调,所述癌 症包括神经胶质瘤、套细胞淋巴瘤、非小细胞肺癌、B细胞非霍奇金淋巴 瘤(B-cell non-Hodgkin's lymphoma)、乳腺癌、结肠直肠癌和前列腺癌。 此外,已经证明Bmi-1是不同组织干细胞自我更新所必需的,所述干细胞 包括神经元干细胞以及神经胶质瘤中的"类干"细胞群体。

[0085] 虽然miRNA介导的细胞功能抑制的可能性已经有了许多的体外证明，但仍然很难像其它分子那样以产生体内作用所必需的效率将这些物质作 为寡聚核苷酸或在病毒载体中递送(例如Li等人Cell Cycle 5:2103-2109 2006)。非复制性载体似乎在任何情况下都不足以有效地将治疗基因递送 到很大一部分肿瘤中。然而，也很难发现如何利用复制性载体递送miRNA 类型的药剂。特定来说，并不清楚如何将额外RNA序列并入可复制逆转录 病毒的RNA基因组中，和如何维持复制效率并保持添加物稳定地并入基因 组中。

[0086] 已经利用复制缺陷性逆转录病毒和慢病毒载体通过例如CMV或 LTR等聚合酶II启动子来稳定地表达pri-mi RNA，且证明可产生成熟 miRNA。然而，这些载体不是必须经历逆转录病毒或慢病毒的完整生命 周期达到复制病毒所必需的次数。基因组必须能经受比简单进入、整合和 转录更多的事件。与使用基于RNA的病毒表达miRNA有关的问题包括：(1)病毒RNA基因组在RNA加工期间转录后步骤的完整性；(2)所 插入盒在复制期间的稳定性；和(3)作为从LTR启动子转录的病毒RNA 的部分的pri-miRNA的适当加工，从而产生成熟miRNA。

[0087] 因此，已经广泛利用在非复制性逆转录病毒和慢病毒载体中并入III型 RNA聚合酶III启动子(例如U6和H1启动子)来表达功能性小干扰RNA (siRNA)，产生短发夹结构RNA(Bromberg-White等人，2004 J Virol 78:9，4914-4916;Sliva等人，2006 Virology 351，218-225;Haga等人，2006，Transplant Proc 38(10):3184-8)。用切酶裂解环序列，产生长度为21-22个 核苷酸的成熟siRNA。shRNA可在细胞中稳定表达以便下调靶基因表达。然而，通过向重组可复制逆转录病毒载体中并入所述盒、表达和切酶加工 来产生成熟miRNA仍然存在问题。

[0088] 在一个实施方案中，本发明提供一种重组可复制逆转录病毒载体，其 含有初级前体miRNA的异源多聚核苷酸序列。

[0089] 在另一个实施方案中，初级前体miRNA是人类来源。在另一个实施 方案中，初级前体RNA序列位于env基因下游。

[0090] 在另一个实施方案中，本发明提供一种重组可复制逆转录病毒载体，其含有位于env基因下游的人类初级前体miR-128-2的异源多聚核苷酸序 列(SEQ ID NO:32)。

[0091] 在癌症中被下调的miRNA可并入用于治疗基因递送的载体中，例如 let-7、miR-26、miR-124和miR-137(Esquela-Kerscher等人，2008 Cell Cycle 7，759-764;Kumar等人，2008 Proc Natl Acad Sci USA 105，3903-3908; Kota等人，2009 Cell 137，1005-1017; Silber等人，2008 BMC Medicine 6:14 1-17)。

[0092] 在另一个实施方案中，本发明提供一种重组可复制逆转录病毒载体，其含有位于env基因下游的与人类H1启动子连接的短发夹结构人类 pre-miR-128的异源多聚核苷酸序列(SEQ ID NO:33和SEQ ID NO:34)。在癌症中被下调的miRNA可并入用于治疗基因递送的载体中，例如let-7、 miR-26、miR-124和miR-137(Esquela-Kerscher等人，2008 Cell Cycle 7，759-764;Kumar等人，2008 Proc Natl Acad Sci USA 105，3903-3908; Kota等人，2009Cell 137，1005-1017;Silber等人，2008 BMC Medicine 6:14 1-17)。

[0093] 反义核酸是与特异性mRNA分子的至少一部分互补的DNA或RNA 分子(Weintraub，Scientific American，262:40，1990)。在细胞中，反义核 酸与对应mRNA杂交，形成双链分子。因为细胞无法翻译双链mRNA，所以反义核酸干扰mRNA的翻译。优选约15个核苷酸的反

义低聚物，这 是因为其容易合成且在引入靶细胞中时相比大分子引起问题的可能性较小。使用反义方法抑制体外基因翻译在所属领域中是众所周知的 (Marcus-Sakura，Anal.Biochem.，172:289，1988)。

[0094] 反义核酸可用于阻断突变蛋白或显性激活基因产物(例如在阿尔茨海 默氏病(Alzheimer′s disease)中积聚的淀粉状前体蛋白)的表达。所述方 法也可用于治疗亨廷顿氏病(Huntington′s disease)、遗传性帕金森综合 症(hereditary Parkinsonism)和其它疾病。特别关注与细胞增殖失调有关 的基因的阻断。反义核酸也可用于抑制与毒性有关的蛋白的表达。

[0095] 使用寡聚核苷酸阻碍转录被称为三螺旋(三螺旋)策略，这是因为寡 聚物在双螺旋DNA周围缠绕，形成三螺旋。因此，这些三螺旋化合物可 被设计用于识别选定基因上的独特位点(Maher等人，Antisense Res.and Dev.，1(3):227，1991;Helene，C.，Anticancer Drug Design，6(6):569，1991)。

[0096] 核酶是能够以与DNA限制性内切酶类似的方式特异性地裂解其它单 链RNA的RNA分子。通过修饰编码这些RNA的核苷酸序列，有可能工 程改造识别RNA分子中的特异性核苷酸序列并将其裂解的分子(Cech，J. Amer.Med.Assn.，260:3030，1988)。所述方法的主要优点在于，因为这 些分子是序列特异性的，所以只有具有特定序列的mRNA才被失活。

[0097] 如本文所用，术语“RNA干扰”(RNAi)是指由短的干扰核酸(siRNA 或微小RNA(miRNA))介导的序列特异性翻译后基因沉默过程。术语 “能介导RNA干扰的介质”是指siRNA以及当在细胞内转录时编码siRNA 的DNA和RNA载体。术语siRNA或miRNA打算涵盖能介导序列特异 性RNA干扰的任何核酸分子，例如短干扰RNA(siRNA)、双链RNA (dsRNA)、微小RNA(miRNA)、短发夹RNA(shRNA)、短干扰寡 聚核苷酸、短干扰核酸、短干扰修饰寡聚核苷酸、化学修饰siRNA、转 录后基因沉默RNA(ptgsRNA)等。

[0098] 用于设计发夹双螺旋的茎长度的合适范围包括20-30个核苷酸、30-50 个核苷酸、50-100个核苷酸、100-150个核苷酸、150-200个核苷酸、200-300 个核苷酸、300-400个核苷酸、400-500个核苷酸、500-600个核苷酸和 600-700个核苷酸的茎长度。用于设计发夹双螺旋的环长度的合适范围包 括4-25个核苷酸、25-50个核苷酸或更长(如果发夹双螺旋的茎长度很长) 的环长度。在某些情况下，具有长度超过21个核苷酸的双螺旋区域的发 夹结构可促进有效的siRNA指导的沉默，与环序列和长度无关。

[0099] 本公开书的复制性逆转录病毒载体可通过表达经工程改造的siRNA 或miRNA(Dennis，Nature，418:1222002)而用于治疗疾病，所述siRNA 或miRNA关闭或降低控制患病细胞(包括肿瘤细胞)增殖或存活的关键 基因的表达。所述靶标包括例如Rad 51等基因，Rad 51是DNA修复中 的关键酶，且在没有它的情况下细胞生长会受到显著限制。其它靶标包括 控制细胞生长的许多信号传导路径分子(Marquez&McCaffrey Hum Gene Ther.19:272008)。siRNA或miRNA可与从本公开书的相同或不同逆转 录病毒载体表达细胞毒性基因组合使用。合适的细胞毒性基因的实例包含 本公开书的胞嘧啶脱氨酶或经修饰的胞嘧啶脱氨酶。

[0100] 在使用中，逆转录病毒载体将通过肿瘤或其它靶组织复制，且在生长 抑制发生之前，病毒首先整合入宿主基因组中并在所述细胞的生长被抑制 之后继续制造病毒。用于筛选功能性miRNA或siRNA序列的方法在所属 领域中是已知的。一般来说，设计有效siRNA或

miRNA序列中的关键特 征通常是避免“脱靶”作用。然而,对于使用对肿瘤细胞具有高特异性的复 制载体(例如本公开书的复制载体)来说,这些副作用不是非常重要,因 为预期细胞最终都会死亡。本公开书的逆转录病毒载体可使用来自对应蛋 白质没有被显著靶向的其它物种的细胞制得。所述细胞包括狗细胞系或鸡 细胞系。或者,病毒可以通过在人类293衍生细胞或允许有效瞬时转染的 其它细胞系上的瞬时转染来制得。对于所述用途来说,病毒不需要利用 IRES,且siRNA或miRNA序列可仅在病毒基因组的适宜位点插入。所 述位点包括包膜下游和复制性逆转录病毒的3'LTR上游的区域。或者, polIII转录单元可以插入具有适当siRNA或miRNA的病毒基因组中,典 型地插入3'包膜基因下游。可插入一些不同的siRNA或miRNA序列以 确保靶基因的有效下调或一个以上基因的下调。合适的序列和靶标可以从 所属领域技术人员已知的来源获得。例如:

[0101] • MIT/ICBP siRNA Database http:(//)web.mit.edu/sirna/-″The MIT [Massachusetts Institute of Technology]/ICBP[Integrative Cancer Biology Program]siRNA Database是大学范围的科研成果,其用于分类这些实验上 被确认的试剂且使所述信息可供MIT团体内部和外部的其它研究者使用 (Massachusetts Institute of Technology)。

[0102] • RNAi Central-http:(//)katahdin.cshl.org:9331/RNAi_web/ scripts/main2.pl RNAi resources,包括siRNA和shRNA设计工具(Hannon Lab,Cold Spring Harbor Laboratory)。

[0103] • The RNAi Web-http:(//)www.rnaiweb.com/General resource。

[0104] • siDIRECT-http:(//)genomics.jp/sidirect/用于哺乳动物RNA干扰的 在线靶特异性siRNA设计(University of Tokyo,Japan)。

[0105] • siRNA Database一含有针对各种生物体中的所有已知mRNA序列 的siRNA靶标的综合性siRNA数据库(Protein Lounge系统生物学网站的 部分)。

[0106] • siRNA Database and Resources for RNA Interference Studies http:(//)www.rnainterference.org/

[0107] • siRNA Selector-http:(//)bioinfo.wistar.upenn.edu/siRNA/siRNA.htm.使用一组规则基于热动力学参数(Khvorova等人,2003;Schwarz等人, 2003)和由Dharmacon开发的序列相关决定因素(Reynolds等人,2004) 来评估siRNA功能性。针对UniGene数据库使用BLAST来确定特异性 (Wistar Institute)。

[0108] • siRNA Target Finder http:(//)www(.)ambion.com/techlib/misc/ siRNA_finder.html(Ambion)。

[0109] 本公开书的复制性逆转录病毒也可表达天然存在的siRNA的靶标, 所述靶标在表达上受限于特定细胞类型,使得载体复制在那些细胞类型中 被显著抑制。基于鼠白血病病毒的重组可复制逆转录病毒载体的产生可允 许高转导水平和因此体内基因递送的高效率。使用可复制逆转录病毒载体 的一个主要问题在于先前所报导的病毒传播失控(Donahue等人,J.Exp Med.1992,176:1124-1135;Calmes等人,Blood 2005,106:2530-2533; Seggewiss等人,Blood 2006,107:3865-3867)。因为病毒的性质,病毒传 播最初可以在淋巴细胞内发生且随后传播到外周组织。对于抗肿瘤目的, 以一定水平天然复制的一些体内正常细胞是造血细胞、肠壁细胞和一些内 皮细胞。这些正常细胞是在循环中可有效

感染病毒的潜在位点。一般来说 这是不合需要的。例如这些正常细胞等细胞的任何离群感染(stray infection)可通过在这些细胞类型中包括天然存在的miRNA或miRNA组 合的靶标来抑制。已经展示了使用miRNA靶标抑制免疫反应的一些可行 性(Brown等人,Nat Biotechnol.200725:1457-67)。这些靶标是与天然存 在的miRNA序列具有同源匹配的小RNA序列。这些序列可以插入本发 明载体的任何适宜位点中,一般不会对载体活力造成显著有害的后果,除 非在期望类型的细胞中。载体可如本文所述来制备和使用。

[0110] 在一个实施方案中,本发明提供一种重组可复制逆转录病毒载体,其 在与IRES连接的转基因(例如yCD2或GFP)下游含有miR-142-3p靶 序列(142-3pT,SEQ ID NO:35)的单拷贝。除了miR181和miR-223以 外,其它组织或细胞富含的miRNA的靶序列也可并入载体中以限制病毒 以特异性组织或细胞型方式传播。举例来说,miR-133和miR206在肌肉 组织中的表达水平很高(Kelly等人,2008 Nature Medicine 14:11 1278-1283)。

[0111] 在另一个实施方案中,本发明提供一种重组可复制逆转录病毒载体,其在与IRES连接的转基因(例如yCD2或GFP)下游含有142-3pT(SEQ ID NO:36)的4个拷贝。除了miR181和miR-223以外,其它组织或细胞 富含的miRNA的靶序列也可并入载体中以限制病毒以特异性组织或细胞 型方式传播。举例来说,miR-133和miR206在肌肉细胞中的表达水平很高。本发明提供miRNA的单个靶序列、多个靶序列或其组合的灵活性,且因此提供限制病毒在体外和体内(例如造血细胞和/或肌肉细胞)以组 织和/或细胞特异性方式的失控传播(Kelly等人,2008 Nature Medicine 14:11 1278-1283)。

[0112] miRNA靶标可以插入转基因的3',但在3'LTR或IRES上游之前, 和包膜3'端之后。一般来说,靶标不会插入蛋白质编码序列中。

[0113] 在其它实施方案中,异源多聚核苷酸可包含细胞因子,例如白细胞介 素、γ干扰素等。可以从本公开书的逆转录病毒载体表达的细胞因子包括 (但不限于)IL-1α、IL-1beta、IL-2、IL-3、IL-4、IL-5、IL-6、IL-7、IL-8、 IL-9、IL-10、IL-11、IL-12、IL-13、IL-14、IL-15、IL-16、IL-17、IL-18、 IL-19、IL-20和IL-21、抗CD40、CD40L、IFN-γ和TNF-α、可溶形式的 TNF-α、淋巴毒素-α(LT-α,也称为TNF-β)、LT-β(在复合异源三聚体 LT-α2-β中发现)、OPGL、FasL、CD27L、CD30L、CD40L、4-1BBL、 DcR3、OX40L、TNF-γ(国际公开案第WO 96/14328号)、AIM-I(国际 公开案第WO 97/33899号)、endokine-α(国际公开案第WO 98/07880 号)、OPG和neutrokine-α(国际公开案第WO 98/18921号)、OX40和 神经生长因子(NGF),和可溶形式的Fas、CD30、CD27、CD40和4-IBB、 TR2(国际公开案第WO 96/34095号)、DR3(国际公开案第WO 97/33904 号)、DR4(国际公开案第WO 98/32856号)、TR5(国际公开案第WO 98/30693号)、TRANK、TR9(国际公开案第WO 98/56892号)、TR10 (国际公开案第WO 98/54202号)、312C2(国际公开案第WO 98/06842 号)和TR12,和可溶形式的CD154、CD70和CD153。血管生成蛋白可 用于一些实施方案中,尤其对于从细胞系产生蛋白质来说。所述血管生成因子包括(但不限于)神经胶质瘤源性生长因子(GDGF)、血小板源性 生长因子-A(PDGF-A)、血小板源性生长因子-B(PDGF-B)、胎盘生 长因子(PIGF)、胎盘生长因子-2(PIGF-2)、血管内皮生长因子(VEGF)、 血管内皮生长因子-A(VEGF-A)、血管内皮生长因子-2(VEGF-2)、血管内皮生长因子B(VEGF-3)、血管内皮生长因子B-186(VEGF-B186)、 血管内皮生长因子-D(VEGF-D)、血管内皮生长因子-D(VEGF-D)和 血管内皮生长因子-E(VEGF-E)。成纤维细胞生长因子可通过本公开书 的载体递送,且包括(但不限于)FGF-1、FGF-2、FGF-3、FGF-4、FGF-

5、FGF-6、FGF-7、FGF-8、FGF-9、FGF-10、FGF-11、FGF-12、FGF-13、FGF-14和FGF-15。造血生长因子可使用本公开书的载体递送，所述生长 因子包括(但不限于)粒细胞巨噬细胞集落刺激因子(GM-CSF)(沙格 司亭(sargramostim))、粒细胞集落刺激因子(G-CSF)(非格司亭(filgrastim))、巨噬细胞集落刺激因子(M-CSF、CSF-1)、红细胞生 成素(阿法依泊汀(epoetin alfa))、干细胞因子(SCF、c-kit配体、青 灰因子)、巨核细胞集落刺激因子、PIXY321(GMCSF/IL-3)融合蛋白 等。

[0114] 本公开书的重组病毒一般能够将核酸序列转移到靶细胞中。

[0115] 术语“调节核酸序列”总体是指启动子序列、聚腺苷酸化信号、转录终 止序列、上游调节域、复制起点、增强子等，这些序列总体用于受体细胞 中编码序列的复制、转录和翻译。并不是所有的这些控制序列都必须存在，只要所选的编码序列能在适当宿主细胞中复制、转录和翻译即可。所属领 域技术人员容易从公开数据库和材料鉴别调节核酸序列。此外，所属领域 技术人员可鉴别适用于预期用途(例如体内、体外或体外)的调节序列。

[0116] 内部核糖体进入位点(“IRES”)是指在通常处于IRES的3’的编码序 列翻译期间促进核糖体进入或滞留的核酸片段。在一些实施方案中，IRES 可包含剪接受体/供体位点，但优选的IRES缺少剪接受体/供体位点。核 糖体通常经由位于所有真核信使RNA的5’端的帽序列进入mRNA。然而， 这一普遍规律存在例外。在一些病毒mRNA中缺乏帽序列表明在这些RNA的内部位点上存在可允许核糖体进入的替代结构。迄今为止，已经 在未加帽病毒mRNA的5’非编码区中鉴别了许多此类结构(根据其功能 命名为IRES)，尤其例如在细小核糖核酸病毒中，例如脊髓灰质炎病毒 (Pelletier等人，1988，Mol.Cell.Biol.，8，1103-1112)和EMCV病毒(脑 心肌炎病毒)(Jang等人，J.Virol.，1988，62，2636-2643)。本公开书提供IRES在可复制逆转录病毒载体的情况中的用途。

[0117] 术语“启动子区”在本文中以其普通意义使用，是指包含DNA调节序 列的核苷酸区域，其中调节序列来源于能结合RNA聚合酶并起始下游(3’ 方向)编码序列的转录的基因。调节序列可与期望基因序列同源或异源。举例来说，可利用多种启动子，包括上文所述的病毒或哺乳动物启动子。

[0118] 异源核酸序列典型地处于病毒LTR启动子-增强子信号或内部启动子 的控制下，且逆转录病毒LTR内保留的信号仍可导致载体有效地整合到 宿主细胞基因组中。因此，本公开书的重组逆转录病毒载体、期望序列、 基因和/或基因片段可在一些位点中和不同调节序列的控制下插入。举例 来说，插入位点可以是病毒增强子/启动子邻近位点(即5’LTR驱动的基 因座位)。或者，期望序列可插入调节序列远端位点(例如env基因3’ 的IRES序列)中或存在两个或两个以上异源序列的地方，一个异源序列 可处于第一调节区的控制下且第二异源序列可处于第二调节区的控制下。其它远端位点包括病毒启动子序列，其中一个或多个期望序列的表达是通 过启动子邻近顺反子的剪接，可使用内部异源启动子(例如SV40或CMV) 或内部核糖体进入位点(IRES)。

[0119] 在一个实施方案中，本公开书的逆转录病毒基因组含有在IRES下游 包含克隆位点以用于插入期望/异源多聚核苷酸的IRES。在一个实施方案 中，IRES位于逆转录病毒载体中env基因的3’，但位于期望异源多聚核 苷酸的5’。因此，编码期望多肽的异源多聚核苷酸可操作地连接于IRES。

[0120] 在另一个实施方案中，包括靶向多聚核苷酸序列作为本公开书的重组 逆转录病

毒载体的部分。靶向多聚核苷酸序列是靶向配体(例如肽激素(例 如调蛋白(heregulin))、单链抗体、受体或受体的配体)、组织特异性或 细胞类型特异性调节元件(例如组织特异性或细胞类型特异性启动子或增 强子)或靶向配体与组织特异性/细胞类型特异性调节元件的组合。优选 地,靶向配体可操作地连接于逆转录病毒的env蛋白,从而产生嵌合逆转 录病毒env蛋白。病毒GAG、病毒POL和病毒ENV蛋白可来源于任何 合适的逆转录病毒(例如MLV或慢病毒源性的)。在另一个实施方案中, 病毒ENV蛋白是非逆转录病毒源性的(例如CMV或VSV)。

[0121] 在一个实施方案中,本公开书的重组逆转录病毒可以将病毒靶向于特 定细胞类型(例如平滑肌细胞、肝细胞、肾细胞、成纤维细胞、角质形成 细胞、间质干细胞、骨髓细胞、软骨细胞、上皮细胞、肠细胞、乳房细胞、 赘生性细胞、神经胶质瘤细胞、神经元细胞和所属领域中已知的其它细胞) 的方式遗传改造:使得逆转录病毒载体的重组基因组被递送到非分裂靶细 胞、分裂靶细胞或具有细胞增殖失调的靶细胞。

[0122] 在一个实施方案中,逆转录病毒载体是通过结合于在细胞外表面上具 有分子的细胞而靶向于细胞。所述靶向逆转录病毒的方法利用靶向配体在 逆转录病毒包膜上表达,以便帮助病毒靶向于具有与逆转录病毒表面上的 靶向配体相互作用的受体或结合分子的细胞或组织。在病毒感染细胞之后, 病毒将其核酸注入细胞且逆转录病毒遗传物质可整合入宿主细胞基因组。

[0123] 在另一个实施方案中,靶向操作利用细胞或组织特异性调节元件,以 促进病毒基因组在主动利用调节元件的被靶向细胞中的表达和转录,如下 文更详细地描述。转移的逆转录病毒遗传物质然后在宿主细胞内被转录和 翻译成蛋白质。靶向调节元件典型地连接于5'和/或3'LTR,从而产生嵌 合LTR。

[0124] 通过将目的异源多聚核苷酸连同编码例如特异性靶细胞上的受体的 配体的另一个基因插入本公开书的病毒载体,载体现在具有靶标特异性。 病毒载体可通过连接例如糖、糖脂或蛋白质而具有靶标特异性。靶向可使 用靶向病毒载体的抗体来完成。所属领域技术人员将了解或容易确定,可 插入与病毒包膜连接的病毒基因组或蛋白中以允许含有目的核酸序列的 病毒载体的靶特异性递送的特异性多聚核苷酸序列。

[0125] 因此,在一个实施方案中,本公开书包括嵌合env蛋白,其包含可操 作地连接于靶向多肽的逆转录病毒ENV蛋白。靶向多肽可以是细胞特异 性受体分子、细胞特异性受体的配体、细胞特异性抗原表位的抗体或抗体 片段,或所属领域中容易鉴别的可结合靶细胞或与靶细胞相互作用的任何 其它配体。靶向多肽或分子的实例包括使用生物素-链霉亲和素作为连接 子的二价抗体(Etienne-Julan等人,J.Of General Virol.,73,3251-3255(1992);Roux等人,Proc.Natl.Acad.Sci USA 86,9079-9083(1989));在 包膜上含有编码针对半抗原的单链抗体可变区的序列的重组病毒(Russell 等人,Nucleic Acids Research,21,1081-1085(1993));将肽激素配体克隆 到逆转录病毒包膜中(Kasahara等人,Science,266,1373-1376(1994)); 嵌合EPO/env构建体(Kasahara等人,1994);针对亲嗜性MLV包膜中 的低密度脂蛋白(LDL)受体的单链抗体,导致表达LDL受体的HeLa 细胞的特异性感染(Somia等人,Proc.Natl.Acad.Sci USA,92,7570-7574 (1995));类似地,ALV的宿主范围可通过并入整合素配体而改变,使得 病毒现在能跨物种以特异性地感染大鼠胶质母细胞瘤细胞 (Valsesia-Wittmann等人,J.Virol.68,4609-4619(1994));且Dornberg

和其合作者(Chu和Dornburg,J.Virol 69,2659-2663(1995))已经报导了 脾坏死病毒(SNV)(一种禽类逆转录病毒)的组织特异性靶向,其使用 含有针对肿瘤标志物导向的单链抗体的包膜。

[0126] 本公开书提供一种产生能感染靶细胞的重组逆转录病毒的方法,其包 括用以下载体转染合适宿主细胞和回收重组病毒:所述载体包含编码病毒 gag、病毒pol和病毒env的多聚核苷酸序列,和可操作地连接于调节核 酸序列的异源多聚核苷酸。

[0127] 本公开书的逆转录病毒和方法提供可复制逆转录病毒,其不需要辅助 病毒或其它核酸序列或蛋白质以便传播和产生病毒粒子。举例来说,本公 开书的逆转录病毒的核酸序列编码例如本文分别讨论的群特异性抗原和 逆转录酶(以及成熟和逆转录所必需的整合酶和蛋白酶)。病毒gag和 pol可来源于慢病毒(例如HIV)或致癌病毒或γ逆转录病毒(例如 MoMLV)。此外,本公开书的逆转录病毒的核酸基因组包括编码病毒包 膜(ENV)蛋白的序列。env基因可来源于任何逆转录病毒。env可为允 许人类和其它物种的细胞转导的双嗜性包膜蛋白,或可为只能转导小鼠和 大鼠细胞的亲嗜性包膜蛋白。此外,可能需要通过将包膜蛋白与抗体或用 于靶向于特定细胞类型的受体的特定配体连接来靶向重组病毒。如上所述, 逆转录病毒载体可通过插入例如糖脂或蛋白质而具有靶标特异性。靶向操 作通常通过使用抗体将逆转录病毒载体靶向于特定细胞类型上的抗原(例 如在某些组织中发现的细胞类型,或癌细胞类型)来实现。所属领域技术 人员将认识到或不经过多实验即可容易地确定,用于将逆转录病毒载体递 送于特定靶标的特定方法。在一个实施方案中,env基因来源于非逆转录 病毒(例如CMV或VSV)。逆转录病毒源性的env基因的实例包括(但 不限于):莫洛尼鼠白血病病毒(MoMuLV)、哈维鼠类肉瘤病毒(Harvey murine sarcoma virus)(HaMuSV)、鼠类乳腺致癌病毒(MuMTV)、 长臂猿白血病病毒(GALV)、人类免疫缺陷病毒(HIV)和劳氏肉瘤病 毒(RSV)。也可使用其他env基因,例如水泡性口炎病毒(VSV)(蛋 白质G)、巨细胞病毒包膜(CMV)或流感病毒血细胞凝集素(HA)。

[0128] 在一个实施方案中,逆转录病毒基因组来源于致癌逆转录病毒,且更 具体地来源于哺乳动物致癌逆转录病毒。“来源”意思是母体多聚核苷酸序 列是已经通过插入或去除天然存在的序列(例如插入IRES、插入编码目 的多肽或抑制核酸的异源多聚核苷酸、用来自不同逆转录病毒或病毒的更 有效的启动子交换野生型启动子等)而被修饰的野生型致癌病毒。

[0129] 不同于本领域标准方法所产生的重组逆转录病毒,这些病毒存在缺陷 且需要帮助以产生感染性载体粒子,本公开书提供可复制逆转录病毒。

[0130] 在另一个实施方案中,本公开书提供使用调节序列靶向的逆转录病毒 载体。可利用细胞或组织特异性调节序列(例如启动子)来靶向基因序列 在特异性细胞群体中的表达。本文其它地方描述用于本公开书的合适哺乳 动物和病毒启动子。因此,在一个实施方案中,本公开书提供在逆转录病 毒基因组的5’端具有组织特异性启动子元件的逆转录病毒。典型地,组 织特异性调节元件/序列位于逆转录病毒基因组的LTR的U3区中,包括 例如用于赘生性细胞的细胞或组织特异性启动子和增强子(例如肿瘤细胞 特异性增强子和启动子),和诱导型启动子(例如四环素)。

[0131] 本公开书的转录控制序列也可包括与编码超抗原、细胞因子或趋化因 子的基因天然连接的天然存在的转录控制序列。

[0132] 在一些情况下,可能需要调节表达。举例来说,取决于期望表达水平,可利用具有不同活性强度的不同病毒启动子。在哺乳动物细胞中,经常使 用CMV立即早期启动子来提供强转录活化。当期望转基因的表达水平降 低时,也已经使用了效力较低的CMV启动子的经修饰形式。当期望转基 因在造血细胞中表达时,可使用逆转录病毒启动子,例如来自MLV或 MMTV的LTR。可使用的其它病毒启动子包括SV40、RSV LTR、HIV-1 和HIV-2LTR、例如来自E1A、E2A或MLP区的腺病毒启动子、AAV LTR、花椰菜花叶病毒、HSV-TK和禽类肉瘤病毒。

[0133] 可使用类似的组织特异性或选择性启动子来实现在特异性组织或细 胞中的转录,以便降低对非靶向组织的潜在毒性或不当影响。举例来说,可使用例如PSA、前列腺碱性蛋白(probasin)、前列腺酸性磷酸酶或前列 腺特异性腺激肽释放酶(hK2)来靶向前列腺中的基因表达。乳清辅助蛋 白(WAP)可用于乳房组织表达(Andres等人,PNAS 84:1299-1303,1987)。表1列举可使用的其它启动子/调节域。

[0134] “组织特异性调节元件”是能驱动基因在一种组织中转录同时在其它 组织类型中大体上保持“沉默”的调节元件(例如启动子)。然而,应了解 组织特异性启动子在其沉默的那些组织中可具有可检测量的“背景”或“基 础”活性。启动子在靶组织中被选择性地活化的程度可用选择性比(靶组 织中的活性/对照织中的活性)表示。就此而言,可用于实施本公开的组 织特异性启动子典型地具有大于约5的选择性比。选择性比优选地大于约 15。

[0135] 在某些适应症中,可能需要在施用本公开书的重组可复制逆转录病毒 (RRCR)之后的特定时间活化转录。这可以用可由激素或细胞因子调节 的启动子进行。举例来说,在适应症是产生特异性类固醇或特异性类固醇 导向的性腺组织的治疗应用中,使用受雄激素或雌激素调节的启动子可能 是有利的。可由激素调节的所述启动子包括MMTV、MT-1、蜕皮激素和 RuBisco。可使用其它受激素调节的启动子,例如响应于甲状腺、垂体和 肾上腺激素的启动子。可使用的细胞因子和免疫蛋白响应启动子包括K 和T激肽原(Kininogen)(Kageyama等人,1987)、c-fos、TNF-α、C 反应蛋白(C-reactive protein)(Arcone等人,1988)、结合珠蛋白 (haptoglobin)(Oliviero等人,1987)、血清淀粉样蛋白A2、C/EBPα、IL-1、IL-6(Poli和Cortese,1989)、补体C3(Wilson等人,1990)、IL-8、α-1酸性糖蛋白(Prowse和Baumann,1988)、α-1抗胰蛋白酶、脂蛋白脂 酶(Zechner等人,1988)、血管紧张素原(angiotensinogen)(Ron等人, 1990)、纤维蛋白原(fibrinogen)、c-jun(可由佛波醇酯(phorbol ester)、 TNF-α、UV辐射、视黄酸和过氧化氢诱导)、胶原酶(由佛波醇酯和视 黄酸诱导)、金属硫蛋白(可由重金属和糖皮质激素诱导)、基质降解酶 (Stromelysin)(可由佛波醇酯、白细胞介素-1和EGF诱导)、α-2巨球 蛋白和α-1抗胰凝乳蛋白酶(antichymotrypsin)。也可使用肿瘤特异性启 动子(例如骨钙素(osteocalcin)、缺氧反应元件(hypoxia-responsive element; HRE)、MAGE-4、CEA、α-甲胎蛋白(α-fetoprotein)、GRP78/BiP和酪 氨酸酶)来调节肿瘤细胞中的基因表达。

[0136] 此外,所述启动子清单不应解释为穷尽性的或限制性的,所属领域技 术人员将知晓可与本文所公开的启动子和方法结合使用的其它启动子。

[0137] 表1组织特异性启动子

组织	启动子
胰腺	胰岛素弹性蛋白淀粉酶
	pdr-1 pdx-1 葡糖激酶
[0138] 肝脏	白蛋白 PEPCK HBV 增强子
	A 甲胎蛋白载脂蛋白 C α-1
	抗胰蛋白酶卵黄蛋白原，NF-AB 转甲状腺素蛋白（Transthyretin）

骨骼肌	肌球蛋白 H 链肌肉肌酸激酶 肌肉肌萎缩蛋白钙蛋白酶 p94（Dystrophin Calpain p94 Skeletal） α-肌动蛋白快速肌钙蛋白 1（α-actin fast troponin 1）
皮肤	角蛋白 K6、角蛋白 K1
肺	CFTR 人细胞角蛋白 18（K18） 肺表面活性蛋白 A、B 和 C CC-10 P1
平滑肌	sm22 α SM-α-肌动蛋白
内皮	内皮素（Endothelin）-1、E-选择素（selectin）、von Willebrand 因子 TIE（Korhonen 等人, 1995）、KDR/flk-1 黑素细胞酪氨酸酶
脂肪组织	脂蛋白脂肪酶（Zechner 等人, 1988）、降脂肪蛋白（Adipsin）（Spiegelman 等人, 1989）、乙酰辅酶 A 羧化酶（Pape 和 Kim, 1989）、磷酸甘油脱氢酶（Dani 等人, 1989）、脂肪细胞 P2（Hunt 等人, 1986）
乳房	乳清酸性蛋白（WAP）（Andres 等人 PNAS 84:1299-1303 1987）
血液	β-球蛋白

[0139]

[0140] 应进一步了解，某些启动子虽然在活性上不受限于单一组织类型，但 仍然可能显示选择性，因为其在一组组织中可能有活性，而在另一组组织 中活性较低或沉默。所述启动子也称为“组织特异性”，且预期用于本公开 书。举例来说，在各种中枢神经系统(CNS)神经元中有活性的启动子可 在治疗上用于防止因中风所致的伤害，中风可影响许多不同的大脑区域。因此，本公开书中所用的组织特异性调节元件适用于调节异源蛋白以及可 用作本发明逆转录病毒载体中的靶向多聚核苷酸序列。

[0141] 在另一个实施方案中，本公开书提供包括重组逆转录病毒源性的构建 体的质粒。质粒可以直接引入靶细胞或细胞培养物中，例如NIH 3T3或 其它组织培养细胞。所得细胞将逆转录病毒载体释放到培养基中。

[0142] 本公开书提供一种多聚核苷酸构建体，其从5’到3’包含：可用于起 始转录的启动子或调节区；psi包装信号；gag编码核酸序列；pol编码核 酸序列；env编码核酸序列；内部核糖体进入位点核酸序列；编码标志物、治疗或诊断多肽的异源多聚核苷酸；和LTR核酸序列。如本文其它地方 所述和如下所示，本公开书的多聚核苷酸构建体的各种链段(例如重组可 复制逆转录病毒多聚核苷酸)部分取决于期望宿主细胞、表达时间或量和 异源多聚核苷酸而经工程改造。本公开书的可复制逆转录病毒构建体可分 为多个结构域，它们可由所属领域技术人员单独修饰。

[0143] 举例来说，启动子可包括具有如SEQ ID NO:19、20或22的核苷酸1 到约核苷酸582中所述的序列的CMV启动子，且可包含对一个或多个(例 如2-5、5-10、10-20、20-30、30-50、50-100个或以上核酸碱基)核酸碱 基所作的修饰，只要经修饰的启动子能够引导和起始转录即可。在一个实 施方案中，启动子或调节区包含CMV-R-U5域多聚核苷酸。CMV-R-U5 域包含来自人类巨细胞病毒的立即早期启动子和MLV R-U5区的融合体。 在一个实施方案中，CMV-R-U5域多聚核苷酸包含如SEQ ID NO:19、20 或22的约核苷酸1到约核苷酸1202所述的序列，或与如SEQ ID NO:19、 19或22所述的序列具有至少95％同一性的序列，其中所述多聚核苷酸促 进与其可操作地连接的核酸分子的转录。多聚核苷酸的gag域可来源于多 种逆转录病毒，但典型地来源于致癌逆转录病毒且更具体地来源于哺乳动 物致癌逆转录病毒。在一个实施方案中，gag域包含来自约核苷酸1203 到约核苷酸2819的序列或与所述序列具有至少95％、98％、99％或99.8％ (四舍五入到十分位)同一性的序列。多聚核苷酸的pol域可来源于多种 逆转录病毒，但典型地来源于致癌逆转录病毒且更具体地来源于哺乳动物 致癌逆转录病毒。在一个实施方案中，pol域包含来自约核苷酸编号2820 到约核苷酸6358的序列或与所述序列具有至少95％、98％、99％或99.9％ (四舍五入到十分位)同一性的序列。多聚核苷酸的env域可来源于多种 逆转录病毒，但典型地来源于致癌逆转录病毒或γ-逆转录病毒且更具体 地来源于哺乳动物致癌逆转录病毒或γ-逆转录病毒。在一些实施方案中， env编码域包括双嗜性env域。在一个实施方案中，env域包括来自约核 苷酸编号6359到约核苷酸8323的序列或与所述序列具有至少95％、98％、 99％或99.8％(四舍五入到十分位)同一性的序列。多聚核苷酸的IRES 域可从任何内部核糖体进入位点获得。在一个实施方案中，IRES来源于 脑心肌炎病毒。在一个实施方案中，IRES域包含来自约核苷酸编号8327 到约核苷酸8876的序列或与所述序列具有至少95％、98％或99％(四舍 五入到十分位)同一性的序列，只要所述域允许核糖体进入即可。异源域 可包含本公开书的胞嘧啶脱氨酶。在一个实施方案中，CD多聚核苷酸包 含经人类密码子优化的序列。在另一个实施方案中，CD多聚核苷酸编码 具有胞嘧啶脱氨酶的突变多肽，其中所述突变赋予增加的热稳定性，这使 得解链温度(Tm)升高10℃，从而允许在更为宽广的温度范围下的持续动力学活性和提高的蛋白质累积水平。在一个实施方案中，胞嘧啶脱氨酶 包含如SEQ ID NO:19或22的约核苷酸编号8877到约9353所述的序列。 异源域之后可紧跟聚嘌呤富集域。3’LTR可来源于任何数目的逆转录病 毒，典型地致癌逆转录病毒且优选地哺乳动物致癌逆转录病毒。在一个实 施方案中，3’LTR包括U3-R-U5域。在另一个实施方案中，LTR包括如SEQ ID NO:19的约核苷酸9405到约9998所述的序列，或与其具有至少 95％、98％或99.5％(四舍五入到十分位)同一性的序列。

[0144] 本公开书还提供一种重组逆转录病毒载体，其从5’到3’依次包含：CMV-R-U5，即

人类巨细胞病毒的立即早期启动子与MLV R-U5区的融 合体;PBS,即逆转录酶的引物结合位点;5'剪接位点;ψ包装信号;gag, 即MLV群特异性抗原的ORF;pol,即MLV聚合酶多聚蛋白的ORF;3' 剪接位点;4070A env,即MLV品系4070A的包膜蛋白的ORF;IRES, 即脑心肌炎病毒的内部核糖体进入位点;经修饰的胞嘧啶脱氨酶(经热稳 定化和密码子优化);PPT,即聚嘌呤束;和U3-R-U5,即MLV长末端 重复序列。这个结构在图3中进一步描绘。

[0145] 本公开书还提供一种包括如SEQ ID NO:19、20或22所述的序列的 逆转录病毒载体。

[0146] 逆转录病毒载体可用于治疗多种疾病和失调,包括许多细胞增殖疾病 和失调(参看例如美国专利第4,405,712和4,650,764号;Friedmann,1989, Science,244:1275-1281;Mulligan,1993,Science,260:926-932,R.Crystal, 1995,Science 270:404-410,其各自以全文引用的方式并入本文中;另外也 参看The Development of Human Gene Therapy, Theodore Friedmann编, Cold Spring Harbor Laboratory Press,Cold Spring Harbor, N.Y.,1999.ISBN 0-87969-528-5,其以全文引用的方式并入本文中)。

[0147] 本公开书还提供用于治疗细胞增殖失调的基因疗法。所述疗法通过将 适当的治疗性多聚核苷酸(例如反义分子、核酶、自杀基因、siRNA)引 入到患有增殖失调的受试对象的细胞中来实现其治疗作用。多聚核苷酸构 建体可使用本公开书的重组逆转录病毒载体来递送,尤其当所述载体是基 于能感染分裂细胞的MLV时。

[0148] 此外,如本文所述的治疗方法(例如基因疗法或基因递送方法)可在 体内或体外进行。优选在基因治疗之前去除大部分肿瘤,例如通过外科手 术或通过放射。在一些方面,逆转录病毒治疗可在外科手术、化学治疗或 放射治疗之后或之前进行。

[0149] 因此,本公开书提供一种能感染非分裂细胞、分裂细胞或赘生性细胞 的重组逆转录病毒,其中重组逆转录病毒包含病毒GAG;病毒POL;病 毒ENV;可操作地连接于IRES的异源核酸;和包装、逆转录和整合所必 需的顺式作用核酸序列。重组逆转录病毒可以是慢病毒(例如HIV), 或可以是致癌病毒。如上文对于产生重组逆转录病毒的方法所述,本公开书的重组逆转录病毒可进一步包括VPR、VIF、NEF、VPX、TAT、REV 和VPU蛋白的至少一种。不希望受特定理论束缚,相信这些基因/蛋白产 物中的一种或一种以上对于增加所产生的重组逆转录病毒(例如NEF) 的病毒滴度至关重要,或可能是感染和包装病毒粒子所必需的。

[0150] 本公开书还提供一种将核酸转移到靶细胞以提供特定核酸(例如异源 序列)的表达的方法。因此,在另一个实施方案中,本公开书提供一种在 靶细胞中引入和表达异源核酸的方法,其包含用本公开书的重组病毒感染 靶细胞,和在靶细胞中表达异源核酸。如上文所述,靶细胞可以是任何细 胞类型,包括分裂、非分裂、赘生性、永生化、经修饰细胞和所属领域技 术人员知晓的其它细胞类型,只要它们能被逆转录病毒感染即可。

[0151] 可能需要通过用本公开书的方法引入核酸序列(例如异源核酸序列) 来调整细胞中的基因表达,其中核酸序列例如产生反义或核酶分子。术语 "调整"表示当基因过度表达时抑制基因表达,或者当基因表达不足时增强 基因表达。当细胞增殖失调与基因表达相关时,可使用在翻译水平上干扰 基因表达的核酸序列。这种方法利用例如反义核酸、核酶或三螺旋介质, 通过用反义核酸或三螺旋介质屏蔽特异性mRNA,或通过用核酶裂解特 异性mRNA来阻断特定mRNA的转录和翻译。

[0152] 可能需要将编码生物反应调节剂(例如细胞因子)的核酸转移到细胞 或受试对象

中。所述类别包括免疫增强剂，包括编码被分类为“白细胞介 素”的许多细胞因子的核酸。这些生物反应调节剂包括例如白细胞介素1 到15，以及本文其它地方所述的其它反应调节剂和因子。虽然不一定根 据相同机制工作，但所述类别也包括干扰素且尤其是γ干扰素、肿瘤坏死 因子(TNF)和粒细胞-巨噬细胞集落刺激因子(GM-CSF)。其它多肽包 括例如血管生成因子和抗血管生成因子。可能需要将所述核酸递送到骨髓 细胞或巨噬细胞以治疗酶缺陷或免疫缺陷。也可将编码生长因子、毒性肽、配体、受体或其它生理学上重要的蛋白质的核酸引入特定靶细胞中。

[0153] 本公开书可用于递送促进药物特异性靶向和作用的异源多聚核苷酸。举例来说，EGF受体家族成员HER2(参看例如SEQ ID NO:23和24)是 药物曲妥珠单抗(trastuzumab)(赫赛汀(Herceptin)™，Genentech)的 结合靶标。曲妥珠单抗是抗体依赖性细胞毒性(ADCC)的介导物。活性 利用免疫组织化学优先被靶向于2+和3+级过度表达的HER2表达细胞，而不是1+级和非表达细胞(Herceptin prescribing information，Crommelin 2002)。通过在低HER2肿瘤中引入表达HER2或截短HER2(只表达胞 外域和跨膜域)的载体来增强HER2的表达可促进ADCC的最佳触发并 克服在临床应用上观察到的快速出现HER2抗性。

[0154] 用yCD2(包含SEQ ID NO:19的约8877到9353)取代HER2的胞 内域可允许HER2的细胞表面表达和yCD2的胞质定位。HER2胞外域 (ECD)和跨膜域(TM)(SEQ ID NO:23的约175到2200位的约2026 bp)可通过PCR扩增(Yamamoto等人，Nature 319:230-234，1986；Chen 等人，Canc.Res.，58:1965-1971，1998)或化学合成(BioBasic Inc.，Markham，Ontario，Canada)，并插入载体pAC3-yCD2 SEQ ID NO:19中的IRES与 yCD2基因之间(例如在SEQ ID NO:19的约核苷酸8876与8877之间)。或者，yCD基因可被切除并用编码HER2多肽或其片段的多聚核苷酸取 代。仅具有ECD和TM域的赫赛汀结合域IV(1910到2200位的约290bp) 的另一种截短HER2可如上文所述被扩增或化学合成并使用(Landgraf 2007；Garrett等人，J.of Immunol.，178:7120-7131，2007)。具有融合于IV 和TM域的天然信号肽(175-237位的约69bp)的所述截短形式的另一 种修饰体可如上文所述被化学合成和使用。所得病毒可与曲妥珠单抗或曲 妥珠单抗和5-FC结合用于治疗受试对象的细胞增殖失调。

[0155] 或者，如上所述的HER2和修饰体可在含有不同ENV基因或其它适 当表面蛋白的单独载体中表达。这种载体可以是编码目的包膜和基因的可 复制性(Logg等人J.Mol Biol.369:12142007)或非复制性“第一代”逆转 录病毒载体(Emi等人J.Virol 65:12021991)。在后一种情况下，已有的 病毒感染会提供互补gag和pol以允许“非复制性”载体从任何先前感染细 胞进行感染传播。替代性ENV和糖蛋白包括能感染人类细胞的异嗜性和 多嗜性ENV和糖蛋白，例如来自MLV的NZB品系的ENV序列和来自 MCF、VSV、GALV和其它病毒的糖蛋白(Palu 2000，Baum等人，Mol. Therapy，13(6):1050-1063，2006)。举例来说，多聚核苷酸可包括如下序 列，其中SEQ ID NO:19的GAG和POL和yCD2基因被缺失，ENV对 应于NZB MLV或VSV-g的异嗜性ENV域，且IRES或启动子(例如RSV) 可操作地直接连接于HER2、HER2 ECDTM、HER2 ECDIVTM或HER2 SECDIVTM。

[0156] 用VSVG假型病毒和双嗜性逆转录病毒混合感染细胞会产生携带被 另一种病毒的包膜蛋白衣壳化的一种病毒的基因组的子代病毒粒子。这一 点同样适用于假型逆转录病毒粒子的其它包膜。举例来说，上述来源的逆 转录病毒所造成的感染会产生能在被感染细胞中编码yCD2和HER2(或 变异体)的子代病毒粒子。所得病毒可与曲妥珠单抗或曲妥珠单

抗和5-FC 结合用于治疗受试对象的细胞增殖失调。

[0157] 最近,已经将γ逆转录病毒XMRV与具有在前列腺组织中显示强复 制优先性的病毒的人类前列腺癌联系起来(R.Schlaberg等人PNAS106: 16351-163562009)。所述病毒似乎非常类似于异嗜性MLV。在本发明 的一个实施方案中,提供非复制性逆转录病毒载体,其具有治疗基因(胞 嘧啶脱氨酶、胸苷激酶、其它前药活化基因、干扰素、IL-2、IL-12、其 它细胞因子、p53其它抗癌基因、抗癌miRNA等)与能由XMRV gag和 gagpol功能物互补的包膜基因,例如双嗜性包膜、GALV包膜、VSVg蛋 白包膜或所属领域技术人员已知的其它包膜。非复制性载体多聚核苷酸作 为DNA或RNA分子利用以下物质中的一种被递送到患者或动物前列腺 癌:非病毒或物理递送系统;异源病毒递送系统,例如腺病毒载体;或以 所制造的逆转录病毒粒子的形式。一旦被递送后,非复制性载体将通过 XMRV互补而传播,且邻近细胞将发生感染,直至到达XMRV感染边界 为止,此时XMRV互补不再可用。治疗基因然后可仅在XMRV感染区域 具有作用(例如在施用前药后)。可使用本发明的复制性逆转录病毒载体 实现相同的拯救作用。这种策略(与治疗基因互补的非复制性载体)可用 于任何逆转录病毒疾病(HIV感染、HTLV1感染、其它与逆转录病毒相 关的癌症)或任何病毒或病毒相关疾病(在宫颈癌、EBV相关淋巴瘤或 癌瘤等中的HPV感染和HPV E6&E7表达)。

[0158] 发展曲妥珠单抗抗性的另一方面涉及干扰曲妥珠单抗活性所必需的 胞内信号传导。抗性细胞显示缺少PTEN和p27kip1表达较低(Fujita,Brit J.Cancer,94:247,2006;Lu等人,Journal of the National Cancer Institute, 93(24):1852-1857,2001;Kute等人,Cytometry Part A 57A:86-93,2004)。举例来说,编码PTEN(SEQ ID NO:25)的多聚核苷酸可被重组产生或化 学合成(BioBasic Inc.,Markham,Canada),并可操作地直接插入载体 pAC3-yCD2 SEQ ID NO:19或22中的yCD2多聚核苷酸之后,或连同先 前所述的连接子序列一起插入,或代替yCD2插入。在另一个实施例中, 编码PTEN的多聚核苷酸可如上合成并插入IRES与yCD2序列之间,或 连同先前所述的连接子一起插入。

[0159] 或者,PTEN可在含有不同ENV基因或其它适当表面蛋白的单独载 体中表达。这种载体可以是编码目的包膜和基因的可复制性(Logg等人 J.Mol Biol.369:12142007)或非复制性“第一代”逆转录病毒载体(Emi等 人J.Virol 65:12021991)。在后一种情况下,已有的病毒感染会提供互 补gag和pol以允许“非复制性”载体从任何先前感染细胞进行感染传播。 替代性ENV和糖蛋白包括能感染人类细胞的异嗜性和多嗜性ENV和糖 蛋白,例如来自MLV的NZB品系的ENV序列和来自MCF、VSV、GALV 和其它病毒的糖蛋白(Palu,Rev Med Virol.2000,Baum,Mol.Ther. 13(6):1050-1063,2006)。举例来说,多聚核苷酸可包括如下序列,其中 SEQ ID NO:19的GAG和POL和yCD2基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV域,且IRES或启动子(例如RSV)可操 作地直接连接于PTEN。

[0160] 用VSVG假型病毒和双嗜性逆转录病毒混合感染细胞会产生带有被 另一种病毒的包膜蛋白衣壳化的一种病毒的基因组的子代病毒粒子[Emi 1991]。这一点同样适用于假型逆转录病毒粒子的其它包膜。举例来说, 上述来源的逆转录病毒所造成的感染会产生能在被感染细胞中编码 yCD2和PTEN(或变异体)或只编码PTEN的子代病毒粒子。所得病毒 可与曲妥珠单抗或曲妥珠单抗和5-FC结合用于治疗受试对象的细胞增殖 失调。

[0161] 类似地,编码p27kip1(SEQ ID NO:27和28)的多聚核苷酸可被化 学合成(BioBasic Inc.,Markham,Canada),并可操作地直接插入载体 pAC3-yCD2 SEQ ID NO:19

中的yCD2基因之后,或连同连接子序列一起 插入。在另一个实施例中,编码p27kip1的多聚核苷酸可如上合成并插入 IRES与yCD2序列之间,或连同先前所述的连接子一起插入或代替yCD2 基因插入。

[0162] 或者,p27kip1可在含有不同ENV基因或其它适当表面蛋白的单独 载体中表达。这种载体可以是编码目的包膜和基因的可复制性(CR.Logg 等人J.Mol Biol.369:12142007)或非复制性"第一代"逆转录病毒载体 (Emi等人J.Virol 65:12021991)。在后一种情况下,已有的病毒感染会 提供互补gag和pol以允许"非复制性"载体从任何先前感染细胞进行感染 传播。替代性ENV和糖蛋白包括能感染人类细胞的异嗜性和多嗜性ENV 和糖蛋白,例如来自MLV的NZB品系的ENV序列和来自MCF、VSV、GALV和其它病毒的糖蛋白(Palu 2000,Baum 2006,上文)。举例来说, 多聚核苷酸可包括如下序列,其中SEQ ID NO:19的GAG和POL和yCD2 基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV域,且IRES 或启动子(例如RSV)可操作地直接连接于p27kip1。

[0163] 用VSVG假型病毒和双嗜性逆转录病毒混合感染细胞会产生一种病 毒的基因组被另一种病毒的包膜蛋白衣壳化的子代病毒粒子[Emi 1991]。这一点同样适用于假型逆转录病毒粒子的其它包膜。举例来说,如上所述 来源于SEQ ID NO:19和22的逆转录病毒所造成的感染会产生能在被感 染细胞中编码yCD2和p27kip1(或变异体)的子代病毒粒子。所得病毒 可与曲妥珠单抗或曲妥珠单抗和5-FC结合用于治疗受试对象的细胞增殖 失调。

[0164] 在另一个实施例中,CD20是药物利妥昔单抗(rituximab)(Rituxan™, Genentech)的结合靶标。利妥昔单抗是补体依赖性细胞毒性(CDC)和 ADCC的介导物。根据流式细胞术测量,具有较高平均荧光强度的细胞 显示提高的利妥昔单抗敏感性(van Meerten等人,Clin Cancer Res 2006; 12(13):4027-4035,2006)。通过在低CD20B细胞中引入表达CD20的载 体来增强CD20表达可促进ADCC的最佳触发。

[0165] 举例来说,编码CD20(SEQ ID NO:29和30)的多聚核苷酸可被化 学合成(BioBasic Inc.,Markham,Canada),并连同先前所述的连接子序 列一起可操作地直接插入载体pAC3-yCD2(-2)SEQ ID NO:19或22中的 yCD2基因之后,或代替yCD2基因插入。在另一个实施例中,编码CD20 的多聚核苷酸可如上合成并插入IRES与yCD2序列之间,或连同先前所 述的连接子一起插入。作为另一个替代方案,CD20序列可在通过Psi1和 Not1消化切除CD基因后插入pAC3-yCD2载体中。

[0166] 在另一个实施例中,编码CD20(SEQ ID NO:29和30)的多聚核苷 酸可被化学合成(BioBasic Inc.,Markham,Canada),并插入含有非双嗜 性ENV基因或其它适当表面蛋白的载体中(Tedder等人,PNAS, 85:208-212,1988)。替代性ENV和糖蛋白包括能感染人类细胞的异嗜性 和多嗜性ENV和糖蛋白,例如来自MLV的NZB品系的ENV序列和来 自MCF、VSV、GALV和其它病毒的糖蛋白(Palu 2000,Baum 2006)。 举例来说,多聚核苷酸可包括如下序列,其中SEQ ID NO:19的GAG和 POL和yCD2基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV域,且IRES或启动子(例如RSV)可操作地直接连接于CD20。

[0167] 用VSVG假型病毒和双嗜性逆转录病毒混合感染细胞会产生含有被 另一种病毒的包膜蛋白衣壳化的一种病毒的基因组的子代病毒粒子[Emi 1991]。这一点同样适用于假型逆转录病毒粒子的其它包膜。举例来说, 如上所述来源于SEQ ID NO:19或22的逆转录病毒所造成的感染会产生 能在被感染细胞中编码yCD2和CD20的子代病毒粒子。所得病毒可与

美 罗华(Rituxan)和/或5-FC结合用于治疗受试对象的细胞增殖失调。类似 地,用仅编码CD20标志物的载体感染肿瘤可导致肿瘤可通过使用美罗华 来治疗。

[0168] 嘧啶合成代谢中所涉及的酶和辅因子的水平可构成限制。与敏感细胞 系相比,OPRT、胸苷激酶(TK)、尿苷磷酸激酶和嘧啶核苷磷酸化酶在 5-FU抗性癌细胞中的表达水平很低(Wang等人,Cancer Res., 64:8167-8176,2004)。大群体分析显示酶水平与疾病结果的相关性(Fukui 等人,Int'l.J.OF Mol.Med.,22:709-716,2008)。可利用CD和其它嘧啶合成代谢酶(PAE)的共表达来增加活性且因此增加氟嘧啶药物的治疗指数。

[0169] 为了进一步增加含yCD2/PAE载体的遗传稳定性(参看例如图5), 可化学合成在序列中具有随机突变的酶编码基因。这些突变可以是基本上 随机的,或可以仅由每个氨基酸的摆动位置上的突变组成。突变序列的文 库如先前对于SEQ ID NO:11和13所述插入yCD2基因下游或代替yCD2 基因插入,以产生质粒文库,其可接着通过瞬时转染293T细胞或同等物 而用于产生感染粒子文库。可用编码融合多肽的逆转录病毒感染敏感细胞 并用适当化学物进行筛选。

[0170] DNA改组或"分子育种"允许改组遗传信息,从而产生具有期望特性 的重组体。已经利用DNA改组改良了不同的蛋白质和酶(Stemmer 1994 Proc Natl Acad Sci USA 91(22):10747-51;Stemmer 1994 Nature 370 (6488):389-91)。基因重组是推动许多病毒进化的主要动力。在逆转录病 毒中,两个共包装逆转录病毒基因组之间的重组可以每复制周期高达40% 的速率发生。每个复制周期中的高重组速率使得遗传信息能被快速改组, 从而在短时间内产生具有新突变模式和表型的重组体。举例来说,含有来 自6种鼠类白血病病毒的重组亲嗜性包膜序列的文库的逆转录病毒的分 子育种产生具有新向性的病毒克隆。使用相同方法,筛选具有提高的稳定 性和加工产率的一些病毒克隆(Soong等人,2000 Nat Genet 25(4):436-9; Powell等人,2000 Nat Biotechnol 18(12):1279-82)。

[0171] 并入本发明中的多聚核苷酸序列有时不稳定,从而导致多聚核苷酸序 列随着时间会从病毒基因组中缺失。其原因尚未完全了解,但清楚使用本 发明筛选已经在异源多聚核苷酸序列内获得最佳重组的重组病毒克隆的 序列依赖性分子育种可用于筛选具有较大载体稳定性的病毒克隆。

[0172] 举例来说,HSV-TK编码序列(SEQ ID NO:37)在一些情形中没有 预期中稳定。重组逆转录病毒载体的分子育种涵盖用于筛选具有较大载体 稳定性的重组载体的HSV-TK简并编码序列的库。可化学合成(Bio Basic Inc,Markham,Canada)随机突变疱疹胸苷激酶(TK)。合成序列可插入 SEQ ID NO:19中的yCD2序列的3',或在切除CD2基因后独自插入pAC3-yCD2载体骨架中。如先前所述包装逆转录病毒载体混合物。用载 体感染小鼠成纤维细胞LMTK细胞或人类143Tk细胞并在HAT培养基 中筛选TK活性(Hiller等人,Mol.Cell Biol.8(8):3298-3302,1988)。在 HAT培养基中再次筛选的新鲜LMTK-/143Tk-细胞的抗性细胞上清液的 连续传代允许筛选表达TK的稳定载体。可分离TK+抗性细胞并用标准 PCR型技术拯救TK序列以用于突变分析(Cowell等人,CDNA Library Protocols,由Humana Press发行,1996)。以此方式,筛选功能蛋白表达 和逆转录病毒载体构建体的基因组稳定的序列。类似策略可用于UPRT (SEQ ID NO:11,13)、OPRT(SEQ ID NO:15,17)(Olah等人,Cancer Res.40:2869-2875,1980;和Suttle,Somatic Cell&Mol.Genet., 15(5):435-443,1989)和其它目的基因。此外,通过PCR在IRES-插入基 因中筛检连续传代的全长插入物之后,连续

传代策略可用于非筛选基因和 基因组DNA。可纯化全长插入物，并重新克隆到病毒载体中，然后进行 再测试。所述程序可进行若干循环以筛选最稳定的基因。所述策略也可用 于具有或不具有肿瘤的动物和甚至患者组织中的传代。

[0173] 或者，OPRT、UPRT、TK或其它PAE可在含有不同ENV基因或其 它适当表面糖蛋白的单独载体中表达。这种载体可以是编码目的包膜和基 因的可复制性(Logg等人J.Mol Biol.369:12142007)或非复制性“第一 代”逆转录病毒载体(Emi等人J.Virol 65:12021991)。在后一种情况下，已有的病毒感染会提供互补gag和pol以允许“非复制性”载体从任何先前 感染细胞进行感染传播。替代性ENV和糖蛋白包括能感染人类细胞的异 嗜性和多嗜性ENV和糖蛋白，例如来自MLV的NZB品系的ENV序列 和来自MCF、VSV、GALV和其它病毒的糖蛋白(Palu 2000，Baum 2006， 上文)。举例来说，多聚核苷酸可包括如下序列，其中GAG和POL基 因被缺失，ENV对应于NZB MLV或VSV-g的异嗜性ENV域，且IRES 或启动子(例如RSV)可操作地直接连接于OPRT、UPRT、TK或其它 PAE基因。

[0174] 用VSV-g假型病毒和双嗜性逆转录病毒混合感染细胞会产生一种病 毒的基因组被另一种病毒的包膜蛋白衣壳化的子代病毒粒子[Emi等人，J. Virol.65:1202，1991]。这一点同样适用于假型逆转录病毒粒子的其它包膜。举例来说，如上所述来源于SEQ ID NO:19和22的逆转录病毒所造成的 感染会产生能在被感染细胞中编码yCD2和OPRT的子代病毒粒子。所得 病毒可与5-FC结合用于治疗受试对象的细胞增殖失调。

[0175] 本公开书的重组逆转录病毒可用于治疗神经元失调，所述逆转录病毒 例如可任选地含有外源基因，例如编码受体的基因或编码配体的基因。如 上所述，所述受体包括响应于多巴胺、GABA、肾上腺素、去甲肾上腺素、 5-羟色胺、谷氨酸盐、乙酰胆碱和其它神经肽的受体，如上所述。可在神 经元失调中提供治疗作用的配体的实例包括多巴胺、肾上腺素、去甲肾上 腺素、乙酰胆碱、γ-氨基丁酸和5-羟色胺。必需配体在被感染供体细胞分 泌之后发生扩散和摄取在受试对象的神经细胞在产生所述基因产物方面 存在缺陷的失调中是有益的。经遗传改造以分泌神经营养因子(例如神经 生长因子(NGF))的细胞可能用于防止胆碱能神经元退化，所述神经元 如果不经治疗可能会死亡。

[0176] 或者，被移植入患有基底神经节失调(例如帕金森症(Parkinson′s disease))的受试对象中的细胞可被修饰以含有编码多巴胺前体L-DOPA 的外源基因。帕金森症的特征在于中脑黑质中缺少以基底神经节作为主要 靶器官的多巴胺神经元。

[0177] 可用本公开书的方法类似地治疗的其它神经元失调包括阿尔茨海默 氏病、亨廷顿氏病、因中风引起的神经元损伤和脊髓损伤。阿尔茨海默氏 病的特征在于基底前脑的胆碱能神经元退化。这些神经元的神经递质是乙 酰胆碱，其是所述神经元的生存所必需的。可用所述的本公开书的方法植 入被含有促进这些神经元生存的因子的外源基因的本公开书的重组逆转 录病毒感染的胆碱能细胞。在中风后，海马以及皮层细胞的CA1中的细 胞发生选择性损失，这可能是这些患者的认知和记忆损失的基础原因。一 旦被鉴别出来，负责CA1细胞死亡的分子可通过本公开书的方法抑制。举例来说，反义序列或编码拮抗剂的基因可被转移到神经元细胞中并植入 大脑海马区。

[0178] 对于因蛋白产物缺乏引起的疾病，基因转移可将正常基因引入被感染 组织以用于取代治疗，以及利用反义突变来产生疾病的动物模型。举例来 说，可能需要将因子IX编码核酸插入用于感染肌肉或肝细胞的逆转录病 毒中。

[0179] 本公开书还提供用于治疗细胞增殖失调或免疫失调的基因疗法。所述 疗法通过将反义或显性阴性编码多聚核苷酸引入具有增殖失调的细胞中 来实现治疗作用,其中所述多聚核苷酸结合于细胞增殖失调的相关基因并 防止其翻译或表达。可使用本公开书的重组逆转录病毒载体递送适用于治 疗或调节细胞增殖失调(例如反义多聚核苷酸)的异源核酸。在另一个实 施方案中,通过引入本公开书的CD多聚核苷酸、表达多聚核苷酸以产生包含胞嘧啶脱氨酶活性的多肽和使细胞与可产生细胞毒性量的5-FU的量 的5-氟胞嘧啶接触可产生细胞毒性量的5-FU的时间来治疗细胞增殖失调。

[0180] 目前已有许多化学治疗剂在市面销售,其具有从完全缓解到暂时缓解 的不同程度的效果和复发率达到预期的延长寿命。市面销售的一些癌症治 疗剂是以肿瘤的血管生成特性为目标。组合物以肿瘤的血管生成为目标, 设法减少供应给肿瘤或癌细胞的血液和营养物,从而减小肿瘤并延长受试 对象的寿命。VEGF是已知在肿瘤生长中起作用的血管生成因子。因此, 已经开发了作为抗癌剂的VEGF拮抗剂。

[0181] 人类VEGF介导正常和恶性脉管系统中的新生血管生成;其在大多 数恶性肿瘤中过度表达,且高含量已经与许多患者的较高转移风险和不良 预后有关。当VEGF在体外血管生成模型中与其受体相互作用时,发生 内皮细胞增殖和新血管形成。在动物模型中,已经证明了VEGF诱导血 管内皮细胞增殖/迁移,维持新形成血管的存活,并提高血管通透性。

[0182] VEGF拮抗剂靶向VEGF信号传导路径或负向调节VEGF信号传导 路径。实例包括VEGF抑制剂(例如直接抑制VEGF(例如VEGF-A、-B、 -C或-D)的介质,例如通过结合VEGF(例如抗VEGF抗体,例如贝伐 单抗(bevacizumab)(AVASTIN®)或兰尼单抗(ranibizumab)(LUCENTIS®),或其它抑制剂,例如派加他尼(pegaptanib)、NEOVASTAT®、AE-941、VEGF Trap和PI-88))、VEGF表达调节剂 (例如INGN-241、口服四硫钼酸盐(oral tetrathiomolybdate)、2-甲氧基 雌二醇、2-甲氧基雌二醇纳米晶体分散液、贝伐西尼钠(bevasiranib sodium)、PTC-299、Veglin)、VEGF受体抑制剂(例如KDR或VEGF 受体III(Flt4),例如抗KDR抗体、VEGFR2抗体(例如CDP-791)、 IMC-1121B、VEGFR2阻断剂(例如CT-322))、VEGFR表达调节剂(例 如VEGFR1表达调节剂Sirna-027)或VEGF受体下游信号传导抑制剂。 在本文所述的一些方面,VEGF拮抗剂是贝伐单抗、派加他尼、兰尼单抗、 索拉非尼(sorafenib)、舒尼替尼(sunitinib)、AE-941、VEGF Trap、 帕唑帕尼(pazopanib)、凡德他尼(vandetanib)、瓦他拉尼(vatalanib)、 西地尼布(cediranib)、芬维A胺(fenretinide)、角鲨胺(squalamine)、 INGN-241、口服四硫钼酸盐、四硫钼酸盐、Panzem NCD、2-甲氧基雌二 醇、AEE-788、AG-013958、贝伐西尼钠、AMG-706、阿西替尼(axitinib)、 BIBF-1120、CDP-791、CP-547632、PI-88、SU-14813、SU-6668、XL-647、 XL-999、IMC-1121B、ABT-869、BAY-57-9352、BAY-73-4506、BMS-582664、 CEP-7055、CHIR-265、CT-322、CX-3542、E-7080、ENMD-1198、OSI-930、 PTC-299、Sirna-027、TKI-258、Veglin、XL-184或ZK-304709。

[0183] 贝伐单抗(AVASTATIN®)(rhuMAb-VEGF)(抗VEGF单克隆抗 体)是针对血管内皮生长因子(VEGF)的重组人类/鼠类嵌合单克隆抗体。贝伐单抗是通过将鼠类抗VEGF单克隆抗体的VEGF结合残基工程改造 到人类免疫球蛋白-1(IgG1)的框架区中来制备(Prod Info Avastin,2004)。只有7%的氨基酸序列来源于鼠类抗体,其余93%来自IgG1。贝伐单

抗通 过识别两种人类VEGF受体类型(flt-1和flk-1)的结合位点来结合和中 和所有人类VEGF形式。在动物模型中,已经显示抗体可通过抑制VEGF 所诱导的血管生成来稳定化已形成的肿瘤或抑制肿瘤生长。

[0184] 贝伐单抗的药物动力学在0.3mg/kg或以上的剂量后是线性的(Anon, 2002)。在晚期癌症患者(n=25)中静脉内输注0.3、1、3和10mg/kg 达90分钟后,贝伐单抗的峰值血清浓度分别在5到9mcg/mL、21到39 mcg/mL、52到92mcg/mL和186到294mcg/mL范围内;在重复剂量(每 周一次)下观察到轻微累积,但这并不显著且药物动力学仍然保持线性。在每周一次、每2周一次或每3周一次施用1到20mg/kg后,于100天 内获得贝伐单抗的稳态水平。

[0185] 贝伐单抗的推荐剂量是每2周一次历经30分钟静脉内输注5毫克/ 千克,直到疾病发展减轻为止。贝伐单抗应该在化学治疗后面进行。单一 介质贝伐单抗的功效尚未确定。贝伐单抗(其可与吉西他滨(gemcitabine) 和多西他赛(docetaxel)共同施用,或在化学治疗之前或之后一周内施用) 是以约1mg/kg到约15mg/kg、优选约5mg/kg静脉内施用。

[0186] 本公开书的方法和组合物可用于包括使用贝伐单抗的疗法的组合疗 法。如本文所述,本公开的包含治疗介质(例如,细胞毒性基因)的复制 逆转录病毒(RCR)可用于治疗细胞细胞增殖失调。本公开的RCR的优点 包括其能够感染复制中的细胞癌细胞。当载体的转基因包括细胞毒性基因 (例如,编码将肺细胞毒性介质转化为细胞毒性介质的基因)时,提供杀死 癌细胞的能力。

[0187] 本公开书提供治疗例如癌症和赘瘤等细胞增殖失调的方法,其包含施 用本公开书的RCR载体,随后用化学治疗剂或抗癌剂进行治疗。在一方 面,RCR载体在施用化学治疗剂或抗癌剂之前施用于受试对象,持续允 许RCR感染和复制的一段时段。然后在可减少增殖或杀死癌细胞的时段 和剂量下用化学治疗剂或抗癌剂治疗受试对象。在一方面,如果化学治疗 剂或抗癌剂治疗减少但不杀死癌/肿瘤(例如部分缓解或暂时缓解),那 么受试对象随后可用非毒性治疗剂(例如5-FC)治疗,所述非毒性治疗 剂可在表达来自RCR的细胞毒性基因(例如胞嘧啶脱氨酶)的细胞中转 化为毒性治疗剂。使用所述方法,本公开书的RXCR载体在肿瘤细胞复 制过程中传播,所述细胞然后可通过用抗癌剂或化学治疗剂处理而被杀死, 且进一步杀死可使用本文所述的RCR处理过程来进行。

[0188] 在本公开书的另一个实施方案中,异源基因可包含靶抗原(例如癌抗 原)的编码序列。在这个实施方案中,用包含编码靶抗原的异源多聚核苷 酸的RCR感染包含细胞增殖失调的细胞以提供靶抗原的表达(例如癌抗 原的过度表达)。然后对受试对象施用包含特异性地与靶抗原相互作用的 靶向同源部分的抗癌剂。靶向同源部分可操作地连接于细胞毒性剂或自身 可以是抗癌剂。因此,被包含靶向抗原编码序列的RCR感染的癌细胞增 加靶标在癌细胞上的表达,从而提高细胞毒性靶向作用的效率/功效。

[0189] 已经显示阻断免疫系统细胞之间的相互作用可具有显著的活化性或 抑制性免疫作用(Waldmann Annu Rev Med.57:652006)。举例来说,已 经显示阻断CTLA-4(CD 152)与B7.1(CD80)的相互作用(其调节T 细胞活化)可引起免疫刺激,推测这是通过阻断此抑制性相互作用而实现 的(Peggs等人Curr.Opin.Immunol.18:206,2006)。所述阻断可能用 针对CTLA-4的抗体或用可溶性B7.1来实现。全身施用这些类型的分子 会具有不合需要的全身作用,其可在最小程度上导致有害副作用或甚至在 一种C28激动剂的情况中导致死亡(Suntharalingam等人NEJM 355 1018 2006)。Pfizer一直在开发一种该类抗CTLA-4阻断抗

体(CP-675,206) 作为抗癌剂,但是因为显著的副作用,最近已经停止开发。局部递送阻断分子(其在感染编码可释放到胞外空间中的所述分子的可复制性载体之后 从肿瘤释放到局部环境中)会提供局部免疫调节并避免这些严重副作用。阻断分子是抗体、单链抗体、天然配体的可溶形式或结合所述受体的其它 肽。

[0190] 在另一个实施方案中,本公开书的RCR可包含含有与同源抗原或配 体特异性地相互作用的结合域(例如抗体、抗体片段、抗体域或受体配体) 的编码序列。包含结合域的编码序列的RCR然后可用于感染患有细胞增 殖失调的受试对象的细胞,例如癌细胞或赘生性细胞。被感染细胞然后表 达结合域或抗体。然后可将可操作地连接于细胞毒性剂或自身具有细胞毒 性的抗原或同源物施用于受试对象。细胞毒性同源物然后选择性地杀死表 达结合域的被感染细胞。或者,结合域自身可以是抗癌剂。

[0191] 如本文所用,术语"抗体"是指包含至少一个免疫球蛋白可变域或免疫 球蛋白可变域序列的蛋白质。举例来说,抗体可包含重(H)链可变区(本 文缩写为VH)和轻(L)链可变区(本文缩写为VL)。在另一个实施例 中,抗体包含两个重(H)链可变区和两个轻(L)链可变区。术语"抗体" 涵盖抗体的抗原结合片段(例如单链抗体、Fab片段、$F(ab')_2$、Fd片段、 Fv片段和dAb片段)以及完整抗体。

[0192] VH和VL区可进一步细分为被称为"互补决定区"(CDR)的高变区, 其间散布有被称为"框架区"(FR)的更保守区域。框架区和CDR的范围 已被精确定义(参看Kabat,E.A.等人(1991)Sequences of Proteins of Immunological Interest,第5版,U.S.Department of Health and Human Services,NIH Publication No.91-3242,和Chothia,C.等人(1987) J.Mol. Biol.196:901-917)。本文使用Kabat定义。每个VH和VL典型地由三个 CDR和四个FR构成,它们从氨基端到羧基端以如下顺序排列:FR1、CDR1、 FR2、CDR2、FR3、CDR3、FR4。

[0193] "免疫球蛋白域"是指来自免疫球蛋白分子的可变域或恒定域的区域。免疫球蛋白域典型地含有2个由约7个β股形成的β折叠,和保守二硫 键(参看例如A.F.Williams和A.N.Barclay 1988Ann.Rev Immunol. 6:381-405)。如Chothia等人(1992) J.Mol.Biol.227:799-817;Tomlinson 等人(1992)J.Mol.Biol.227:776-798);和Tomlinson等人(1995)EMBO J. 14(18):4628-38中所述,可自免疫球蛋白可变域的序列推断其高变环的典 型结构。

[0194] 如本文所用,"免疫球蛋白可变域序列"是指可形成免疫球蛋白可变域 结构的氨基酸序列。举例来说,所述序列可包含天然存在的可变域的全部 或部分氨基酸序列。举例来说,序列可省略1个、2个或以上N或C端 氨基酸、内部氨基酸,可包含一个或多个插入或其它末端氨基酸,或可包 含其它改变。在一个实施方案中,包含免疫球蛋白可变域序列的多肽可与 另一个免疫球蛋白可变域序列结合以形成靶结合结构(或"抗原结合位 点"),例如与Tie1相互作用(例如结合或抑制Tie1)的结构。

[0195] 抗体的VH或VL链可进一步包含全部或部分重链或轻链恒定区,从 而分别形成重免疫球蛋白链或轻免疫球蛋白链。在一个实施方案中,抗体 是两个重免疫球蛋白链和两个轻免疫球蛋白链的四聚体,其中重免疫球蛋 白链和轻免疫球蛋白链通过例如二硫键互相连接。重链恒定区包含三个域 CH1、CH2和CH3。轻链恒定区包含CL域。重链和轻链的可变区含有与 抗原相互作用的结合域。抗体的恒定区典型地介导抗体与宿主组织或因子 (包括免疫系统的各种细胞(例如效应细胞)和经典补体系统的第一组分 (C1q))的结合。术语"抗

体”包括IgA、IgG、IgE、IgD、IgM型(以及 其子类型)的完整免疫球蛋白。免疫球蛋白的轻链可为κ或λ型。在一个 实施方案中,抗体被糖基化。抗体可对于抗体依赖性细胞毒性和/或补体 介导的细胞毒性发挥功能。

[0196] 术语“单特异性抗体”是指对于特定靶标(例如表位)显示单一结合特 异性和亲和力的抗体。所述术语包括“单克隆抗体”,其是指例如从同系分 离细胞群体作为单一分子物种产生的抗体。“单克隆抗体组合物”是指抗体 或其片段在包含抗体的单一分子物种的组合物中的制剂。在一个实施方案 中,单克隆抗体由哺乳动物细胞产生。可组合一种或一种以上单克隆抗体 物种。

[0197] 抗体的一个或多个区域可以是人源的或有效地是人源的。举例来说, 一个或多个可变区可以是人源的或有效地是人源的。举例来说,一个或多 个CDR可以是人源的,例如HC CDR1、HC CDR2、HC CDR3、LC CDR1、 LC CDR2和LC CDR3。每个轻链CDR可以是人源的。HC CDR3可以是 人源的。一个或多个框架区可以是人源的,例如HC或LC的FR1、FR2、 FR3和FR4。在一个实施方案中,所有框架区都是人源的,例如来源于人 类体细胞,例如产生免疫球蛋白的造血细胞或非造血细胞。在一个实施方 案中,人类序列是例如由生殖系核酸编码的生殖系序列。一个或多个恒定 区可以是人源的或有效地是人源的。在另一个实施方案中,至少70、75、 80、85、90、92、95或98%的框架区(例如,总地来说FR1、FR2和FR3, 或总地来说FR1、FR2、FR3和FR4)或整个抗体可以是人源的或有效地 是人源的。举例来说,FR1、FR2和FR3总地来说可与编码重链或轻链序 列的座位的人类生殖系V链段所编码人类序列具有至少70、75、80、85、 90、92、95、98或99%的同一性。

[0198] 全部或部分抗体可由免疫球蛋白基因或其链段编码。示例性人类免疫 球蛋白基因包括κ、λ、α(IgA1和IgA2)、γ (IgG1、IgG2、IgG3、IgG4)、 δ、ε和μ恒定区基因,以及多种免疫球蛋白可变区基因。全长免疫球蛋 白轻链(约25Kd或214个氨基酸)是由NH2端的可变区基因(约110 个氨基酸)和COOH端的κ或λ恒定区基因编码。全长免疫球蛋白重链 (约50Kd或446个氨基酸)类似地由可变区基因(约116个氨基酸)和 一种其它上述恒定区基因(例如γ(编码约330个氨基酸))编码。轻链 是指包含轻链可变域的任何多肽。重链是指包含重链可变域的任何多肽。

[0199] 如本文所用的术语全长抗体的“抗原结合片段”(或简称为“抗体部分” 或“片段”)是指保留与目的靶标特异性地结合的能力的全长抗体的一个或 多个片段。术语全长抗体的“抗原结合片段”所涵盖的结合片段的实例包括 (i)Fab片段,其是由VL、VH、CL和CH1域组成的单价片段;(ii) $F(ab')_2$片段,其是包含由铰链区的二硫桥连接的两个Fab片段的二价片段; (iii)Fd片段,其是由VH和CH1域组成;(iv)Fv片段,其是由抗体 单臂的VL和VH域组成;(v)dAb片段(Ward等人,(1989)Nature 341:544-546),其是由VH域组成;和(vi)保留功能的分离互补决定区 (CDR)。此外,虽然Fv片段的两个域VL和VH由单独基因编码,但 VL和VH可使用重组方法通过合成连接子接合,使得其可作为单一蛋白 链制备,其中VL和VH区域配对形成被称为单链Fv(scFv)的单价分子。参看例如Bird等人(1988)Science 242:423-426;和Huston等人(1988) Proc.Natl.Acad.Sci.USA 85:5879-5883。

[0200] 抗体片段可使用包括所属领域技术人员已知的常规技术的任何适当 技术来获得。术语“单特异性抗体”是指对于特定靶标(例如表位)显示单 一结合特异性和亲和力的抗体。所述术语包括“单克隆抗体”或“单克隆抗 体组合物”,其在本文中是指单一分子组合

物的抗体或其片段的制剂。如 本文所用,“同种型”是指由重链恒定区基因编码的抗体类别(例如IgM 或IgG1)。

[0201] 本公开书提供一种治疗患有细胞增殖失调的受试对象的方法。受试对 象可以是任何哺乳动物,且优选是人。用本公开书的重组可复制逆转录病 毒载体接触受试对象。接触可在体内或体外进行。施用本公开书的逆转录 病毒载体的方法在所属领域中已知,且包括例如全身施用、局部施用、腹 膜内施用、肌肉内施用、颅内施用、脑脊髓施用以及在肿瘤或细胞增殖失 调部位直接施用。其它施用途径也在所属领域中已知。

[0202] 因此,本公开书包括可用于治疗细胞增殖失调的各种医药组合物。

[0203] 本公开书的医药组合物是通过使用载剂、赋形剂和添加剂或助剂,将 根据本公开书的含有可用于治疗或调节细胞增殖失调的异源多聚核苷酸 序列的逆转录病毒载体制成适合于施用于受试对象的形式来制备。经常使 用的载剂或助剂包括碳酸镁、二氧化钛、乳糖、甘露醇和其它糖、滑石粉、 乳蛋白、明胶、淀粉、维生素、纤维素和其衍生物、动物和植物油、聚乙 二醇和溶剂,例如无菌水、醇类、甘油和多元醇。静脉内运载体包括流体 和营养补充物。防腐剂包括抗微生物剂、抗氧化剂、螯合剂和惰性气体。其它医药学上可接受的载剂包括水溶液、无毒赋形剂(包括盐)、防腐剂、 缓冲剂等,如例如Remington′s Pharmaceutical Sciences,第15版Easton: Mack Publishing Co.,1405-1412,1461-1487(1975)和The National Formulary XIV.,第14版Washington:American Pharmaceutical Association(1975)中所述,其内容以引用的方式并入本文中。医药组合物 的各种组分的pH和确切浓度根据所属领域的常规技术来调节。参看 Goodman和Gilman的The Pharmacological Basis for Therapeutics(第7版)。

[0204] 举例来说且不构成限制,可用于治疗细胞增殖失调的逆转录病毒载体 包含双嗜性ENV蛋白、GAG和POL蛋白、U3区逆转录病毒基因组中的 启动子序列,和逆转录病毒基因组在靶细胞中复制、包装和整合所必需的 所有顺式作用序列。

[0205] 以下实施例打算说明但不限制本公开书。虽然所述实施例代表可能被 使用的实施例,但仍可替代性地利用所属领域技术人员已知的其它程序。

[0206] 实施例

[0207] 实施例1.将pACE-GFPemd的载体骨架修饰成pAC3-GFPemd并 插入胞嘧啶脱氨酶基因序列以代替GFP.用图3E所示的3种方式修饰 pACE-GFPemd质粒的先前骨架(US6899871,Wang等人Hum Gene Ther 14:117 2003)。所述修饰是通过PCR介导的寡聚核苷酸引导的诱变来进 行(Logg等人,J.Mol Biol 369:1214,2007;也参看“Molecular Biology and Biotechnology”J M.Walker,R.Rapley编,Royal Society of Chemistry, London UK,2000)。进行以下修饰。1)双嗜性env基因3’端的p15区的 核酸序列起初来源于亲嗜性包膜-用来自4070A双嗜性包膜的对应序列取 代此序列;所编码的包膜氨基酸在两个构建体中相同。2)修饰IRES序 列3’端以允许插入有PstI1位点的选定转基因能更容易地插入,并去除IRES转基因位点任一端的微小不完全重复序列。3)去除3’LTR下游的 残余病毒序列。所得质粒是pACE-emdGFP(aka pACE-GFP、pACE-eGFP 和T5.0006),其用作编码胞嘧啶脱氨酶和其变异体的载体的基础。最初 使用两种插入编码盒的方法。第一种方法产生序列5’TTATAAT3’,且第 二种方法紧邻ATG起始密码子上游产生5’TTATAA3’。第二种方法比较 简单,因为其涉及载体中的简单PstI1和Not1酶切割序列和合成胞嘧啶 脱氨酶基因,以及随

后的再连接。利用这两种方法用CDopt(CD1)和 CDopt+3pt(CD2)(参看图2)编码序列制备具有胞嘧啶脱氨酶插入物 的载体,且如实施例3中所述通过瞬时转染293T细胞来制备感染性病毒 制剂。然后在培养物中以0.1的MOI感染U87细胞,且细胞生长直到100% 被感染为止。如实施例6中所述,分析100%被感染细胞的细胞提取物的 胞嘧啶脱氨酶活性,并发现所述酶的比活性对于具有任一种上游序列但在 其它方面相同的构建体是相等的。因此,在图2的表中,pACE-eGFP (T5.0006)和pACE-yCD(T5.0007)具有第一个上游序列,而进一步测 试的所有其它构建体具有第二个上游序列。随后用简单PStI1和Not 1切 割常规构建具有不同基因插入物的载体。

[0208] 关于初始转染的可复制逆转录病毒的病毒构建体的示意图,请参看以 下图3A。CMV是人类CMV立即早期启动子,U3、R和U5是病毒长末 端重复序列(LTR)的对应区域。Gag、pol和env是病毒蛋白编码区。图3B和3D显示质粒结构和本公开书的序列。

[0209] 本公开书的载体与Tai等人,Mol.Ther.12:842,2005的载体相比提供 许多差别。Tai等人的载体已被改变以消除3'LTR下游的MLV序列的约 70bp。包膜中的ClaI位点下游的DNA序列被变为双嗜性包膜序列。这种 改变不会改变包膜的氨基酸序列。此外,已经消除了IRES-CD盒任一侧 的小重复序列以避免因同源重组所致的不稳定。这些改变还出乎意料地使 载体在宿主细胞中复制和传代期间的稳定性提高(图5)。

[0210] 已经认识到,在逆转录和第一次整合到被处理细胞中之后,DNA前 病毒和任何后续子代逆转录病毒在每一端具有来自MLV的常规LTR结 构。已经显示这种构形在多个感染周期后是稳定的(参看以下图5)。

[0211] 实施例2对野生型酵母胞嘧啶脱氨酶基因的遗传增强

[0212] 已经进行了两组改变:(1)3个位置突变,其改变3个氨基酸(A23L、I140L和V108I)以增加酵母胞嘧啶脱氨酶蛋白的热稳定性;和(2)其它 基因序列修饰,用于增强人类密码子使用序列以提高在人类细胞中的蛋白 翻译效率而不会对氨基酸序列造成其它改变。

[0213] CD的序列设计包括CD优化型、CD-UPRT(+/-连接子)和 CD-OPRTase(+/-连接子)。最终胞嘧啶脱氨酶编码序列可在5'端包含 PSI1位点(全长)且在3'端包含Not1位点和用于PSI1/Not1盒策略的poly A尾巴。盒是从商业供应商(BioBasic Inc.,Markham,Ontario,Canada) 定购并由其提供。

[0214] 包含酵母胞嘧啶脱氨酶的以下序列用于克隆、优化和突变(带框核酸 包含限制位点Psi1和Not1,用于随后克隆方法中):

[0215] AACACGA TTATAA ATGGTGACAGGGGGAATGGCAAGCAAGTG GGATCAGAAGGGTATGGACAT TGCCTATGAGGAGGCGGCCTTAGG TTACAAAGAGGGTGGTGTTCCTATTGGCGGATGTCTTATCAATAAC AAAG ACGGAAGTGTTCTCGGTCGTGGTCACAACATGAGATTTCAA AAGGGATCCGCCACACTACATGGTGAGATCTCCA CTTTGGAAAACT GTGGGAGATTAGAGGGCAAAGTGTACAAAGATACCACTTTGTATA CGACGCTGTCTCCATGC GACATGTGTACAGGTGCCATCATCATGTA TGGTATTCCACGCTGTGTTGTCGGTGAGAACGTTAATTTCAAAAGT AAGGGCGAGAAATATTTACAAACTAGAGGTCACGAGGTTGTTGTT GTTGACGATGAGAGGTGTAAAAAGATCATG AAACAATTTATCGAT GAAAGACCTCAGGATTGGTTTGAAGATATTGGTGAGTAG GCGGCC GC GCCATAG ATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAG GGGGG(SEQ ID NO:32)

[0216] 下表总结所制备的基因和所得质粒载体和其名称。

[0217] 表:载体构建体和名称

[0218]

Tocagen代码	参考名称	原始名称	5'LTR Prom	包膜	载体	IRES	转基因	3'LTR
T5.0000	pACE-yCD	pACE-CD (Tai等人2005)	CMV	Ampho (4070A)	pACE	EMCV	野生型酵母CD	MLV U3
T5.0001	pAC3-yCD1	CDopt序列	CMV	Ampho (4070A)	pAC3	EMCV	经修饰CD	MLV U3
T5.0002	pAC3-yCD2	CDopt+3pt	CMV	Ampho (4070A)	pAC3	EMCV	经修饰CD	MLV U3
T5.0003	pAC3-yCD2-U	Cdopt+3pt-UPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-UPRT	MLV U3
T5.0004	pAC3-yCD2-O	CDopt+3pt-OPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-OPRT	MLV U3
T5.0005	pAC3-yCD2-LO	CDopt+3pt-LINK-OPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-L-OPRT	MLV U3
T5.0006	pAC3-eGFP	pAC3-emd, pAC3GFP	CMV	Ampho (4070A)	pAC3	EMCV	Emerald GFP	MLV U3
T5.0007	pAC3-yCD	pAC3-yCD	CMV	Ampho (4070A)	pAC3	EMCV	野生型酵母CD	MLV U3

[0219] Kasahara等人pACE-CD(美国专利第6,899,871号,其公开内容以 引用的方式并入本文中)所述的可复制逆转录病毒载体用作其它修饰的基 础。对载体(pAC3-yCD)进行修饰以表达如本文所述的经修饰酵母胞嘧 啶脱氨酶基因并将所述载体用于构建体中。关于初始转染的可复制逆转录 病毒的病毒构建体的示意图,请参看以下图3A。CMV是人类CMV立即早期启动子,U3、R和U5是病毒长末端重复序列(LTR)的对应区域。Gag、pol和env是病毒蛋

白编码区。图3B和3D显示本公开书的质粒结 构和序列。

[0220] 在由承包商(Bio Basic Inc.,Markham,Ontario,Canada)合成基因后，将其插入pAC3载体骨架的Psi1-Not1位点(图3)。一般通过用Psi1和 Not1切割质粒pAC3-eGFP并纯化编码质粒和逆转录病毒骨架的大(约 11kb)片段来产生质粒骨架。

[0221] A.人源化密码子优化CD基因(CDopt,aka CD1,T5.0001)Conrad 等人与PCT WO 99/60008的人类密码子优化胞嘧啶脱氨酶的比较表明，两种胞嘧啶脱氨酶中总共优化91个密码子,其中36个密码子相同,47 个密码子具有第三碱基对改变(都编码相同氨基酸)，且9个密码子不同 (但其编码相同氨基酸)。不同的9个密码子中：

[0222] AGC(Ser)与TCC(Ser)

[0223] CGT(Arg)与AGG(Arg)

[0224] CCA(Pro)与CCT(Pro)

[0225] 它们都具有相等的GC含量且编码相同氨基酸。上述天然酵母基因序 列单独进行密码子优化以得到以下CD基因(CD1)，且在插入编码具有 IRES的可复制逆转录病毒(RCR)的质粒载体pAC3中时被称作T5.0001。

[0226]

TTATAAATGGTGACCGGCGGCATGGCCTCCAAGTGGGATC
AAAAGGGCATGGATATCGCTTACGAGGAGGCCGCCCTGGGCTA
CAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAAC
AAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTC
CAGAAGGGCTCCGCCACCCTGCACGGCGAGATCTCCACCCTGG
AGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCA
CCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGC
TATCATCATGTACGGCATCCCTAGGTGTGTGGTGGGCGAGAAC
GTGAACTTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGG
GCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGAT
CATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAG
GATATCGGCGAGTGATAAGCGGCCGCAGATAAAATAAAAGATTT
TATTTAGTCTCCAGAAAAAGGGGGG.

[0227] B.热稳定化CD基因进行其它修饰以增强胞嘧啶脱氨酶的稳定性。对野生型酵母胞嘧啶脱氨酶基因进行遗传增强以包含3个位置突变,其改 变3个氨基酸(A23L、I140L和V108I)以提高酵母胞嘧啶脱氨酶蛋白的 热稳定性。

[0228] 使用以下引物对产生本公开书的胞嘧啶脱氨酶多肽的基因：

[0229] 定点诱变引物:有义引物:5′-tcgaggatatcggcgagtgaaacccgttattctttttggc-3′(SEQ ID NO:33)

[0230] 反义引物:5′-gccaaaaagaataacgggtttcactcgccgatatcctcga-3′(SEQ ID NO:34)

[0231] 有义引物:

[0232] 5′tcggcgagtgatccggcggcggcgcctccggcggcggcgcctccggcggcggcgcctccggcggcggcgccaacccgttatt-3′(SEQ ID NO:35)

[0233] 反义引物:

[0234] 5′-aataacgggttggcgccgccgccggaggcgccgccgccggaggcgccgccgccggaggcgccgccgccggatcactcgccga-3′(SEQ ID NO:36)

[0235] 为了提高天然酵母CD蛋白的稳定性,将3个氨基酸取代工程改造到 蛋白中。这些取代是单独使用或与人类密码子优化组合使用。

[0236] 这3个氨基酸取代是:A23L、V108I、I140L。以下显示编码这些取 代的序列。

[0237] ATGGTGACAGGGGGAATGGCAAGCAAGTGGGATCAGAAGGGTATG GACATTGCCTATGAGGAGGCGTTATTAGGTTACAAAGAGGGTGGT GTTCCTATTGGCGGATGTCTTATCAATAACAAAGACGGAAGTGTTC TCGGTCGTGGTCACAACATGAGATTTCAAAAGGGATCCGCCACACT ACATGGTGAGATCTCCACTTTGGAAAACTGTGGGAGATTAGAGGG CAAAGTGTACAAAGATACCACTTTGTATACGACGCTGTCTCCATGC GACATGTGTACAGGTGCCATCATCATGTATGGTATTCCACGCTGTG TCATCGGTGAGAACGTTAATTTCAAAAGTAAGGGCGAGAAATATTTACAAACTAGAGGTCACGAGGTTGTTGTTGTTGACGATGAGAGGTG TAAAAAGTTAATGAAACAATTTATCGATGAAAGACCTCAGGATTG GTTTGAAGATATTGGTGAGTAG GCGGCCGC GCCATAGATAAAATA AAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG(SEQ ID NO:3)

[0238] 所编码的多肽包含以下序列(被取代氨基酸以下划线表示):

[0239] 1 MVTGGMASKWDQKGMDIAYEEALLGYKEGGVPIGGCLINNKDGSVLGRGHNMRFQKGSAT

[0240] 61 LHGEISTLENCGRLEGKVYKDTTLYTTLSPCDMCTGAIIMYGIPRCVIGENVNFKSKGEK

[0241] 121 YLQTRGHEVVVVDDERCKKLMKQFIDERPQDWFEDIGE-(SEQ ID NO:4)

[0242] 当插入到编码具有IRES的RCR的质粒载体pAC3中时,利用优选 密码子整合3个氨基酸取代A23L/V108I/I140L并对完整序列(所述基因 称为CDopt+3pt[aka CD2](SEQ ID NO:37)和T5.0002)使用优选人类密 码子最终构建体设计:

[0243]
```
  1 ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTTACGAG
 61 GAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAAC
121 AAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACC
181 CTGCACGGCGAGATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAG
241 GACACCACCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATG
301 TACGGCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCCAAGGGCGAGAAG
361 TACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTG
421 ATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTGATAA
```

[0244] 加下划线的密码子表示氨基酸取代的优选密码子。

[0245] 蛋白翻译序列比对表明优选密码子改变且氨基酸取代产生期望蛋白 结构:

[0246] CD优化序列设计(人类密码子优选+3个氨基酸取代)

[0247] AACACGA TTATAA ATGGTGACCGGCGGCATGGCCTCCAAGTGGGA TCAAAAGGGCATGGATATCGCTTACGAGGAGGCCCTGCTGGGCTA CAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAA GG

ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAA GGGCTCCGCCACCCTGCACGGCGAGATCTCCA CCCTGGAGAACTGT GGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACC ACCCTGTCCCCTTGT GACATGTGTACCGGCGCTATCATCATGTACG GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCC AA GGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT TGACGATGAGAGGTGTAAGAAGCTGAT GAAGCAGTTCATCGACGA GAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTAA GCGGCCGC GCCATA GATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGG GGG(SEQ ID NO:38)

[0248] C.构建CD-UPRT融合基因(CDopt+3pt-UPRT,[aka CDopt-UPRT和CD2-UPRT],在pAC3质粒RCR载体中称作T5.0003)也使用如图 10A所述的方案I形成融合构建体,其包含与UPRT多肽连接的上述CD 多肽以产生优化CD-UPRT序列。使用以下引物缺失CD与UPRT之间的终止-起始序列。

[0249] 引物序列:

[0250]

引物名称	引物序列（5'到3'）			
del118-123	5'-tcgaggatatcggcgagtgaaacccgttattctttttggc-3'			
del118-123-反义	5'-gccaaaaagaataacgggtttcactcgccgatatcctcga-3'			
引物名称	长度 (nt.)	Tm	68℃下的双螺旋能量	错配的能量消耗
del118-123	40	79.06℃	-44.37 kcal/mole	21.1%
del118-123-反义	40	79.06℃	-47.95 kcal/mole	20.3%

[0251]

引物名称	引物-模板双螺旋
del118-123	5'-tcgaggatatcggcgagtga------aacccgttattctttttggc-3' |||||||||||||||||||| |||||||||||||||||||| ccaagctcctatagccgctcactatctacttgggcaataagaaaaaccgaag
del118-123-反义	ggttcgaggatatcggcgagtgatagatgaacccgttattctttttggcttc |||||||||||||||||||| |||||||||||||||||||| 3'-agctcctatagccgctcact------ttgggcaataagaaaaaccg

[0252] 所得融合多聚核苷酸包含1296bp和下文所述的序列:

[0253] AACACGA TTATAA ATGGTGACCGGCGGCATGGCCTCCAAGTGGGA TCAAAAGGGCATGGATAT

CGCTTACGAGGAGGCCCTGCTGGGCTA CAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAA GG ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAA GGGCTCCGCCACCCTGCACGGCGAGATCTCCA CCCTGGAGAACTGT GGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACC ACCCTGTCCCCTTGT GACATGTGTACCGGCGCTATCATCATGTACG GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCC AA GGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT TGACGATGAGAGGTGTAAGAAGCTGAT GAAGCAGTTCATCGACGA GAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGAACCCGTTATTC TTTTTGGCTT CTCCATTCTTGTACCTTACATATCTTATATATTATCCA AACAAAGGGTCTTTCGTTAGCAAACCTAGAAATCTGC AAAAAATG TCTTCGGAACCATTTAAGAACGTCTACTTGCTACCTCAAACAAACC AATTGCTGGGTTTGTACACC ATCATCAGAAATAAGAATACAACTAG ACCTGATTTCATTTTCTACTCCGATAGAATCATCAGATTGTTGGTTG A AGAAGGTTTGAACCATCTACCTGTGCAAAAGCAAATTGTGGAAA CTGACACCAACGAAAACTTCGAAGGTGTCTC ATTCATGGGTAAAAT CTGTGGTGTTTCCATTGTCAGAGCTGGTGAATCGATGGAGCAAGGA TTAAGAGACTGTT GTAGGTCTGTGCGTATCGGTAAAATTTTAATTC AAAGGGACGAGGAGACTGCTTTACCAAAGTTATTCTACGAAA AAT TACCAGAGGATATATCTGAAAGGTATGTCTTCCTATTAGACCCAAT GCTGGCCACCGGTGGTAGTGCTATC ATGGCTACAGAAGTCTTGATT AAGAGAGGTGTTAAGCCAGAGAGAATTTACTTCTTAAACCTAATCT GTAGTAA GGAAGGGATTGAAAAATACCATGCCGCCTTCCCAGAGG TCAGAATTGTTACTGGTGCCCTCGACAGAGGTCTAGA TGAAAACAA GTATCTAGTTCCAGGGTTGGGTGACTTTGGTGACAGATACTACTGT GTTTAA**GCGGCCGC**

GCCATAGATAAAATAAAAGATTTTATTTAGTC TCCAGAAAAAGGGGGG

[0254] D.构建CD-连接子UPRT融合基因(CDopt+3pt-LINK-UPRT[aka CDopt-LINKER-UPRT和CD2-L-UPRT])也通过利用如图10B所述的方 案II克隆CD多肽与UPRT多肽之间和同框的连接子 $(Ser\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Gly)_4$域以产生优化CD-连接子-UPRT序列来形成融合 构建体。使用以下引物来插入连接子。

[0255]

引物名称	引物序列（5'到3'）
ins_60nt_after_477	5'-tcggcgagtgatccggcggcggcgcctccggcggcggcgcctccggcggcggcgcctccggcggcggcgccaacccgttatt-3'
ins_60nt_after_477-反义	5'-aataacgggttggcgccgccgccggaggcgccgccgccggaggcgccgccgccggaggcgccgccgccggatcactcgccga-3'

引物名称	长度（nt.）	Tm	68℃下的双螺旋能量	错配的能量消耗
ins_60nt_after_477	82	79.77℃	-30.19 kcal/mole	83.3%
ins_60nt_after_477-反义	82	79.77℃	-32.31 kcal/mole	82.2%

[0256] 所得构建体具有1356bp的尺寸和以下序列：

[0257] AACACGA**TTATAA**ATGGTGACCGGCGGCATGGCCTCCAAGTGGGA TCAAAAGGGCATGGATAT CGCTTACGAGGAGGCCCTGCTGGGCTA CAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAA GG ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAA GGGCTCCGCCACCCTGCACGGCGAGATCTCCA CCCTGGAGAACTGT GGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACC ACCCTGTCCCCTTGT GACATGTGTACCGGCGCTATCATCATGTACG GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCC AA GGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT TGACGATGAGAGGTGTAAGAAGCTGAT GAAGCAGTTCATCGACGA GAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTCCGGCGGCGGCGCCTCCGGCGG CGGCGCCTCCGGCGGCGGCGCCTCCGGCGGCGGCGCCAACCCGTTATTCTTTTTGGCTTCTCCATTCTTGTACCTT ACAT ATCTTATATATTATCCAAACAAAGGGTCTTTCGTTAGCAAACCTAG AAATCTGCAAAAAATGTCTTCGGA ACCATTTAAGAACGTCTACTTG CTACCTCAAACAAACCAATTGCTGGGTTTGTACACCATCATCAGAA ATAAGA ATACAACTAGACCTGATTTCATTTTCTACTCCGATAGAAT CATCAGATTGTTGGTTGAAGAAGGTTTGAACCATC TACCTGTGCAA AAGCAAATTGTGGAAACTGACACCAACGAAAACTTCGAAGGTGTC TCATTCATGGGTAAAATC TGTGGTGTTTCCATTGTCAGAGCTGGTG AATCGATGGAGCAAGGATTAAGAGACTGTTGTAGGTCTGTGCGTAT CGGTAAAATTTTAATTCAAAGGGACGAGGAGACTGCTTTACCAAA GTTATTCTACGAAAAATTACCAGAGGATAT ATCTGAAAGGTATGTC TTCCTATTAGACCCAATGCTGGCCACCGGTGGTAGTGCTATCATGG CTACAGAAGTCT TGATTAAGAGAGGTGTTAAGCCAGAGAGAATTT ACTTCTTAAACCTAATCTGTAGTAAGGAAGGGATTGAAAAAT ACCA TGCCGCCTTCCCAGAGGTCAGAATTGTTACTGGTGCCCTCGACAGA GGTCTAGATGAAAACAAGTATCTA

GTTCCAGGGTTGGGTGACTTTG GTGACAGATACTACTGTGTTTAA GCGGCCGCGCCATAGATAAAATAA AAGATTTTATTTAGTCTCCAGA AAAAGGGGGG

[0258] E.构建CD-OPRT融合基因(CDopt+3pt-OPRT[aka CDopt-OPRT和CD2-OPRT],在插入pAC3质粒RCR载体中时称为T5.0004)也使用如 图10C所示的方案III形成融合构建体,其包含与OPRT多肽连接上述 CD多肽以产生优化CD-OPRTase(CD人源化+3pt突变+OPRTase人类功能域)。

[0259] 所得构建体包含1269bp的尺寸和以下序列:

[0260] AACACGATTATAAATGGTGACCGGCGGCATGGCCTCCAAGTGGGA TCAAAAGGGCATGGATAT CGCTTACGAGGAGGCCCTGCTGGGCTA CAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAA GG ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAA GGGCTCCGCCACCCTGCACGGCGAGATCTCCA CCCTGGAGAACTGT GGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACC ACCCTGTCCCCTTGT GACATGTGTACCGGCGCTATCATCATGTACG GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCC AA GGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT TGACGATGAGAGGTGTAAGAAGCTGAT GAAGCAGTTCATCGACGA GAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGGCGGTCGCTC GTGcagctttgg ggccattggtgacgggtctgtacgacgtgcaggctttcaagtttggggacttcgtgctgaa gagcgggctttcct cccccatctacatcgatctgcggggcatcgtgtctcgaccgcgtcttctgagtcaggttgc agatattttattcc aaactgcccaaaatgcaggcatcagttttgacaccgtgtgtggagtgccttatacagctttgc cattggctacag ttatctgttcaaccaatcaaattccaatgcttattagaaggaaagaaacaaaggattatggaact aagcgtcttg tagaaggaactattaatccaggagaaacctgtttaatcattgaagatgttgtcaccagtggatcta gtgttttgg aaactgttgaggttcttcagaaggagggcttgaaggtcactgatgccatagtgctgttggacagag agcagggag gcaaggacaagttgcaggcgcacgggatccgcctccactcagtgtgtacattgtccaaaatgct ggagattctcg agcagcagaaaaaagttgatgctgagacagttgggagagtgaagaggtttattcaggagaat gtctttgtggcag cgaatcataatggttctccccctttctataaaggaagcacccaaagaactcaGCTTCGG TGCACGTGCAGAGCTGC CCAGGATCCACCCAGTTGCATCGAAGTA AGCGGCCGCGCCATAGATAAAATAAAAGATTTTATTTAGTCT CCAG AAAAAGGGGGG

[0261] F.构建CD-连接子-OPRT融合基因(CDopt+3pt-LINK-OPRT,[aka CDopt-LINKER-OPRT和CD2-L-OPRT],在pAC3质粒RCR载体中T5.0005)也通过利用如图10D所述的方案IV克隆CD多肽与OPRT多 肽之间和同框的连接子(Ser-Gly-Gly-Gly-Gly)4域以产生优化CD-连接子 -OPRT序列来形成融合构建体。

[0262] 所得构建体包含1329bp的尺寸和以下序列:

[0263] AACACGATTATAAATGGTGACCGGCGGCATGGCCTCCAAGTGGGA TCAAAAGGGCATGGATAT CGCTTACGAGGAGGCCCTGCTGGGCTA CAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAA GG ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAA GGGCTCCGCCACCCTGCACGGCGAGATCTCCA CCCTGGAGAACTGT GGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACC ACCCTGTCCCCTTGT GACATGTGTACCGGCGCTATCATCATGTACG GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCC AA GGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT TGACGATGAGAGGTGTAAGAAGCTGAT GAAGCAGTTCATCGACGA GAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTCCGGCGGCGGCGCCTCCGGCGG

CGGCGCCTCCGGCGGCGGCGCCTCCGGCGGCGG CGCCGCGGTCGCTCGTGcagctttggggccattggtgacggg tctgtacgacgtgcaggctttc aagtttggggacttcgtgctgaagagcgggctttcctcccccatctacatcga tctgcggggcatcgtgtctcga ccgcgtcttctgagtcaggttgcagatattttattccaaactgcccaaaatgc aggcatcagttttgacaccgtgtg tggagtgccttatacagctttgccattggctacagttatctgttcaaccaa tcaaattccaatgcttattagaaggaa agaaacaaaggattatggaactaagcgtcttgtagaaggaactattaa tccaggagaaacctgtttaatcattgaa gatgttgtcaccagtggatctagtgttttggaaactgttgaggttct tcagaaggagggcttgaaggtcactgatg ccatagtgctgttggacagagagcagggaggcaaggacaagttgca ggcgcacgggatccgcctccactca gtgtgtacattgtccaaaatgctggagattctcgagcagcagaaaaaagt tgatgctgagacagttgggagagt gaagaggtttattcaggagaatgtctttgtggcagcgaatcataatggttc tccccttttctataaaggaagcaccc aaagaactcaGCTTCGGTGCACGTGCAGAGCTGCCCAGGATCCACCCAG TTGCATCGAAGTAA GCGGCCGC GCCATAGATAAAATAAAAGATTT TATTTAGTCTCCAGAAAAAGGGGGG

[0264] 图4说明与被感染U-87细胞中的野生型yCD蛋白相比,具有较高稳 定性的yCD蛋白中经3个突变优化的人类密码子的水平较高。

[0265] 实施例3通过瞬时转染产生载体

[0266] 载体可以许多方式产生,但第一步是将DNA载体引入细胞中以便产 生感染粒子,然后从细胞上清液收集所述感染粒子。一旦产生了感染粒子,所属领域技术人员就可实施其它产生方法。通过瞬时转染293T细胞来产 生载体粒子(Pear等人Proc Natl Acad Sci U S A.90:8392-83961993)。

[0267] 将293T细胞解冻并放入培养物中,然后在含有15mL DMEM培养 基的T-75烧瓶中传代两次,所述DMEM培养基是通过混合DMEM高葡 萄糖培养基(Hyclone#30081,500mL)与FBS(Hyclone#SH30070,50 mL)、L-谷氨酰胺(Cellgro#25-005-CI,5mL)、NEAA(Hyclone#SH30238, 5mL)和青霉素-链霉素(Cellgro#30-002-CI,5mL)来制备。烧瓶在37℃ 和5%CO_2下培养。在第3次传代后,将细胞接种于6个T-25中,每个 T-25含有5mL培养基,细胞接种密度为1.8×10^6个细胞/T-25(或7.2×10^4个细胞/cm^2)。在接种T-25后一天,使用购自Promega的磷酸钙转染试 剂盒(目录号E1200)用表达病毒载体的T5.0002质粒转染细胞。在转染 后18小时,用含有10mM NaB的新鲜培养基替换一组烧瓶(每组3个烧 瓶)中的培养基。不替换第2组烧瓶中的培养基,其用作对照(无NaB)。 在NaB处理后8小时,用不含NaB的新鲜培养基替换所有烧瓶中的培养 基。继续两组烧瓶中的表达直到下一天(22小时持续时间)为止。收集 两组烧瓶的上清液,并通过qPCR分析其滴度,表示为转导单位(TU)/ 毫升(参看实施例4)。

[0268] 下表显示滴度结果。

[0269]

条件	第一滴度	第二滴度(在-80℃下储存68天后)

[0270]

无 NaB	1.5 (±0.05) × 10^6 TU/mL	1.2 (±0.2) × 10^6 TU/mL
10 mM NaB	1.4 (±0.3) × 10^6 TU/mL	7.0 (±0.14) × 10^5 TU/mL

[0271] TU＝转导单位

[0272] 以此方式，但不使用丁酸钠，制备后续载体制剂。已经用相同方式使 用其它载体质粒(表2)来产生滴度在1E5TU/ml与1E7TU/ml之间的载 体制剂。必要时，所述物质可如下文所述被进一步纯化和浓缩。也参看 US 5792643;T.Rodriguez等人J Gene Med 9:2332007;美国申请 61218063。

[0273] 在本公开书的某些实施方案中，依据脑重量的克数计算剂量。在所述 实施方案中，可用于治疗方法中的本公开书的可复制逆转录病毒载体的剂 量可在每克脑重量10^4到10^6TU的范围内。

[0274] 实施例4定量PCR滴定分析

[0275] 使用基于定量PCR(qPCR)的方法测定功能载体浓度或滴度。在这 种方法中，通过用标准体积的载体感染可转导宿主细胞系(例如PC-3人 类前列腺癌细胞，ATCC Cat#CRL-1435)并测量在转导后于宿主细胞内 存在的前病毒的所得量来滴定载体。细胞和载体在标准培养条件(37℃， 5%CO_2)下培养24小时以允许完全感染，随后添加抗逆转录病毒AZT 以终止载体复制。接着从培养皿收集细胞，使用Invitrogen Purelink gDNA 纯化试剂盒来纯化基因组DNA(gDNA)，并用无RNase-/DNase的无菌 水将其从纯化柱洗脱。在分光光度计上测量A_{260}/A_{280}吸光率比以确定样 品的浓度和相对纯度。用另外的无RNase-/DNase水将gDNA浓度标准化 为任何指定组的gDNA制剂的最低浓度，使得qPCR的输入DNA对于分 析的所有样品是恒定的。通过在经溴化乙锭染色的0.8%琼脂糖凝胶上电 泳每个样品的等分试样来进一步分析基因组DNA纯度。如果样品具有 1.8-2.0的A_{260}/A_{280}吸光率范围且显示单一gDNA带，那么样品可用于 qPCR分析载体的前病毒拷贝数。使用查询前病毒LTR区(逆转录的载体DNA和整合到宿主gDNA中的载体DNA)的引物，进行qPCR以估 算当使用已知体积的载体转导已知数目的细胞时所发生的转导事件的总 数。从标准曲线计算每个反应的转导事件数目，所述标准曲线利用从10^7个拷贝连续稀释成10个拷贝并在相同qPCR条件下作为样品来测量的已 知拷贝数的靶标携带质粒。通过知道每个qPCR反应使用多少基因组等效 物(根据先前测定的浓度)以及每个反应中发生多少转导事件，我们可基 于在转导时存在的细胞总数来确定所发生的转导事件总数。这个值是载体 在初始转导期间稀释到含有细胞的培养基中之后的滴度。为了计算经校正 的滴度值，稀释度通过乘以培养物体积并除以滴定体积来校正。这些实验 在重复培养物中进行，并通过qPCR，对于每个条件使用三次测量来进行 分析，以测定平均滴度和与其有关的标准偏差和变异系数。

[0276] 实施例5载体测试

[0277] 为了变得有效，载体构建体和由其衍生的感染粒子需要:1)通过瞬 时转染而获得良好的病毒滴度(参看实施例3和4);2)在多次传代后 稳定;3)在5-FC的存在下有效地杀死细胞;和4)在感染靶细胞后表达 酶活性。实施例3显示可从载体获得有用的滴度。

[0278] 病毒载体的遗传稳定性为了证明稳定性，进行以下实验。起初利用 病毒载体以

0.01的MOI感染约10^6个天然U-87细胞，且生长到被完全感 染以完成单一感染周期。然后将上清液再转移到未被感染的细胞上并重复 所述周期。在本实验中，已经在双重复试验中完成了12个周期(图5显 示双重复试验中的一个，另一重复得到类似结果)。使用侧接转基因插入 位点的MLV特异引物，通过从被感染细胞PCR扩增被整合的前病毒来 分析yCD2或其它转基因序列的基因组稳定性。任何小于全长扩增子的带 的出现将指示载体不稳定性。图5证明本公开书的载体(T5.0007，包含 被修饰载体和CD异源多聚核苷酸)与pACE-CD相比，在传代更多的情 况下仍保持稳定性。此外，T5.0003有点不太稳定，而T5.0004和T5差 不多与pACE-CD同样稳定。pACE-CD已经用于小鼠肿瘤研究中并在小 鼠模型中显示良好的抗肿瘤作用。然而，更稳定的病毒基因组在动物和人 类治疗中将更为有效和更持久，特别是在需要多个5-FC治疗周期的时候。T5.0001和T5.0002甚至比T5.0007都要稳定得多，显示蛋白编码序列中 的沉默改变或导致点突变的微小改变可产生意想不到的优越特性以及更 稳定的载体基因组。

[0279] 细胞杀死实验细胞滴度96水性单溶液细胞增殖分析(MTS)是用 于测定增殖分析中的活细胞数目的比色方法。我们已经利用了这种分析来 测定细胞生长动力学，以及测定各种细胞系对5-氟胞嘧啶(5-FC)和5- 氟尿嘧啶(5-FU)的剂量反应。

[0280] 将经载体100％感染的细胞以1000个细胞/孔接种于96孔板中。在用 各种浓度的5-FC处理(5-FU用作对照)后，历经8天监测所述细胞。利 用细胞滴度96水性单溶液试剂(MTS)每两天分析其细胞生长。简言之， 将20μl MTS与100μl培养基(根据制造商推荐)混合并添加到96孔板 的样品中。样品在37℃/5％CO_2培养箱中培养60分钟。然后在平板读取 器上以490nm波长读取吸光率读数。然后丢弃平板。

[0281] 图6A显示实验结果，所述实验通过对经AC3-yCD2(载体)或 AC3-yCD(载体)感染的U-87细胞进行5-FC滴定，证明表达yCD2蛋 白的细胞中的胞嘧啶脱氨酶至少与在表达野生型yCD蛋白的细胞中活性 相同。简言之，在用AC3-yCD(野生型CD)载体或AC3-yCD2(热稳定化和密码子优化)载体以0.1的感染复数(即100％被感染)感染5天后， U-87细胞接触增加量的5-FC或作为阳性对照的0.1mM 5-FU，维持8天。 在第8天，使用MTS分析(Promega细胞滴度96水性单溶液增殖分析) 评定细胞培养物的活力。数据显示在增加剂量的5-FC处理下，两种逆转 录病毒载体的杀死作用相当。

[0282] 在使用RG2细胞(ATCC Cat#CRL-2433)的类似体外细胞培养实验 中，用5种不同载体(pACE-CD、T5.0001、T5.0002、T5.0004和T5.0007) 转导RG2细胞系。随后使其接触增加浓度的5-FC(5-FU用作对照)持 续8天，并如上所述进行监测。结果展示于图2中。0.01mM的浓度足以 在测试的所有载体中诱导完全杀死，但野生型胞嘧啶脱氨酶(pACE-yCD) 除外。其不太敏感且需要1.0mM的5-FC才能实现完全杀死。

[0283] CD表达分析以0.1的感染复数(MOI)转导U87细胞，培养5天 以允许病毒传播，并收集转导5天后的细胞。然后通过在800x g下离心 5分钟来收集细胞。从细胞团块中吸出上清液，用5mL磷酸盐缓冲盐水 (PBS)洗涤，并在800x g下再次离心5分钟。将所得细胞团块溶解于 1.5mL PBS中，通过吸管尖来再悬浮，并置于-20℃的冷冻器中。用冷冻/ 解冻方法溶解细胞。使先前再悬浮的细胞在室温下解冻，通过吸管尖，与 蛋白酶抑制剂混合液混合并在-20℃下再次冷冻。在酶分析之前，将样品 在室温下再次解冻并在吸管尖中通过。悬浮液然后在台式离心机中以 14,000rpm离心5分钟。从团块中轻轻倒出上清液，并置于新鲜

eppendorf 管中且置于冰上。

[0284] 使用HPLC分析法来评定yCD酶活性。HPLC分析是在与光阵列检 测器和自动进样器串联连接的Shimadzu LC20AT单元上进行。固相是 Hypersil BDS C_{18}HPLC柱,其具有5μm球体大小和4.0×250mm柱尺寸。 流动相是50mM磷酸铵(pH 2.1),其含有0.01%四丁基高氯酸铵和5% 甲醇;所述系统在22℃下平衡。所有试剂是ACS级且溶剂是HPLC级。 制备PBS中的最终浓度为0.125mg/mL 5FC(1mM)的800μL反应混合 物,并置于1.5mL自动进样器小瓶中。然后通过添加200μL每种细胞裂 解液来起始反应。将反应管/自动进样器管置于自动进样器中,并注入5μL 反应混合物。通过从每个反应管获取5μL等分试样并在HPLC柱上分析 来周期性地选取时间点。通过比较峰面积与来自先前产生的5FU标准曲 线的已知量来计算5FC至5FU的转化速率。通过对所产生的5FU量(以 nmol计)与其对应时间间隔作图来推导5FC至5FU的转化速率。推导细 胞样品的蛋白浓度,并通过用转化速率(nmol/min)除以分析中的蛋白用 量(mg)来计算细胞裂解液样品的比活性。图6B显示以0.1的MOI转 导5天后各种载体的比活性。数据证明在组织培养细胞中的胞嘧啶脱氨酶 的比活性方面,pACE-yCD(T5.0000)＜pAC3-yCD1(T5.0001)＜pAC3-CD2 (T5.0002)。

[0285] 实施例6载体纯化和浓缩

[0286] 通过在293T细胞上瞬时转染(实施例3),随后收集细胞上清液,过 滤,用benzonase处理,透滤/浓缩和透析来制造本公开书的载体。可包含所 属领域技术人员已知的另一个色谱柱步骤(参看例如US5792643;T. Rodriguez等人J Gene Med 9:2332007;美国申请61218063)。已基于标准 测试(例如无菌性、支原体和内毒素、plus产物特定效力、强度和同一性 测试)公开临床材料。通过靶细胞中的整合病毒载体DNA的PCR定量来测 定滴度,以转导单位(TU)表示(实施例4)。选定滴度高达10^9TU/ml的 最终产物,其是在等张Tris缓冲蔗糖溶液中配制成无菌注射液。

[0287] 一般来说,为了准确且精确地测定载体批次的强度,已经开发了基于 定量PCR的滴度分析(如实施例4中概括地描述)。所述分析程序的细 节由以下步骤组成:

[0288] 转导.使用来源于PC-3细胞系的稳定人类前列腺腺癌,以12孔板型式 进行转导。对于每个测试样品,使用三个未滴定的载体制剂的稀释液在一 式三份的孔中转导PC-3细胞。在转导后24小时用叠氮胸苷(azidothymidine; AZT)终止病毒复制。在收集和基因组DNA纯化前另外维持细胞24-64小 时。

[0289] 基因组DNA纯化.根据制造商方案,使用Qiagen DNeasy DNA Minikits 从收集的被转导细胞制备基因组DNA。通过直接吸光度测量,使用UV/vis 分光光度测定法以测定A260和A260/A280比率,从而评定DNA浓度和质量。

[0290] 实时定量PCR.使用BioRad CFX96实时PCR仪器或等效物进行定量 PCR。通过使用特异性DNA寡聚核苷酸引物结合TaqMan探针(其经设计用 于相对于CMV/Psi质粒启动子扩增被整合的或前逆转录病毒U3/Psi包装), 测量每个测试样品中所存在的前病毒拷贝数。相对于拷贝数标准曲线来表 示载体滴度。为了产生载体拷贝数标准曲线,在来自PC-3细胞的基因组 DNA中外加含有前逆转录病毒U3/Psi序列的唯一质粒。通过计算多个稀释 液(每个稀释液使用多个反应)中被转导基因组的数目来获得载体测试样 品滴度。

[0291] 对于每个滴度评定,一式三份地执行非模板对照(含有除质粒或基因 组DNA之外的所有组分的孔)和阴性对照(包含来自非转导PC-3细胞的等 效基因组DNA的所有组分)。

滴度值以转导单位/毫升(TU/mL)表示。

[0292] 本公开书的载体的效能依赖于载体复制和在靶细胞中的所得胞嘧啶 脱氨酶(CD)活性。因此,效能分析测量到随着载体感染在组织培养中的 先前未感染的细胞系中传播,CD活性增加。所述分析通过使用反相HPLC 分离以及UV检测来监控5-氟胞嘧啶(5-FC)至5-氟尿嘧啶(5-FU)的转化,从而测量在感染早期、中期和晚期所转移的yCD2蛋白在转导细胞中的酶活 性。在感染过程中的CD活性增加是随着时间被感染的细胞百分比的函数,且表示TOCA 511载体复制的能力。在每个时间点测量基于5-FC至5-FU转 化率的CD活性,以CD单位/毫克蛋白表示(比活性,SA)。然后随时间对 SA增加作图,SA增加反映因为培养物中的病毒复制所导致的被转导细胞 的百分比的增加,和所得CD活性转移。已经使用来自多个分析和载体批次 的积累数据来确定CD的所述SA增加对于产物释放的适当规格。所述分析 具有在以约0.1的MOI初始感染和未感染对照之后的1、3和5天时间点。

[0293] 比较各批次之间的来自晚期感染细胞(5天时间点)的CD活性以评估 所述活性用作同一性测试的用途。所述分析包含以下步骤:

[0294] 转导.对于U87细胞以多孔板型式进行转导。对于每个转导,使用三个 合适的稀释液且每个转导是一式三份地进行。在转导后0、1、3和5天收集 细胞。

[0295] 建立CD反应.裂解细胞并使用以BSA作为标准的BCA蛋白分析来测 定总蛋白浓度。对于yCD2酶分析,将适量的细胞裂解液添加到含有5-FC 的缓冲液中,使得F-FU形成速率在37℃下载1-2小时内保持线性。反应混 合物的最终体积是100μL。2h后,通过添加三氯乙酸来终止酶反应,短暂 涡旋并用于随后HPLC分析。将来自非转导细胞的细胞裂解液用作阴性对 照,而使用100%感染细胞的样品的类似分析则用作阳性对照。

[0296] HPLC分析.将被终止的反应混合物在室温下在微离心管中以12,000 rpm离心5分钟。然后从团块中轻轻倒出上清液,并使其通过0.2μ滤器以进 一步去除颗粒物,随后注射到先前用含有磷酸盐缓冲液(pH约为4.0)的水 性流动相平衡的反相HPLC柱上。在260nm和280nm下跟踪色谱图以监控底 物消耗和产物形成。使用GraphPad的制图和分析能力,通过将底物或产物 的浓度与从已知底物或产物浓度计算的先前产生的标准曲线比较来确定底物或产物的浓度。使用经1-2h产生的5-FU的量来确定CD活性单位(1个 CD活性单位是定义为每分钟形成1nmol的5-FU),且将此数字除以分析中 的蛋白用量(来自细胞裂解液)以计算比活性。

[0297] 实施例7构建和使用编码针对CTLA-4(CD 152)的单链抗体的载体。

[0298] 单链抗体来源于具有期望作用的已知完整抗体序列。可获得所述序列 (例如WO2006048749、US2006165706、US7034121、Genbank登记号 DJ437648、CS441506、CS441500、CS441494、CS441488,其公开内容 以引用的方式并入本文中)。通过所属领域技术人员已知的常用方法将所 述常规抗体基因序列转化为单链抗体(scFv)序列(参看例如Gilliland等人“Rapid and reliable cloning of antibody variable regions and generation of recombinant single chain antibody fragments.”Tissue Antigens 47,1-20,1996)。也可从筛选噬菌体scFv文库直接获得针对CTLA-4的噬菌体单 链抗体,并可将其直接用于插入到本发明的复制逆转录病毒载体中 (Pistillo等人Tissue Antigens 55:2292000)。无论序列的来源如何,scFv 典型地是约700-900bp长且如先前所述由商业供应商(BioBasic)合成, 其在5’端具有Psi1位点且在3’端具有相容的Not1位点。所述序列然后

在 去除yCD2序列后插入pAC3-yCD2的复制逆转录病毒骨架的Psi1-Not1 位点。如上所述产生和滴定载体,且在必要时如上所述进行进一步纯化。人类和小鼠CTLA4的序列非常同源,且首先在所属领域技术人员众所周 知的合适的同系免疫活性小鼠模型(例如CT26/Balb/c模型和s91小鼠黑 素瘤模型)中测试本发明的复制逆转录病毒(参看例如Hodge等人J.Immunol.174:5994200)。结果是由一种或一种以上下列因素衡量:肿瘤 生长的调节;细胞毒性的缺乏;抗肿瘤反应的产生;表明系统免疫性的远 端病变收缩。小鼠中的剂量在10^3到10^7TU的范围内。在患者中,载体 是通过病变部位内注射入肿瘤,或通过施用于循环中且然后由循环将病毒 带入肿瘤中来施用。剂量在10^5到10^{11}TU的范围内。

[0299] 实施例8小鼠黑素瘤模型中的使用抗CD152单链抗体表达载体的抗黑素瘤功效研究

[0300] 目标

[0301] 此项研究的目标在于评定当通过在具有皮下黑素瘤(Cloudman S91) 的DBA/2小鼠中肿瘤内(IT)注射来递送时,携带基于针对细胞毒性T 淋巴细胞抗原4(CTLA-4,也称为分化群(CD152)序列(pAC3-αCD152)) 的单链抗体的新颖MLV型可复制逆转录病毒载体对黑素瘤生长的影响。

[0302] 小鼠

[0303] 雌性DBA/2小鼠(～8周大)是从Jackson Laboratories(Bar Harbor,ME) 购得。驯化小鼠7天,随后开始研究。

[0304] 细胞

[0305] Cloudman S91细胞(ATCC,Manassas VA)是从DBA/2小鼠获得的 自发出现的黑素瘤。将细胞在含有10%胎牛血清、丙酮酸钠和Glutamax (Hyclone,Logan UT和Invitrogen,San Diego CA)的杜贝卡改良伊格氏 培养基(Dulbecco's modified Eagles medium)中培养。将细胞再悬浮于 PBS(Hyclone,Logan UT)中以用于植入。将S91细胞(1E6,100μL) 注入DBA/2小鼠右肋。

[0306] 载体

[0307] 通过瞬时转染(或从HT1080细胞中的产病毒细胞系,参看美国申请) 制备载体制剂,滴度为约7E6TU/ml。在初始研究中,没有进一步纯化或 浓缩载体。在用于测定全剂量反应曲线的跟踪实验中,制备经纯化的高滴 度材料,滴度预期为约10E8/ml。载体以50-100μL的体积肿瘤内施用并 以100μL的体积静脉内施用,每只小鼠的总剂量为约7E5TU/小鼠。将表达αCD152的载体标示为TocaαCD152。

[0308] 肿瘤植入和载体注射

[0309] 对5组雌性DBA/2(55只小鼠,9-10周大)皮下植入S91黑素瘤细 胞(第0天),然后施用(第4-7天,取决于S91肿瘤的生长速率,约 50-100mm^3)运载体(第1组)、对照载体[AC3-GFP(V)](第2组)、肿瘤内(IT)TocaαCD152载体注射液(第3组)或静脉内TocaαCD152 载体注射液(第4组)。第5组小鼠没有植入肿瘤,且只静脉内注射Toca αCD152。

[0310] 数据分析

[0311] 执行肿瘤生长分析直到2000mm^3或第60天,这要依据哪种情况先发 生。将来自每组的10只小鼠的肿瘤尺寸与时间作图。使用方差分析 (ANOVA)来确定统计显著性。＜0.05的P值被视为在所有分析中具有 统计显著性,所述分析是用Prism 5统计软件(GraphPad软

件)或等效物 执行。还采用生活观察以评定治疗期间αCD152表达的任何不利事件。

[0312] 结果

[0313] 与对照相比,通过复制性MLV肿瘤内递送αCD152可显示统计上显 著的生长迟滞。与对照相比,通过复制性MLV静脉内递送αCD152显示 统计上显著的生长迟滞,消除来自DBA/2-Cloudman S91小鼠黑素瘤模型 的黑素瘤负担。

[0314] 其它动物研究如下进行:

[0315] 实施例9ACE-yCD2病毒载体在裸小鼠的颅内人类异种移植物(U87)中具有治疗作用

[0316] 建立使用U87人类神经胶质瘤细胞系的颅内异种移植物模型,以测试 RCR载体传播和生物分布,以及RCR载体介导的胞嘧啶脱氨酶自杀基因疗 法于裸小鼠宿主中的治疗功效。

[0317] 在驯化之后,将小鼠随机分配给8个治疗组中的1组(参看下文组描述)。在第0天对7个组的右侧纹状体颅内施用1×10^5个U87细胞/小鼠。第8组小鼠 未植入肿瘤。第5天,对小鼠只注射制剂缓冲液,或注射9×10^5/5μl、9×10^4/5 μl或9×10^3/5μl的RCR载体。令不接受载体或接受9×10^5/5μl或9×10^3/5μl 的载体的小鼠随机接受5-FC(500mg/kg/天)(通过单次腹腔注射施用, 在第19天开始),或不接受5-FC。接受中等剂量载体的小鼠都接受5-FC(即 对于此剂量没有单独的对照)。5-FC施用每天一次,持续连续7天,随后 15天不进行治疗。药物加测试的循环重复高达4个循环。将来自除第8组以 外的每组的10只小鼠随机分配给存活分析类别。剩余的小鼠根据预定方案 处死。

[0318] 组分配和剂量水平

[0319]

组	测试物品	体积	药物 TX	N	N/分析类别	
					(A)存活分析	(B)预定处死

[0320]

1	形式缓冲液	5 ul	无	4		在第一个药物周期前 4 只
2	形式缓冲液	5 ul	5FC	10	10	
3	T5.0002	9e5/5ul	PBS	10	10	
4	T5.0002	9e5/5ul	5FC	25	10	在每个周期开始前 3 只，总共 15 只
5	T5.0002	9e4/5ul	5FC	10	10	
6	T5.0002	9e3/5ul	5FC	25	10	在每个周期开始前 3 只，总共 15 只
7	T5.0002	9e3/5ul	PBS	10	10	
8 无肿瘤	无		5FC	15		在每个周期开始前 3 只，总共 15 只
动物总数				**109**	60	49

[0321] 静脉内给药通过注射入尾部静脉来进行。腹腔给药通过小心注射入腹 部同时避免注射入膀胱来进行。在颅内注射中，用异氟烷麻醉小鼠并将其 安置于具有钝耳棒(blunt ear bar)的立体定位装置中。将皮肤剃毛并使用 聚维酮碘(betadine)处理头皮以制备外科部位。将动物置于加热垫上并在 无菌条件下使用手术刀在皮肤中制造中线切口。可以通过切口部位的皮肤 回缩和筋膜反射来观察头骨。穿过头骨中的小钻孔插入具有3mm突出物

的 引导套管(其配备具有3.5mm突出物的盖子),并用牙科粘固粉和3个小 螺钉将所述套管与头骨连接。在粘固粉硬化后,用缝线缝合皮肤。映射的 立体定位坐标是AP＝0.5-1.0mm,ML＝1.8-2.0mm,DV＝3.0mm。可通过注 射染料并确定其位置,在预实验(2-3只动物)中确定所接受的动物组的精 确立体定位坐标。将在麻醉恢复期间监控动物。在程序结束前皮下(SC) 施用止痛剂丁丙诺啡(buprenorphine),然后大约每12小时施用丁丙诺啡, 持续长达3天。每天都监控动物。通过注射套管(具有穿过引导套管插入 的3.5mm突出物)颅内输注细胞或载体。用配备Hamilton注射器和软管的 注射泵控制速率。对于细胞注射,以0.2微升/分钟的流动速率(总共5分钟) 递送1微升细胞。对于载体注射,以0.33微升/分钟的流动速率(总共15分 钟)递送5微升载体。

[0322] 将载体递送到小鼠,且以每克脑重量的转化单位(TU)来计算。使用 所述计算,可计算其它哺乳动物(包括人)的剂量翻译。图8显示此小鼠 模型中对于载体剂量的影响。

[0323] 实施例10 AC3-yCD2(V)在同系小鼠脑癌模型中具有治疗作用

[0324] 在同系动物模型中进行其它实验以证明本公开书的方法和组合物。

[0325] 产生同系BALB/c小鼠中的使用CT26结肠直肠癌细胞系的颅内植入物 模型,以测试RCR载体传播和生物分布,以及RCR载体介导的胞嘧啶脱氨 酶自杀基因疗法的治疗功效和其免疫学影响。

[0326] 此项研究包含129只动物,0雄性,119只雌性和10只偶然动物(10只 雌性)。在驯化之后,将小鼠随机分配给8个治疗组中的1组(参看下文组 描述)。在第0天对7个组的右侧纹状体颅内施用1×10^4个CT26细胞/小鼠。 第8组小鼠未植入肿瘤。第4天,对小鼠只注射制剂缓冲液,或注射$9\times10^5/5$ μl、$9\times10^4/5\mu l$或$9\times10^3/5\mu l$的载体。令不接受载体或接受$9\times10^5/5\mu l$或$9\times$ $10^3/5\mu l$的载体的小鼠随机接受5-FC(500mg/kg/BID)(通过单次腹腔注 射施用,在第13天开始),或不接受5-FC。接受中等剂量载体的小鼠都接 受5-FC(即对于此剂量没有单独的对照)。5-FC施用每天一次,持续连续 7天,随后10天不进行治疗。药物加测试的周期重复高达4个周期。将来自 除第8组以外的每组的10只小鼠随机指派给存活分析类别。剩余的小鼠根 据预定方案处死。

[0327] 将监测稚小鼠和预定处死动物共同圈养,并在相同时间点取出以评定 通过脱皮的载体传输。

[0328] 组分配和剂量水平

[0329]

组	测试物品	体积	药物TX	N	N/分析类别		
					(A)存活分析	(B)预定处死	(C)标记
1	形式缓冲液	5 ul	PBS	4		在第一个药物周期前 4	

[0330]

						只	
2	形式缓冲液	5 ul	5FC	10	10		
3	T5.0002	9e5/5ul	PBS	10	10		
4	T5.0002	9e5/5ul	5FC	25	10	在每个周期开始前3只，总共15只	在每个周期开始前1只，总共5只
5	T5.0002	9e4/5ul	5FC	10	10		
6	T5.0002	9e3/5ul	5FC	25	10	在每个周期开始前3只，总共15只	在每个周期开始前1只，总共5只
7	T5.0002	9e3/5ul	PBS	10	10		
8 NO TUMOR	none		5FC	15		在每个周期开始前3只，总共15只	
动物总数				**119**	60	49	10

[0331] 静脉内给药通过注射入尾部静脉来进行。腹腔给药通过小心注射入腹 部同时避免注射入膀胱来进行。在颅内施用中，使用具有引导套管的小鼠，所述套管具有植入右侧纹状体的3.2mm突出物且配备具有3.7mm突出物 的盖子。映射的立体定位坐标是AP＝0.5-1.0mm，ML＝1.8-2.0mm，DV＝3.2 mm(从前囟开始)。通过注射套管(具有穿过引导套管插入的3.7mm突 出物)颅内输注细胞或载体。用配备Hamilton注射器和软管的注射泵控制 速

率。

[0332] 对于细胞注射,以0.2微升/分钟的流动速率(总共5分钟)递送1微升 细胞。对于载体注射,以0.33微升/分钟的流动速率(总共15分钟)递送5 微升载体。

[0333] 将载体递送到小鼠,且依据每克脑重量的转化单位(TU)来计算。使 用所述计算,可计算其它哺乳动物(包括人)的剂量转化量。图9显示当 颅内递送载体时,此小鼠模型中对于载体剂量的影响。

[0334] 实施例11构建和测试表达miR-128-1和miR128-2的RCR载体

[0335] A.构建含有人类pri-miRNA-128-1的异源多聚核苷酸序列的重组可 复制逆转录病毒载体

[0336] 可复制逆转录病毒载体(表达miR-128-1的pAC3-miR-128-1)来源 于一个实施方案中所述的pAC3-yCD2的骨架。通过用Mlu I和Not I对 pAC3-yCD2质粒DNA进行核酸内切酶消化以去除IRES-yCD2多聚核苷 酸序列来分离pAC3-miR-128-1载体的pAC3骨架。从pEP-mir-128-1表 达载体(Cell BioLabs Inc.)(SEQ ID NO:31)的产品清单获得pri-miR-128-1 的多聚核苷酸DNA序列。合成在双链DNA片段的5'端具有Mlu I限制 酶位点且在3'端具有Not I限制酶位点的pri-miR-128-1的DNA序列,以 用于随后在经Mlu I和Not I消化的上述pAC3-yCD2质粒DNA中的对应 位点插入。所得构建体pAC3-miR-128-1编码3种基因gag、pol和env, 以及非编码pri-miR-128-1序列(图11)。

[0337] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0338] B.构建含有人类pri-miRNA-128-2的异源多聚核苷酸序列的重组可 复制逆转录病毒载体

[0339] 可复制逆转录病毒载体(表达miR-128-2的pAC3-miR-128-2)来源 于一个实施方案中所述的pAC3-yCD2的骨架。通过用Mlu I和Not I对 pAC3-yCD2质粒DNA进行核酸内切酶消化以去除IRES-yCD2多聚核苷 酸序列来分离pAC3-miR-128-1载体的pAC3骨架。从表达pri-miR128-2 的表达载体Lenti-miR-128-2(System Biosciences)(SEQ ID NO:32)的序列分析获得pri-miR-128-2的多聚核苷酸DNA序列。合成在双链DNA片 段的5'端具有Mlu I限制酶位点且在3'端具有Not I限制酶位点的 pri-miR128-2的DNA序列,以用于随后在经Mlu I和Not I消化的上述 pAC3-yCD2质粒DNA中的对应位点插入。所得构建体pAC3-miR-128-1 编码3种基因gag、pol和env,以及非编码pri-miR-128-2序列(图1)。

[0340] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0341] C.构建含有人类H1启动子和人类pre-miRNA-128-的异源多聚核苷 酸序列的重组可复制逆转录病毒载体

[0342] 可复制逆转录病毒载体(表达miR-128-2的pAC3-miR-128-2)来源 于一个实施方案中所述的pAC3-yCD2的骨架。通过用Mlu I和Not I对 pAC3-yCD2质粒DNA进行核酸内切酶消化以去除IRES-yCD2多聚核苷 酸序列来分离pAC3-miR-128-1载体的pAC3骨架。从pSilencer 3.1 H1 hygro表达载体(Ambion)的产品信息获得人类H1启动子的多聚核苷酸DNA序列,并从http://www.mirbase.org/获得短发夹结构pre-miR-128-1 的多聚核苷酸DNA序列。合成在双链DNA片段的5'端具有Mlu I限制 酶位点且在3'端具有Not I限制酶位点的与人类H1启动子连接的 pri-miR128-1的DNA序列(SEQ ID NO:33),以用于随后在经Mlu I和 Not I消化的上述pAC3-yCD2质粒DNA中的对应位点插入。所得构建体 pAC3-H1-shRNAmiR128编码3种基因gag、pol和env,以及非编码短发 夹结构pre-miR-128-1序列(图11)。

[0343] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0344] D.通过qPCR分析来分析重组可复制逆转录病毒载体的复制动力学

[0345] 目前存在至少两种常用方式来获得重组可复制逆转录病毒载体的复 制动力学:(1)通过流式细胞分析法来分析表达GFP蛋白的载体的GFP 表达,和(2)通过测量从被转导细胞的培养基收集的载体储液的逆转录 酶活性来分析逆转录酶。与通过RNA和转基因表达获得的滴度(其中滴 度分别被高估和低估)相比,由被转导细胞的DNA分析所评定的滴度提供功能滴度的最可靠的评价(Sastry等人,2002Gene Therapy 9, 1155-1162)。病毒传播的复制动力学与所转导的细胞群体的百分比相关, 其中可通过qPCR并利用对培养物中测试的病毒序列具有特异性的引物 来检测病毒基因组的被整合前病毒DNA。如果测试复制动力学的比较结 果,那么本分析需要在同一载体内和各种载体之间从所有时间点开始输入 等量的基因组DNA。通过以C(t)值的倒数对时间点作图来产生复制动力 学图表。图12A和图12B显示各种重组可复制逆转录病毒载体的复制动 力学的比较结果。所述分析很容易揭露各种载体的复制动力学的细小差别。

[0346] E.测试含有miR-128的重组可复制逆转录病毒载体在培养物中的复 制动力学

[0347] 为了证实在本发明中分别并入的pri-miR-128-1、pri-miR-128-2和 H1-pre-miR-128-1序列正常复制,使用从上述瞬时转染收集的经计算体积 的每种载体储液以0.1的MOI分别感染新鲜人类纤维肉瘤细胞HT1080 和人类神经胶质瘤细胞U87-MG。在感染后3、6和9天使被转导细胞传 代。在每个时间点,收集一部分细胞进行基因组DNA提取,以便用于qPCR。 制备基因组DNA的稀释液以产生在qPCR中的等量基因组输入中具有相 同浓度的基因组DNA的等分试样。通过以C(t)值的倒数对时间点作图来 产生每个载体的复制动力学。图12A和图12B显示测试的所有载体以与 对照MLV病毒类似的动力学复制。

[0348] E.测试从用含有miR-128的重组可复制逆转录病毒载体转导的细胞 的成熟miR-

128表送

[0349] 为了证实用含有重组可复制逆转录病毒载体的miR-128转导的细胞 中的miR-128表达,使来自感染后9天的细胞(此时达到最大感染性) (图12A和图12B)繁殖,并进行收集以提取总RNA用于检测成熟miRNA 表达。来自Taqman微小RNA分析的结果显示,与未转导的细胞相比, 用pAC3-miR-128-1、pAC3-miR-128-2和pAC3-H1-shRNAmiR128载体转 导的HT1080和U87-MG细胞中发生成熟miR-128的过度表达(图13)。在两种细胞系中,用pAC3-miR-128-1和pAC3-H1-shRNAmiR128载体转 导的细胞中的成熟miR-128表达水平高于用pAC3-miR-128-2载体转导的 细胞中的表达水平。

[0350] F.测试从用含有miR-128的重组可复制逆转录病毒载体转导的细胞中 的Bmi-1表达以证明miR-12的靶标接合

[0351] 已经观察到Bmi-1表达在包括胶质母细胞瘤在内的各种人类癌症中得 到上调,且已经显示Bmi-1表达是miR-128的靶标。为了证实miR-128的靶 标接合,通过qRT-PCR来检测分别用pAC3-miR-128-1、pAC3-miR-128-2和 pAC3-H1-shRNAmiR128载体转导的细胞中的Bmi-1表达。图14显示分别用 pAC3-miR-128-1、pAC3-miR-128-2和pAC3-H1-shRNAmiR128载体转导的 U87-MG细胞中的Bmi-1表达水平低于未转导细胞,而在HT1080细胞中, 被转导细胞与未转导细胞之间没有观察到显著差别。数据支持miR-128在 中枢神经系统中扮演重要功能角色的观点。

[0352] 实例12:构建和测试含有IRES、yCD2、人类H1启动子和人类pre-miR128-1的异源多聚核苷酸序列的重组可复制逆转录病毒载体

[0353] A.构建

[0354] 可复制逆转录病毒载体pAC3-yCD2-H1-shRNAmiR128来源于一个实 施方案中所述的pAC3-yCD2的骨架。通过用Not I对pAC3-yCD2质粒 DNA进行核酸内切酶消化来分离pAC3-yCD2-H1-shRNAmiR128载体中 的pAC3-yCD2骨架。从pSilencer 3.1 H1 hygro表达载体(Ambion)的产 品信息获得人类H1启动子的多聚核苷酸DNA序列,并从 http://www.mirbase.org/获得短发夹结构pre-miR-128-1的多聚核苷酸DNA 序列。合成在双链DNA片段的两端具有Not I限制酶位点的与人类H1 启动子连接的pre-miR128-1的DNA序列(SEQ ID NO:34),以用于随 后插入经Not I消化的上述pAC3-yCD2质粒DNA中的对应位点。所得构建体pAC3-H1-shRNAmiR128编码4种基因gag、pol和env和yCD2,以 及非编码短发夹结构pre-miR-128-1序列(图11)。

[0355] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0356] B.测试pAC3-yCD2-H1-shRNAmiR128重组可复制逆转录病毒载体在 培养物中的复制动力学

[0357] 为了证实在本发明中并入的H1-pre-miR-128-1正常复制,使用从上 述瞬时转染收集的经计算体积的每种载体储液以0.1的MOI分别感染新 鲜人类纤维肉瘤细胞HT1080和

人类神经胶质瘤细胞U87-MG。在感染后 3、6和9天使被转导细胞传代。在每个时间点，收集一部分细胞进行基 因组DNA提取，以便用于qPCR。制备基因组DNA的稀释液以产生用于qPCR的等量基因组输入的具有相同浓度的基因组DNA的等分试样。通 过以C(t)值的倒数对时间点作图来产生每个载体的复制动力学。结果显示 载体在与对照MLV病毒类似的动力学下复制。

[0358] C.测试用pAC3-yCD2-H1-shRNAmiR128重组可复制逆转录病毒载体 转导的细胞中的成熟miR-128表达

[0359] 为了证实用pAC3-yCD2-H1-shRNAmiR128重组可复制逆转录病毒载 体转导的细胞中的miR-128表达，使来自感染后9天的细胞(此时达到 最大感染性)繁殖，并进行收集以提取总RNA用于检测成熟miRNA表 达。来自Taqman微小RNA分析的结果显示，与未转导的细胞相比，用 pAC3-yCD2-H1-shRNAmiR128转导的HT1080和U87-MG细胞中发生成 熟miR-128的过度表达

[0360] D.测试用pAC3-yCD2-H1-shRNAmiR128重组可复制逆转录病毒载 体转导的细胞中的Bmi-1表达以证明miR-128的靶标接合

[0361] 已经观察到Bmi-1表达在包括胶质母细胞瘤在内的各种人类癌症中得 到上调，且已经显示Bmi-1表达是miR-128的靶标。为了证实miR-128的靶 标接合，通过qRT-PCR检测用pAC3-yCD2-H1-shRNAmiR128转导的细胞中 的Bmi-1表达。结果显示用pAC3-yCD2-H1-shRNAmiR128转导的U87-MG 细胞中的Bmi-1表达水平低于未转导细胞，而在HT1080细胞中，被转导细 胞与未转导细胞之间没有观察到显著差别。数据支持miR-128在中枢神经 系统中扮演重要功能角色的观点。

[0362] E.通过免疫印迹测试用pAC3-yCD2-H1-shRNAmiR128转导的细胞中 的yCD2表达

[0363] 为了证实用pAC3-yCD2-H1-shRNAmiR128重组可复制逆转录病毒载 体转导的细胞中的yCD2表达，使来自感染后9天的细胞(此时达到最大 感染性)繁殖，并进行收集以提取总蛋白用于检测yCD2表达。免疫印迹 结果显示与用pAC3-yCD2转导的细胞相比，用 pAC3-yCD2-H1-shRNAmiR128转导的HT1080与U87-MG细胞中的yCD2 正常表达。

[0364] 实例13:测试含有miR-128的重组逆转录病毒载体在培养物中的载体稳定性

[0365] 测试含有miR-128的重组逆转录病毒载体(pAC3-miR-128-1、pAC3-miR-128-2、pAC3-H1-shRNAmiR128和 pAC3-yCD2-H1-shRNAmiR128)的多重连续感染周期以评定载体稳定性。最初用载体以低MOI感染HT1080和U87-MG细胞，并使所述细胞在培养物 中传播。收集每个感染周期的载体储液，过滤，稀释以用于感染新鲜细胞。在每个感染周期结束时，收集细胞并提取基因组DNA，以便通过标准PCR 使用与env基因的3’和3’LTR上游的载体未翻译区的3’结合的引物来评定转 基因稳定性。结果显示测试的所有载体在至少8-20个周期后仍保持稳定。

[0366] 实例14:在小鼠/人类异种移植物模型中的利用miRNA表达载体的抗肿瘤功效研究

[0367] 目标

[0368] 此项研究的目标在于评定当通过在具有人类神经胶质瘤异种移植物 的裸小鼠中以3种Toca 511剂量水平颅内(IC)注射来递送时，携带 miR128序列(AC3-miR128-1(V)；AC3-miR128-2(V)；AC3-miR128-3(V)) 的基于新颖MLV型可复制逆转录病毒载体对存活的影响。

[0369] 小鼠

[0370] 雌性无胸腺裸Foxn1^nu(裸)小鼠(约8周大)是从Harlan(Indianapolis IN)获得。小鼠在到达后驯化7天。通过外科手术将留置引导套管安置 在小鼠中,所述引导套管具有植入右侧纹状体的3.0mm突出物且配备含 有3.5mm突出物的盖子。立体定位坐标是Ap=+0.5mm,ML=-1.8mm(从 前囟开始)。

[0371] 细胞

[0372] U-87MG细胞(ATCC,Manassas VA)来源于44岁高加索女性的恶 性神经胶质瘤。将细胞在含有10%胎牛血清、丙酮酸钠和Glutamax (Hyclone,Logan UT和Invitrogen,San Diego CA)的杜贝卡改良伊格氏 培养基(Dulbecco's modified Eagles medium)中培养。将细胞再悬浮于 PBS(Hyclone,Logan UT)中以用于植入。通过具有穿过注射套管插入的3.5mm突出物的注射套管以0.2μL/min(5分钟,随后保持5分钟)的速 率颅内输注U-87MG细胞(1E5,1μL)。

[0373] 通过瞬时转染(或从产病毒细胞系,参看临时申请)制备载体制剂, 滴度均为约5E6TU/ml。在初始研究中,没有进一步纯化或浓缩载体。在 用于测定全剂量反应曲线的跟踪实验中,制备经纯化的高滴度材料,滴度 为约10E8/ml。以5μl或以下的体积颅内施用载体,每只小鼠的最小总剂 量为约2.5E4TU/小鼠。

[0374] 肿瘤植入和载体注射

[0375] 对6组雌性无胸腺裸Foxn1^nu小鼠(65只小鼠,9-10周大)颅内植 入U-87肿瘤细胞(第0天),然后颅内施用(取决于U87细胞的生长速 率,第4-7天)运载体(第1组)、施用对照载体(AC3-GFP(V),第2 组)或颅内施用AC3-miR128-1(V)、AC3-miR128-2(V)、AC3-miR128-3(V) (第3-5组)。第6组小鼠没有植入肿瘤或载体。

[0376] 数据分析

[0377] 对各自来自第1-6组的10只小鼠执行截止到第60天的存活分析,并 绘制成Kaplan Meyer曲线。通过时序检验(log-rank test)来比较存活曲 线。<0.05的P值被视为在所有分析中具有统计显著性,所述分析是用 Prism 5统计软件(GraphPad软件)或等效物执行。

[0378] 来自载体治疗组的结果显示了,与用对照载体或仅用运载体治疗相比, 此人类神经胶质瘤异种移植物模型中的统计上显著的存活优势。

[0379] 实例15:在小鼠/人类异种移植物模型中的利用yCD2和miRNA表达载体的抗肿瘤功效研究

[0380] 此项研究的目标在于评定当通过在具有"干细胞"样的富集人类神经 胶质瘤异种移植物的裸小鼠中以3种剂量水平颅内(IC)注射来递送时, 表达yCD2和miR-128的新颖MLV型可复制逆转录病毒载体(命名为 AC3-yCD2-H1-shRNAmiR128(V))对存活的影响。

[0381] 雌性无胸腺裸Foxn1^nu(裸)小鼠(约8周大)是从Harlan(Indianapolis IN)获得。小鼠在到达后驯化7天。通过外科手术将留置引导套管安置 在小鼠中,所述引导套管具有植入右侧纹状体的3.0mm突出物且配备含 有3.5mm突出物的盖子。立体定位坐标是AP=+0.5mm,ML=-1.8mm(从 前囟开始)。

[0382] 在含有EGF,bFGF and B27培养物的无血清培养基中培养原代人类 神经胶质瘤(Carol Kruse博士,Burnham Biomedical Res Inst,San Diego, CA)的早期传代物。将细胞再悬浮于PBS(Hyclone,Logan UT)中以用 于植入。通过具有穿过注射套管插入的3.5mm

突出物的注射套管以0.2 μL/min(5分钟,随后保持5分钟)的速率颅内输注约1000个细胞(1 μL)。

[0383] 通过瞬时转染(或从产病毒细胞系,参看临时申请)制备载体制剂，滴度均为约5E6TU/ml。在初始研究中,没有进一步纯化或浓缩载体。在 用于测定全剂量反应曲线的跟踪实验中,制备经纯化的高滴度材料,滴度 为约10E8/ml。以5μl或以下的体积颅内施用载体,每只小鼠的最小总剂 量为约2.5E4TU/小鼠。

[0384] 肿瘤植入和载体注射

[0385] 对6组雌性无胸腺裸Foxn1^nu小鼠(66只小鼠,9-10周大)颅内植 入“干细胞样”细胞(第0天),然后取决于细胞的生长速率,在肿瘤植入 后7-14天如下颅内施用:第1组:运载体;第2组:对照载体AC3-GFP(V)； 第3组:TOCA511;第4组:AC3-yH1-shRNAmiR128(V);第5组: AC3-yCD2-H1-shRNAmiR128(V);和第6组:没有植入肿瘤或载体的未 处理小鼠。

[0386] 数据分析

[0387] 对各自来自第1-5组的10只小鼠执行截止到第60天的存活分析,并 绘制成Kaplan Meyer曲线。通过时序检验来比较存活曲线。＜0.05的P值 被视为在所有分析中具有统计显著性,所述分析是用Prism 5统计软件 (GraphPad软件)或等效物执行。

[0388] 结果

[0389] 来自载体治疗组的结果显示,与用对照载体或仅用运载体治疗相比，此人类神经胶质瘤异种移植物模型中的统计上显著的存活优势。

[0390] 实例16:构建具有miR靶序列的载体

[0391] A.构建表达GFP且含有142-3p靶序列的单拷贝的可复制逆转录病 毒载体

[0392] 编码GFP的可复制逆转录病毒载体pAC3-emd-142-3pT来源于上述 (当前现有专利)pAC3-emd的骨架。通过用Not I对pAC3-emd质粒DNA 进行核酸内切酶消化来分离pAC3-emd-142-3pT载体中的pAC3-emd骨架。从公开文献(Brown等人,2006Nature Medicine 12:5585-591)获得 miR142-3pT的完全互补靶序列。合成在双链DNA片段的每一端具有Not I限制酶位点的miR-142-3p的靶序列(SEQ ID NO:35),以用于随后在 pAC3-emd质粒DNA中的对应位点插入。通过测序分析来证实142-3pT 插入物的定向。所得构建体pAC3-emd-142-3pT编码4种基因gag、pol、 env和emd,以及非编码142-3pT序列(图15)。

[0393] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体，通过qPCR测定载体储液的滴度。

[0394] B.构建表达yCD2且含有142-3p靶序列的单拷贝的可复制逆转录病 毒载体

[0395] 编码yCD2基因的可复制逆转录病毒载体pAC3-yCD2-142-3pT来源 于上述(当前现有专利)pAC3-yCD2的骨架。通过用Not I对pAC3-yCD2 质粒DNA进行核酸内切酶消化来分离pAC3-yCD2-142-3pT载体中的 pAC3-yCD2骨架。从公开文献(Brown等人,2006Nature Medicine 12:5 585-591)获得miR142-3pT的完全互补靶序列。合成在双链DNA片段的 每一端具有Not I限制酶位点的miR-142-3p的靶序列(SEQ ID NO:35)，以用于随后在pAC3-

yCD2质粒DNA中的对应位点插入。所得构建体 pAC3-yCD2-142-3pT编码4种基因gag、pol、env和yCD2,以及非编码 142-3pT序列(图16)。

[0396] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0397] C.构建表达GFP且含有142-3p靶序列的4个拷贝的可复制逆转录 病毒载体

[0398] 编码GFP的可复制逆转录病毒载体pAC3-emd-142-3pT4X来源于上 述(当前现有专利)pAC3-emd的骨架。通过用Not I对pAC3-emd质粒 DNA进行核酸内切酶消化来分离pAC3-emd-142-3pT 4X载体中的 pAC3-emd骨架。从公开文献(Brown等人,2006Nature Medicine 12:5 585-591)获得miR142-3pT4X的完全互补靶序列的4个串联重复序列。合成在双链DNA片段的每一端具有Not I限制酶位点的miR-142-3p4X的 靶序列(SEQ ID NO:36),以用于随后在pAC3-emd质粒DNA中的对 应位点插入。通过测序分析来确认142-3pT插入物的定向。所得构建体 pAC3-emd-142-3pT4X编码4种基因gag、pol、env和emd,以及非编码 142-3pT4X序列(图15)。

[0399] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0400] D.构建表达yCD2且含有142-3p靶序列的4个拷贝的可复制逆转录 病毒载体

[0401] 编码GFP的可复制逆转录病毒载体pAC3-yCD2-142-3pT4X来源于上 述pAC3-emd的骨架。通过用Not I对pAC3-yCD2质粒DNA进行核酸内 切酶消化来分离pAC3-yCD2-142-3pT 4X载体中的pAC3-emd骨架。从公 开文献(Brown等人,2006Nature Medicine 12:5585-591)获得 miR142-3pT4X的完全互补靶序列的4个串联重复序列。合成在双链DNA 片段的每一端具有Not I限制酶位点的miR-142-3pT4X的靶序列(SEQ ID NO:36),以用于随后在pAC3-yCD2质粒DNA中的对应位点插入。通 过测序分析来确认142-3pT4X插入物的定向。所得构建体 pAC3-emd-142-3pT4X编码4种基因gag、pol、env和emd,以及非编码 142-3pT4X序列(图16)。

[0402] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来制备载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时,向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时,提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体, 通过qPCR测定载体储液的滴度。

[0403] 实例17:通过qPCR分析对重组逆转录病毒的复制动力学制图

[0404] 目前存在至少两种常用方式来获得重组可复制逆转录病毒载体的复 制动力学：(1)通过流式细胞分析法来分析表达GFP蛋白的载体的GFP 表达，和(2)通过测量从被转导细胞的培养基收集的载体储液的逆转录 酶活性来分析逆转录酶。与通过RNA和转基因表达获得的滴度(其中滴 度分别被高估和低估)相比，由被转导细胞的DNA分析所评定的滴度提供功能滴度的最可靠的评价(Sastry等人，2002Gene Therapy 9，1155-1162)。病毒传播的复制动力学与所转导的细胞群体的百分比相关，其中可通过qPCR并利用对培养物中测试的病毒序列具有特异性的引物 来检测病毒基因组的被整合前病毒DNA。如果测试复制动力学的比较结 果，那么本分析需要在同一载体内和各种载体之间从所有时间点开始输入 等量的gDNA。通过以C(t)值的倒数对时间点作图来产生复制动力学图表。图17A和图17A显示各种重组可复制逆转录病毒载体的复制动力学的比 较结果。所述分析很容易揭露各种载体的复制动力学的细小差别。

[0405] 实例18：测试复制动力学

[0406] 非造血人类细胞系中的含有142-3pT的重组逆转录病毒载体

[0407] 为了证实在本发明中并入的142-3pT序列与其母体载体类似地复制，使用从上述瞬时转染收集的经计算体积的每种载体储液以0.1的MOI分 别感染新鲜人类纤维肉瘤细胞HT1080和人类神经胶质瘤细胞U87-MG。在感染后3、6和9天使被转导细胞传代。在每个时间点，收集一部分细 胞进行基因组DNA提取，以便用于qPCR。制备基因组DNA的稀释液以 产生在qPCR中的等量基因组输入中具有相同浓度的基因组DNA的等分 试样。

[0408] 通过以C(t)值的倒数对时间点作图来产生每个载体的复制动力学。图 17A和图18A显示测试的所有载体以与其母体载体(pAC3-emd和 pAC3-yCD2)类似的动力学复制。也通过流式细胞分析法来评定表达GFP 蛋白的载体(pAC3-emd、pAC3-emd-142-3pT和pAC3-emd-142-3pT4X) 的复制动力学。图17B和图18B显示pAC3-emd-142-3pT和 pAC3-emd-142-3pT4X载体以与其母体载体pAC3-emd类似的动力学复制。

[0409] 测试含有142-3pT的重组逆转录病毒载体在人类和小鼠造血细胞中 的复制动力学

[0410] 通过Taqman微小RNA分析，使用对小鼠和人类miR-142-3p具有特 异性的引物组(因为这两种物种的成熟miR-142-3p的序列是相同的)，首先确定小鼠T淋巴细胞系ELA、人类T淋巴细胞系SUP-T1和人类单核 细胞系U937中的成熟miR-142-3p表达。在以2的MOI初始感染的所有 三种细胞系中测试表达GFP的重组逆转录病毒载体的复制动力学。图19 显示pAC3-emd载体在人类T淋巴细胞和单核细胞(SUP-T1和U937) 中有效地复制，因为病毒传播截止到感染后28天分别达到细胞群体的65％ 和95％。EL4细胞中的载体传播在所测试的时间范围期间保持小于5％。

[0411] 实例18：测试含有142-3pT的重组逆转录病毒载体在人类和小鼠造血细胞中的载体传播

[0412] 通过流式细胞分析法测试在表达GFP的重组逆转录病毒载体中 miR-142-3p对于抑制GFP表达的功能作用。使用从上述瞬时转染收集的 经计算体积的pAC3-emd、pAC3-emd-142-3pT和pAC3-emd-142-3pT4X 载体储液以2的MOI感染新鲜EL4、SUP-T1和U937细胞。在感染后第 6、12和18天收集一部分细胞，以用于通过流式细胞分析法分析GFP表 达。图20A显示与母体载体pAC3-emd相比，用pAC3-emd-142-3pT和 pAC3-emd-142-3pT4X载体转导的ELA

细胞中的GFP表达被抑制到背景 水平。GFP抑制作用在所测试的时间范围内持久保持。图20B和图20C 显示与母体载体pAC3-emd相比，分别用pAC3-emd-142-3pT和 pAC3-emd-142-3pT4X载体转导的SUP-T1和U937细胞中的GFP表达被 显著抑制。结果证实小鼠和人类造血细胞中的miR-142-3p表达有效地抑 制被含有142-3pT序列的重组逆转录病毒载体转导的细胞中的GFP表达。在U937细胞中，结果表明含有4个142-3pT拷贝的载体 (pAC3-emd-142-3pT4X)可比含有单个142-3pT拷贝的载体 (pAC3-emd-142-3pT4X)更有效地抑制GFP表达。

[0413] 实例19:测试病毒RNA基因组

[0414] 尚不清楚miR-142-3pT抑制上述GFP表达的功能作用是因为在翻译 水平直接抑制GFP表达还是因为在转录后水平降解病毒基因组。在实验 时间点最后收集一部分细胞，以便使用标准分子生物学方法提取总RNA。使用对来源于逆转录病毒RNAd的cDNA具有特异性的引物(例如pol 引物组和env2引物组)进行qRT-PCR以评定被转导细胞的病毒负荷。结 果显示分别用pAC3-emd-142-3pT和pAC3-emd-142-3pT4X转导的细胞 的病毒负荷显著低于用pAC3-emd载体转导的细胞。结果支持以下观点： 小鼠和人类造血细胞中的miR-142-3p在转录后水平有效地降解病毒RNA 基因组，从而限制小鼠和人类造血细胞中的载体传播。

[0415] 实例20.通过qPCR测试整合的前病毒DNA

[0416] 尚不清楚miR-142-3pT抑制上述GFP表达的功能作用是因为在翻译 水平直接抑制GFP表达还是因为降解病毒基因组和因此整合前病毒DNA。 在实验时间点最后收集一部分细胞，以便使用标准分子生物学方法提取基 因组DNA。使用对整合的前病毒DNA具有特异性的引物进行qPCR以评 定每个细胞中整合的前病毒DNA的拷贝数。结果显示分别用pAC3-emd-142-3pT和pAC3-emd-142-3pT4X转导的细胞的每个细胞的拷贝数显著 低于用pAC3-emd载体转导的细胞。结果支持以下观点:小鼠和人类造血 细胞中的miR-142-3p在转录后水平有效地降解病毒RNA基因组，从而限 制小鼠和人类造血细胞中的载体传播。

[0417] 实例21:测试含有142-3pT的重组逆转录病毒载体在培养物中的载体稳定性

[0418] 测试含有142-3pT的重组逆转录病毒载体的多重连续感染周期以评 定载体稳定性。最初用载体以低MOI感染HT1080和U87-MG细胞，并 使所述细胞在培养物中传播。收集每个感染周期的载体储液，过滤，稀释 以用于感染新鲜细胞。在每个感染周期结束时，收集细胞并提取基因组 DNA，以便通过标准PCR使用与env基因的3’和的载体中连接于IRES 的异源多聚核苷酸序列下游的未翻译区的3’结合的引物来评定转基因稳 定性。结果显示含有142-3pT单拷贝的载体(pAC3-emd-142-3pT和 pAC3-yCD2-142-3pT)在至少10-20个周期后仍保持稳定，而含有142-3pT 的4个串联重复序列的载体(pAC3-emd-142-3pT4X和 pAC3-yCD2-142-3pT4X)显示在早期感染周期中142-3pT序列的缺失。

[0419] 实例22:测试含有142-3pT的重组逆转录病毒载体在体内的受控载体传播

[0420] 此实验是用与以下实施例27相同的设计来进行。通过在植入有U87 异种移植物的8周大的裸Balb/C小鼠中静脉内注射重组逆转录病毒载体 (pAC3-emd、pAC3-emd-142-3pT和pAC3-emd-142-3pT4X)来体内测试 miR-142-3p限制病毒经由造血细胞传播的功能作用。每种载体存在3组 小鼠，各自代表1个时间点(例如在病毒载体施用后30天、60天和90 天)。通过静脉内注射对所有动物施用1E4到1E7TU剂量的每种载体储 液。处死来自每个时间点的动物以收集脾、淋巴结和骨髓和肿瘤。收集细 胞子群(例如CD4+、CD8+等)的GFP表达并通过流式细胞分析法进行 分析。在双重复实验中，处死来自每个时间点的动物并收集组织

(例如肝、肾、脾、肿瘤等)以用于基因组DNA提取。进行qPCR以评定所收集的 组织中被整合的前病毒DNA的存在情况。结果显示含有142-3pT的载体 的载体传播在造血组织中显著减少,从而证明载体复制在这些组织中减少。 同时,在肿瘤仍然观察到GFP和PCR信号,显示miR靶序列在淋巴样组 织中的传播减少,但仍显示在肿瘤组织中的传播。

[0421] 实施例23.具有自发复发恶性神经胶质瘤并用T5.0002载体加5-FC处理的病患狗中的延长存活

[0422] 用T5.0002病毒(经纯化且在与甘露醇和蔗糖等张的等张Tris/NaCl pH7.2、1mg/ml HSA、0.1mg/ml抗坏血酸盐中配制(参看US 5792643; T.Rodriugez等人J Gene Med 9:2332007;美国申请61218063))结合 5-FC处理在完全外科切除后3个月出现复发性间变性少突胶质细胞瘤的 雄性35kg拳师犬(Boxer dog)。肿瘤为约13cm^3,且引起严重侧脑室挤压和显著的中线偏移(参看图1)。由于肿瘤尺寸较大,通过2个独立导 管(400μL和480μL)使用增强对流递送(Convection Enhanced Delivery; CED)来输注Toca 511。施用的总Toca 511剂量为每克脑约4.1×10^6TU。 在注射前将ProHance®(钆特醇(gadoteridol))添加到Toca 511中以允 许通过MRI来可视化递送。载体的分布体积据估计为肿瘤体积的约 10-12%。

[0423] 图21A和图21B是在肿瘤内CED输注Toca 511和钆期间从病患狗 获得的MRI影像的静止帧。应注意影像左侧的大肿瘤挤压脑的两侧并使 中线结构向右位移。白色区域是钆-Toca 511输注物。图21B显示在肿瘤 中安置两个导管。

[0424] 使Toca 511传播8天。以130mg/kg/天的分剂量(每天三次,和食 物一起,)用5-FC经口处理狗持续5天。在接下来的2天将剂量增加到 160mg/kg/天(5-FC总共持续7天)。跟踪MRI显示肿瘤尺寸无变化且 肿瘤内部区域有一些可能的变化。在病毒传播21天后,用160mg/kg/天 的较高剂量(分剂量,每天三次,和食物一起)开始5-FC的第二个周期。 由于在施用第5天出现了皮疹,停止给药。

[0425] 在5-FC的第一疗程后第2周和5-FC的第二疗程后第2周(在治疗 开始后第7周)进行的MRI显示肿瘤体积达到稳定水平,而肿瘤生长速 率则下降。患者变得更为警惕和活跃,且在载体注射后13周仍保持临床 稳定。直到记录时狗仍然存活。狗的估计寿命从初次注射载体算起不超过 3-4周。根据存活期是主治兽医最初估计的3到4倍,显示了Toca 511/5-FC组合在所述病患狗中的功效。

[0426] 实施例24.构建γ干扰素载体

[0427] 构建和测试编码人类IFN-γ基因的可复制逆转录病毒载体

[0428] 编码人类IFN-γ基因的可复制逆转录病毒载体pAC3-hIFNg来源于上 述pAC3-yCD2的骨架。通过用Psi I和Not I对pAC3-yCD2质粒DNA进 行核酸内切酶消化来分离pAC3-hIFNg载体中的pAC3骨架。在从不同登 记号(AM903379、BC070256和NM000619)获得的三种序列中鉴别和 确定人类IFN-γ基因的cDNA序列。序列比对显示这三种序列中的相同序列。合成在DNA片段的每一端具有Psi I和Not I限制酶位点的人类IFN-γ 的开放阅读框(SEQ ID NO:38),以用于随后在pAC3骨架中的对应位 点插入。所得构建体pAC3-hIFNg编码4种基因:gag、pol、env和人类 IFN-γ(图22)。

[0429] 通过使用FUGENE HD转染试剂将载体编码质粒DNA瞬时转染到 HT1080细胞中来制备载体储液。在转染后48小时,收集含有载体的上 清液,将其滤过0.45μm滤器,并立即使用

或作为等分试样储存在-80℃ 下以备随后使用。使用特定体积的未稀释载体储液来感染新鲜的75％融 合HT1080。在感染后第4天和第5天，收集含有载体的上清液，将其滤 过0.45μm滤器，并立即使用或作为等分试样储存在-80℃下以备随后使 用。使用24微升所收集的载体储液来感染人类前列腺癌细胞PC3。在感 染后24小时，向细胞中添加AZT以抑制病毒进一步复制。在感染后48 小时，提取被感染PC3细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体，通过qPCR测定载体储液的滴度。

[0430] 首先测试RNA水平上的人类IFN-γ表达。在感染后第5天，在第二 次感染中，使用标准RNA提取方法从被转导HT1080细胞中提取总RNA。进行RT-PCR以检测人类IFN-γ表达。在RT反应中使用50ng总RNA 来产生cDNA。随后将1/10体积的RT反应物用于PCR，所述PCR使用 对人类IFN-γ具有特异性的PCR引物组。RT-PCR结果显示人类IFN-γ 在用pAC3-hIFNg载体感染的HT1080细胞中发生表达。

[0431] 通过标准ELISA测试分泌的人类IFN-γ蛋白的表达。在ELISA分析 中连续稀释从感染后第4天和第5天收集的载体储液，以便获得蛋白浓度 与稀释因子之间的线性范围。结果显示感染后第5天HT1080细胞所分泌 的人类IFN-γ蛋白的浓度高于感染后第4天细胞所分泌的蛋白浓度(图 24)。感染后第5天细胞分泌约325-355pg/mL人类IFN-γ蛋白。

[0432] 构建和测试编码小、鼠IFN-γ基因的可复制逆转录病毒载体

[0433] 编码小鼠IFN-γ基因的可复制逆转录病毒载体pAC3-mIFNg来源于上 述pAC3-yCD2的骨架。通过用Psi I和Not I对pAC3-yCD2质粒DNA进 行核酸内切酶消化来分离pAC3-mIFNg载体中的pAC3骨架。在从不同登 记号(BC119063、BC119065和NM008337)获得的三种序列中鉴别和确 定小鼠IFN-γ基因的cDNA序列。序列比对显示这三种序列中的相同序列。合成在DNA片段的每一端具有Psi I和Not I限制酶位点的小鼠IFN-γ的 开放阅读框(SEQ ID NO:39)，以用于随后在pAC3骨架中的对应位点 插入。所得构建体pAC3-mIFNg编码4种基因:gag、pol、env和小鼠IFN-γ (图22)。

[0434] 通过使用FUGENE HD转染试剂将载体编码质粒DNA瞬时转染到 HT1080细胞中来制备载体储液。在转染后48小时，收集含有载体的上 清液，将其滤过0.45微米滤器，并立即使用或作为等分试样储存在-80℃ 下以备随后使用。使用特定体积的未稀释载体储液来感染新鲜的75％融 合HT1080。在感染后第4天和第5天，收集含有载体的上清液，将其滤 过0.45微米滤器，并立即使用或作为等分试样储存在-80℃下以备随后使 用。使用24微升所收集的载体储液来感染人类前列腺癌细胞PC3。在感 染后24小时，向细胞中添加AZT以抑制病毒进一步复制。在感染后48 小时，提取被感染PC3细胞的基因组DNA以用于滴定分析。利用已知拷 贝数的标准物的包涵体，通过qPCR测定载体储液的滴度。

[0435] 首先测试RNA水平上的小鼠IFN-γ表达。在感染后第5天，在第二 次感染中，使用标准RNA提取方法从被转导HT1080细胞中提取总RNA。进行RT-PCR以检测小鼠IFN-γ表达。在RT反应中使用50ng总RNA 来产生cDNA。随后将1/10体积的RT反应物用于PCR，所述PCR使用 对小鼠IFN-γ具有特异性的PCR引物组(图23)。RT-PCR结果显示小 鼠IFN-γ在用pAC3-mIFNg载体感染的HT1080细胞中发生表达。

[0436] 通过标准ELISA测试分泌的人类IFN-γ蛋白的表达。在ELISA分析 中连续稀释从感染后第4天和第5天收集的载体储液，以便获得蛋白浓度 与稀释因子之间的线性范围。结果显示感染后第5天HT1080细胞所分泌 的小鼠IFN-γ蛋白的浓度高于感染后第4天细胞所

分泌的蛋白浓度(图 25)。感染后第5天细胞分泌约33-42ng/mL小鼠IFN-γ蛋白。

[0437] 实施例25实施例D:在小鼠皮下肿瘤模型中的利用γ干扰素表达载体的抗肿瘤功效研究

[0438] 目标

[0439] 此项研究的目标在于评定当通过在具有皮下结肠癌(CT26.WT)的 BALB/c小鼠中经由肿瘤内(IT)注射来递送时,携带鼠类γ干扰素序列 (pAC3-mIFNg)的基于新颖MLV的可复制逆转录病毒载体对肿瘤生长 的影响。

[0440] 小鼠

[0441] 雌性BALB/c小鼠(～8周大)是从Jackson Laboratories(Bar Harbor, ME)购得。小鼠在到达后驯化7天,随后开始研究。

[0442] 细胞

[0443] CT26.WT细胞(ATCC,Manassas VA)是经N-亚硝基-N-甲基氨基甲 酸乙酯(NNMU)诱导的未分化结肠癌细胞系。将细胞在含有10％胎牛 血清、丙酮酸钠和Glutamax(Hyclone,Logan UT和Invitrogen,San Diego CA)的杜贝卡改良伊格氏培养基中培养。将细胞再悬浮于PBS(Hyclone, Logan UT)中以用于植入。将CT26.WT细胞(2E5,100μL)注入BALB/c 小鼠的右肋。

[0444] 载体

[0445] 通过瞬时转染(或来自用来自初始转染的感染病毒感染另一种细胞系 后的生产细胞系)制备载体制剂,滴度为约3E6TU/ml。在初始研究中, 没有进一步纯化或浓缩载体。在用于测定全剂量反应曲线的跟踪实验中, 制备经纯化的高滴度材料,滴度预期为约10E8/ml。为了获得高滴度物质, 选择犬细胞系CF2用于制备,因为γ干扰素的跨物种反应性很低,且使 用异种细胞系可预防γ干扰素对生产细胞的抑制作用。如本说明书中所述对载体进行纯化和浓缩。载体以100μL体积肿瘤内施用,每只小鼠的总 剂量为约3E3、3E4和3E5TU/小鼠。将表达γ干扰素的载体标示为Toca 621。

[0446] 肿瘤植入和载体注射

[0447] 对9组雌性BALB/c(99只小鼠,9-10周大)皮下植入CT26.WT肿 瘤细胞(第0天),然后施用(第4-7天,取决于CT26肿瘤的生长速率, 约50-100mm^3)运载体(第1组)、对照载体[AC3-GFP(V)](第2组)、肿瘤内Toca 621载体注射液(第3-5组)或静脉内Toca 621载体注射液 (第6-8组)。第9组小鼠没有植入肿瘤,且只静脉内注射载体。

[0448] 数据分析

[0449] 执行肿瘤生长分析直到2000mm^3或第60天,这要依据哪种情况先发 生。将来自每组的10只小鼠的肿瘤尺寸与时间作图。使用方差分析 (ANOVA)来确定统计显著性。＜0.05的P值被视为在所有分析中具有 统计显著性,所述分析是用Prism 5统计软件(GraphPad软件)或等效物 执行。还采用生活观察以评定Toca 621施用的任何不利事件。

[0450] 结果

[0451] 随时间的肿瘤尺寸测量结果显示用表达γIFN的载体处理的肿瘤的 生长与接受对照载体或运载体的动物体内的肿瘤的生长存在统计上显著 的差异。

[0452] 实施例26实施例D:在小鼠皮下肿瘤模型中的利用γ干扰素表达载 体的抗肿瘤功效研究

[0453] 目标

[0454] 此项研究的目标在于评定当通过在具有皮下结肠癌(CT26.WT)的 BALB/c小鼠中经由肿瘤内(IT)注射来递送时,携带鼠类γ干扰素序列 (pAC3-mIFNg)的新颖MLV型可复制逆转录病毒载体对肿瘤生长的影 响。

[0455] 小鼠

[0456] 雌性BALB/c小鼠(～8周大)是从Jackson Laboratories(Bar Harbor, ME)购得。小鼠在到达后驯化7天,随后开始研究。

[0457] 细胞

[0458] CT26.WT细胞(ATCC,Manassas VA)是经N-亚硝基-N-甲基氨基甲 酸乙酯(NNMU)诱导的未分化结肠癌细胞系。将细胞在含有10％胎牛 血清、丙酮酸钠和Glutamax(Hyclone,Logan UT和Invitrogen,San Diego CA)的杜贝卡改良伊格氏培养基中培养。将细胞再悬浮于PBS(Hyclone, Logan UT)中以用于植入。将CT26.WT细胞(2E5,100μL)注入BALB/c 小鼠的右肋。

[0459] 载体

[0460] 通过瞬时转染(或从经具有来自初始转染的感染病毒的另一种细胞系 感染的产病毒细胞系)制备载体制剂,滴度为约3E6TU/ml。在初始研究 中,没有进一步纯化或浓缩载体。在用于测定全剂量反应曲线的跟踪实验 中,制备经纯化的高滴度材料,滴度预期为约10E8/ml。为了获得高滴度 物质,选择犬细胞系CF2用于制备,因为γ干扰素的跨物种反应性很低, 且使用异种细胞系可预防γ干扰素对产生细胞的抑制作用。如本说明书中 所述对载体进行纯化和浓缩。载体以100μL体积肿瘤内施用,每只小鼠 的总剂量为约3E3、3E4和3E5TU/小鼠。将表达γ干扰素的载体标示为 Toca 621。

[0461] 肿瘤植入和载体注射

[0462] 对9组雌性BALB/c(99只小鼠,9-10周大)皮下植入CT26.WT肿 瘤细胞(第0天),然后施用(第4-7天,取决于CT26肿瘤的生长速率, 约50-100mm^3)运载体(第1组)、对照载体[AC3-GFP(V)](第2组)、肿瘤内Toca 621载体注射液(第3-5组)或静脉内Toca 621载体注射液 (第6-8组)。第9组小鼠没有植入肿瘤,且只静脉内注射载体。

[0463] 数据分析

[0464] 执行肿瘤生长分析直到2000mm^3或第60天,这要依据哪种情况先发 生。将来自每组的10只小鼠的肿瘤尺寸与时间作图。使用方差分析 (ANOVA)来确定统计显著性。＜0.05的P值被视为在所有分析中具有 统计显著性,所述分析是用Prism 5统计软件(GraphPad软件)或等效物 执行。还采用生活观察以评定Toca 621施用的任何不利事件。

[0465] 结果

[0466] 随时间的肿瘤尺寸测量结果显示用表达γIFN的载体处理的肿瘤的 生长与接受对照载体或运载体的动物体内的肿瘤的生长存在统计上显著 的差异。

[0467] 实施例27使用复制性逆转录病毒载体的静脉内基因递送

[0468] 目标

[0469] 此项研究的目标在于评定静脉内递送携带标志绿色荧光蛋白 (AC3-GFP(V))的新颖MLV型可复制逆转录病毒载体对植入裸小鼠脑 内的U87神经胶质瘤的作用。

[0470] 小鼠

[0471] 雌性无胸腺裸Foxn 1ˆnu(裸)小鼠(约8周大)是从Harlan(Indianapolis IN)获得。小鼠在到达后驯化7天。通过外科手术将留置引导套管安置 在小鼠中,所述引导套管具有植入右侧纹状体的3.0mm突出物且配备含 有3.5mm突出物的盖子。立体定位坐标是Ap=+0.5mm,ML=-1.8mm(从 前囟开始)。

[0472] 细胞

[0473] U-87MG细胞(ATCC,Manassas VA)来源于44岁高加索女性的恶 性神经胶质瘤。将细胞在含有10%胎牛血清、丙酮酸钠和Glutamax (Hyclone,Logan UT和Invitrogen,San Diego CA)的杜贝卡改良伊格氏 培养基(Dulbecco's modified Eagles medium)中培养。将细胞再悬浮于 PBS(Hyclone,Logan UT)中以用于植入。通过具有穿过注射套管插入的3.5mm突出物的注射套管以0.2μL/min(5分钟,随后保持5分钟)的速 率颅内(IC)输注U-87MG细胞(1E5,1μL)。

[0474] 载体

[0475] 通过瞬时转染制备载体制剂,其滴度均为约2.8E7TU/ml。以5μl或 以下的体积肿瘤内(IT)施用载体,每只小鼠的最小总剂量为约1.4E45TU/ 小鼠。通过尾部静脉来静脉内注射2.8E6/100μL。

[0476] 肿瘤植入和载体注射

[0477] 对5组雌性无胸腺裸Foxn1ˆnu小鼠(16只小鼠,9-10周大)颅内植 入U-87肿瘤细胞(第0天),然后肿瘤内或静脉内施用以下物质(第4-7 天,取决于U87肿瘤的生长速率):运载体(静脉内)(第1组)、载 体(静脉内)(第2组)、血脑屏障破坏物伐地那非(Vardenafil)和载体(肿瘤内)(第3组)、伐地那非和载体(肿瘤内)(第4组)或载体 (肿瘤内)(第5组)。在载体注射后14天处死小鼠并分离肿瘤,且分 析肿瘤中的GFP表达。

[0478] 数据分析

[0479] 通过流式细胞术分析从第2-5组小鼠分离的被破坏肿瘤的U87细胞 的GFP阳性百分比(图26)。也对第1、3和5组进行直方图分析以测 量所分离的U87细胞中的GFP信号分布(图27)。

[0480] 结果

[0481] 静脉内递送GFP与在颅内植入裸小鼠中的U87神经胶质瘤细胞中肿 瘤内注射同样有效。

[0482] 实施例28构建编码人类IL-2基因的可复制逆转录病毒载体

[0483] 编码人类IL2基因的可复制逆转录病毒载体pAC3-hIL2来源于上述 pAC3-yCD2的骨架。通过用Psi I和Not I对pAC3-yCD2质粒DNA进行 核酸内切酶消化来分离pAC3-hIL2载体编码质粒DNA中的pAC3骨架。使用从不同登记号(BC066255和BC066257)获得的序列鉴别和确定人 类IFN-γ基因的cDNA序列。这两种序列的序列比对揭露相同序列。合成 在DNA片段的每一端具有Psi I和Not I限制酶位点的人类IFN-γ的开放 阅读框(SEQ ID NO:40),以用于随后在pAC3骨架中的对应位点插入。所得构建体pAC3-hIL2编码4种基因:gag、pol、env和人类IL2(图28)。

[0484] 通过使用磷酸钙方法将载体编码质粒DNA瞬时转染到293T细胞中 来产生载体储液。在转染后18小时,用新鲜培养基更换培养物。在培养 基更换后24小时,收集含有载体的上清液,将其滤过0.45μm滤器,并 立即使用或作为等分试样储存在-80℃下以备随后使用。

使用24微升所收 集的载体储液来感染人类前列腺癌细胞PC3。在感染后24小时，向细胞 中添加AZT以抑制病毒进一步复制。在感染后48小时，提取被感染PC3 细胞的基因组DNA以用于滴定分析。利用已知拷贝数的标准物的包涵体， 通过qPCR测定载体储液的滴度。

[0485] 首先测试RNA水平上的人类IL-2表达。在感染后第5天，使用标准 RNA提取方法从被转导的CF2TH和HT1080细胞提取总RNA。进行 RT-PCR以检测人类IL-2表达。在RT反应中使用50ng总RNA来产生 eDNA。随后将1/10体积的RT反应物用于PCR，所述PCR使用对人类 IL-2具有特异性的PCR引物组。RT-PCR结果显示人类IL-2在用 pAC3-hIL2载体转导的HT1080细胞中发生表达。

[0486] 通过标准ELISA测试分泌的人类IL-2蛋白的表达。在ELISA分析中 连续稀释从感染后第5天收集的载体储液，以便获得蛋白浓度与稀释因子 之间的线性范围。结果显示在感染后第5天CF2TH细胞所分泌的人类IL-2 蛋白的浓度高于HT1080。

[0487] 实施例29在小鼠皮下肿瘤模型中的利用白细胞介素2表达载体的抗肿瘤功效研究

[0488] 目标

[0489] 此项研究的目标在于评定当通过在具有皮下结肠癌(CT26.WT)的 BALB/c小鼠中经由肿瘤内(IT)注射来递送时，携带鼠类白细胞营养激 素白细胞介素2(IL-2)序列(pAC3-mIL2)的新颖MLV型可复制逆转 录病毒载体对肿瘤生长的影响。

[0490] 小鼠

[0491] 雌性BALB/c小鼠(～8周大)是从Jackson Laboratories(Bar Harbor， ME)购得。小鼠在到达后驯化7天，随后开始研究。

[0492] 细胞

[0493] CT26.WT细胞(ATCC，Manassas VA)是经N-亚硝基-N-甲基氨基甲 酸乙酯(NNMU)诱导的未分化结肠癌细胞系。将细胞在含有10％胎牛 血清、丙酮酸钠和Glutamax(Hyclone，Logan UT和Invitrogen，San Diego CA)的杜贝卡改良伊格氏培养基中培养。将细胞再悬浮于PBS(Hyclone， Logan UT)中以用于植入。将CT26.WT细胞(2E5，100μL)注入BALB/c 小鼠的右肋。

[0494] 载体

[0495] 通过瞬时转染(或从产病毒细胞系，参看临时申请)制备载体制剂， 滴度为约6E6TU/ml。在初始研究中，没有进一步纯化或浓缩载体。在用 于测定全剂量反应曲线的跟踪实验中，制备经纯化的高滴度材料，滴度预 期为约10E8/ml。载体以100μL体积肿瘤内施用，每只小鼠的总剂量为 约6E5TU/小鼠。将表达γ干扰素的载体标示为Toca IL2。

[0496] 肿瘤植入和载体注射

[0497] 对5组雌性BALB/c(55只小鼠，9-10周大)皮下植入CT26.WT肿 瘤细胞(第0天)，然后施用(第4-7天，取决于CT26肿瘤的生长速率， 约50-100mm^3)运载体(第1组)、对照载体[AC3-GFP(V)](第2组)、肿瘤内Toca IL2载体注射液(第3组)或静脉内Toca IL2载体注射液(第 4组)。第5组小鼠没有植入肿瘤，且只静脉内注射载体。

[0498] 数据分析

[0499] 执行肿瘤生长分析直到2000mm^3或第60天，这要依据哪种情况先发 生。将来自每组的10只小鼠的肿瘤尺寸与时间作图。使用方差分析 (ANOVA)来确定统计显著性。＜0.05的P值被视为在所有分析中具有 统计显著性，所述分析是用Prism 5统计软件(GraphPad软

件)或等效物 执行。还采用生活观察以评定治疗期间IL-2表达的任何不利事件。

[0500] 结果

[0501] 通过复制MLV递送IL-2会减少且在一些情况下清除BALB/c CT26小鼠 模型的肿瘤负荷。

[0502] 已经描述了许多本公开书的实施方案。然而,应了解可在不偏离本公 开书的精神和范围的情况下进行各种修改。因此,其它实施方案属于以下 权利要求书的范围内。

序列表

<110> 托卡根公司
哈利·E·格鲁贝尔.
道格拉斯·乔利
奥马尔·佩雷斯

<120> 重组载体

<130> 00014-007WO2

<140> 未知
<141> 2009-09-26

<150> US 61/100,666
<151> 2008-09-26

<150> US 61/120,618
<151> 2008-12-08

<150> 61/186,823
<151> 2009-06-13

<160> 30

<170> PatentIn 3.5版

[0001]

<210> 1
<211> 477
<212> DNA
<213> 酿酒酵母

<220>
<221> CDS
<222> (1)..(477)

<400> 1

```
atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac          48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1               5                   10                  15

att gcc tat gag gag gcg gcc tta ggt tac aaa gag ggt ggt gtt cct          96
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
            20                  25                  30

att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt         144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
        35                  40                  45

ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag         192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
    50                  55                  60

atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa         240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
```

65 70 75 80

gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt 288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

gcc atc atc atg tat ggt att cca cgc tgt gtt gtc ggt gag aac gtt 336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtt gtt gtt gtt gac gat gag agg tgt aaa aag atc atg aaa caa ttt 432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 2
<211> 158
<212> PRT
<213> 酿酒酵母

[0002]

<400> 2

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 3
<211> 477
<212> DNA
<213> 人工序列

<220>
<223> 经工程改造的胞嘧啶脱氨酶

<220>
<221> CDS
<222> (1)..(477)

<400> 3

[0003]

```
atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac   48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1               5                   10                  15

att gcc tat gag gag gcg tta tta ggt tac aaa gag ggt ggt gtt cct   96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
            20                  25                  30

att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt  144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
        35                  40                  45

ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag  192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
    50                  55                  60

atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa  240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65                  70                  75                  80

gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt  288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
                85                  90                  95

gcc atc atc atg tat ggt att cca cgc tgt gtc atc ggt gag aac gtt  336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
            100                 105                 110
```

aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtt gtt gtt gtt gac gat gag agg tgt aaa aag tta atg aaa caa ttt 432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 4
<211> 158
<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<400> 4

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

[0004] Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe

130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 5
<211> 480
<212> DNA
<213> 人工序列

<220>
<223> 经人类密码子优化的胞嘧啶脱氨酶

<220>
<221> CDS
<222> (1)..(480)

<400> 5

[0005]

```
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat          48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1               5                   10                  15

atc gct tac gag gag gcc gca ctg ggc tac aag gag ggc ggc gtg cct          96
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
            20                  25                  30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg         144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
        35                  40                  45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag         192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
    50                  55                  60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag         240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65                  70                  75                  80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc         288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
                85                  90                  95

gct atc atc atg tac ggc atc cct agg tgt gtg gtc ggc gag aac gtg         336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
            100                 105                 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag         384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
        115                 120                 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag atc atg aag cag ttc         432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
    130                 135                 140
```

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tga taa 480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 6
<211> 158
<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<400> 6

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

[0006]

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 7
<211> 756
<212> DNA
<213> 酿酒酵母

<220>
<221> CDS
<222> (1)..(756)

<400> 7

```
atg aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat          48
Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
1               5                  10                  15

ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat          96
Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
            20                  25                  30

ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct         144
Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
        35                  40                  45

caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat         192
Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
    50                  55                  60

aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga ttg         240
Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
65                  70                  75                  80

ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg         288
Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
                85                  90                  95

gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa         336
Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
            100                 105                 110

atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga         384
Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly
        115                 120                 125

tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa         432
Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
    130                 135                 140

agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta cca         480
Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro
145                 150                 155                 160

gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc         528
Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala
                165                 170                 175

acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga ggt         576
```

[0007]

Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly
180 185 190

gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag gaa 624
Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu
195 200 205

ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act 672
Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr
210 215 220

ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg 720
Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly
225 230 235 240

ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa 756
Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
245 250

<210> 8
<211> 251
<212> PRT
<213> 酿酒酵母

<400> 8

[0008] Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
1 5 10 15

Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
20 25 30

Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
35 40 45

Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
50 55 60

Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
65 70 75 80

Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
85 90 95

Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
100 105 110

Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly

115 120 125

Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
130 135 140

Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro
145 150 155 160

Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala
165 170 175

Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly
180 185 190

Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu
195 200 205

Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr
210 215 220

[0009] Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly
225 230 235 240

Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
245 250

<210> 9
<211> 1443
<212> DNA
<213> 智人

<220>
<221> CDS
<222> (1)..(1443)

<400> 9
atg gct gtt gct cgt gct gct ctt ggt cct ctt gtt act ggt ctt tat 48
Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr
1 5 10 15

gat gtt caa gct ttt aaa ttt ggt gat ttt gtt ctt aaa tct ggt ctt 96
Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu
20 25 30

tct tct cct att tat att gat ctt cgt ggt att gtt tct cgt cct cgt 144
Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg

35 40 45

ctt ctt tct caa gtt gct gat att ctt ttt caa act gct caa aat gct 192
Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala
50 55 60

ggt att tct ttt gat act gtt tgt ggt gtt cct tat act gct ctt cct 240
Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro
65 70 75 80

ctt gct act gtt att tgt tct act aat caa att cct atg ctt att cgt 288
Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg
85 90 95

cgt aaa gaa act aaa gat tat ggt act aaa cgt ctt gtt gaa ggt act 336
Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr
100 105 110

att aat cct ggt gaa act tgt ctt att att gaa gat gtt gtt act tct 384
Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser
115 120 125

ggt tct tct gtt ctt gaa act gtt gaa gtt ctt caa aaa gaa ggt ctt 432
Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu
130 135 140

[0010]
aaa gtt act gat gct att gtt ctt ctt gat cgt gaa caa ggt ggt aaa 480
Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys
145 150 155 160

gat aaa ctt caa gct cat ggt att cgt ctt cat tct gtt tgt act ctt 528
Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu
165 170 175

tct aaa atg ctt gaa att ctt gaa caa caa aaa aaa gtt gat gct gaa 576
Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu
180 185 190

act gtt ggt cgt gtt aaa cgt ttt att caa gaa aat gtt ttt gtt gct 624
Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala
195 200 205

gct aat cat aat ggt tct cct ctt tct att aaa gaa gct cct aaa gaa 672
Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu
210 215 220

ctt tct ttt ggt gct cgt gct gaa ctt cct cgt att cat cct gtt gct 720
Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala
225 230 235 240

tct aaa ctt ctt cgt ctt atg caa aaa aaa gaa act aat ctt tgt ctt 768
Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu
245 250 255

tct gct gat gtt tct ctt gct cgt gaa ctt ctt caa ctt gct gat gct 816
Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala

```
                260                 265                 270

ctt ggt cct tct att tgt atg ctt aaa act cat gtt gat att ctt aat      864
Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn
        275                 280                 285

gat ttt act ctt gat gtt atg aaa gaa ctt att act ctt gct aaa tgt      912
Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys
    290                 295                 300

cat gaa ttt ctt att ttt gaa gat cgt aaa ttt gct gat att ggt aat      960
His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn
305                 310                 315                 320

act gtt aaa aaa caa tat gaa ggt ggt att ttt aaa att gct tct tgg     1008
Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp
                325                 330                 335

gct gat ctt gtt aat gct cat gtt gtt cct ggt tct ggt gtt gtt aaa     1056
Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys
            340                 345                 350

ggt ctt caa gaa gtt ggt ctt cct ctt cat cgt ggt tgt ctt ctt att     1104
Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile
        355                 360                 365

gct gaa atg tct tct act ggt tct ctt gct act ggt gat tat act cgt     1152
Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg
    370                 375                 380

gct gct gtt cgt atg gct gaa gaa cat tct gaa ttt gtt gtt ggt ttt     1200
Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe
385                 390                 395                 400

att tct ggt tct cgt gtt tct atg aaa cct gaa ttt ctt cat ctt act     1248
Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr
                405                 410                 415

cct ggt gtt caa ctt gaa gct ggt ggt gat aat ctt ggt caa caa tat     1296
Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr
            420                 425                 430

aat tct cct caa gaa gtt att ggt aaa cgt ggt tct gat att att att     1344
Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile
        435                 440                 445

gtt ggt cgt ggt att att tct gct gct gat cgt ctt gaa gct gct gaa     1392
Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu
    450                 455                 460

atg tat cgt aaa gct gct tgg gaa gct tat ctt tct cgt ctt ggt gtt     1440
Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
465                 470                 475                 480

taa                                                                1443
```

[0011]

<210> 10
<211> 480
<212> PRT
<213> 智人

<400> 10

Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr
1 5 10 15

Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu
20 25 30

Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg
35 40 45

Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala
50 55 60

Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro
65 70 75 80

[0012]

Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg
85 90 95

Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr
100 105 110

Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser
115 120 125

Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu
130 135 140

Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys
145 150 155 160

Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu
165 170 175

Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu
180 185 190

Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala
195 200 205

Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu
210 215 220

Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala
225 230 235 240

Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu
245 250 255

Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala
260 265 270

Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn
275 280 285

Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys
290 295 300

[0013]

His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn
305 310 315 320

Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp
325 330 335

Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys
340 345 350

Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile
355 360 365

Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg
370 375 380

Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe
385 390 395 400

Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr
405 410 415

Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr
420 425 430

Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile
435 440 445

Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu
450 455 460

Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
465 470 475 480

<210> 11
<211> 1227
<212> DNA
<213> 人工序列

<220>
<223> 融合构建体CDopt-UPRT

<220>
<221> CDS
<222> (1)..(1227)

[0014]

<400> 11

```
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat   48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1               5                   10                  15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct   96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
            20                  25                  30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg  144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
        35                  40                  45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag  192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
    50                  55                  60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag  240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65                  70                  75                  80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc  288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
                85                  90                  95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg  336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
```

```
                100                 105                 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag   384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
        115                 120                 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc   432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
    130                 135                 140

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag aac ccg   480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro
145                 150                 155                 160

tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat ctt ata tat   528
Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr
                165                 170                 175

tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat ctg caa aaa   576
Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys
            180                 185                 190

atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct caa aca aac   624
Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
        195                 200                 205

caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat aca act aga   672
Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg
    210                 215                 220

cct gat ttc att ttc tac tcc gat aga atc atc aga ttg ttg gtt gaa   720
Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
225                 230                 235                 240

gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg gaa act gac   768
Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
                245                 250                 255

acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa atc tgt ggt   816
Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
            260                 265                 270

gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga tta aga gac   864
Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
        275                 280                 285

tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa agg gac gag   912
Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
    290                 295                 300

gag act gct tta cca aag tta ttc tac gaa aaa tta cca gag gat ata   960
Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
305                 310                 315                 320

tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc acc ggt ggt  1008
Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
```

[0015]

325 330 335

agt gct atc atg gct aca gaa gtc ttg att aag aga ggt gtt aag cca 1056
Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
340 345 350

gag aga att tac ttc tta aac cta atc tgt agt aag gaa ggg att gaa 1104
Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
355 360 365

aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act ggt gcc ctc 1152
Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
370 375 380

gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg ttg ggt gac 1200
Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
385 390 395 400

ttt ggt gac aga tac tac tgt gtt taa 1227
Phe Gly Asp Arg Tyr Tyr Cys Val
405

<210> 12
<211> 408
<212> PRT
<213> 人工序列

[0016]

<220>
<223> 合成构建体

<400> 12

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro
145 150 155 160

Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr
165 170 175

Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys
180 185 190

Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
195 200 205

[0017]

Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg
210 215 220

Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
225 230 235 240

Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
245 250 255

Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
260 265 270

Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
275 280 285

Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
290 295 300

Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
305 310 315 320

Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
325 330 335

Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
340 345 350

Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
355 360 365

Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
370 375 380

Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
385 390 395 400

Phe Gly Asp Arg Tyr Tyr Cys Val
405

<210> 13
<211> 1287
<212> DNA
[0018] <213> 人工序列

<220>
<223> 融合构建体 - CDopt - 连接子 - UPRT

<220>
<221> CDS
<222> (1)..(1287)

<400> 13
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag 240

```
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65                  70                  75                  80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc      288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
                85                  90                  95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg      336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
            100                 105                 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag      384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
        115                 120                 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc      432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
    130                 135                 140

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc      480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
145                 150                 155                 160

ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc      528
Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
                165                 170                 175

[0019] ggc gcc aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca  576
Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr
            180                 185                 190

tat ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga      624
Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg
        195                 200                 205

aat ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta      672
Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu
    210                 215                 220

cct caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag      720
Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys
225                 230                 235                 240

aat aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga      768
Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg
                245                 250                 255

ttg ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att      816
Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile
            260                 265                 270

gtg gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt      864
Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly
        275                 280                 285

aaa atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa      912
```

Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln
290 295 300

gga tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att 960
Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile
305 310 315 320

caa agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta 1008
Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu
325 330 335

cca gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg 1056
Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu
340 345 350

gcc acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga 1104
Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg
355 360 365

ggt gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag 1152
Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys
370 375 380

gaa ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt 1200
Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val
385 390 395 400

[0020] act ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca 1248
Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro
405 410 415

ggg ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa 1287
Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
420 425

<210> 14
<211> 428
<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<400> 14

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
145 150 155 160

[0021]

Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
165 170 175

Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr
180 185 190

Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg
195 200 205

Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu
210 215 220

Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys
225 230 235 240

Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg
245 250 255

Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile
260 265 270

Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly
275 280 285

Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln
290 295 300

Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile
305 310 315 320

Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu
325 330 335

Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu
340 345 350

Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg
355 360 365

Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys
370 375 380

[0022]

Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val
385 390 395 400

Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro
405 410 415

Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
420 425

<210> 15
<211> 1200
<212> DNA
<213> 人工序列

<220>
<223> 融合构建体 - CDopt3 - OPRT

<220>
<221> CDS
<222> (1)..(1200)

<400> 15

```
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat          48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1               5                   10                  15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct          96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
            20                  25                  30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg         144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
        35                  40                  45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag         192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
    50                  55                  60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag         240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65                  70                  75                  80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc         288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
                85                  90                  95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg         336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
            100                 105                 110
```

[0023]

```
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag         384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
        115                 120                 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc         432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
    130                 135                 140

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag gcg gtc         480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val
145                 150                 155                 160

gct cgt gca gct ttg ggg cca ttg gtg acg ggt ctg tac gac gtg cag         528
Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln
                165                 170                 175

gct ttc aag ttt ggg gac ttc gtg ctg aag agc ggg ctt tcc tcc ccc         576
Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro
            180                 185                 190

atc tac atc gat ctg cgg ggc atc gtg tct cga ccg cgt ctt ctg agt         624
Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser
        195                 200                 205

cag gtt gca gat att tta ttc caa act gcc caa aat gca ggc atc agt         672
Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser
    210                 215                 220
```

```
ttt gac acc gtg tgt gga gtg cct tat aca gct ttg cca ttg gct aca        720
Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr
225                 230                 235                 240

gtt atc tgt tca acc aat caa att cca atg ctt att aga agg aaa gaa        768
Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu
                245                 250                 255

aca aag gat tat gga act aag cgt ctt gta gaa gga act att aat cca        816
Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro
            260                 265                 270

gga gaa acc tgt tta atc att gaa gat gtt gtc acc agt gga tct agt        864
Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser
        275                 280                 285

gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc ttg aag gtc act        912
Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr
    290                 295                 300

gat gcc ata gtg ctg ttg gac aga gag cag gga ggc aag gac aag ttg        960
Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu
305                 310                 315                 320

cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca ttg tcc aaa atg       1008
Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met
                325                 330                 335
```

[0024]

```
ctg gag att ctc gag cag cag aaa aaa gtt gat gct gag aca gtt ggg       1056
Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly
            340                 345                 350

aga gtg aag agg ttt att cag gag aat gtc ttt gtg gca gcg aat cat       1104
Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His
        355                 360                 365

aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa gaa ctc agc ttc       1152
Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe
    370                 375                 380

ggt gca cgt gca gag ctg ccc agg atc cac cca gtt gca tcg aag taa       1200
Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys
385                 390                 395
```

<210> 16
<211> 399
<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<400> 16

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp

1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

[0025] Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val
145 150 155 160

Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln
165 170 175

Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro
180 185 190

Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser
195 200 205

Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser
210 215 220

Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr

225 230 235 240

Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu
245 250 255

Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro
260 265 270

Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser
275 280 285

Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr
290 295 300

Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu
305 310 315 320

Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met
325 330 335

[0026] Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly
340 345 350

Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His
355 360 365

Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe
370 375 380

Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys
385 390 395

<210> 17
<211> 1260
<212> DNA
<213> 人工序列

<220>
<223> 融合构建体 - CDopt3 - 连接子 - OPRT

<220>
<221> CDS
<222> (1)..(1260)

<400> 17

```
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat          48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1               5                   10                  15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct          96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
            20                  25                  30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg         144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
        35                  40                  45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag         192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
    50                  55                  60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag         240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65                  70                  75                  80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc         288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
                85                  90                  95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg         336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
            100                 105                 110
```

[0027]

```
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag         384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
        115                 120                 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc         432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
    130                 135                 140

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc         480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
145                 150                 155                 160

ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc         528
Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
                165                 170                 175

ggc gcc gcg gtc gct cgt gca gct ttg ggg cca ttg gtg acg ggt ctg         576
Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu
            180                 185                 190

tac gac gtg cag gct ttc aag ttt ggg gac ttc gtg ctg aag agc ggg         624
Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly
        195                 200                 205

ctt tcc tcc ccc atc tac atc gat ctg cgg ggc atc gtg tct cga ccg         672
Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro
    210                 215                 220
```

```
cgt ctt ctg agt cag gtt gca gat att tta ttc caa act gcc caa aat      720
Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn
225                 230                 235                 240

gca ggc atc agt ttt gac acc gtg tgt gga gtg cct tat aca gct ttg      768
Ala Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu
                245                 250                 255

cca ttg gct aca gtt atc tgt tca acc aat caa att cca atg ctt att      816
Pro Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile
            260                 265                 270

aga agg aaa gaa aca aag gat tat gga act aag cgt ctt gta gaa gga      864
Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly
        275                 280                 285

act att aat cca gga gaa acc tgt tta atc att gaa gat gtt gtc acc      912
Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr
    290                 295                 300

agt gga tct agt gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc      960
Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly
305                 310                 315                 320

ttg aag gtc act gat gcc ata gtg ctg ttg gac aga gag cag gga ggc     1008
Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly
                325                 330                 335
```

[0028]

```
aag gac aag ttg cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca     1056
Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr
            340                 345                 350

ttg tcc aaa atg ctg gag att ctc gag cag cag aaa aaa gtt gat gct     1104
Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala
        355                 360                 365

gag aca gtt ggg aga gtg aag agg ttt att cag gag aat gtc ttt gtg     1152
Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val
    370                 375                 380

gca gcg aat cat aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa     1200
Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys
385                 390                 395                 400

gaa ctc agc ttc ggt gca cgt gca gag ctg ccc agg atc cac cca gtt     1248
Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val
                405                 410                 415

gca tcg aag taa                                                    1260
Ala Ser Lys
```

<210> 18
<211> 419

<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<400> 18

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

[0029]

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
145 150 155 160

Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
165 170 175

Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu
180 185 190

Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly
195 200 205

Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro
210 215 220

Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn
225 230 235 240

Ala Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu
245 250 255

Pro Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile
260 265 270

Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly
275 280 285

Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr
290 295 300

[0030]

Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly
305 310 315 320

Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly
325 330 335

Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr
340 345 350

Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala
355 360 365

Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val
370 375 380

Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys
385 390 395 400

Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val
405 410 415

Ala Ser Lys

<210> 19
<211> 11892
<212> DNA
<213> 人工序列

<220>
<223> RCR载体 - pAC3-yCD2

<400> 19

[0031]

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360

catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 540

acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600

actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca tccgacttgt 660

ggtctcgctg ttccttggga gggtctcctc tgagtgattg actacccgtc agcgggggtc 720

tttcatttgg gggctcgtcc gggatcggga gacccctgcc cagggaccac cgacccacca 780

ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840

tgattttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggcggacccg 900

tggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960

cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020

gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080

tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140

tctgctgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaga 1200

atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260

agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320

[0032]

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

tcatcaccca ggttaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440

tcccctacat cgtgacctgg gaagccttgg cttttgaccc ccctccctgg gtcaagccct 1500

ttgtacaccc taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560

ctcctcgttc gaccccgcct cgatcctccc tttatccagc cctcactcct tctctaggcg 1620

ccaaacctaa acctcaagtt ctttctgaca gtgggggggcc gctcatcgac ctacttacag 1680

aagacccccc gccttatagg gacccaagac cacccccttc cgacagggac ggaaatggtg 1740

gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800

ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag 1860

gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920

ataataaccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgttc 1980

tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100

gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctcctagcgg 2220

gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340

ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc 2400

agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460

ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520

gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtcccaaga 2760

aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg 2820

gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac 2940

ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga 3000

[0033]

ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactctttc ctccatgtac 3060

cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120

actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180

ccctaaatat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcctc aggcctgggc ggaaaccggg ggcatgggac 3300

tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

tgttggacca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480

ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540

gggtggaaga catccacccc accgtgccca acccttacaa cctcttgagc gggctcccac 3600

cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720

gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca cttctgagct agactgccaa caaggtactc 3900

gggccctgtt acaaacccta gggaacctcg ggtatcgggc ctcggccaag aaagcccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgaggccag aaaagagact gtgatggggc agcctactcc gaagacccct cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgcctac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag 4500

aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560

aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg 4620

ctacgctgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

[0034]

aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740

ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040

taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100

gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160

tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcatcaccc tacacctcag 5220

aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg 5280

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520

gaactagggt ccgcgggcat cggcccggca ctcattggga gatcgatttc accgagataa 5580

agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640

tagaagcctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gagtcagaca gtggccgatc tgttggggat tgattggaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt 5880

taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940

tgtaccgagc ccgcaacacg ccgggccccc atggcctcac cccatatgag atcttatatg 6000

gggcaccccc gcccccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060

cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg 6180

gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga 6300

tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

ggcgcgttca acgctctcaa aacccccтca agataagatt aacccgtgga agcccttaat 6420

agtcatggga gtcctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt 6480

aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg 6540

aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga 6600

gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720

tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca ccccctggga 6840

cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900

cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcactgatgc 6960

aggaaaaaag gctaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg 7020

aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt 7080

ccccataggg cccaacccag tattacccga ccaaagactc ccttcctcac caatagagat 7140

tgtaccggct ccacagccac ctagcccccт caataccagt tacccccctt ccactaccag 7200

[0035]

tacaccctca acctccccta caagtccaag tgtcccacag ccacccccag gaactggaga 7260

tagactacta gctctagtca aaggagccta tcaggcgctt aacctcacca atcccgacaa 7320

gacccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gagtagcggt 7380

cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttcccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccacccaaag cgccggctca ggatcctact accttgcagc 7560

acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt 7620

gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680

ctcccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat 7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc *atgccgctat* 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

[0036]

gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttaa tagatccccc tggtttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220

caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280

tcagcaatat caccagctaa aacccataga gtacgagcca tgaacgcgtt actggccgaa 8340

gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt attttccacc atattgccgt 8400

cttttggcaa tgtgagggcc cggaaacctg gccctgtctt cttgacgagc attcctaggg 8460

gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc 8520

ctctggaagc ttcttgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaacc 8580

ccccacctgg cgacaggtgc ctctgcggcc aaaagccacg tgtataagat acacctgcaa 8640

aggcggcaca accccagtgc cacgttgtga gttggatagt tgtggaaaga gtcaaatggc 8700

tctcctcaag cgtattcaac aaggggctga aggatgccca gaaggtaccc cattgtatgg 8760

gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac 8820

gtctaggccc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg 8880

tgaccggcgg catggcctcc aagtgggatc aaaagggcat ggatatcgct tacgaggagg 8940

ccctgctggg ctacaaggag ggcggcgtgc ctatcggcgg ctgtctgatc aacaacaagg 9000

acggcagtgt gctgggcagg ggccacaaca tgaggttcca gaagggctcc gccaccctgc 9060

acggcgagat ctccaccctg gagaactgtg gcaggctgga gggcaaggtg tacaaggaca 9120

ccaccctgta caccaccctg tccccttgtg acatgtgtac cggcgctatc atcatgtacg 9180

gcatccctag gtgtgtgatc ggcgagaacg tgaacttcaa gtccaagggc gagaagtacc 9240

tgcaaaccag gggccacgag gtggtggttg ttgacgatga gaggtgtaag aagctgatga 9300

agcagttcat cgacgagagg cctcaggact ggttcgagga tatcggcgag taagcggccg 9360

cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc 9420

tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcatggaa aaatacataa 9480

ctgagaatag agaagttcag atcaaggtca ggaacagatg gaacagctga atatgggcca 9540

aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca 9600

gctgaatatg ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctcagggcca 9660

agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg 9720

[0037]

tttccagggt gccccaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag 9780

ttcgcttctc gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa 9840

cccctcactc ggggcgccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca 9900

ataaaccctc ttgcagttgc atccgacttg tggtctcgct gttccttggg agggtctcct 9960

ctgagtgatt gactacccgt cagcgggggt ctttcattac atgtgagcaa aaggccagca 10020

aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc 10080

tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata 10140

aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 10200

gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcaatgctc 10260

acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 10320

acccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 10380

ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 10440

gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 10500

gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 10560

ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt tttttgtttt gcaagcagca 10620

gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 10680

cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 10740

cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 10800

gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 10860

tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga 10920

gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc 10980

agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac 11040

tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc 11100

agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt cacgctcgtc 11160

gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 11220

catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt 11280

ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 11340

atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 11400

tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag 11460

cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 11520

cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc 11580

atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 11640

aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 11700

ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 11760

aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 11820

aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct *11880*

tcaagaattc at 11892

<210> 20
<211> 11892
<212> DNA
<213> 人工序列

<220>
<223> RCR载体 - pAC3-yCD

[0038] <400> 20
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360

catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 540

acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600

actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca tccgacttgt 660

ggtctcgctg ttccttggga gggtctcctc tgagtgattg actacccgtc agcgggggtc 720

tttcatttgg gggctcgtcc gggatcggga gacccctgcc cagggaccac cgacccacca 780

ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840

tgattttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggcggacccg 900

tggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960

cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg *1020*

gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080

tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140

tctgctgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaga 1200

atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260

agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct *1320*

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

tcatcaccca ggttaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440

tcccctacat cgtgacctgg gaagccttgg cttttgaccc ccctccctgg gtcaagccct 1500

ttgtacaccc taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560

ctcctcgttc gaccccgcct cgatcctccc tttatccagc cctcactcct tctctaggcg 1620

[0039] ccaaacctaa acctcaagtt ctttctgaca gtgggggggcc gctcatcgac ctacttacag 1680

aagacccccc gccttatagg gacccaagac cacccccttc cgacagggac ggaaatggtg *1740*

gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800

ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag 1860

gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920

ataataaccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgttc 1980

tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg *2040*

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100

gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctcctagcgg 2220

gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340

ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc 2400

agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460

ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520

[0040]

gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtcccaaga 2760

aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg 2820

gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac 2940

ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga 3000

ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactctttc ctccatgtac 3060

cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120

actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180

ccctaaatat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcctc aggcctgggc ggaaaccggg ggcatgggac 3300

tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

tgttggacca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480

ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540

gggtggaaga catccacccc accgtgccca acccttacaa cctcttgagc gggctcccac 3600

cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720

gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca cttctgagct agactgccaa caaggtactc 3900

gggccctgtt acaaacccta gggaacctcg ggtatcgggc ctcggccaag aaagcccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgaggccag aaaagagact gtgatggggc agcctactcc gaagacccct cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagcccccct gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

[0041]

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgcctac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag 4500

aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560

aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg 4620

ctacgctgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740

ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040

taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100

gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160

tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcatcaccc tacacctcag 5220

aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg 5280

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520

gaactagggt ccgcgggcat cggcccggca ctcattggga gatcgatttc accgagataa 5580

agcccggatt gtatggctat aaatatcttc tagttttat agataccttt tctggctgga 5640

tagaagcctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gagtcagaca gtggccgatc tgttggggat tgattggaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt 5880

[0042]

taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940

tgtaccgagc ccgcaacacg ccgggccccc atggcctcac cccatatgag atcttatatg 6000

gggcaccccc gcccccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060

cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg 6180

gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga 6300

tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

ggcgcgttca acgctctcaa aacccccctca agataagatt aacccgtgga agcccttaat 6420

agtcatggga gtcctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt 6480

aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg 6540

aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga 6600

gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720

tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca ccccctggga 6840

cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900

cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcactgatgc 6960

aggaaaaaag gctaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg 7020

aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt 7080

ccccataggg cccaacccag tattacccga ccaaagactc ccttcctcac caatagagat 7140

tgtaccggct ccacagccac ctagcccccct caataccagt tacccccctt ccactaccag 7200

tacaccctca acctccccta caagtccaag tgtcccacag ccacccccag gaactggaga 7260

tagactacta gctctagtca aaggagccta tcaggcgctt aacctcacca atcccgacaa 7320

gacccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gagtagcggt 7380

cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttcccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccacccaaag cgccggctca ggatcctact accttgcagc 7560

[0043]

acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt 7620

gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680

ctccccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat 7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttaa tagatccccc tggtttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220

caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280

tcagcaatat caccagctaa aacccataga gtacgagcca tgaacgcgtt actggccgaa 8340

gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt attttccacc atattgccgt 8400

cttttggcaa tgtgagggcc cggaaacctg gccctgtctt cttgacgagc attcctaggg 8460

gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc 8520

ctctggaagc ttcttgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaacc 8580

ccccacctgg cgacaggtgc ctctgcggcc aaaagccacg tgtataagat acacctgcaa 8640

aggcggcaca accccagtgc cacgttgtga gttggatagt tgtggaaaga gtcaaatggc 8700

tctcctcaag cgtattcaac aaggggctga aggatgccca gaaggtaccc cattgtatgg 8760

gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac 8820

gtctaggccc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg 8880

tgacaggggg aatggcaagc aagtgggatc agaagggtat ggacattgcc tatgaggagg 8940

cggccttagg ttacaaagag ggtggtgttc ctattggcgg atgtcttatc aataacaaag 9000

acggaagtgt tctcggtcgt ggtcacaaca tgagatttca aaagggatcc gccacactac 9060

atggtgagat ctccactttg gaaaactgtg ggagattaga gggcaaagtg tacaaagata 9120

ccactttgta tacgacgctg tctccatgcg acatgtgtac aggtgccatc atcatgtatg 9180

gtattccacg ctgtgttgtc ggtgagaacg ttaatttcaa aagtaagggc gagaaatatt 9240

tacaaactag aggtcacgag gttgttgttg ttgacgatga gaggtgtaaa aagatcatga 9300

aacaatttat cgatgaaaga cctcaggatt ggtttgaaga tattggtgag taggcggccg 9360

cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc 9420

tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcatggaa aaatacataa 9480

ctgagaatag agaagttcag atcaaggtca ggaacagatg gaacagctga atatgggcca 9540

aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca 9600

gctgaatatg ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctcagggcca 9660

agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg 9720

tttccagggt gccccaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag 9780

ttcgcttctc gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa 9840

cccctcactc ggggcgccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca 9900

ataaaccctc ttgcagttgc atccgacttg tggtctcgct gttccttggg agggtctcct 9960

ctgagtgatt gactacccgt cagcgggggt ctttcattac atgtgagcaa aaggccagca 10020

[0044] aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc 10080

tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata 10140

aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 10200

gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcaatgctc 10260

acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 10320

acccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 10380

ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 10440

gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 10500

gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 10560

ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca 10620

gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 10680

cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 10740

cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 10800

gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 10860

tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga 10920

gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc 10980

agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac 11040

tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc 11100

agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt cacgctcgtc 11160

gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 11220

catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt 11280

ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 11340

atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 11400

tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag 11460

cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 11520

cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc 11580

atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 11640

aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 11700

[0045] ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 11760

aaataaacaa atagggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 11820

aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct 11880

tcaagaattc at 11892

<210> 21
<211> 12007
<212> DNA
<213> 人工序列

<220>
<223> RCR载体 - pACE-CD

<400> 21
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360

[0046]

catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 540

acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600

actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca tccgacttgt 660

ggtctcgctg ttccttggga gggtctcctc tgagtgattg actacccgtc agcgggggtc 720

tttcatttgg gggctcgtcc gggatcggga gacccctgcc cagggaccac cgacccacca 780

ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840

tgattttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggcggacccg 900

tggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960

cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020

gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080

tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140

tctgctgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaga 1200

atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260

agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

tcatcaccca ggttaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440

tcccctacat cgtgacctgg gaagccttgg cttttgaccc ccctccctgg gtcaagccct 1500

ttgtacaccc taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560

ctcctcgttc gaccccgcct cgatcctccc tttatccagc cctcactcct tctctaggcg 1620

ccaaacctaa acctcaagtt ctttctgaca gtgggggggcc gctcatcgac ctacttacag 1680

aagaccccccc gccttatagg gacccaagac cacccccttc cgacagggac ggaaatggtg 1740

gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800

ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag 1860

gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920

ataataaccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgttc 1980

tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

[0047]

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100

gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctcctagcgg 2220

gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340

ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc 2400

agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460

ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520

gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtcccaaga 2760

aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg 2820

gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac 2940

ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga 3000

ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactctttc ctccatgtac 3060

cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120

actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180

ccctaaatat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcctc aggcctgggc ggaaaccggg ggcatgggac 3300

tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

tgttggacca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480

ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540

gggtggaaga catccacccc accgtgccca acccttacaa cctcttgagc gggctcccac 3600

cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720

[0048]

gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca cttctgagct agactgccaa caaggtactc 3900

gggccctgtt acaaacccta gggaacctcg ggtatcgggc ctcggccaag aaagcccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgaggccag aaaagagact gtgatggggc agcctactcc gaagacccct cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgcctac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag 4500

aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560

aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg 4620

ctacgctgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740

ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040

taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100

gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160

tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcatcaccc tacacctcag 5220

aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg 5280

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

[0049]

tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520

gaactagggt ccgcgggcat cggcccggca ctcattggga gatcgatttc accgagataa 5580

agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640

tagaagcctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gagtcagaca gtggccgatc tgttggggat tgattggaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt 5880

taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940

tgtaccgagc ccgcaacacg ccgggcccccc atggcctcac cccatatgag atcttatatg 6000

gggcaccccc gcccccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060

cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg 6180

gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga 6300

tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

ggcgcgttca acgctctcaa aacccccctca agataagatt aacccgtgga agcccttaat 6420

agtcatggga gtcctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt 6480

aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg 6540

aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga 6600

gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720

tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca ccccctggga 6840

cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900

cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcactgatgc 6960

aggaaaaaag gctaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg 7020

aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt 7080

ccccataggg cccaacccag tattacccga ccaaagactc ccttcctcac caatagagat 7140

tgtaccggct ccacagccac ctagcccccc caataccagt tacccccctt ccactaccag 7200

tacaccctca acctccccta caagtccaag tgtcccacag ccacccccag gaactggaga 7260

tagactacta gctctagtca aaggagccta tcaggcgctt aacctcacca atcccgacaa 7320

gacccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gagtagcggt 7380

cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttcccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccacccaaag cgccggctca ggatcctact accttgcagc 7560

acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt 7620

gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680

ctcccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat 7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat 7860

[0050] ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttaa tagatccccc tggtttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220

caatcgatta gtccaatttg ttaaagacag gatatcagtg gtccaggctc tagttttgac 8280

tcaacaatat caccagctga agcctataga gtacgagcca tgacgtacgt tactggccga 8340

agccgcttgg aataaggccg gtgtgcgttt gtctatatgt tattttccac catattgccg 8400

tcttttggca atgtgagggc ccggaaacct ggccctgtct tcttgacgag cattcctagg 8460

ggtctttccc ctctcgccaa aggaatgcaa ggtctgttga atgtcgtgaa ggaagcagtt 8520

cctctggaag cttcttgaag acaaacaacg tctgtagcga ccctttgcag gcagcggaac 8580

cccccacctg gcgacaggtg cctctgcggc caaaagccac gtgtataaga tacacctgca 8640

aaggcggcac aaccccagtg ccacgttgtg agttggatag ttgtggaaag agtcaaatgg 8700

ctctcctcaa gcgtattcaa caaggggctg aaggatgccc agaaggtacc ccattgtatg 8760

[0051]

ggatctgatc tggggcctcg gtgcacatgc tttacatgtg tttagtcgag gttaaaaaaa 8820

cgtctaggcc ccccgaacca cggggacgtg gttttccttt gaaaaacacg ataataccat 8880

ggtgacaggg ggaatggcaa gcaagtggga tcagaagggt atggacattg cctatgagga 8940

ggcggcctta ggttacaaag agggtggtgt tcctattggc ggatgtctta tcaataacaa 9000

agacggaagt gttctcggtc gtggtcacaa catgagattt caaaagggat ccgccacact 9060

acatggtgag atctccactt tggaaaactg tgggagatta gagggcaaag tgtacaaaga 9120

taccactttg tatacgacgc tgtctccatg cgacatgtgt acaggtgcca tcatcatgta 9180

tggtattcca cgctgtgttg tcggtgagaa cgttaatttc aaaagtaagg gcgagaaata 9240

tttacaaact agaggtcacg aggttgttgt tgttgacgat gagaggtgta aaaagatcat 9300

gaaacaattt atcgatgaaa gacctcagga ttggtttgaa gatattggtg agtaggcggc 9360

cgcgccatag ataaaataaa agattttatt tagtctccag aaaaaggggg gaatgaaaga 9420

ccccacctgt aggtttggca agctagctta agtaacgcca ttttgcaagg catggaaaaa 9480

tacataactg agaatagaga agttcagatc aaggtcagga acagatggaa cagctgaata 9540

tgggccaaac aggatatctg tggtaagcag ttcctgcccc ggctcagggc caagaacaga 9600

tggaacagct gaatatgggc caaacaggat atctgtggta agcagttcct gccccggctc 9660

agggccaaga acagatggtc cccagatgcg gtccagccct cagcagtttc tagagaacca 9720

tcagatgttt ccagggtgcc ccaaggacct gaaatgaccc tgtgccttgt ttaaactaac 9780

caatcagttc gcttctcgct tctgttcgcg cgcttctgct ccccgagctc aataaaagag 9840

cccacaaccc ctcactcggg gcgccagtcc tccgattgac tgagtcgccc gggtacccgt 9900

gtatccaata aaccctcttg cagttgcatc cgacttgtgg tctcgctgtt ccttgggagg 9960

gtctcctctg agtgattgac tacccgtcag cgggggtctt tcatttgggg gctcgtccgg 10020

gatcgggaga cccctgccca gggaccaccg acccaccacc gggaggtaag ctggctgcct 10080

cgcgcgtttc ggtgatgacg gtgaaaacct ctgacatgtg agcaaaaggc cagcaaaagg 10140

ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccccctgacg 10200

agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 10260

accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta 10320

ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct 10380

gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 10440

ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa 10500

gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 10560

taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag 10620

tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 10680

gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 10740

cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 10800

agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 10860

cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 10920

cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 10980

ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 11040

taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 11100

tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat 11160

ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta 11220

[0052] atagtttgcg caacgttgtt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 11280

gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt 11340

tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg 11400

cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 11460

taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 11520

ggcgaccgag ttgctcttgc ccggcgtcaa cacgggataa taccgcgcca catagcagaa 11580

ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac 11640

cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt 11700

ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 11760

gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 11820

gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 11880

aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 11940

ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtcttcaag 12000

aattcat 12007

<210> 22
<211> 11893
<212> DNA
<213> 人工序列

<220>
<223> RCR载体 - pAC3-yCD2

<400> 22

[0053]

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360

catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 540

acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600

actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca tccgacttgt 660

ggtctcgctg ttccttggga gggtctcctc tgagtgattg actacccgtc agcgggggtc 720

tttcatttgg gggctcgtcc gggatcggga gacccctgcc cagggaccac cgacccacca 780

ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840

tgattttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggcggacccg 900

tggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960

cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020

gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080

tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140

tctgctgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaaa 1200

atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260

agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

tcatcaccca ggttaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440

[0054]

tcccctacat cgtgacctgg gaagccttgg cttttgaccc ccctccctgg gtcaagccct 1500

ttgtacaccc taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560

ctcctcgttc gaccccgcct cgatcctccc tttatccagc cctcactcct tctctaggcg 1620

ccaaacctaa acctcaagtt ctttctgaca gtgggggggcc gctcatcgac ctacttacag 1680

aagacccccc gccttatagg gacccaagac cacccccttc cgacagggac ggaaatggtg 1740

gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800

ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag 1860

gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920

ataataaccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgtcc 1980

tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100

gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctcctagcgg 2220

gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340

ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc 2400

agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460

ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520

gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtcccaaga 2760

aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg 2820

gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc gggggcaac 2880

ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac 2940

ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga 3000

ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactctttc ctccatgtac 3060

cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120

[0055]

actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180

ccctaaatat agaagatgag tatcggctac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcctc aggcctgggc ggaaaccggg ggcatgggac 3300

tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

tgttggacca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480

ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540

gggtggaaga catccacccc accgtgccca acccttacaa cctcttgagc gggctcccac 3600

cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720

gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca cttctgagct agactgccaa caaggtactc 3900

gggccctgtt acaaacccta gggaacctcg ggtatcgggc ctcggccaag aaagcccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccccct cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgcctac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag 4500

aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560

aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg 4620

ctacgctgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740

ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

[0056]

ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040

taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100

gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160

tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcatcaccc tacacctcag 5220

aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg 5280

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520

gaactagggt ccgcgggcat cggcccggca ctcattggga gatcgatttc accgagataa 5580

agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640

tagaagcctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gagtcagaca gtggccgatc tgttggggat tgattggaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt 5880

taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940

tgtaccgagc ccgcaacacg ccgggcccca atggcctcac cccatatgag atcttatatg 6000

gggcaccccc gcccccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060

cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg 6180

gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga 6300

tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

ggcgcgttca acgctctcaa aacccccctca agataagatt aacccgtgga agcccttaat 6420

agtcatggga gtcctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt 6480

[0057]

aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg 6540

aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga 6600

gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720

tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca ccccctggga 6840

cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900

cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcactgatgc 6960

aggaaaaaag gctaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg 7020

aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt 7080

ccccataggg cccaacccag tattacccga ccaaagactc ccttcctcac caatagagat 7140

tgtaccggct ccacagccac ctagcccccc caataccagt tacccccctt ccactaccag 7200

tacaccctca acctccccta caagtccaag tgtcccacag ccacccccag gaactggaga 7260

tagactacta gctctagtca aaggagccta tcaggcgctt aacctcacca atcccgacaa 7320

gacccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gagtagcggt 7380

cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttcccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccacccaaag cgccggctca ggatcctact accttgcagc 7560

acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt 7620

gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680

ctcccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat 7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttaa tagatccccc tggtttacca ccttaatctc 8160

[0058]

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220

caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280

tcagcaatat caccagctaa aacccataga gtacgagcca tgaacgcgtt actggccgaa 8340

gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt attttccacc atattgccgt 8400

cttttggcaa tgtgagggcc cggaaacctg gccctgtctt cttgacgagc attcctaggg 8460

gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc 8520

ctctggaagc ttcttgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaacc 8580

ccccacctgg cgacaggtgc ctctgcggcc aaaagccacg tgtataagat acacctgcaa 8640

aggcggcaca accccagtgc cacgttgtga gttggatagt tgtggaaaga gtcaaatggc 8700

tctcctcaag cgtattcaac aaggggctga aggatgccca gaaggtaccc cattgtatgg 8760

gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac 8820

gtctaggccc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg 8880

tgaccggcgg catggcctcc aagtgggatc aaaagggcat ggatatcgct tacgaggagg 8940

ccctgctggg ctacaaggag ggcggcgtgc ctatcggcgg ctgtctgatc aacaacaagg 9000

acggcagtgt gctgggcagg ggccacaaca tgaggttcca gaagggctcc gccaccctgc 9060

acggcgagat ctccaccctg gagaactgtg gcaggctgga gggcaaggtg tacaaggaca 9120

ccaccctgta caccaccctg tccccttgtg acatgtgtac cggcgctatc atcatgtacg 9180

gcatccctag gtgtgtgatc ggcgagaacg tgaacttcaa gtccaagggc gagaagtacc 9240

tgcaaaccag gggccacgag gtggtggttg ttgacgatga gaggtgtaag aagctgatga 9300

agcagttcat cgacgagagg cctcaggact ggttcgagga tatcggcgag taagcggccg 9360

cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc 9420

tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcatggaa aaatacataa 9480

ctgagaatag agaagttcag atcaaggtca ggaacagatg gaacagctga atatgggcca 9540

aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca 9600

gctgaatatg ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctcagggcca 9660

agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg 9720

tttccagggt gccccaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag 9780

ttcgcttctc gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa 9840

[0059]

ccctcactc ggggcgccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca 9900

ataaaccctc ttgcagttgc atccgacttg tggtctcgct gttccttggg agggtctcct 9960

ctgagtgatt gactacccgt cagcgggggt ctttcattac atgtgagcaa aaggccagca 10020

aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc 10080

tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata 10140

aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 10200

gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc 10260

acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 10320

acccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 10380

ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 10440

gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 10500

gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 10560

ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca 10620

gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 10680

cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 10740

cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 10800

gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 10860

tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga 10920

gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc 10980

agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac 11040

tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc 11100

agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt cacgctcgtc 11160

gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 11220

catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt 11280

ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 11340

atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 11400

tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag 11460

cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 11520

cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc 11580

atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 11640

aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 11700

ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 11760

aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 11820

aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct 11880

tcaagaattc cat 11893

<210> 23
<211> 4473
<212> DNA
<213> 智人

<220>
<221> CDS
<222> (175)..(3942)

<400> 23
aaggggaggt aaccctggcc cctttggtcg gggccccggg cagccgcgcg ccccttccca 60

[0060]

cggggccctt tactgcgccg cgcgcccggc ccccaccccc cgcagcaccc cgcgccccgc 120

gccctcccag ccgggtccag ccggagccat ggggccggag ccgcagtgag cacc atg 177
Met
1

gag ctg gcg gcc ttg tgc cgc tgg ggg ctc ctc ctc gcc ctc ttg ccc 225
Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu Pro
5 10 15

ccc gga gcc gcg agc acc caa gtg tgc acc ggc aca gac atg aag ctg 273
Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys Leu
20 25 30

cgg ctc cct gcc agt ccc gag acc cac ctg gac atg ctc cgc cac ctc 321
Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His Leu
35 40 45

tac cag ggc tgc cag gtg gtg cag gga aac ctg gaa ctc acc tac ctg 369
Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr Leu
50 55 60 65

ccc acc aat gcc agc ctg tcc ttc ctg cag gat atc cag gag gtg cag 417
Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val Gln
70 75 80

ggc tac gtg ctc atc gct cac aac caa gtg agg cag gtc cca ctg cag 465
Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu Gln

```
Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu Gln
            85                  90                  95

agg ctg cgg att gtg cga ggc acc cag ctc ttt gag gac aac tat gcc      513
Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr Ala
        100                 105                 110

ctg gcc gtg cta gac aat gga gac ccg ctg aac aat acc acc cct gtc      561
Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro Val
    115                 120                 125

aca ggg gcc tcc cca gga ggc ctg cgg gag ctg cag ctt cga agc ctc      609
Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser Leu
130                 135                 140                 145

aca gag atc ttg aaa gga ggg gtc ttg atc cag cgg aac ccc cag ctc      657
Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln Leu
                150                 155                 160

tgc tac cag gac acg att ttg tgg aag gac atc ttc cac aag aac aac      705
Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn Asn
            165                 170                 175

cag ctg gct ctc aca ctg ata gac acc aac cgc tct cgg gcc tgc cac      753
Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys His
        180                 185                 190

[0061]
ccc tgt tct ccg atg tgt aag ggc tcc cgc tgc tgg gga gag agt tct      801
Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser Ser
    195                 200                 205

gag gat tgt cag agc ctg acg cgc act gtc tgt gcc ggt ggc tgt gcc      849
Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys Ala
210                 215                 220                 225

cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat gag cag tgt gct      897
Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys Ala
                230                 235                 240

gcc ggc tgc acg ggc ccc aag cac tct gac tgc ctg gcc tgc ctc cac      945
Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu His
            245                 250                 255

ttc aac cac agt ggc atc tgt gag ctg cac tgc cca gcc ctg gtc acc      993
Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val Thr
        260                 265                 270

tac aac aca gac acg ttt gag tcc atg ccc aat ccc gag ggc cgg tat     1041
Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg Tyr
    275                 280                 285

aca ttc ggc gcc agc tgt gtg act gcc tgt ccc tac aac tac ctt tct     1089
Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu Ser
290                 295                 300                 305

acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg cac aac caa gag     1137
```

```
Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln Glu
                310                 315                 320

gtg aca gca gag gat gga aca cag cgg tgt gag aag tgc agc aag ccc   1185
Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys Pro
            325                 330                 335

tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac ttg cga gag gtg   1233
Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu Val
        340                 345                 350

agg gca gtt acc agt gcc aat atc cag gag ttt gct ggc tgc aag aag   1281
Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys Lys
    355                 360                 365

atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt gat ggg gac cca   1329
Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp Pro
370                 375                 380                 385

gcc tcc aac act gcc ccg ctc cag cca gag cag ctc caa gtg ttt gag   1377
Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe Glu
                390                 395                 400

act ctg gaa gag atc aca ggt tac cta tac atc tca gca tgg ccg gac   1425
Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro Asp
            405                 410                 415

[0062]
agc ctg cct gac ctc agc gtc ttc cag aac ctg caa gta atc cgg gga   1473
Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg Gly
        420                 425                 430

cga att ctg cac aat ggc gcc tac tcg ctg acc ctg caa ggg ctg ggc   1521
Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu Gly
    435                 440                 445

atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg ggc agt gga ctg   1569
Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly Leu
450                 455                 460                 465

gcc ctc atc cac cat aac acc cac ctc tgc ttc gtg cac acg gtg ccc   1617
Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val Pro
                470                 475                 480

tgg gac cag ctc ttt cgg aac ccg cac caa gct ctg ctc cac act gcc   1665
Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr Ala
            485                 490                 495

aac cgg cca gag gac gag tgt gtg ggc gag ggc ctg gcc tgc cac cag   1713
Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His Gln
        500                 505                 510

ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc acc cag tgt gtc   1761
Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys Val
    515                 520                 525

aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg gag gaa tgc cga   1809
```

```
Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys Arg
530                 535                 540                 545

gta ctg cag ggg ctc ccc agg gag tat gtg aat gcc agg cac tgt ttg      1857
Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys Leu
                550                 555                 560

ccg tgc cac cct gag tgt cag ccc cag aat ggc tca gtg acc tgt ttt      1905
Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys Phe
            565                 570                 575

gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac tat aag gac cct      1953
Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp Pro
        580                 585                 590

ccc ttc tgc gtg gcc cgc tgc ccc agc ggt gtg aaa cct gac ctc tcc      2001
Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu Ser
    595                 600                 605

tac atg ccc atc tgg aag ttt cca gat gag gag ggc gca tgc cag cct      2049
Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln Pro
610                 615                 620                 625

tgc ccc atc aac tgc acc cac tcc tgt gtg gac ctg gat gac aag ggc      2097
Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly
                630                 635                 640

[0063]
tgc ccc gcc gag cag aga gcc agc cct ctg acg tcc atc atc tct gcg      2145
Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser Ala
            645                 650                 655

gtg gtt ggc att ctg ctg gtc gtg gtc ttg ggg gtg gtc ttt ggg atc      2193
Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly Ile
        660                 665                 670

ctc atc aag cga cgg cag cag aag atc cgg aag tac acg atg cgg aga      2241
Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg Arg
    675                 680                 685

ctg ctg cag gaa acg gag ctg gtg gag ccg ctg aca cct agc gga gcg      2289
Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Ala
690                 695                 700                 705

atg ccc aac cag gcg cag atg cgg atc ctg aaa gag acg gag ctg agg      2337
Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu Arg
                710                 715                 720

aag gtg aag gtg ctt gga tct ggc gct ttt ggc aca gtc tac aag ggc      2385
Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly
            725                 730                 735

atc tgg atc cct gat ggg gag aat gtg aaa att cca gtg gcc atc aaa      2433
Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile Lys
        740                 745                 750

gtg ttg agg gaa aac aca tcc ccc aaa gcc aac aaa gaa atc tta gac      2481
```

```
Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp
    755                 760                 765

gaa gca tac gtg atg gct ggt gtg ggc tcc cca tat gtc tcc cgc ctt      2529
Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg Leu
770                 775                 780                 785

ctg ggc atc tgc ctg aca tcc acg gtg cag ctg gtg aca cag ctt atg      2577
Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu Met
                790                 795                 800

ccc tat ggc tgc ctc tta gac cat gtc cgg gaa aac cgc gga cgc ctg      2625
Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg Leu
            805                 810                 815

ggc tcc cag gac ctg ctg aac tgg tgt atg cag att gcc aag ggg atg      2673
Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly Met
        820                 825                 830

agc tac ctg gag gat gtg cgg ctc gta cac agg gac ttg gcc gct cgg      2721
Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala Arg
    835                 840                 845

aac gtg ctg gtc aag agt ccc aac cat gtc aaa att aca gac ttc ggg      2769
Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly
850                 855                 860                 865
```

[0064]

```
ctg gct cgg ctg ctg gac att gac gag aca gag tac cat gca gat ggg      2817
Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp Gly
                870                 875                 880

ggc aag gtg ccc atc aag tgg atg gcg ctg gag tcc att ctc cgc cgg      2865
Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg Arg
            885                 890                 895

cgg ttc acc cac cag agt gat gtg tgg agt tat ggt gtg act gtg tgg      2913
Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp
        900                 905                 910

gag ctg atg act ttt ggg gcc aaa cct tac gat ggg atc cca gcc cgg      2961
Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala Arg
    915                 920                 925

gag atc cct gac ctg ctg gaa aag ggg gag cgg ctg ccc cag ccc ccc      3009
Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro
930                 935                 940                 945

atc tgc acc att gat gtc tac atg atc atg gtc aaa tgt tgg atg att      3057
Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile
                950                 955                 960

gac tct gaa tgt cgg cca aga ttc cgg gag ttg gtg tct gaa ttc tcc      3105
Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe Ser
            965                 970                 975

cgc atg gcc agg gac ccc cag cgc ttt gtg gtc atc cag aat gag gac      3153
```

Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu Asp
980 985 990

ttg ggc cca gcc agt ccc ttg gac agc acc ttc tac cgc tca ctg ctg 3201
Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu Leu
995 1000 1005

gag gac gat gac atg ggg gac ctg gtg gat gct gag gag tat ctg 3246
Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu
1010 1015 1020

gta ccc cag cag ggc ttc ttc tgt cca gac cct gcc ccg ggc gct 3291
Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala
1025 1030 1035

ggg ggc atg gtc cac cac agg cac cgc agc tca tct acc agg agt 3336
Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser
1040 1045 1050

ggc ggt ggg gac ctg aca cta ggg ctg gag ccc tct gaa gag gag 3381
Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu
1055 1060 1065

gcc ccc agg tct cca ctg gca ccc tcc gaa ggg gct ggc tcc gat 3426
Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp
1070 1075 1080

[0065] gta ttt gat ggt gac ctg gga atg ggg gca gcc aag ggg ctg caa 3471
Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln
1085 1090 1095

agc ctc ccc aca cat gac ccc agc cct cta cag cgg tac agt gag 3516
Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu
1100 1105 1110

gac ccc aca gta ccc ctg ccc tct gag act gat ggc tac gtt gcc 3561
Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val Ala
1115 1120 1125

ccc ctg acc tgc agc ccc cag cct gaa tat gtg aac cag cca gat 3606
Pro Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro Asp
1130 1135 1140

gtt cgg ccc cag ccc cct tcg ccc cga gag ggc cct ctg cct gct 3651
Val Arg Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro Ala
1145 1150 1155

gcc cga cct gct ggt gcc act ctg gaa agg ccc aag act ctc tcc 3696
Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Pro Lys Thr Leu Ser
1160 1165 1170

cca ggg aag aat ggg gtc gtc aaa gac gtt ttt gcc ttt ggg ggt 3741
Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala Phe Gly Gly
1175 1180 1185

gcc gtg gag aac ccc gag tac ttg aca ccc cag gga gga gct gcc 3786

Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly Ala Ala
1190 1195 1200

cct cag ccc cac cct cct cct gcc ttc agc cca gcc ttc gac aac 3831
Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp Asn
1205 1210 1215

ctc tat tac tgg gac cag gac cca cca gag cgg ggg gct cca ccc 3876
Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro Pro
1220 1225 1230

agc acc ttc aaa ggg aca cct acg gca gag aac cca gag tac ctg 3921
Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr Leu
1235 1240 1245

ggt ctg gac gtg cca gtg tga accagaaggc caagtccgca gaagccctga 3972
Gly Leu Asp Val Pro Val
1250 1255

tgtgtcctca gggagcaggg aaggcctgac ttctgctggc atcaagaggt gggagggccc 4032

tccgaccact tccaggggaa cctgccatgc caggaacctg tcctaaggaa ccttccttcc 4092

tgcttgagtt cccagatggc tggaaggggt ccagcctcgt tggaagagga acagcactgg 4152

ggagtctttg tggattctga ggccctgccc aatgagactc tagggtccag tggatgccac 4212

[0066] agcccagctt ggccctttcc ttccagatcc tgggtactga aagccttagg gaagctggcc 4272

tgagagggga agcggcccta agggagtgtc taagaacaaa agcgacccat tcagagactg 4332

tccctgaaac ctagtactgc cccccatgag gaaggaacag caatggtgtc agtatccagg 4392

ctttgtacag agtgcttttc tgtttagttt ttacttttt tgttttgttt ttttaaagat 4452

gaaataaaga cccaggggga g 4473

<210> 24
<211> 1255
<212> PRT
<213> 智人

<400> 24

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

[0067]

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
340 345 350

Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
355 360 365

Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
370 375 380

[0068]

Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385 390 395 400

Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
405 410 415

Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
420 425 430

Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
435 440 445

Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
450 455 460

Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465 470 475 480

Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
485 490 495

Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
500 505 510

Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
515 520 525

Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
530 535 540

Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545 550 555 560

Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
565 570 575

Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
580 585 590

Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
595 600 605

[0069]

Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
610 615 620

Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
625 630 635 640

Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser
645 650 655

Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
660 665 670

Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
675 680 685

Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
690 695 700

Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
705 710 715 720

Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
725 730 735

Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
740 745 750

Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
755 760 765

Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
770 775 780

Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
785 790 795 800

Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
805 810 815

Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
820 825 830

[0070]

Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
835 840 845

Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
850 855 860

Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865 870 875 880

Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
885 890 895

Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
900 905 910

Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
915 920 925

Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
930 935 940

Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945 950 955 960

Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
965 970 975

Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
980 985 990

Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
995 1000 1005

Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr
1010 1015 1020

Leu Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly
1025 1030 1035

Ala Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg
1040 1045 1050

[0071]

Ser Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu
1055 1060 1065

Glu Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser
1070 1075 1080

Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu
1085 1090 1095

Gln Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser
1100 1105 1110

Glu Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val
1115 1120 1125

Ala Pro Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro
1130 1135 1140

Asp Val Arg Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro
1145 1150 1155

```
Ala Ala  Arg Pro Ala Gly Ala  Thr Leu Glu Arg Pro  Lys Thr Leu
    1160                 1165                 1170

Ser Pro  Gly Lys Asn Gly Val  Val Lys Asp Val Phe  Ala Phe Gly
    1175                 1180                 1185

Gly Ala  Val Glu Asn Pro Glu  Tyr Leu Thr Pro Gln  Gly Gly Ala
    1190                 1195                 1200

Ala Pro  Gln Pro His Pro Pro  Pro Ala Phe Ser Pro  Ala Phe Asp
    1205                 1210                 1215

Asn Leu  Tyr Tyr Trp Asp Gln  Asp Pro Pro Glu Arg  Gly Ala Pro
    1220                 1225                 1230

Pro Ser  Thr Phe Lys Gly Thr  Pro Thr Ala Glu Asn  Pro Glu Tyr
    1235                 1240                 1245

Leu Gly  Leu Asp Val Pro Val
    1250                 1255
```

[0072]

<210> 25
<211> 1212
<212> DNA
<213> 智人

<220>
<221> CDS
<222> (1)..(1212)

<400> 25

```
atg aca gcc atc atc aaa gag atc gtt agc aga aac aaa agg aga tat          48
Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
1               5                  10                  15

caa gag gat gga ttc gac tta gac ttg acc tat att tat cca aac att          96
Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
            20                  25                  30

att gct atg gga ttt cct gca gaa aga ctt gaa ggc gta tac agg aac         144
Ile Ala Met Gly Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn
        35                  40                  45

aat att gat gat gta gta agg ttt ttg gat tca aag cat aaa aac cat         192
Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys His Lys Asn His
    50                  55                  60
```

```
tac aag ata tac aat ctt tgt gct gaa aga cat tat gac acc gcc aaa      240
Tyr Lys Ile Tyr Asn Leu Cys Ala Glu Arg His Tyr Asp Thr Ala Lys
65                  70                  75                  80

ttt aat tgc aga gtt gca caa tat cct ttt gaa gac cat aac cca cca      288
Phe Asn Cys Arg Val Ala Gln Tyr Pro Phe Glu Asp His Asn Pro Pro
                85                  90                  95

cag cta gaa ctt atc aaa ccc ttt tgt gaa gat ctt gac caa tgg cta      336
Gln Leu Glu Leu Ile Lys Pro Phe Cys Glu Asp Leu Asp Gln Trp Leu
            100                 105                 110

agt gaa gat gac aat cat gtt gca gca att cac tgt aaa gct gga aag      384
Ser Glu Asp Asp Asn His Val Ala Ala Ile His Cys Lys Ala Gly Lys
        115                 120                 125

gga cga act ggt gta atg ata tgt gca tat tta tta cat cgg ggc aaa      432
Gly Arg Thr Gly Val Met Ile Cys Ala Tyr Leu Leu His Arg Gly Lys
    130                 135                 140

ttt tta aag gca caa gag gcc cta gat ttc tat ggg gaa gta agg acc      480
Phe Leu Lys Ala Gln Glu Ala Leu Asp Phe Tyr Gly Glu Val Arg Thr
145                 150                 155                 160

aga gac aaa aag gga gta act att ccc agt cag agg cgc tat gtg tat      528
Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg Tyr Val Tyr
                165                 170                 175
```

[0073]

```
tat tat agc tac ctg tta aag aat cat ctg gat tat aga cca gtg gca      576
Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro Val Ala
            180                 185                 190

ctg ttg ttt cac aag atg atg ttt gaa act att cca atg ttc agt ggc      624
Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe Ser Gly
        195                 200                 205

gga act tgc aat cct cag ttt gtg gtc tgc cag cta aag gtg aag ata      672
Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val Lys Ile
    210                 215                 220

tat tcc tcc aat tca gga ccc aca cga cgg gaa gac aag ttc atg tac      720
Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe Met Tyr
225                 230                 235                 240

ttt gag ttc cct cag ccg tta cct gtg tgt ggt gat atc aaa gta gag      768
Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile Lys Val Glu
                245                 250                 255

ttc ttc cac aaa cag aac aag atg cta aaa aag gac aaa atg ttt cac      816
Phe Phe His Lys Gln Asn Lys Met Leu Lys Lys Asp Lys Met Phe His
            260                 265                 270

ttt tgg gta aat aca ttc ttc ata cca gga cca gag gaa acc tca gaa      864
Phe Trp Val Asn Thr Phe Phe Ile Pro Gly Pro Glu Glu Thr Ser Glu
        275                 280                 285
```

```
aaa gta gaa aat gga agt cta tgt gat caa gaa atc gat agc att tgc      912
Lys Val Glu Asn Gly Ser Leu Cys Asp Gln Glu Ile Asp Ser Ile Cys
    290                 295                 300

agt ata gag cgt gca gat aat gac aag gaa tat cta gta ctt act tta      960
Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu Val Leu Thr Leu
305                 310                 315                 320

aca aaa aat gat ctt gac aaa gca aat aaa gac aaa gcc aac cga tac     1008
Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys Ala Asn Arg Tyr
                325                 330                 335

ttt tct cca aat ttt aag gtg aag ctg tac ttc aca aaa aca gta gag     1056
Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr Lys Thr Val Glu
            340                 345                 350

gag ccg tca aat cca gag gct agc agt tca act tct gta aca cca gat     1104
Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser Val Thr Pro Asp
        355                 360                 365

gtt agt gac aat gaa cct gat cat tat aga tat tct gac acc act gac     1152
Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser Asp Thr Thr Asp
    370                 375                 380

tct gat cca gag aat gaa cct ttt gat gaa gat cag cat aca caa att     1200
Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu Asp Gln His Thr Gln Ile
385                 390                 395                 400
```

[0074]

```
aca aaa gtc tga                                                    1212
Thr Lys Val
```

<210> 26
<211> 403
<212> PRT
<213> 智人

<400> 26

```
Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
1               5                   10                  15

Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
            20                  25                  30

Ile Ala Met Gly Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn
        35                  40                  45

Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys His Lys Asn His
    50                  55                  60
```

Tyr Lys Ile Tyr Asn Leu Cys Ala Glu Arg His Tyr Asp Thr Ala Lys
65 70 75 80

Phe Asn Cys Arg Val Ala Gln Tyr Pro Phe Glu Asp His Asn Pro Pro
85 90 95

Gln Leu Glu Leu Ile Lys Pro Phe Cys Glu Asp Leu Asp Gln Trp Leu
100 105 110

Ser Glu Asp Asp Asn His Val Ala Ala Ile His Cys Lys Ala Gly Lys
115 120 125

Gly Arg Thr Gly Val Met Ile Cys Ala Tyr Leu Leu His Arg Gly Lys
130 135 140

Phe Leu Lys Ala Gln Glu Ala Leu Asp Phe Tyr Gly Glu Val Arg Thr
145 150 155 160

Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg Tyr Val Tyr
165 170 175

[0075]

Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro Val Ala
180 185 190

Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe Ser Gly
195 200 205

Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val Lys Ile
210 215 220

Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe Met Tyr
225 230 235 240

Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile Lys Val Glu
245 250 255

Phe Phe His Lys Gln Asn Lys Met Leu Lys Lys Asp Lys Met Phe His
260 265 270

Phe Trp Val Asn Thr Phe Phe Ile Pro Gly Pro Glu Glu Thr Ser Glu
275 280 285

Lys Val Glu Asn Gly Ser Leu Cys Asp Gln Glu Ile Asp Ser Ile Cys
290 295 300

Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu Val Leu Thr Leu
305 310 315 320

Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys Ala Asn Arg Tyr
325 330 335

Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr Lys Thr Val Glu
340 345 350

Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser Val Thr Pro Asp
355 360 365

Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser Asp Thr Thr Asp
370 375 380

Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu Asp Gln His Thr Gln Ile
385 390 395 400

[0076] Thr Lys Val

<210> 27
<211> 597
<212> DNA
<213> 智人

<220>
<221> CDS
<222> (1)..(597)

<400> 27
atg tca aac gtg cga gtg tct aac ggg agc cct agc ctg gag cgg atg 48
Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
1 5 10 15

gac gcc agg cag gcg gag cac ccc aag ccc tcg gcc tgc agg aac ctc 96
Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu
20 25 30

ttc ggc ccg gtg gac cac gaa gag tta acc cgg gac ttg gag aag cac 144
Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His
35 40 45

tgc aga gac atg gaa gag gcg agc cag cgc aag tgg aat ttc gat ttt 192

Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
50 55 60

cag aat cac aaa ccc cta gag ggc aag tac gag tgg caa gag gtg gag 240
Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
65 70 75 80

aag ggc agc ttg ccc gag ttc tac tac aga ccc ccg cgg ccc ccc aaa 288
Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
85 90 95

ggt gcc tgc aag gtg ccg gcg cag gag agc cag gat gtc agc ggg agc 336
Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
100 105 110

cgc ccg gcg gcg cct tta att ggg gct ccg gct aac tct gag gac acg 384
Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr
115 120 125

cat ttg gtg gac cca aag act gat ccg tcg gac agc cag acg ggg tta 432
His Leu Val Asp Pro Lys Thr Asp Pro Ser Asp Ser Gln Thr Gly Leu
130 135 140

gcg gag caa tgc gca gga ata agg aag cga cct gca acc gac gat tct 480
Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
145 150 155 160

[0077] tct act caa aac aaa aga gcc aac aga aca gaa gaa aat gtt tca gac 528
Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
165 170 175

ggt tcc cca aat gcc ggt tct gtg gag cag acg ccc aag aag cct ggc 576
Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly
180 185 190

ctc aga aga cgt caa acg taa 597
Leu Arg Arg Arg Gln Thr
195

<210> 28
<211> 198
<212> PRT
<213> 智人

<400> 28

Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
1 5 10 15

Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu
20 25 30

Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His

35 40 45

Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
50 55 60

Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
65 70 75 80

Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
85 90 95

Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
100 105 110

Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr
115 120 125

His Leu Val Asp Pro Lys Thr Asp Pro Ser Asp Ser Gln Thr Gly Leu
130 135 140

[0078] Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
145 150 155 160

Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
165 170 175

Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly
180 185 190

Leu Arg Arg Arg Gln Thr
195

<210> 29
<211> 894
<212> DNA
<213> 智人

<220>
<221> CDS
<222> (1)..(894)

<400> 29
atg aca aca ccc aga aat tca gta aat ggg act ttc ccg gca gag cca 48
Met Thr Thr Pro Arg Asn Ser Val Asn Gly Thr Phe Pro Ala Glu Pro

```
1               5                  10                  15

atg aaa ggc cct att gct atg caa tct ggt cca aaa cca ctc ttc agg      96
Met Lys Gly Pro Ile Ala Met Gln Ser Gly Pro Lys Pro Leu Phe Arg
            20                  25                  30

agg atg tct tca ctg gtg ggc ccc acg caa agc ttc ttc atg agg gaa     144
Arg Met Ser Ser Leu Val Gly Pro Thr Gln Ser Phe Phe Met Arg Glu
        35                  40                  45

tct aag act ttg ggg gct gtc cag att atg aat ggg ctc ttc cac att     192
Ser Lys Thr Leu Gly Ala Val Gln Ile Met Asn Gly Leu Phe His Ile
    50                  55                  60

gcc ctg ggg ggt ctt ctg atg atc cca gca ggg atc tat gca ccc atc     240
Ala Leu Gly Gly Leu Leu Met Ile Pro Ala Gly Ile Tyr Ala Pro Ile
65                  70                  75                  80

tgt gtg act gtg tgg tac cct ctc tgg gga ggc att atg tat att att     288
Cys Val Thr Val Trp Tyr Pro Leu Trp Gly Gly Ile Met Tyr Ile Ile
                85                  90                  95

tcc gga tca ctc ttg gca gca acg gag aaa aac tct agg aag tgt ttg     336
Ser Gly Ser Leu Leu Ala Ala Thr Glu Lys Asn Ser Arg Lys Cys Leu
            100                 105                 110

gtc aaa gga aaa atg ata atg aat tca ttg agc ctc ttt gct gcc att     384
Val Lys Gly Lys Met Ile Met Asn Ser Leu Ser Leu Phe Ala Ala Ile
        115                 120                 125

tct gga atg att ctt tca atc atg gac ata ctt aat att aaa att tcc     432
Ser Gly Met Ile Leu Ser Ile Met Asp Ile Leu Asn Ile Lys Ile Ser
    130                 135                 140

cat ttt tta aaa atg gag agt ctg aat ttt att aga gct cac aca cca     480
His Phe Leu Lys Met Glu Ser Leu Asn Phe Ile Arg Ala His Thr Pro
145                 150                 155                 160

tat att aac ata tac aac tgt gaa cca gct aat ccc tct gag aaa aac     528
Tyr Ile Asn Ile Tyr Asn Cys Glu Pro Ala Asn Pro Ser Glu Lys Asn
                165                 170                 175

tcc cca tct acc caa tac tgt tac agc ata caa tct ctg ttc ttg ggc     576
Ser Pro Ser Thr Gln Tyr Cys Tyr Ser Ile Gln Ser Leu Phe Leu Gly
            180                 185                 190

att ttg tca gtg atg ctg atc ttt gcc ttc ttc cag gaa ctt gta ata     624
Ile Leu Ser Val Met Leu Ile Phe Ala Phe Phe Gln Glu Leu Val Ile
        195                 200                 205

gct ggc atc gtt gag aat gaa tgg aaa aga acg tgc tcc aga ccc aaa     672
Ala Gly Ile Val Glu Asn Glu Trp Lys Arg Thr Cys Ser Arg Pro Lys
    210                 215                 220

tct aac ata gtt ctc ctg tca gca gaa gaa aaa aaa gaa cag act att     720
Ser Asn Ile Val Leu Leu Ser Ala Glu Glu Lys Lys Glu Gln Thr Ile
```

[0079]

225 230 235 240

gaa ata aaa gaa gaa gtg gtt ggg cta act gaa aca tct tcc caa cca 768
Glu Ile Lys Glu Glu Val Val Gly Leu Thr Glu Thr Ser Ser Gln Pro
245 250 255

aag aat gaa gaa gac att gaa att att cca atc caa gaa gag gaa gaa 816
Lys Asn Glu Glu Asp Ile Glu Ile Ile Pro Ile Gln Glu Glu Glu Glu
260 265 270

gaa gaa aca gag acg aac ttt cca gaa cct ccc caa gat cag gaa tcc 864
Glu Glu Thr Glu Thr Asn Phe Pro Glu Pro Pro Gln Asp Gln Glu Ser
275 280 285

tca cca ata gaa aat gac agc tct cct taa 894
Ser Pro Ile Glu Asn Asp Ser Ser Pro
290 295

<210> 30
<211> 297
<212> PRT
<213> 智人

<400> 30

[0080]

Met Thr Thr Pro Arg Asn Ser Val Asn Gly Thr Phe Pro Ala Glu Pro
1 5 10 15

Met Lys Gly Pro Ile Ala Met Gln Ser Gly Pro Lys Pro Leu Phe Arg
20 25 30

Arg Met Ser Ser Leu Val Gly Pro Thr Gln Ser Phe Phe Met Arg Glu
35 40 45

Ser Lys Thr Leu Gly Ala Val Gln Ile Met Asn Gly Leu Phe His Ile
50 55 60

Ala Leu Gly Gly Leu Leu Met Ile Pro Ala Gly Ile Tyr Ala Pro Ile
65 70 75 80

Cys Val Thr Val Trp Tyr Pro Leu Trp Gly Gly Ile Met Tyr Ile Ile
85 90 95

Ser Gly Ser Leu Leu Ala Ala Thr Glu Lys Asn Ser Arg Lys Cys Leu
100 105 110

Val Lys Gly Lys Met Ile Met Asn Ser Leu Ser Leu Phe Ala Ala Ile
115 120 125

Ser Gly Met Ile Leu Ser Ile Met Asp Ile Leu Asn Ile Lys Ile Ser
130 135 140

His Phe Leu Lys Met Glu Ser Leu Asn Phe Ile Arg Ala His Thr Pro
145 150 155 160

Tyr Ile Asn Ile Tyr Asn Cys Glu Pro Ala Asn Pro Ser Glu Lys Asn
165 170 175

Ser Pro Ser Thr Gln Tyr Cys Tyr Ser Ile Gln Ser Leu Phe Leu Gly
180 185 190

Ile Leu Ser Val Met Leu Ile Phe Ala Phe Phe Gln Glu Leu Val Ile
195 200 205

[0081] Ala Gly Ile Val Glu Asn Glu Trp Lys Arg Thr Cys Ser Arg Pro Lys
210 215 220

Ser Asn Ile Val Leu Leu Ser Ala Glu Glu Lys Lys Glu Gln Thr Ile
225 230 235 240

Glu Ile Lys Glu Glu Val Val Gly Leu Thr Glu Thr Ser Ser Gln Pro
245 250 255

Lys Asn Glu Glu Asp Ile Glu Ile Ile Pro Ile Gln Glu Glu Glu Glu
260 265 270

Glu Glu Thr Glu Thr Asn Phe Pro Glu Pro Pro Gln Asp Gln Glu Ser
275 280 285

Ser Pro Ile Glu Asn Asp Ser Ser Pro
290 295

```
SEQ ID NO:2:   1  MVTGGMASKWDQKGMDIAYEEAALGYKEGGVPIGGCLINNKDGSVLGRGHNMRFQKGSAT
SEQ ID NO:4:   1  MVTGGMASKWDQKGMDIAYEEALLGYKEGGVPIGGCLINNKDGSVLGRGHNMRFQKGSAT

SEQ ID NO:2    61  LHGEISTLENCGRLEGKVYKDTTLYTTLSPCDMCTGAIIMYGIPRCVVGENVNFKSKGEK
SEQ ID NO:4    61  LHGEISTLENCGRLEGKVYKDTTLYTTLSPCDMCTGAIIMYGIPRCVIGENVNFKSKGEK

SEQ ID NO:2:  121  YLQTRGHEVVVVDDERCKKIMKQFIDERPQDWFEDIGE-
SEQ ID NO:4:  121  YLQTRGHEVVVVDDERCKKLMKQFIDERPQDWFEDIGE-

SEO ID NO: 31
长度:    596 nt
类型: DNA pri-miR-128-1
生物体：人类
ACGCGTactggagtcaatgaaagcaaCTATTTCAAAAGATCAGATTACTTACCAGTTTCACTAATAAAGATTTAT
TACTTTAAACCTTTATCATAAAATGTATGCTTTGAATACTGTGAAGTACACTGCATATAAGGAGTGTGGTATAGT
ATAAAGAAACTTTCTGCAGGTAGTAATTATAGTGAAGATTTTAGGTTTACAAAGCCCTAGCTGTTTTCTGTGTAG
CTTTTATTATTCTTATGACTCTTGACAAGTTTGTAGCTTCACCATATACATTTAATATTTTGCAATAATTGGCCT
TGTTCCTGAGCTGTTGGATTCGGGGCCGTAGCACTGTCTGAGAGGTTTACATTTCTCACAGTGAACCGGTCTCTT
TTTCAGCTGCTTCCTGGCTTCTTTTTACTCAGGTTTCCACTGCTTTTTTGCTTTTTTTAATGCTGTATGAAGGTG
TTAACATTTGTTTATATTTTTCATTAATTGTAATACCTTTAAATCATGCATCATACTCAGAAATAGGGATTAGAA
TTTAAGTGACATCTTTGGCCTAATATAATTTACCTGTTAAAaatttgtgaaagctattgcttaGCGGCCGC

SEO ID No.32
长度:    511 nt
类型: DNA pri-miR-128-2
生物体：人类
ACGCGTCCATGTCCGTACCTTTCTAGTTCATACCTTCTTTTAATTTTTTTTTTCTTTTCAATTTGAAGAGAGTGC
TTCCTCTGTTCTTAAGGCTAGGGAACCAAATTAGGTTGTTTCAATATCGTGCTAAAAGATACTGCCTTTAGAAGA
AGGCTATTGACAATCCAGCGTGTCTCGGTGGAACTCTGACTCCATGGTTCACTTTCATGATGGCCACATGCCTCC
TGCCCAGAGCCCGGCAGCCACTGTGCAGTGGGAAGGGGGGCCGATACACTGTACGAGAGTGAGTAGCAGGTCTCA
CAGTGAACCGGTCTCTTTCCCTACTGTGTCACACTCCTAATGGAATGCCGTTATCCAAAGAGCAGCACGAACCCG
ACAGGGCTGAGTGGCTTGTGCTAGGGAGAGGTTTGTGTCATTCCTGCTGACCAAACTGCAGGAAAAACTGCTAAT
TGTCATGCTGAAGACTGCCTGACGGGGAGACTCTGCCTTCTGTAAGTAGGTCAGCGGCCGC
SEO ID NO;33
长度:    203 nt
类型: 与人类H1启动子连接的DNA pre-miR128-1
生物体：人类
acgcgtAATTCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGG
AATCTTATAAGTTCTGTATGAGACCACTCGGAtgagctgttggattcggggccgtagcactgtctgagaggttta
catttctcacagtgaaccggtctcttttttcagctgcttcTTTTTTgcggccgc
```

图1A

SEQ ID NO: 34
长度： 205 nt
类型：与人类H1启动子连接的DNA pre-miR128-1
生物体：人类
gcggccgcAATTCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTG
GGAATCTTATAAGTTCTGTATGAGACCACTCGGAtgagctgttggattcggggccgtagcactgtctgagaggtt
tacatttctcacagtgaaccggtctctttttcagctgcttcTTTTTTgcggccgc

序列 ID No. 35
长度： 45 bp
类型：miR-142-3p的DNA靶序列
生物体： N/A
gcggccgcGTCGACTCCATAAAGTAGGAAACACTACAgcggccgc

图1A(续)

序 列 ID No. 36

长度： 128 bp
类型：DNA靶序列四次重复序列miR-142-3pT4X
生物体： N/A
gcggccgcGTCGACTCCATAAAGTAGGAAACACTACAaccggtTCCAT
AAAGTAGGAAACACTACATCACTCCATAAAGTAGGAAACACTACAgcggccgc

序列 ID No. 37
长度： 1131 bp
类型： DNA HSV-TK
生物体：单纯疱疹病毒-1
ATGGCTTCGTACCCCGGCCATCAGCACGCGTCTGCGTTCGACCAGGCTGCGCGTTCTCGCGGCCATAGCAACCGA
CGTACGGCGTTGCGCCCTCGCCGGCAGCAAGAAGCCACGGAAGTCCGCCCGGAGCAGAAAATGCCCACGCTACTG
CGGGTTTATATAGACGGTCCCCACGGGATGGGGAAAACCACCACCACGCAACTGCTGGTGGCCCTGGGTTCGCGC
GACGATATCGTCTACGTACCCGAGCCGATGACTTACTGGCAGGTGCTGGGGGCTTCCGAGACAATCGCGAACATC
TACACCACACAACACCGCCTCGACCAGGGTGAGATATCGGCCGGGGACGCGGCGGTGGTAATGACAAGCGCCCAG
ATAACAATGGGCATGCCTTATGCCGTGACCGACGCCGTTCTGGCTCCTCATATCGGGGGGGAGGCTGGGAGCTCA
CATGCCCCGCCCCCGGCCCTCACCCTCATCTTCGACCGCCATCCCATCGCCGCCCTCCTGTGTTACCCGGCCGCG
CGATACCTTATGGGCAGCATGACCCCCCAGGCCGTGCTGGCGTTCGTGGCCCTCATCCCGCCGACCTTGCCCGGC
ACAAACATCGTGTTGGGGGCCCTTCCGGAGGACAGACACATCGACCGCCTGGCCAAACGCCAGCGCCCCGGCGAG
CGGCTTGACCTGGCTATGCTGGCCGCGATTCGCCGCGTTTACGAGCTGCTTGCCAATACGGTGCGGTATCTGCAG
GGCGGCGGGTCGTGGCGGGAGGATTGGGGACAGCTTTCGGGGACGGCCGTGCCGCCCCAGGGTGCCGAGCCCCAG
AGCAACGCGGGCCCACGACCCCATATCGGGGACACGTTATTTACCCTGTTTCGGGCCCCCGAGTTGCTGGCCCCC
AACGGCGACCTGTATAACGTGTTTGCCTGGGCCTTGGACGTCTTGGCCAAACGCCTCCGTCCCATGCACGTCTTT
ATCCTGGATTACGACCAATCGCCCGCCGGCTACCGGGACGCCCTGCTGCAACTTACCTCCGGGATGGTCCAGACC
CACGTCACCACCCCCGGCTCCATACCGACGATCTGCGACCTGGCGCGCACGTTTGCCCGGGAGATGGGGGAGGCT
AACTAA
序列 ID No. 38
长度： 501 bp
类型：DNA γ IFN
生物体：人类
ATGAAATATACAAGTTATATCTTGGCTTTTCAGCTCTGCATCGTTTTGGGTTCTCTTGGCTGTTACTGCCAGGAC
CATATGTAAAAGAAGCAGAAAACCTTAAGAAATATTTTAATGCAGGTCATTCAGATGTAGCGGATAATGGAACTC
TTTTCTTAGGCATTTTGAAGAATTGGAAAGAGGAGAGTGACAGAAAAATAATGCAGAGCCAAATTGTCTCCTTTT
ACTTCAAACTTTTTAAAAACTTTAAAGATGACCAGAGCATCCAAAAGAGTGTGGAGACCATCAAGGAAGACATGA
ATGTAAGTTTTTCAATAGCAACAAAAAGAAACGAGATGACTTCGAAAAGCTGACTAATTATTCGGTAACTGACTT
GAATGTCCAACGCAAAGCAATACATGAACTCATCCAAGTGATGGCTGAACTGTCGCCAGCAGCTAAAACAGGGAA
GCGAAAAAGGAGTCAGATGCTGTTTCGAGGTCGAAGAGCATCCCAGTAA

图1B

序列 ID No. 39
长度： 468 bp
类型：DNA γ IFN
生物体：小鼠

ATGAACGCTACACACTGCATCTTGGCTTTGCAGCTCTTCCTCATGGCTGTTTCTGGCTGTTACTGCCACGGCACA
GTCATTGAAAGCCTAGAAAGTCTGAATAACTATTTTAACTCAAGTGGCATAGATGTGGAAGAAAAGAGTCTCTTC
TTGGATATCTGGAGGAACTGGCAAAAGGATGGTGACATGAAAATCCTGCAGAGCCAGATTATCTCTTTCTACCTC
AGACTCTTTGAAGTCTTGAAAGACAATCAGGCCATCAGCAACAACATAAGCGTCATTGAATCACACCTGATTACT
ACCTTCTTCAGCAACAGCAAGGCGAAAAAGGATGCATTCATGAGTATTGCCAAGTTTGAGGTCAACAACCCACAG
GTCCAGCGCCAAGCATTCAATGAGCTCATCCGAGTGGTCCACCAGCTGTTGCCGGAATCCAGCCTCAGGAAGCGG
AAAAGGAGTCGCTGCTGA

图1B(续)

序列 ID No. 40

长度： 462 bp
类型： DNA IL-2
生物体：人类

ATGTACAGGATGCAACTCCTGTCTTGCATTGCACTAAGTCTTGCACTTGTCACAAACAGTGCACCTACTTCAAGT
TCTACAAAGAAAACACAGCTACAACTGGAGCATTTACTGCTGGATTTACAGATGATTTTGAATGGAATTAATAAT
TACAAGAATCCCAAACTCACCAGGATGCTCACATTTAAGTTTTACATGCCCAAGAAGGCCACAGAACTGAAACAT
CTTCAGTGTCTAGAAGAAGAACTCAAACCTCTGGAGGAAGTGCTAAATTTAGCTCAAAGCAAAAACTTTCACTTA
AGACCCAGGGACTTAATCAGCAATATCAACGTAATAGTTCTGGAACTAAAGGGATCTGAAACAACATTCATGTGT
GAATATGCTGATGAGACAGCAACCATTGTAGAATTTCTGAACAGATGGATTACCTTTTGTCAAAGCATCATCTCA
ACACTGACTTGA

图1C

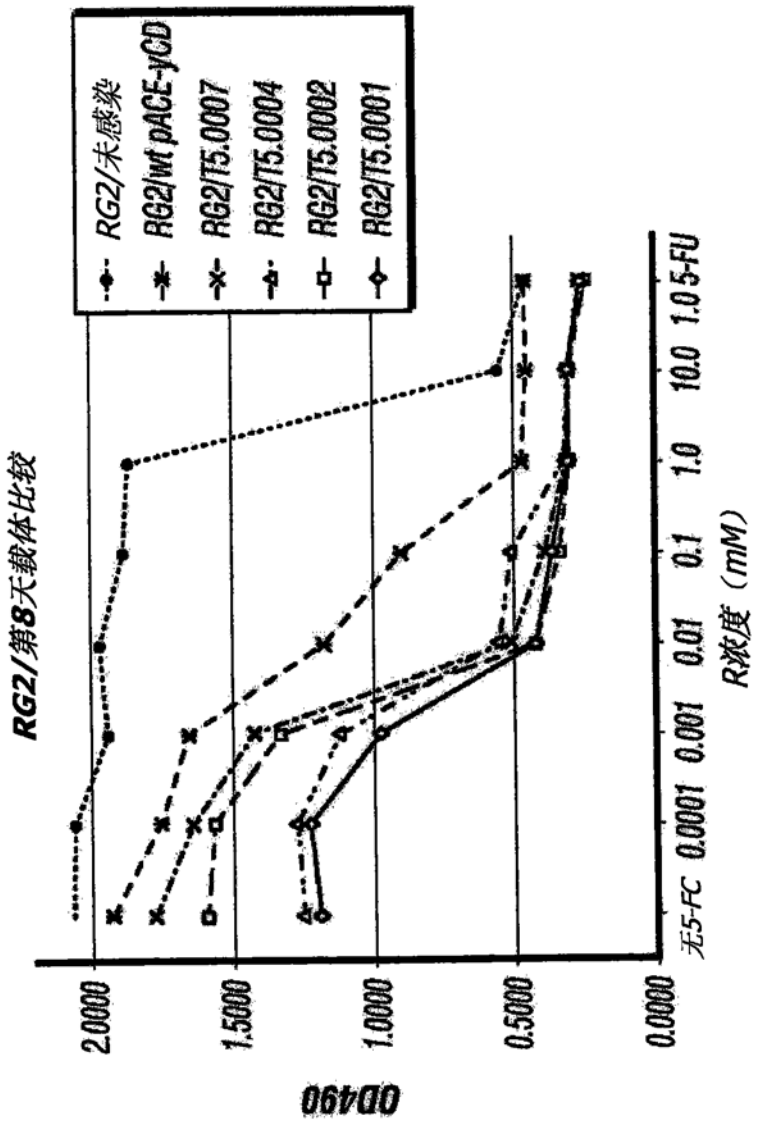
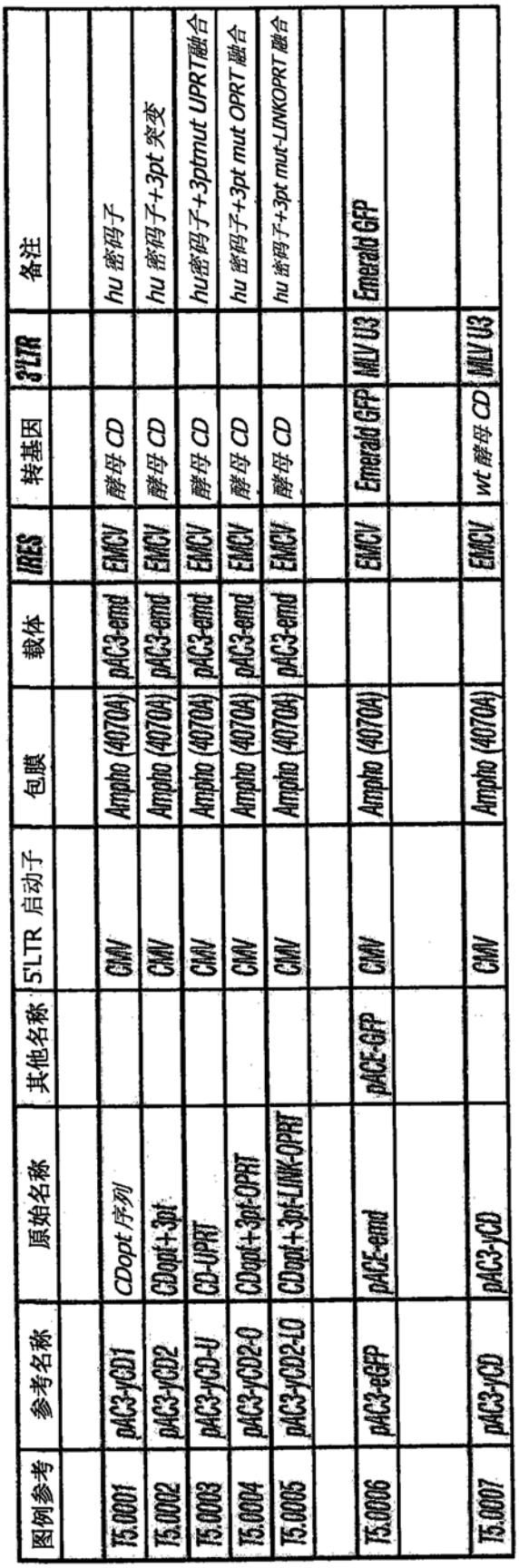

RG2/未感染
RG2/wt pACE-yCD
RG2/T5.0007
RG2/T5.0004
RG2/T5.0002
RG2/T5.0001
RG2/第8天载体比较
OD490
2.0000
1.5000
1.0000
0.5000
0.0000
无5-FC 0.0001 0.001 0.01 0.1 1.0 10.0 1.05-FU
R浓度（mM）

图2



图例参考	参考名称	原始名称	其他名称	5'LTR 启动子	包膜	载体	IRES	转基因	3'LTR	备注
T5.0001	pAC3-yCD1	CDopt 序列		CMV	Ampho (4070A)	pAC3-emd	EMCV	酵母 CD		hu 密码子
T5.0002	pAC3-yCD2	CDopt+3pt		CMV	Ampho (4070A)	pAC3-emd	EMCV	酵母 CD		hu 密码子+3pt 突变
T5.0003	pAC3-yCD-U	CD-UPRT		CMV	Ampho (4070A)	pAC3-emd	EMCV	酵母 CD		hu密码子+3ptmut UPRT融合
T5.0004	pAC3-yCD2-O	CDopt+3pt-OPRT		CMV	Ampho (4070A)	pAC3-emd	EMCV	酵母 CD		hu 密码子+3pt mut OPRT 融合
T5.0005	pAC3-yCD2-LO	CDopt+3pt-LINK-OPRT		CMV	Ampho (4070A)	pAC3-emd	EMCV	酵母 CD		hu 密码子+3pt mut-LINKOPRT 融合
T5.0006	pAC3-eGFP	pACE-emd	pACE-GFP	CMV	Ampho (4070A)		EMCV	Emerald GFP	MLV U3	Emerald GFP
T5.0007	pAC3-yCD	pAC3-yCD		CMV	Ampho (4070A)		EMCV	wt 酵母 CD	MLV U3	

图2(续)

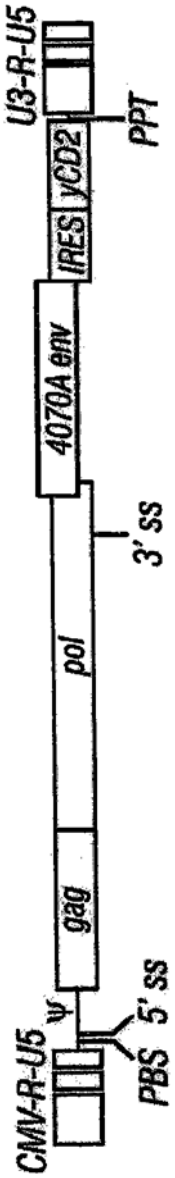


图3A

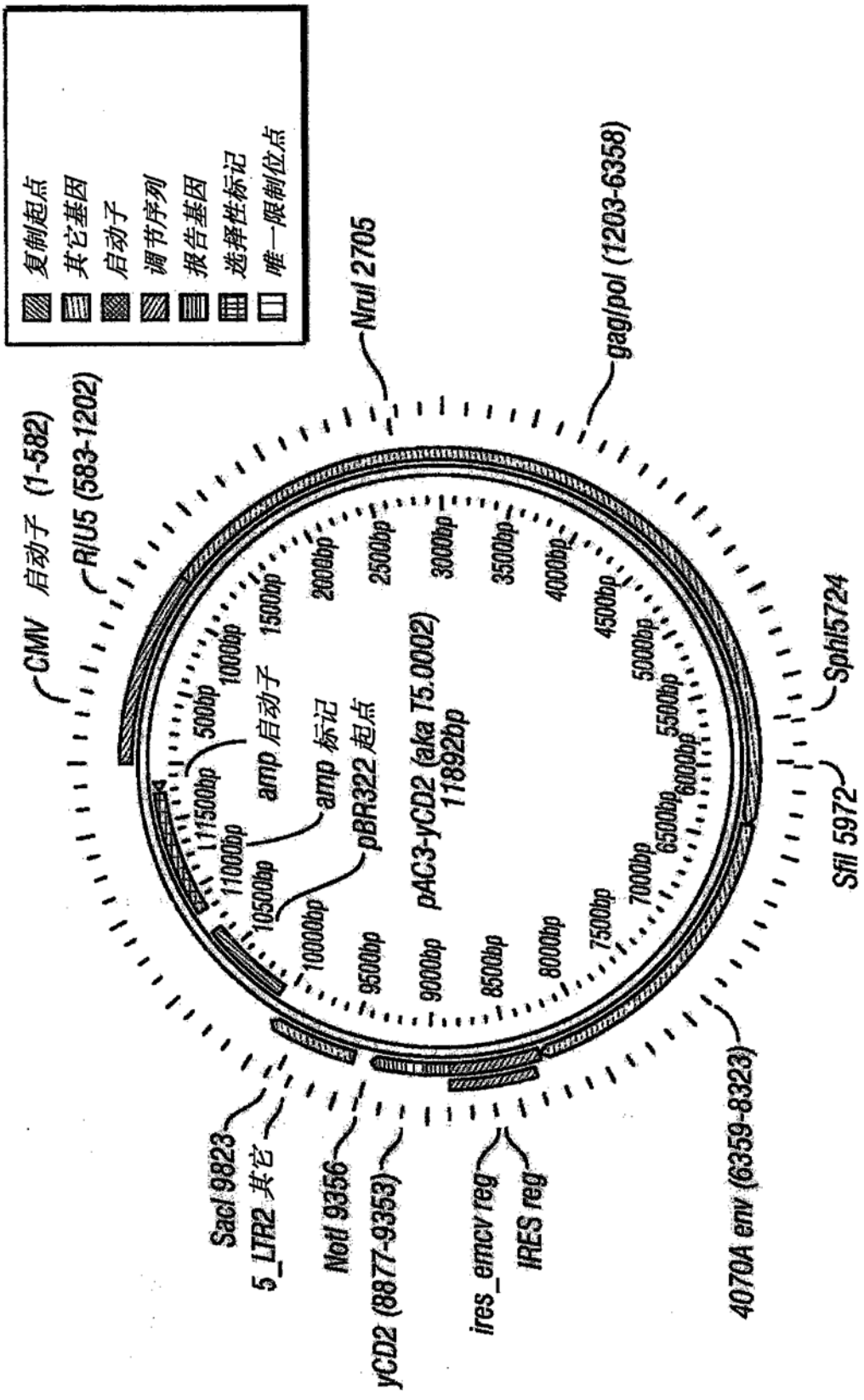

复制起点
其它基因
启动子
调节序列
报告基因
选择性标记
唯一限制位点
CMV 启动子 (1-582)
R/U5 (583-1202)
NruI 2705
gag/pol (1203-6358)
SphI5724
SfiI 5972
4070A env (6359-8323)
ires_emcv reg
IRES reg
yCD2 (8877-9353)
NotI 9356
5_LTR2 其它
SacI 9823
amp 启动子
amp 标记
pBR322 起点
pAC3-yCD2 (aka T5.0002)
11892bp
500bp
1000bp
1500bp
2000bp
2500bp
3000bp
3500bp
4000bp
4500bp
5000bp
5500bp
6000bp
6500bp
7000bp
7500bp
8000bp
8500bp
9000bp
9500bp
10000bp
10500bp
11000bp
11500bp

图3B

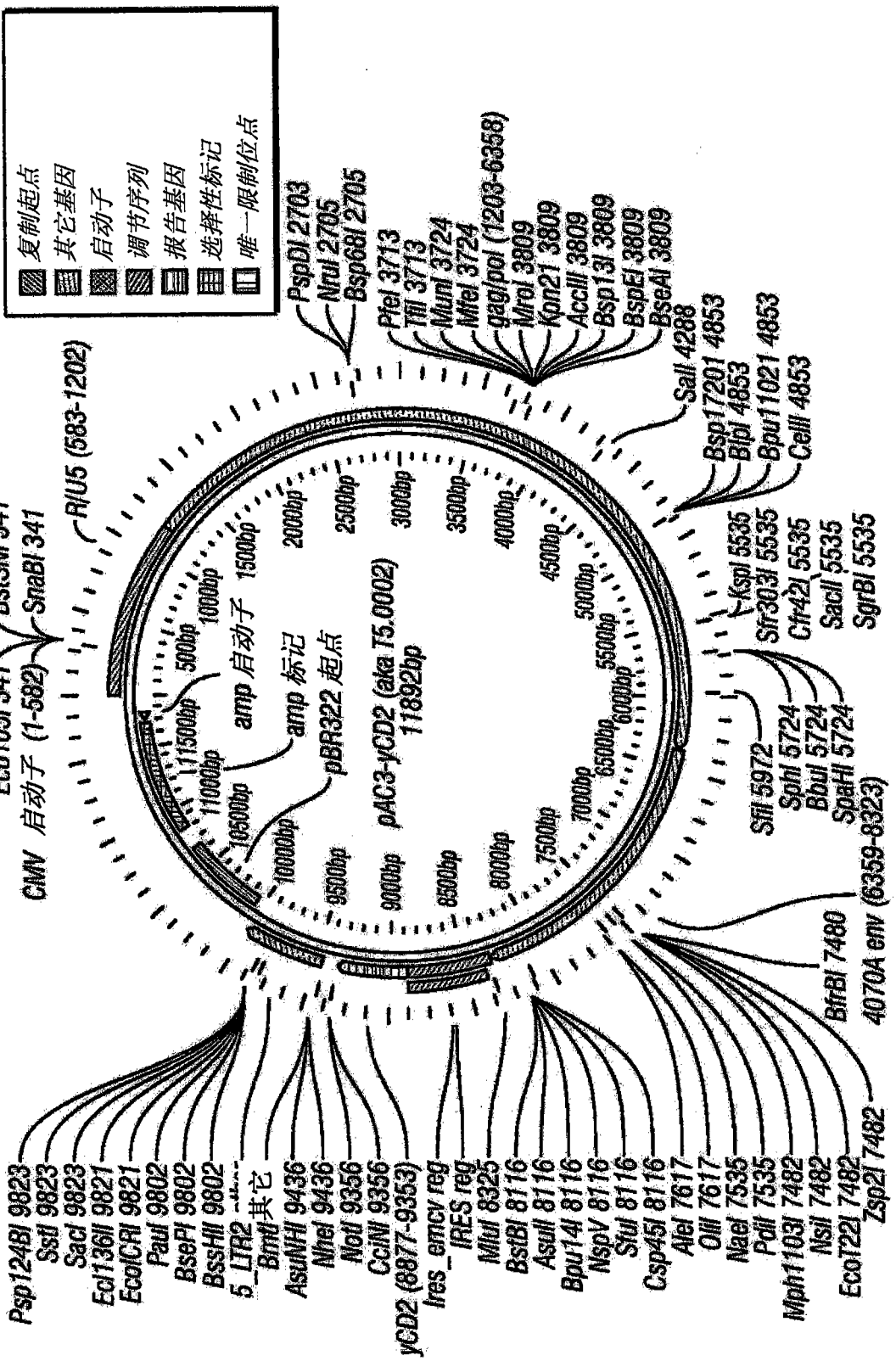

复制起点
其它基因
启动子
调节序列
报告基因
选择性标记
唯一限制位点
pAC3-yCD2 (aka T5.0002)
11892bp
amp 启动子
amp 标记
pBR322 起点
CMV 启动子 (1-582)
R/U5 (583-1202)
5_LTR2
Ires_emcv reg
IRES reg
yCD2 (8877-9353)
4070A env (6359-8323)
Zsp2I 7482
BfrBI 7480
SpaHI 5724
BbuI 5724
SphI 5724
SfiI 5972
SgrBI 5535
SacII 5535
Cfr42I 5535
Sfr303I 5535
KspI 5535
Cell 4853
Bpu11021 4853
BlpI 4853
Bsp17201 4853
SalI 4288
BseAI 3809
BspEI 3809
Bsp13I 3809
KpnI 3809
AccIII 3809
MreI 3809
gag/pol (1203-6358)
MunI 3724
MfeI 3724
TfiI 3713
PfeI 3713
Bsp68I 2705
NruI 2705
PspDI 2703
SnaBI 341
BstSNI 341
Eco105I 341
Psp124BI 9823
SstI 9823
SacI 9823
Eci136II 9821
EcoICRI 9821
PauI 9802
BsePI 9802
BssHII 9802
BtrmI其它
AsuNHI 9436
NheI 9436
NotI 9356
CciNI 9356
MluI 8325
BstBI 8116
AsuII 8116
Bpu14I 8116
NspV 8116
StuI 8116
Csp45I 8116
AleI 7617
OliI 7617
NaeI 7535
PolI 7535
Mph1103I 7482
NsiI 7482
EcoT22I 7482

图3C

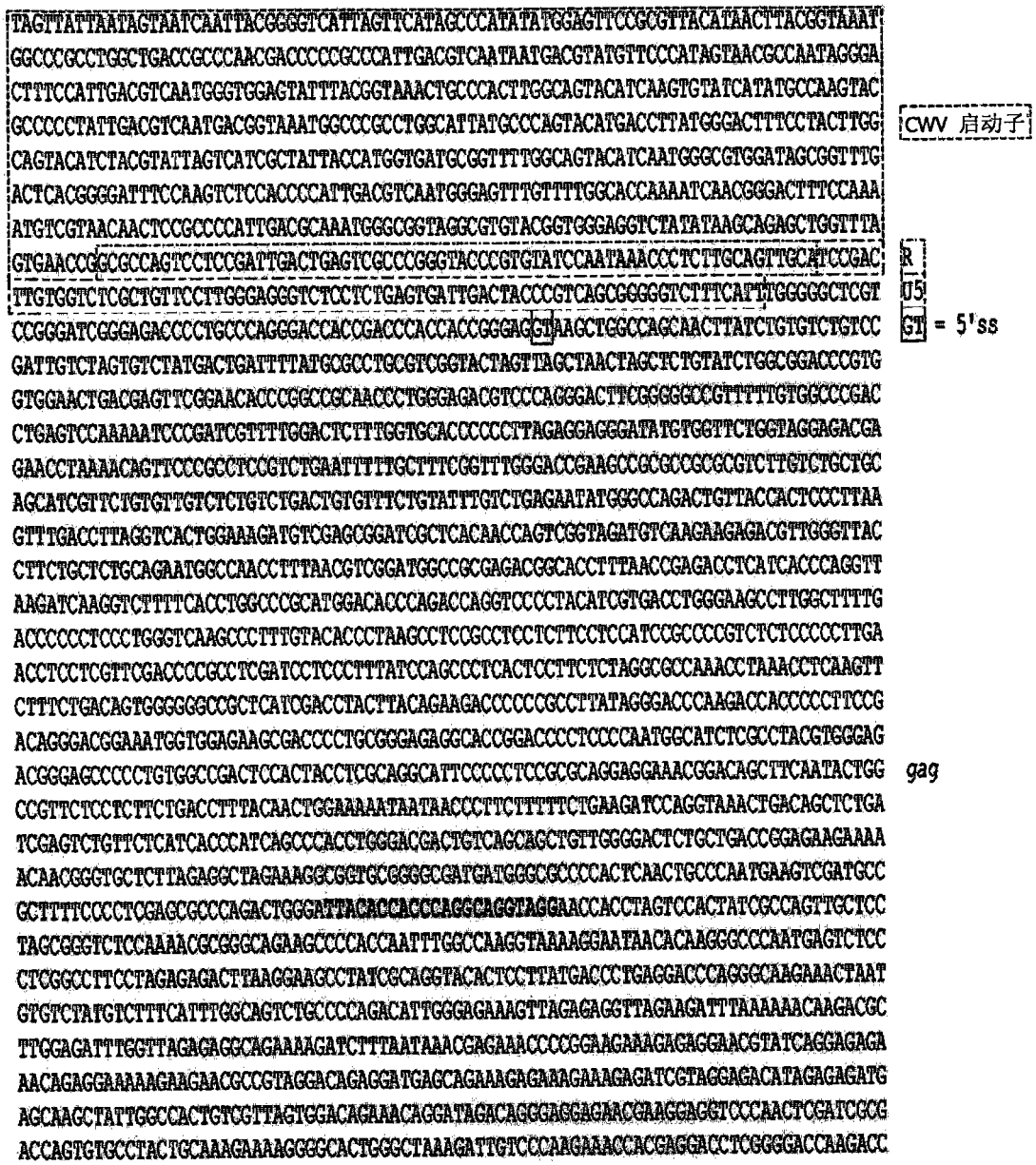
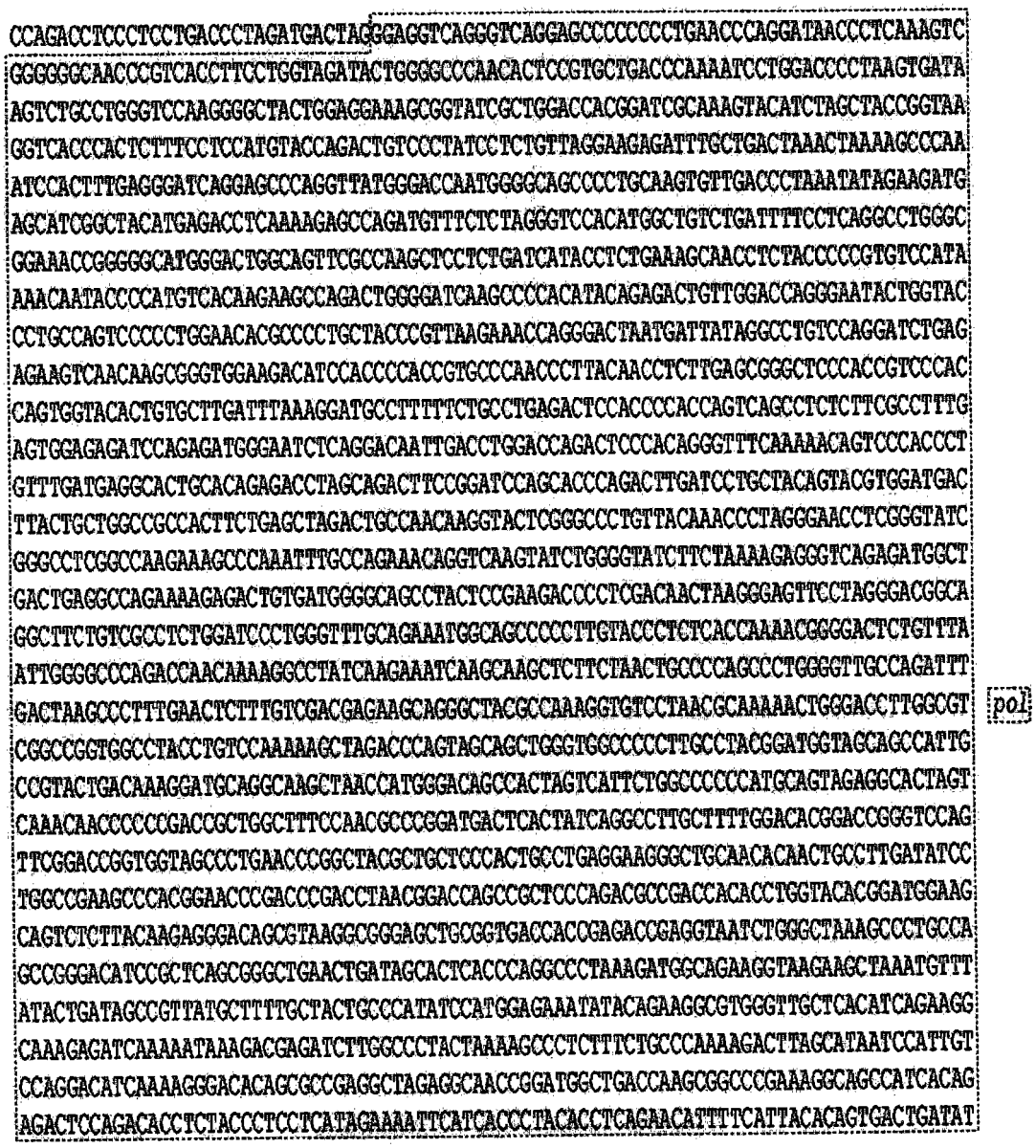


图3D-1


CCAGACCTCCCTCCTGACCCTAGATGACTAGGGAGGTCAGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTC
GGGGGGCAACCCGTCACCTTCCTGGTAGATACTGGGGCCCAACACTCCGTGCTGACCCAAAATCCTGGACCCCTAAGTGATA
AGTCTGCCTGGGTCCAAGGGGCTACTGGAGGAAAGCGGTATCGCTGGACCACGGATCGCAAAGTACATCTAGCTACCGGTAA
GGTCACCCACTCTTTCCTCCATGTACCAGACTGTCCCTATCCTCTGTTAGGAAGAGATTTGCTGACTAAACTAAAAGCCCAA
ATCCACTTTGAGGGATCAGGAGCCCAGGTTATGGGACCAATGGGGCAGCCCCTGCAAGTGTTGACCCTAAATATAGAAGATG
AGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTTCTCTAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGC
GGAAACCGGGGGCATGGGACTGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCTACCCCCGTGTCCATA
AAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGACTGTTGGACCAGGGAATACTGGTAC
CCTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCGTTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAG
AGAAGTCAACAAGCGGGTGGAAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTTGAGCGGGCTCCCACCGTCCCAC
CAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTTCTGCCTGAGACTCCACCCCACCAGTCAGCCTCTCTTCGCCTTTG
AGTGGAGAGATCCAGAGATGGGAATCTCAGGACAATTGACCTGGACCAGACTCCCACAGGGTTTCAAAAACAGTCCCACCCT
GTTTGATGAGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTACAGTACGTGGATGAC
TTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAAGGTACTCGGGCCCTGTTACAAACCCTAGGGAACCTCGGGTATC
GGGCCTCGGCCAAGAAAGCCCAAATTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCTAAAAGAGGGTCAGAGATGGCT
GACTGAGGCCAGAAAAGAGACTGTGATGGGGCAGCCTACTCCGAAGACCCCTCGACAACTAAGGGAGTTCCTAGGGACGGCA
GGCTTCTGTCGCCTCTGGATCCCTGGGTTTGCAGAAATGGCAGCCCCCTTGTACCCTCTCACCAAAACGGGGACTCTGTTTA
ATTGGGGCCCAGACCAACAAAAGGCCTATCAAGAAATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAGATTT
GACTAAGCCCTTTGAACTCTTTGTCGACGAGAAGCAGGGCTACGCCAAAGGTGTCCTAACGCAAAAACTGGGACCTTGGCGT
CGGCCGGTGGCCTACCTGTCCAAAAAGCTAGACCCAGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGGTAGCAGCCATTG
CCGTACTGACAAAGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTCTGGCCCCCCATGCAGTAGAGGCACTAGT
CAAACAACCCCCCGACCGCTGGCTTTCCAACGCCCGGATGACTCACTATCAGGCCTTGCTTTTGGACACGGACCGGGTCCAG
TTCGGACCGGTGGTAGCCCTGAACCCGGCTACGCTGCTCCCACTGCCTGAGGAAGGGCTGCAACACAACTGCCTTGATATCC
TGGCCGAAGCCCACGGAACCCGACCCGACCTAACGGACCAGCCGCTCCCAGACGCCGACCACACCTGGTACACGGATGGAAG
CAGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGACCACCGAGACCGAGGTAATCTGGGCTAAAGCCCTGCCA
GCCGGGACATCCGCTCAGCGGGCTGAACTGATAGCACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTT
ATACTGATAGCCGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGCGTGGGTTGCTCACATCAGAAGG
CAAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTACTAAAAGCCCTCTTTCTGCCCAAAAGACTTAGCATAATCCATTGT
CCAGGACATCAAAAGGGACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAAGGCAGCCATCACAG
AGACTCCAGACACCTCTACCCTCCTCATAGAAAATTCATCACCCTACACCTCAGAACATTTTCATTACACAGTGACTGATAT

pol

图3D-1(续)

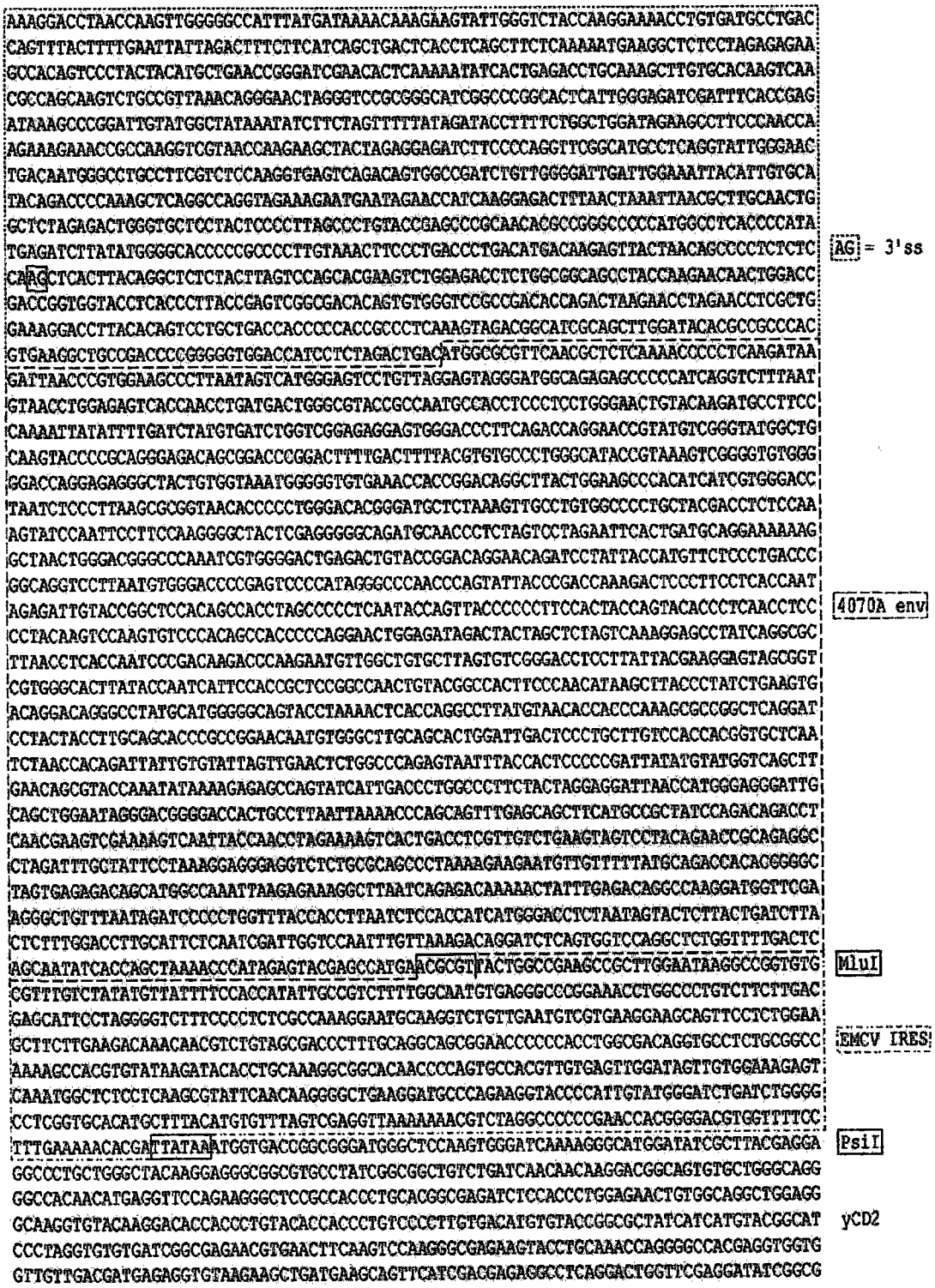


图3D-2

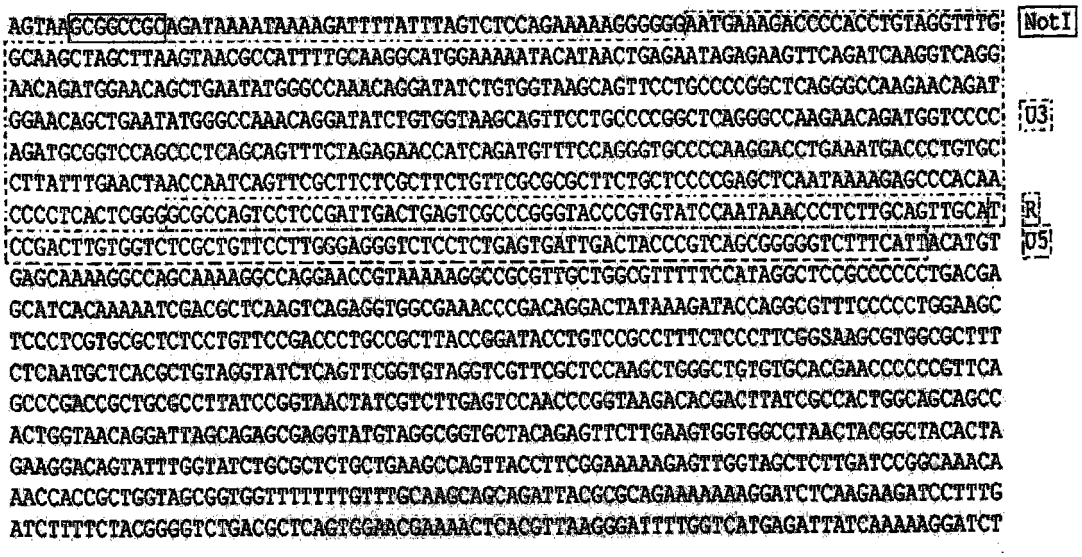


图3D-2(续)

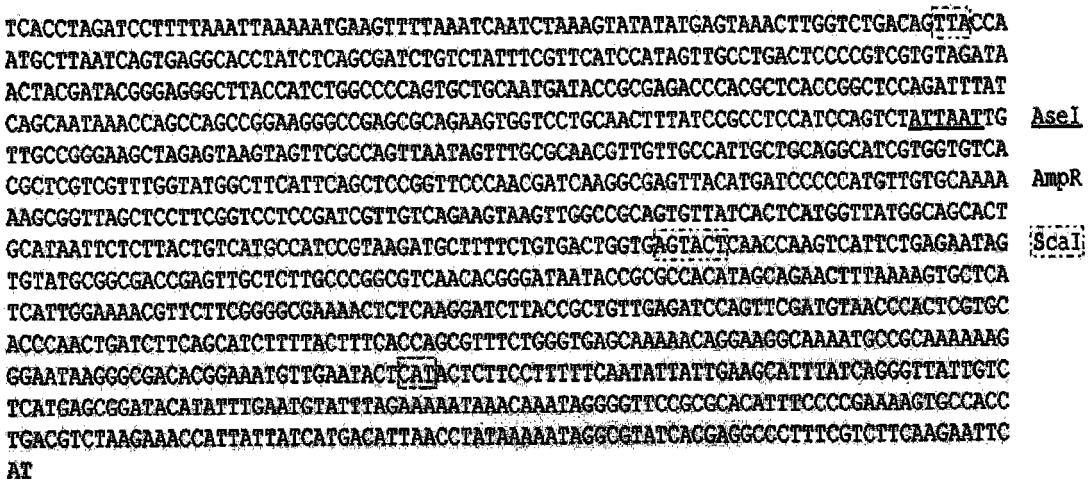


图3D-3

CMV 启动子 (1-582)>>>

```
1   TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG 60
    ATCAATAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGTATATACCTCAAGGC

61  CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATT 120
    GCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGCGGGTAA

121 GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCA 180
    CTGCAGTTATTACTGCATACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAACTGCAGT

181 ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC 240
    TACCCACCTCATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGG

241 AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA 300
    TTCATGCGGGGGATAACTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCAT
```

Eco105I
SnaBI
BstSNI

```
301 CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC 360
    GTACTGGAATACCCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGTAGCGATAATG

361 CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGG 420
    GTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGCCAAACTGAGTGCCCC

421 ATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG 480
    TAAAGGTTCAGAGGTGGGGTAACTGCAGTTACCCTCAAACAAAACCGTGGTTTTAGTTGC

481 GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT 540
    CCTGAAAGGTTTTACAGCATTGTTGAGGCGGGGTAACTGCGTTTACCCGCCATCCGCACA
```

R 区 (583-650)>>>

```
541 ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGGCGCCAGTCCTCCGATTG 600
    TGCCACCCTCCAGATATATTCGTCTCGACCAAATCACTTGGCCGCGGTCAGGAGGCTAAC
```

图3D-4

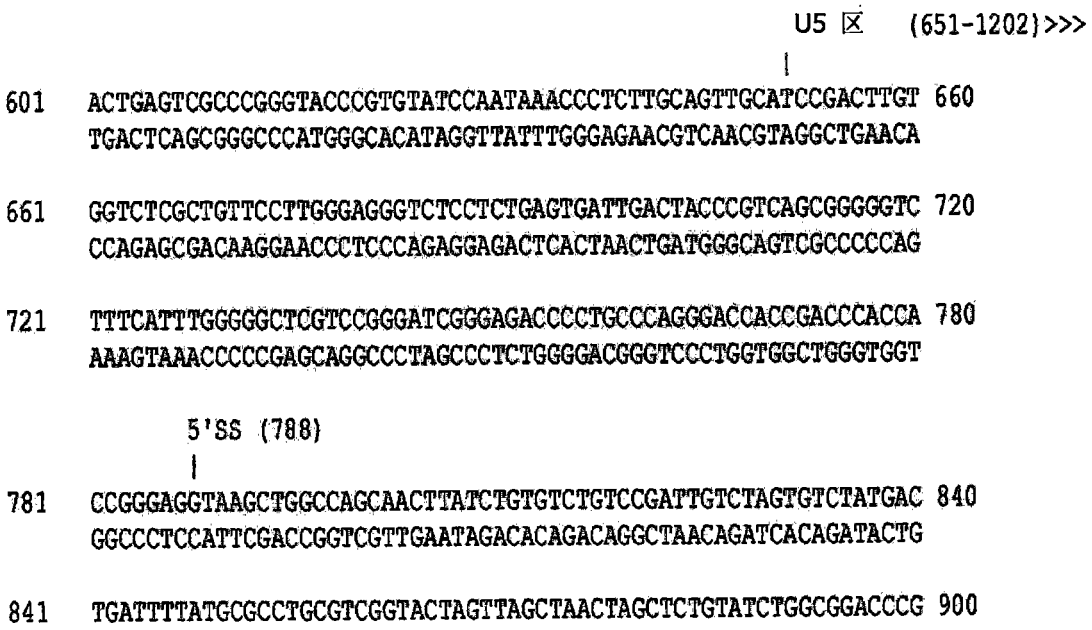

U5 区 (651-1202)>>>
601 ACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGT 660
TGACTCAGCGGGCCCATGGGCACATAGGTTATTTGGGAGAACGTCAACGTAGGCTGAACA
661 GGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGTCAGCGGGGGTC 720
CCAGAGCGACAAGGAACCCTCCCAGAGGAGACTCACTAACTGATGGGCAGTCGCCCCCAG
721 TTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCAGGGACCACCGACCCACCA 780
AAAGTAAACCCCCGAGCAGGCCCTAGCCCTCTGGGGACGGGTCCCTGGTGGCTGGGTGGT
5'SS (788)
781 CCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGAC 840
GGCCCTCCATTCGACCGGTCGTTGAATAGACACAGACAGGCTAACAGATCACAGATACTG
841 TGATTTTATGCGCCTGCGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCG 900

图3D-4(续)

```
      ACTAAAATACGCGGACGCAGCCATGATCAATCGATTGATCGAGACATAGACCGCCTGGGC

901   TGGTGGAACTGACGAGTTCGGAACACCCGGCCGCAACCCTGGGAGACGTCCCAGGGACTT 960
      ACCACCTTGACTGCTCAAGCCTTGTGGGCCGGCGTTGGGACCCTCTGCAGGGTCCCTGAA

961   CGGGGGCCGTTTTTGTGGCCCGACCTGAGTCCAAAAATCCCGATCGTTTTGGACTCTTTG 1020
      GCCCCCGGCAAAAACACCGGGCTGGACTCAGGTTTTTAGGGCTAGCAAAACCTGAGAAAC

1021  GTGCACCCCCCTTAGAGGAGGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAACAGT 1080
      CACGTGGGGGGAATCTCCTCCCTATACACCAAGACCATCCTCTGCTCTTGGATTTTGTCA

1081  TCCCGCCTCCGTCTGAATTTTTGCTTTCGGTTTGGGACCGAAGCCGCGCCGCGCGTCTTG 1140
      AGGGCGGAGGCAGACTTAAAAACGAAAGCCAAACCCTGGCTTCGGCGCGGCGCGCAGAAC

1141  TCTGCTGCAGCATCGTTCTGTGTTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAGA 1200
      AGACGACGTCGTAGCAAGACACAACAGAGACAGACTGACACAAAGACATAAACAGACTCT

        gag(1203,2819)>>>
        |
1201  ATATGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTCG 1260
      TATACCCGGTCTGACAATGGTGAGGGAATTCAAACTGGAATCCAGTGACCTTTCTACAGC

1261  AGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGTTGGGTTACCTTCTGCT 1320
      TCGCCTAGCGAGTGTTGGTCAGCCATCTACAGTTCTTCTCTGCAACCCAATGGAAGACGA

1321  CTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACCTTTAACCGAGACC 1380
      GACGTCTTACCGGTTGGAAATTGCAGCCTACCGGCGCTCTGCCGTGGAAATTGGCTCTGG

1381  TCATCACCCAGGTTAAGATCAAGGTCTTTTCACCTGGCCCGCATGGACACCCAGACCAGG 1440
      AGTAGTGGGTCCAATTCTAGTTCCAGAAAAGTGGACCGGGCGTACCTGTGGGTCTGGTCC

1441  TCCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCCTGGGTCAAGCCCT 1500
      AGGGGATGTAGCACTGGACCCTTCGGAACCGAAAACTGGGGGGAGGGACCCAGTTCGGGA

1501  TTGTACACCCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCTCTCCCCCTTGAAC 1560
      AACATGTGGGATTCGGAGGCGGAGGAGAAGGAGGTAGGCGGGGCAGAGAGGGGGAACTTG

1561  CTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGCG 1620
      GAGGAGCAAGCTGGGGCGGAGCTAGGAGGGAAATAGGTCGGGAGTGAGGAAGAGATCCGC

1621  CCAAACCTAAACCTCAAGTTCTTTCTGACAGTGGGGGGCCGCTCATCGACCTACTTACAG 1680
      GGTTTGGATTTGGAGTTCAAGAAAGACTGTCACCCCCCGGCGAGTAGCTGGATGAATGTC

1681  AAGACCCCCCGCCTTATAGGGACCCAAGACCACCCCCTTCCGACAGGGACGGAAATGGTG 1740
      TTCTGGGGGGGCGGAATATCCCTGGGTTCTGGTGGGGGAAGGCTGTCCCTGCCTTTACCAC
```

图3D-5

```
1741  GAGAAGCGACCCCTGCGGGAGAGGCACCGGACCCCTCCCCAATGGCATCTCGCCTACGTG 1800
      CTCTTCGCTGGGGACGCCCTCTCCGTGGCCTGGGGAGGGGTTACCGTAGAGCGGATGCAC

1801  GGAGACGGGAGCCCCCTGTGGCCGACTCCACTACCTCGCAGGCATTCCCCCTCCGCGCAG 1860
      CCTCTGCCCTCGGGGGACACCGGCTGAGGTGATGGAGCGTCCGTAAGGGGGAGGCGCGTC

1861  GAGGAAACGGACAGCTTCAATACTGGCCGTTCTCCTCTTCTGACCTTTACAACTGGAAAA 1920
      CTCCTTTGCCTGTCGAAGTTATGACCGGCAAGAGGAGAAGACTGGAAATGTTGACCTTTT

1921  ATAATAACCCTTCTTTTTCTGAAGATCCAGGTAAACTGACAGCTCTGATCGAGTCTGTTC 1980
      TATTATTGGGAAGAAAAAGACTTCTAGGTCCATTTGACTGTCGAGACTAGCTCAGACAAG

1981  TCATCACCCATCAGCCCACCTGGGACGACTGTCAGCAGCTGTTGGGGACTCTGCTGACCG 2040
      AGTAGTGGGTAGTCGGGTGGACCCTGCTGACAGTCGTCGACAACCCCTGAGACGACTGGC
```

图3D-5(续)

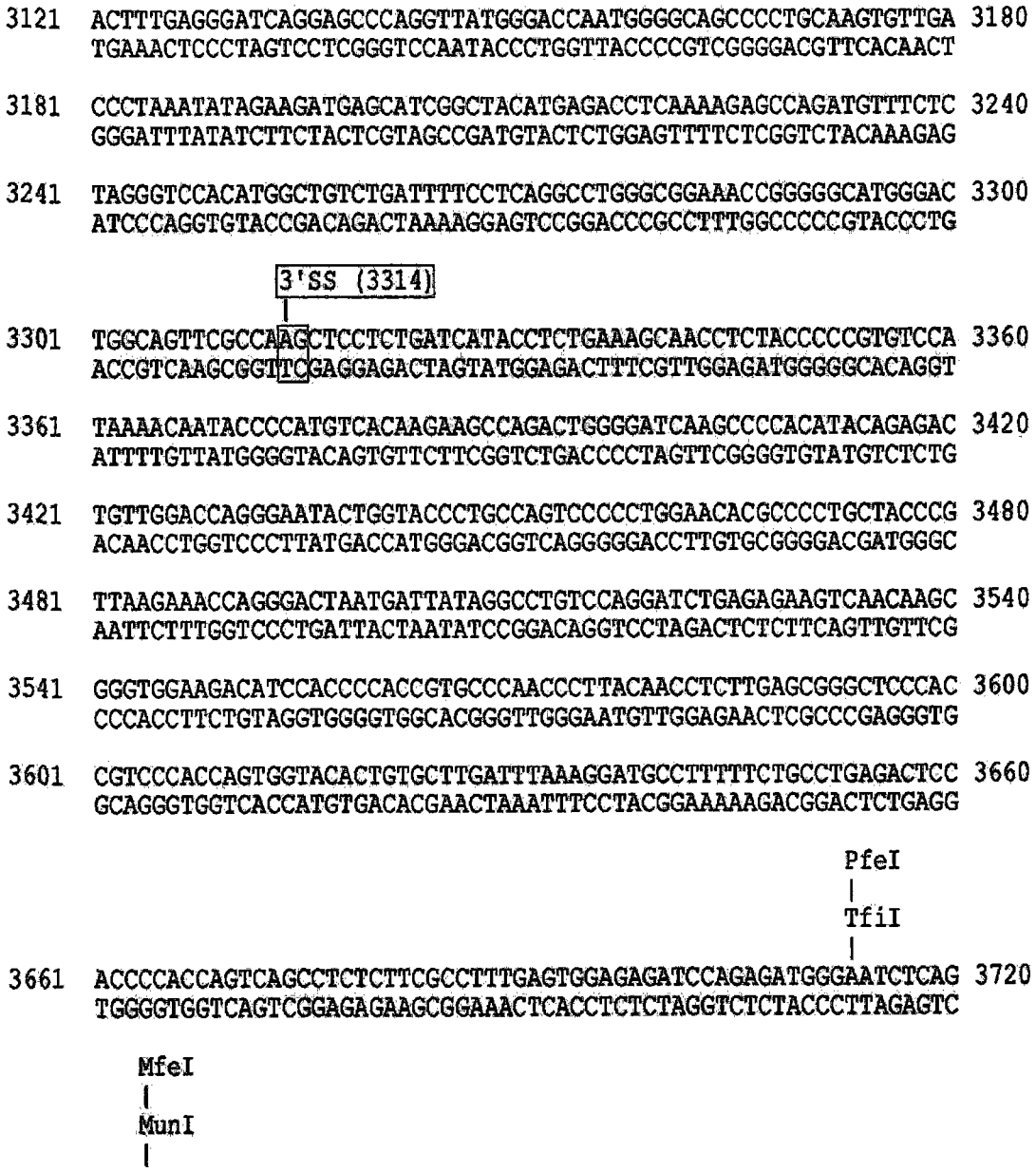

3121 ACTTTGAGGGATCAGGAGCCCAGGTTATGGGACCAATGGGGCAGCCCCTGCAAGTGTTGA 3180
TGAAACTCCCTAGTCCTCGGGTCCAATACCCTGGTTACCCCGTCGGGGACGTTCACAACT
3181 CCCTAAATATAGAAGATGAGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTTCTC 3240
GGGATTTATATCTTCTACTCGTAGCCGATGTACTCTGGAGTTTTCTCGGTCTACAAAGAG
3241 TAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGCGGAAACCGGGGGCATGGGAC 3300
ATCCCAGGTGTACCGACAGACTAAAAGGAGTCCGGACCCGCCTTTGGCCCCCGTACCCTG
3'SS (3314)
3301 TGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCTACCCCCGTGTCCA 3360
ACCGTCAAGCGGTTCGAGGAGACTAGTATGGAGACTTTCGTTGGAGATGGGGGCACAGGT
3361 TAAAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGAC 3420
ATTTTGTTATGGGGTACAGTGTTCTTCGGTCTGACCCCTAGTTCGGGGTGTATGTCTCTG
3421 TGTTGGACCAGGGAATACTGGTACCCTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCG 3480
ACAACCTGGTCCCTTATGACCATGGGACGGTCAGGGGGACCTTGTGCGGGGACGATGGGC
3481 TTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAAGTCAACAAGC 3540
AATTCTTTGGTCCCTGATTACTAATATCCGGACAGGTCCTAGACTCTCTTCAGTTGTTCG
3541 GGGTGGAAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTTGAGCGGGCTCCCAC 3600
CCCACCTTCTGTAGGTGGGGTGGCACGGGTTGGGAATGTTGGAGAACTCGCCCGAGGGTG
3601 CGTCCCACCAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTTCTGCCTGAGACTCC 3660
GCAGGGTGGTCACCATGTGACACGAACTAAATTTCCTACGGAAAAAGACGGACTCTGAGG
PfeI
TfiI
3661 ACCCCACCAGTCAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
TGGGGTGGTCAGTCGGAGAGAAGCGGAAACTCACCTCTCTAGGTCTCTACCCTTAGAGTC
MfeI
MunI

图3D-6

```
3721  GACAATTGACCTGGACCAGACTCCCACAGGGTTTCAAAAACAGTCCCACCCTGTTTGATG 3780
      CTGTTAACTGGACCTGGTCTGAGGGTGTCCCAAAGTTTTTGTCAGGGTGGGACAAACTAC

                                  MroI
                                  |
                                  BseAI
                                  |
                                  Bsp13I
                                  |
                                  BspEI
                                  |
                                  Kpn2I
                                  |
                                  AccIII
                                  |
3781  AGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTAC 3840
      TCCGTGACGTGTCTCTGGATCGTCTGAAGGCCTAGGTCGTGGGTCTGAACTAGGACGATG

3841  AGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAAGGTACTC 3900
      TCATGCACCTACTGAATGACGACCGGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG
```

图3D-6(续)

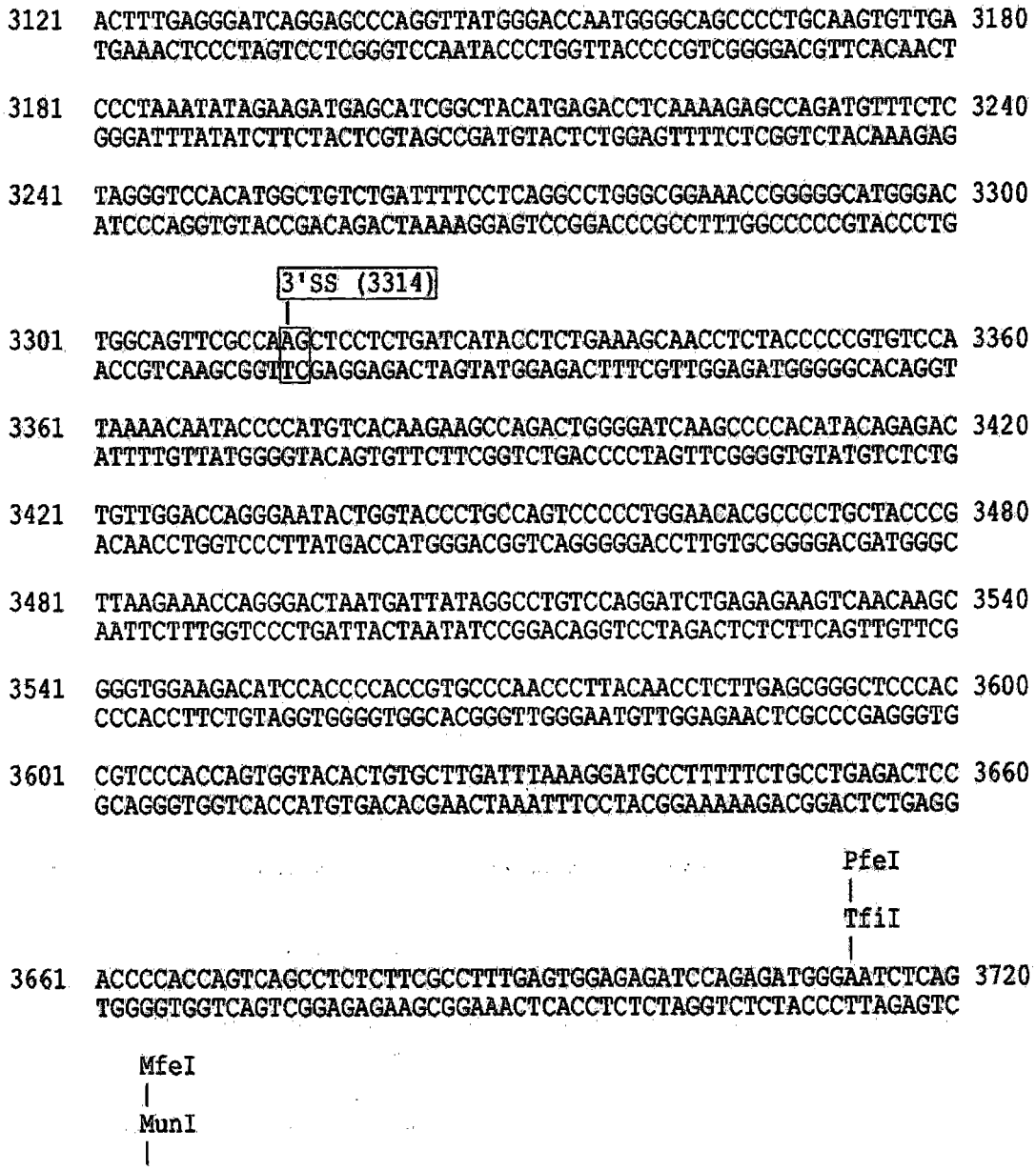

3121 ACTTTGAGGGATCAGGAGCCCAGGTTATGGGACCAATGGGGCAGCCCCTGCAAGTGTTGA 3180
TGAAACTCCCTAGTCCTCGGGTCCAATACCCTGGTTACCCCGTCGGGGACGTTCACAACT
3181 CCCTAAATATAGAAGATGAGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTTCTC 3240
GGGATTTATATCTTCTACTCGTAGCCGATGTACTCTGGAGTTTTCTCGGTCTACAAAGAG
3241 TAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGCGGAAACCGGGGGCATGGGAC 3300
ATCCCAGGTGTACCGACAGACTAAAAGGAGTCCGGACCCGCCTTTGGCCCCCGTACCCTG
3'SS (3314)
3301 TGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCTACCCCCGTGTCCA 3360
ACCGTCAAGCGGTTCGAGGAGACTAGTATGGAGACTTTCGTTGGAGATGGGGGCACAGGT
3361 TAAAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGAC 3420
ATTTTGTTATGGGGTACAGTGTTCTTCGGTCTGACCCCTAGTTCGGGGTGTATGTCTCTG
3421 TGTTGGACCAGGGAATACTGGTACCCTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCG 3480
ACAACCTGGTCCCTTATGACCATGGGACGGTCAGGGGGACCTTGTGCGGGGACGATGGGC
3481 TTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAAGTCAACAAGC 3540
AATTCTTTGGTCCCTGATTACTAATATCCGGACAGGTCCTAGACTCTCTTCAGTTGTTCG
3541 GGGTGGAAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTTGAGCGGGCTCCCAC 3600
CCCACCTTCTGTAGGTGGGGTGGCACGGGTTGGGAATGTTGGAGAACTCGCCCGAGGGTG
3601 CGTCCCACCAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTTCTGCCTGAGACTCC 3660
GCAGGGTGGTCACCATGTGACACGAACTAAATTTCCTACGGAAAAAGACGGACTCTGAGG
PfeI
TfiI
3661 ACCCCACCAGTCAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
TGGGGTGGTCAGTCGGAGAGAAGCGGAAACTCACCTCTCTAGGTCTCTACCCTTAGAGTC
MfeI
MunI

图3D-7

```
3721  GACAATTGACCTGGACCAGACTCCCACAGGGTTTCAAAAACAGTCCCACCCTGTTTGATG 3780
      CTGTTAACTGGACCTGGTCTGAGGGTGTCCCAAAGTTTTTGTCAGGGTGGGACAAACTAC

                                  MroI
                                  |
                                  BseAI
                                  |
                                  Bsp13I
                                  |
                                  BspEI
                                  |
                                  Kpn2I
                                  |
                                  AccIII
                                  |
3781  AGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTAC 3840
      TCCGTGACGTGTCTCTGGATCGTCTGAAGGCCTAGGTCGTGGGTCTGAACTAGGACGATG

3841  AGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAAGGTACTC 3900
      TCATGCACCTACTGAATGACGACCGGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG
```

图3D-7(续)

```
3901  GGGCCCTGTTACAAACCCTAGGGAACCTCGGGTATCGGGCCTCGGCCAAGAAAGCCCAAA 3960
      CCCGGGACAATGTTTGGGATCCCTTGGAGCCCATAGCCCGGAGCCGGTTCTTTCGGGTTT

3961  TTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCTAAAAGAGGGTCAGAGATGGCTGA 4020
      AAACGGTCTTTGTCCAGTTCATAGACCCCATAGAAGATTTTCTCCCAGTCTCTACCGACT

4021  CTGAGGCCAGAAAAGAGACTGTGATGGGGCAGCCTACTCCGAAGACCCCTCGACAACTAA 4080
      GACTCCGGTCTTTTCTCTGACACTACCCCGTCGGATGAGGCTTCTGGGGAGCTGTTGATT

4081  GGGAGTTCCTAGGGACGGCAGGCTTCTGTCGCCTCTGGATCCCTGGGTTTGCAGAAATGG 4140
      CCCTCAAGGATCCCTGCCGTCCGAAGACAGCGGAGACCTAGGGACCCAAACGTCTTTACC

4141  CAGCCCCCTTGTACCCTCTCACCAAAACGGGGACTCTGTTTAATTGGGGCCCAGACCAAC 4200
      GTCGGGGGAACATGGGAGAGTGGTTTTGCCCCTGAGACAAATTAACCCCGGGTCTGGTTG

4201  AAAAGGCCTATCAAGAAATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAG 4260
      TTTTCCGGATAGTTCTTTAGTTCGTTCGAGAAGATTGACGGGGTCGGGACCCCAACGGTC

                                  SalI
                                   |
4261  ATTTGACTAAGCCCTTTGAACTCTTTGTCGACGAGAAGCAGGGCTACGCCAAAGGTGTCC 4320
      TAAACTGATTCGGGAAACTTGAGAAACAGCTGCTCTTCGTCCCGATGCGGTTTCCACAGG

4321  TAACGCAAAAACTGGGACCTTGGCGTCGGCCGGTGGCCTACCTGTCCAAAAAGCTAGACC 4380
      ATTGCGTTTTTGACCCTGGAACCGCAGCCGGCCACCGGATGGACAGGTTTTTCGATCTGG

4381  CAGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGGTAGCAGCCATTGCCGTACTGACAA 4440
      GTCATCGTCGACCCACCGGGGGAACGGATGCCTACCATCGTCGGTAACGGCATGACTGTT

4441  AGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTCTGGCCCCCCATGCAGTAG 4500
      TCCTACGTCCGTTCGATTGGTACCCTGTCGGTGATCAGTAAGACCGGGGGGTACGTCATC

4501  AGGCACTAGTCAAACAACCCCCCGACCGCTGGCTTTCCAACGCCCGGATGACTCACTATC 4560
      TCCGTGATCAGTTTGTTGGGGGGCTGGCGACCGAAAGGTTGCGGGCCTACTGAGTGATAG

4561  AGGCCTTGCTTTTGGACACGGACCGGGTCCAGTTCGGACCGGTGGTAGCCCTGAACCCGG 4620
      TCCGGAACGAAAACCTGTGCCTGGCCCAGGTCAAGCCTGGCCACCATCGGGACTTGGGCC

4621  CTACGCTGCTCCCACTGCCTGAGGAAGGGCTGCAACACAACTGCCTTGATATCCTGGCCG 4680
      GATGCGACGAGGGTGACGGACTCCTTCCCGACGTTGTGTTGACGGAACTATAGGACCGGC

4681  AAGCCCACGGAACCCGACCCGACCTAACGGACCAGCCGCTCCCAGACGCCGACCACACCT 4740
      TTCGGGTGCCTTGGGCTGGGCTGGATTGCCTGGTCGGCGAGGGTCTGCGGCTGGTGTGGA

4741  GGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGA 4800
      CCATGTGCCTACCTTCGTCAGAGAATGTTCTCCCTGTCGCATTCCGCCCTCGACGCCACT
```

图3D-8

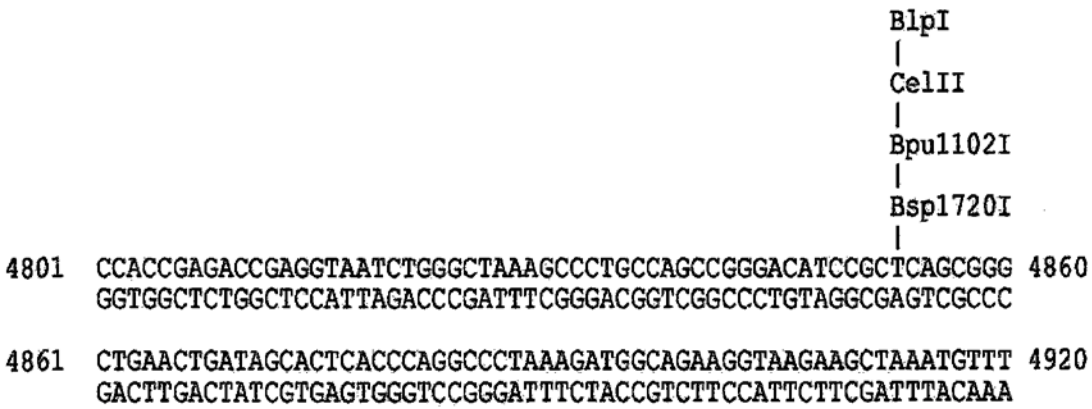

BlpI
CelII
Bpu1102I
Bsp1720I
4801 CCACCGAGACCGAGGTAATCTGGGCTAAAGCCCTGCCAGCCGGGACATCCGCTCAGCGGG 4860
GGTGGCTCTGGCTCCATTAGACCCGATTTCGGGACGGTCGGCCCTGTAGGCGAGTCGCCC
4861 CTGAACTGATAGCACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTT 4920
GACTTGACTATCGTGAGTGGGTCCGGGATTTCTACCGTCTTCCATTCTTCGATTTACAAA

图3D-8(续)

```
4921  ATACTGATAGCCGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGC 4980
      TATGACTATCGGCAATACGAAAACGATGACGGGTATAGGTACCTCTTTATATGTCTTCCG

4981  GTGGGTTGCTCACATCAGAAGGCAAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTAC 5040
      CACCCAACGAGTGTAGTCTTCCGTTTCTCTAGTTTTTATTTCTGCTCTAGAACCGGGATG

5041  TAAAAGCCCTCTTTCTGCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAAGG 5100
      ATTTTCGGGAGAAAGACGGGTTTTCTGAATCGTATTAGGTAACAGGTCCTGTAGTTTTCC

5101  GACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAAGGCAGCCA 5160
      CTGTGTCGCGGCTCCGATCTCCGTTGGCCTACCGACTGGTTCGCCGGGCTTTCCGTCGGT

5161  TCACAGAGACTCCAGACACCTCTACCCTCCTCATAGAAAATTCATCACCCTACACCTCAG 5220
      AGTGTCTCTGAGGTCTGTGGAGATGGGAGGAGTATCTTTTAAGTAGTGGGATGTGGAGTC

5221  AACATTTTCATTACACAGTGACTGATATAAAGGACCTAACCAAGTTGGGGGCCATTTATG 5280
      TTGTAAAAGTAATGTGTCACTGACTATATTTCCTGGATTGGTTCAACCCCCGGTAAATAC

5281  ATAAAACAAAGAAGTATTGGGTCTACCAAGGAAAACCTGTGATGCCTGACCAGTTTACTT 5340
      TATTTTGTTTCTTCATAACCCAGATGGTTCCTTTTGGACACTACGGACTGGTCAAATGAA

5341  TTGAATTATTAGACTTTCTTCATCAGCTGACTCACCTCAGCTTCTCAAAAATGAAGGCTC 5400
      AACTTAATAATCTGAAAGAAGTAGTCGACTGAGTGGAGTCGAAGAGTTTTTACTTCCGAG

5401  TCCTAGAGAGAAGCCACAGTCCCTACTACATGCTGAACCGGGATCGAACACTCAAAAATA 5460
      AGGATCTCTCTTCGGTGTCAGGGATGATGTACGACTTGGCCCTAGCTTGTGAGTTTTTAT

5461  TCACTGAGACCTGCAAAGCTTGTGCACAAGTCAACGCCAGCAAGTCTGCCGTTAAACAGG 5520
      AGTGACTCTGGACGTTTCGAACACGTGTTCAGTTGCGGTCGTTCAGACGGCAATTTGTCC
```

SacII
|
SgrBI
|
Cfr42I
|
Sfr303I
|
KspI
|

图3D-9

```
5521  GAACTAGGGTCCGCGGGCATCGGCCCGGCACTCATTGGGAGATCGATTTCACCGAGATAA 5580
      CTTGATCCCAGGCGCCCGTAGCCGGGCCGTGAGTAACCCTCTAGCTAAAGTGGCTCTATT

5581  AGCCCGGATTGTATGGCTATAAATATCTTCTAGTTTTTATAGATACCTTTTCTGGCTGGA 5640
      TCGGGCCTAACATACCGATATTTATAGAAGATCAAAAATATCTATGGAAAAGACCGACCT

5641  TAGAAGCCTTCCCAACCAAGAAAGAAACCGCCAAGGTCGTAACCAAGAAGCTACTAGAGG 5700
      ATCTTCGGAAGGGTTGGTTCTTTCTTTGGCGGTTCCAGCATTGGTTCTTCGATGATCTCC

                                PaeI
                                |
                                BbuI
                                |
                                SpaHI
                                |
                                SphI
                                |
5701  AGATCTTCCCCAGGTTCGGCATGCCTCAGGTATTGGGAACTGACAATGGGCCTGCCTTCG 5760
      TCTAGAAGGGGTCCAAGCCGTACGGAGTCCATAACCCTTGACTGTTACCCGGACGGAAGC
```

图3D-9(续)

```
6901  CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCTCTAGTCCTAGAATTCACTGATGC  6960
      GAAGGTTCCCCGATGAGCTCCCCCGTCTACGTTGGGAGATCAGGATCTTAAGTGACTACG

6961  AGGAAAAAAGGCTAACTGGGACGGGCCCAAATCGTGGGGACTGAGACTGTACCGGACAGG  7020
      TCCTTTTTTCCGATTGACCCTGCCCGGGTTTAGCACCCCTGACTCTGACATGGCCTGTCC

7021  AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCCTTAATGTGGGACCCCGAGT  7080
      TTGTCTAGGATAATGGTACAAGAGGGACTGGGCCGTCCAGGAATTACACCCTGGGGCTCA

7081  CCCCATAGGGCCCAACCCAGTATTACCCGACCAAAGACTCCCTTCCTCACCAATAGAGAT  7140
      GGGGTATCCCGGGTTGGGTCATAATGGGCTGGTTTCTGAGGGAAGGAGTGGTTATCTCTA

7141  TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACCAGTTACCCCCCTTCCACTACCAG  7200
      ACATGGCCGAGGTGTCGGTGGATCGGGGGAGTTATGGTCAATGGGGGGAAGGTGATGGTC

7201  TACACCCTCAACCTCCCCTACAAGTCCAAGTGTCCCACAGCCACCCCCAGGAACTGGAGA  7260
      ATGTGGGAGTTGGAGGGGATGTTCAGGTTCACAGGGTGTCGGTGGGGGTCCTTGACCTCT

7261  TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCGCTTAACCTCACCAATCCCGACAA  7320
      ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCGCGAATTGGAGTGGTTAGGGCTGTT

7321  GACCCAAGAATGTTGGCTGTGCTTAGTGTCGGGACCTCCTTATTACGAAGGAGTAGCGGT  7380
      CTGGGTTCTTACAACCGACACGAATCACAGCCCTGGAGGAATAATGCTTCCTCATCGCCA

7381  CGTGGGCACTTATACCAATCATTCCACCGCTCCGGCCAACTGTACGGCCACTTCCCAACA  7440
      GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTTGACATGCCGGTGAAGGGTTGT
```

EcoT22I
NsiI
Mph1103I
Zsp2I
BfrBI

图3D-10

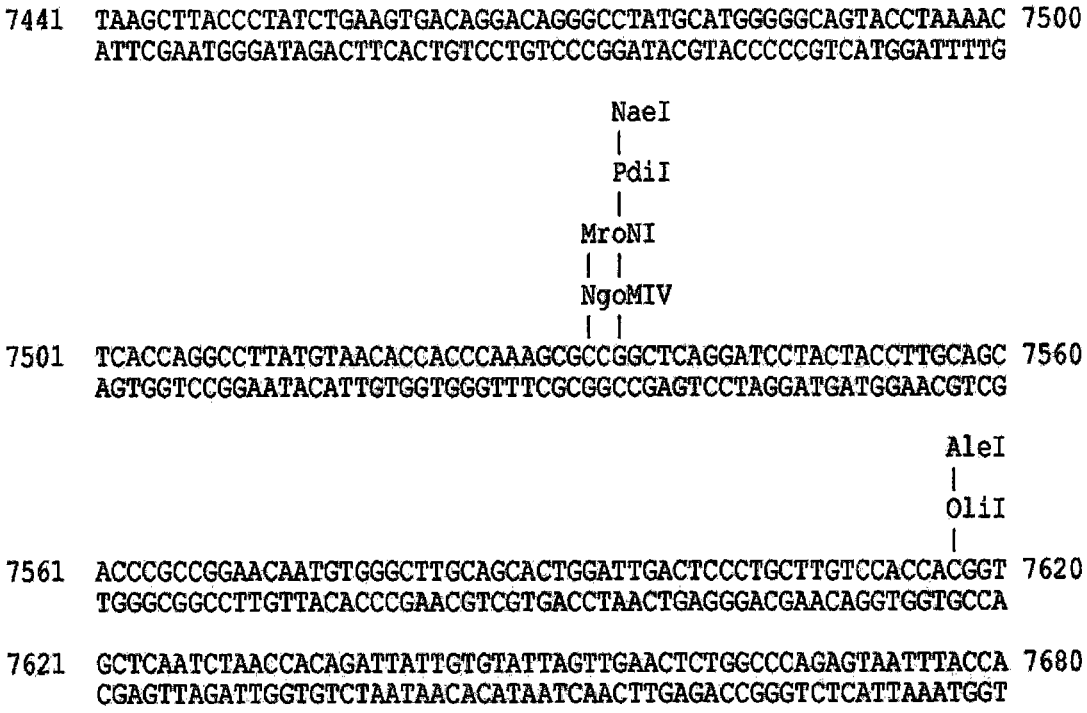

7441 TAAGCTTACCCTATCTGAAGTGACAGGACAGGGCCTATGCATGGGGGCAGTACCTAAAAC 7500
ATTCGAATGGGATAGACTTCACTGTCCTGTCCCGGATACGTACCCCCGTCATGGATTTTG
NaeI
PdiI
MroNI
NgoMIV
7501 TCACCAGGCCTTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTTGCAGC 7560
AGTGGTCCGGAATACATTGTGGTGGGTTTCGCGGCCGAGTCCTAGGATGATGGAACGTCG
AleI
OliI
7561 ACCCGCCGGAACAATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCACGGT 7620
TGGGCGGCCTTGTTACACCCGAACGTCGTGACCTAACTGAGGGACGAACAGGTGGTGCCA
7621 GCTCAATCTAACCACAGATTATTGTGTATTAGTTGAACTCTGGCCCAGAGTAATTTACCA 7680
CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAAATGGT

图3D-10(续)

```
6901  CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCTCTAGTCCTAGAATTCACTGATGC 6960
      GAAGGTTCCCCGATGAGCTCCCCCGTCTACGTTGGGAGATCAGGATCTTAAGTGACTACG

6961  AGGAAAAAAGGCTAACTGGGACGGGCCCAAATCGTGGGGACTGAGACTGTACCGGACAGG 7020
      TCCTTTTTTCCGATTGACCCTGCCCGGGTTTAGCACCCCTGACTCTGACATGGCCTGTCC

7021  AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCCTTAATGTGGGACCCCGAGT 7080
      TTGTCTAGGATAATGGTACAAGAGGGACTGGGCCGTCCAGGAATTACACCCTGGGGCTCA

7081  CCCCATAGGGCCCAACCCAGTATTACCCGACCAAAGACTCCCTTCCTCACCAATAGAGAT 7140
      GGGGTATCCCGGGTTGGGTCATAATGGGCTGGTTTCTGAGGGAAGGAGTGGTTATCTCTA

7141  TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACCAGTTACCCCCCTTCCACTACCAG 7200
      ACATGGCCGAGGTGTCGGTGGATCGGGGGAGTTATGGTCAATGGGGGGAAGGTGATGGTC

7201  TACACCCTCAACCTCCCCTACAAGTCCAAGTGTCCCACAGCCACCCCCAGGAACTGGAGA 7260
      ATGTGGGAGTTGGAGGGGATGTTCAGGTTCACAGGGTGTCGGTGGGGGTCCTTGACCTCT

7261  TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCGCTTAACCTCACCAATCCCGACAA 7320
      ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCGCGAATTGGAGTGGTTAGGGCTGTT

7321  GACCCAAGAATGTTGGCTGTGCTTAGTGTCGGGACCTCCTTATTACGAAGGAGTAGCGGT 7380
      CTGGGTTCTTACAACCGACACGAATCACAGCCCTGGAGGAATAATGCTTCCTCATCGCCA

7381  CGTGGGCACTTATACCAATCATTCCACCGCTCCGGCCAACTGTACGGCCACTTCCCAACA 7440
      GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTTGACATGCCGGTGAAGGGTTGT
```

```
EcoT22I
|
NsiI
|
Mph1103I
|
Zsp2I
|
BfrBI
| |
```

图3D-11

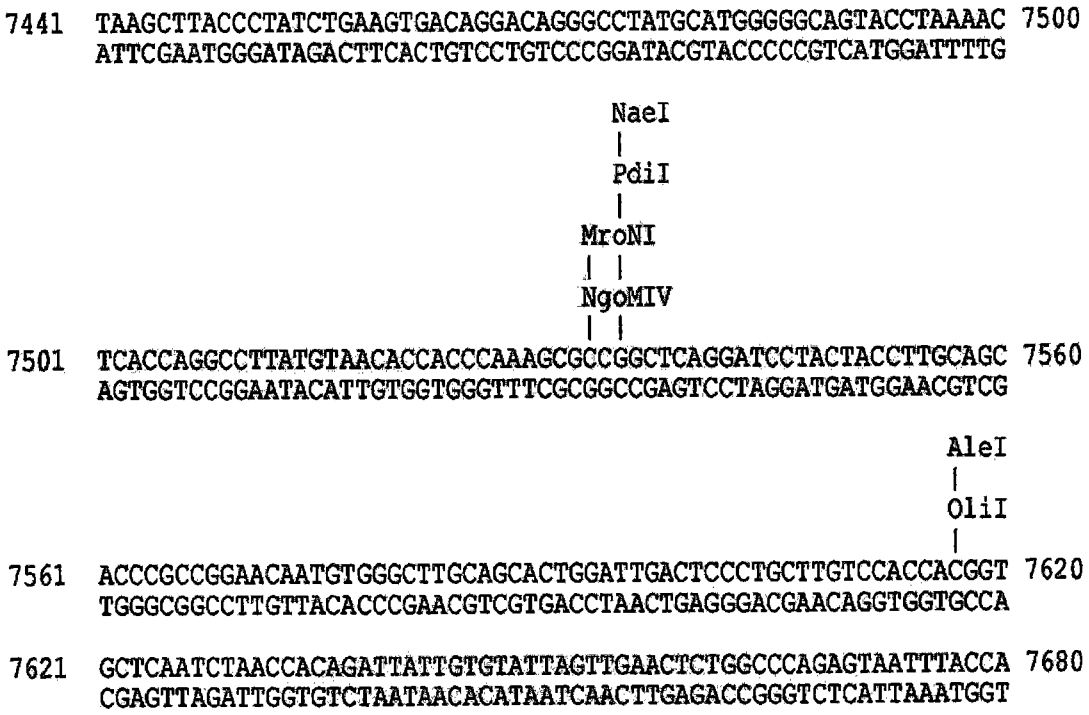

7441 TAAGCTTACCCTATCTGAAGTGACAGGACAGGGCCTATGCATGGGGGCAGTACCTAAAAC 7500
ATTCGAATGGGATAGACTTCACTGTCCTGTCCCGGATACGTACCCCCGTCATGGATTTTG
NaeI
PdiI
MroNI
NgoMIV
7501 TCACCAGGCCTTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTTGCAGC 7560
AGTGGTCCGGAATACATTGTGGTGGGTTTCGCGGCCGAGTCCTAGGATGATGGAACGTCG
AleI
OliI
7561 ACCCGCCGGAACAATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCACGGT 7620
TGGGCGGCCTTGTTACACCCGAACGTCGTGACCTAACTGAGGGACGAACAGGTGGTGCCA
7621 GCTCAATCTAACCACAGATTATTGTGTATTAGTTGAACTCTGGCCCAGAGTAATTTACCA 7680
CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAAATGGT

图3D-11(续)

```
7681  CTCCCCCGATTATATGTATGGTCAGCTTGAACAGCGTACCAAATATAAAAGAGAGCCAGT 7740
      GAGGGGGCTAATATACATACCAGTCGAACTTGTCGCATGGTTTATATTTTCTCTCGGTCA

7741  ATCATTGACCCTGGCCCTTCTACTAGGAGGATTAACCATGGGAGGGATTGCAGCTGGAAT 7800
      TAGTAACTGGGACCGGGAAGATGATCCTCCTAATTGGTACCCTCCCTAACGTCGACCTTA

7801  AGGGACGGGGACCACTGCCTTAATTAAAACCCAGCAGTTTGAGCAGCTTCATGCCGCTAT 7860
      TCCCTGCCCCTGGTGACGGAATTAATTTTGGGTCGTCAAACTCGTCGAAGTACGGCGATA

7861  CCAGACAGACCTCAACGAAGTCGAAAAGTCAATTACCAACCTAGAAAAGTCACTGACCTC 7920
      GGTCTGTCTGGAGTTGCTTCAGCTTTTCAGTTAATGGTTGGATCTTTTCAGTGACTGGAG

7921  GTTGTCTGAAGTAGTCCTACAGAACCGCAGAGGCCTAGATTTGCTATTCCTAAAGGAGGG 7980
      CAACAGACTTCATCAGGATGTCTTGGCGTCTCCGGATCTAAACGATAAGGATTTCCTCCC

7981  AGGTCTCTGCGCAGCCCTAAAAGAAGAATGTTGTTTTTATGCAGACCACACGGGGCTAGT 8040
      TCCAGAGACGCGTCGGGATTTTCTTCTTACAACAAAAATACGTCTGGTGTGCCCCGATCA

8041  GAGAGACAGCATGGCCAAATTAAGAGAAAGGCTTAATCAGAGACAAAAACTATTTGAGAC 8100
      CTCTCTGTCGTACCGGTTTAATTCTCTTTCCGAATTAGTCTCTGTTTTTGATAAACTCTG
```

NspV
|
BstBI
|
Bsp119I
|
AsuII
|
Csp45I
|
SfuI
|
Bpu14I
|
BspT104I
|

```
8101  AGGCCAAGGATGGTTCGAAGGGCTGTTTAATAGATCCCCCTGGTTTACCACCTTAATCTC 8160
      TCCGGTTCCTACCAAGCTTCCCGACAAATTATCTAGGGGGACCAAATGGTGGAATTAGAG

8161  CACCATCATGGGACCTCTAATAGTACTCTTACTGATCTTACTCTTTGGACCTTGCATTCT 8220
      GTGGTAGTACCCTGGAGATTATCATGAGAATGACTAGAATGAGAAACCTGGAACGTAAGA

8221  CAATCGATTGGTCCAATTTGTTAAAGACAGGATCTCAGTGGTCCAGGCTCTGGTTTTGAC 8280
      GTTAGCTAACCAGGTTAAACAATTTCTGTCCTAGAGTCACCAGGTCCGAGACCAAAACTG
```

图3D-12

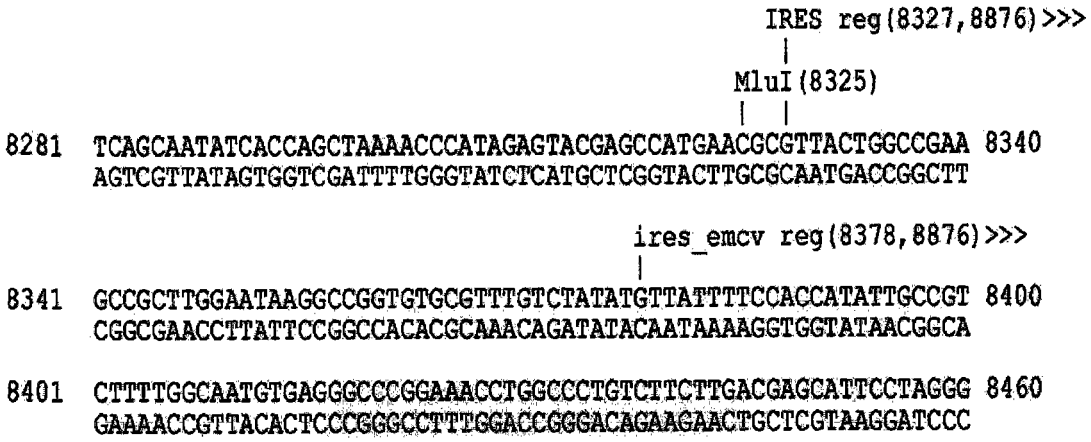

IRES reg(8327,8876)>>>
MluI(8325)
8281 TCAGCAATATCACCAGCTAAAACCCATAGAGTACGAGCCATGAACGCGTTACTGGCCGAA 8340
AGTCGTTATAGTGGTCGATTTTGGGTATCTCATGCTCGGTACTTGCGCAATGACCGGCTT
ires_emcv reg(8378,8876)>>>
8341 GCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGT 8400
CGGCGAACCTTATTCCGGCCACACGCAAACAGATATACAATAAAAGGTGGTATAACGGCA
8401 CTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGG 8460
GAAAACCGTTACACTCCCGGGCCTTTGGACCGGGACAGAAGAACTGCTCGTAAGGATCCC

图3D-12(续)

```
8461  GTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTC 8520
      CAGAAAGGGGAGAGCGGTTTCCTTACGTTCCAGACAACTTACAGCACTTCCTTCGTCAAG

8521  CTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACC 8580
      GAGACCTTCGAAGAACTTCTGTTTGTTGCAGACATCGCTGGGAAACGTCCGTCGCCTTGG

8581  CCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAA 8640
      GGGGTGGACCGCTGTCCACGGAGACGCCGGTTTTCGGTGCACATATTCTATGTGGACGTT

8641  AGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGC 8700
      TCCGCCGTGTTGGGGTCACGGTGCAACACTCAACCTATCAACACCTTTCTCAGTTTACCG

8701  TCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGG 8760
      AGAGGAGTTCGCATAAGTTGTTCCCCGACTTCCTACGGGTCTTCCATGGGGTAACATACC

8761  GATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAAC 8820
      CTAGACTAGACCCCGGAGCCACGTGTACGAAATGTACACAAATCAGCTCCAATTTTTTTG

                                                    yCD2 (8877,9353)>>>
                                                          |
                                                      PsiI(8874)
                                                      |   |
8821  GTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATTATAAATGG 8880
      CAGATCCGGGGGGGCTTGGTGCCCCTGCACCAAAAGGAAACTTTTTGTGCTAATATTTACC

8881  TGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTTACGAGGAGG 8940
      ACTGGCCGCCGTACCGGAGGTTCACCCTAGTTTTCCCGTACCTATAGCGAATGCTCCTCC

8941  CCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAGG 9000
      GGGACGACCCGATGTTCCTCCCGCCGCACGGATAGCCGCCGACAGACTAGTTGTTGTTCC

9001  ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGC 9060
      TGCCGTCACACGACCCGTCCCCGGTGTTGTACTCCAAGGTCTTCCCGAGGCGGTGGGACG

9061  ACGGCGAGATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACA 9120
      TGCCGCTCTAGAGGTGGGACCTCTTGACACCGTCCGACCTCCCGTTCCACATGTTCCTGT

9121  CCACCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACG 9180
      GGTGGGACATGTGGTGGGACAGGGGAACACTGTACACATGGCCGCGATAGTAGTACATGC
```

图3D-13

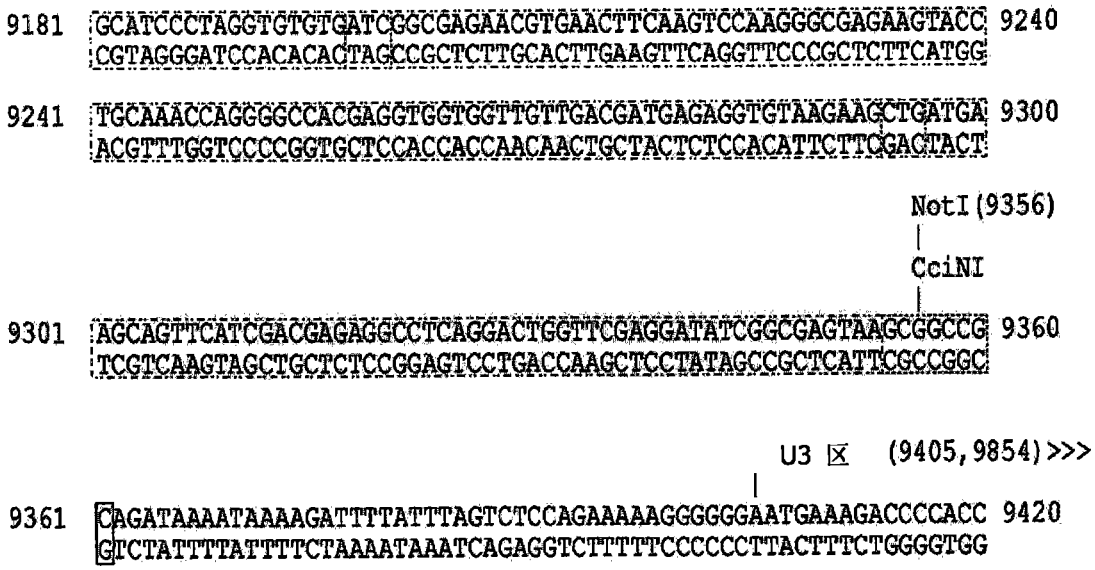

9181
GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCCAAGGGCGAGAAGTACC 9240
CGTAGGGATCCACACACTAGCCGCTCTTGCACTTGAAGTTCAGGTTCCCGCTCTTCATGG
9241
TGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTGATGA 9300
ACGTTTGGTCCCCGGTGCTCCACCACCAACAACTGCTACTCTCCACATTCTTCGACTACT
NotI(9356)
CciNI
9301
AGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTAAGCGGCCG 9360
TCGTCAAGTAGCTGCTCTCCGGAGTCCTGACCAAGCTCCTATAGCCGCTCATTCGCCGGC
U3 区 (9405,9854)>>>
9361
CAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACC 9420
GTCTATTTTATTTTCTAAAATAAATCAGAGGTCTTTTTCCCCCCTTACTTTCTGGGGTGG

图3D-13(续)

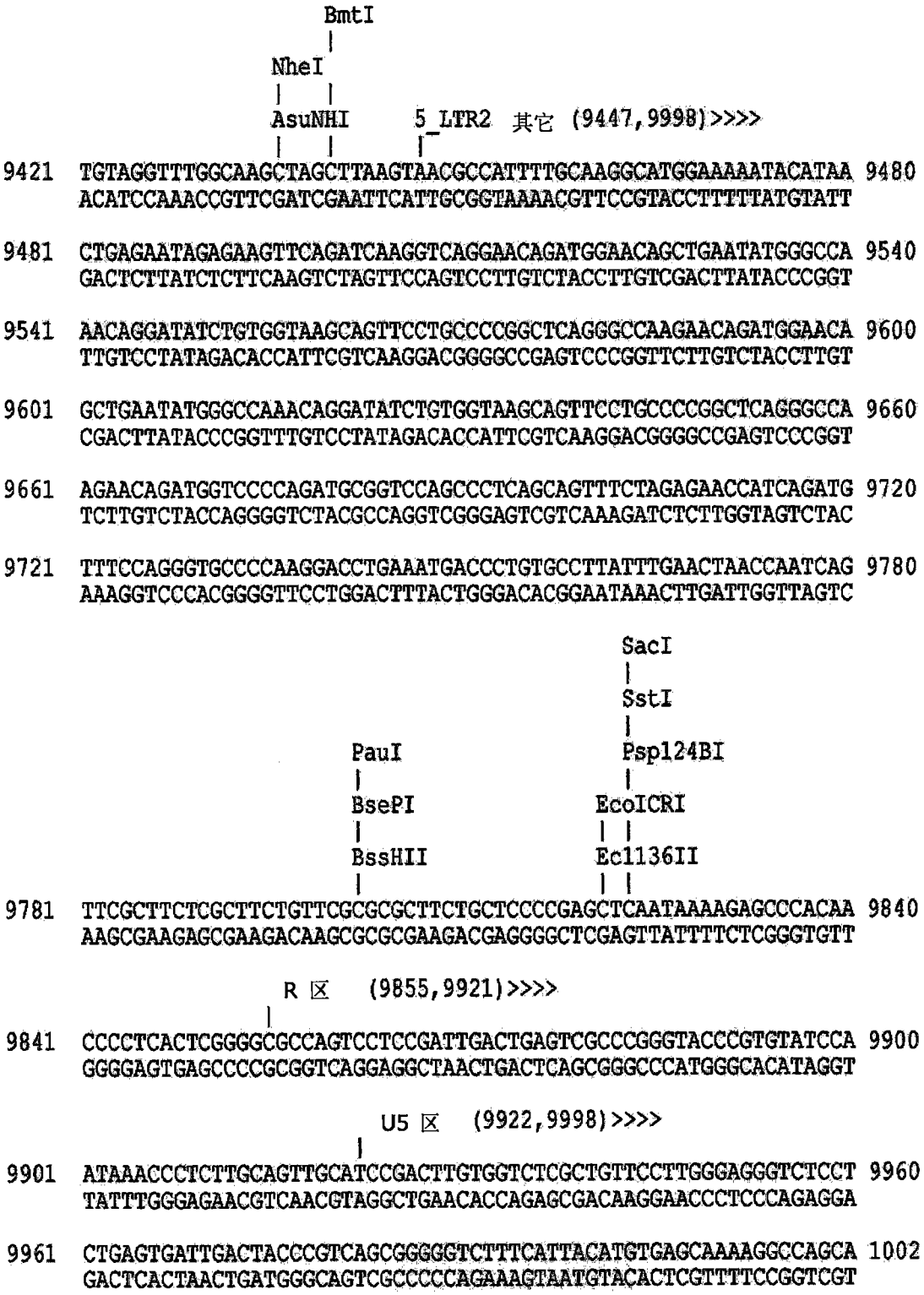

BmtI
NheI
AsuNHI
5_LTR2 其它 (9447,9998)>>>>
9421 TGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGGCATGGAAAAATACATAA 9480
ACATCCAAACCGTTCGATCGAATTCATTGCGGTAAAACGTTCCGTACCTTTTTATGTATT
9481 CTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCA 9540
GACTCTTATCTCTTCAAGTCTAGTTCCAGTCCTTGTCTACCTTGTCGACTTATACCCGGT
9541 AACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACA 9600
TTGTCCTATAGACACCATTCGTCAAGGACGGGGCCGAGTCCCGGTTCTTGTCTACCTTGT
9601 GCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCA 9660
CGACTTATACCCGGTTTGTCCTATAGACACCATTCGTCAAGGACGGGGCCGAGTCCCGGT
9661 AGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATG 9720
TCTTGTCTACCAGGGGTCTACGCCAGGTCGGGAGTCGTCAAAGATCTCTTGGTAGTCTAC
9721 TTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAG 9780
AAAGGTCCCACGGGGTTCCTGGACTTTACTGGGACACGGAATAAACTTGATTGGTTAGTC
SacI
SstI
PauI
Psp124BI
BsePI
EcoICRI
BssHII
Ec1136II
9781 TTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAA 9840
AAGCGAAGAGCGAAGACAAGCGCGCGAAGACGAGGGGCTCGAGTTATTTTCTCGGGTGTT
R 区 (9855,9921)>>>>
9841 CCCCTCACTCGGGGCGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCA 9900
GGGGAGTGAGCCCCGCGGTCAGGAGGCTAACTGACTCAGCGGGCCCATGGGCACATAGGT
U5 区 (9922,9998)>>>>
9901 ATAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCT 9960
TATTTGGGAGAACGTCAACGTAGGCTGAACACCAGAGCGACAAGGAACCCTCCCAGAGGA
9961 CTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTCATTACATGTGAGCAAAAGGCCAGCA 1002
GACTCACTAACTGATGGGCAGTCGCCCCCAGAAAGTAATGTACACTCGTTTTCCGGTCGT

图3D-14


pBR322 起点 (10045,10664)<<<
10021 AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC 10080
TTTCCGGTCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGG
10081 TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 10140
ACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGCTGTCCTGATAT
10141 AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC 10200
TTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGG

图3D-14(续)

```
10201 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTC 10260
      CGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAG

10261 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA 10320
      TGCGACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCT

10321 ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC 10380
      TGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCATTGATAGCAGAACTCAGGTTGGG

10381 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 10440
      CCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCTAATCGTCTCGCTC

10441 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG 10500
      CATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTC

10501 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 10560
      CTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATC

10561 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA 10620
      GAGAACTAGGCCGTTTGTTTGGTGGCGACCATCGCCACCAAAAAAACAAACGTTCGTCGT

10621 GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA 10680
      CTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACTAGAAAAGATGCCCCAGACT

10681 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT 10740
      GCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTCCTA

10741 CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA 10800
      GAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACT

                    amp 标志 (10819,11679)<<<
                    |
10801 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG 10860
      CATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGAC

10861 TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA 10920
      AGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCT

10921 GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC 10980
      CCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGG

10981 AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC 11040
      TCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTG
```

图3D-15

```
11041 TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC 11100
      AAATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGG

11101 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTC 11160
      TCAATTATCAAACGCGTTGCAACAACGGTAACGACGTCCGTAGCACCACAGTGCGAGCAG

11161 GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC 11220
      CAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGG

11221 CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT 11280
      GTACAACACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTCATTCAA

11281 GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC 11340
      CCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAGAGAATGACAGTACGG

11341 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 11400
      TAGGCATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCAC
```

图3D-15(续)

```
11401 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAG 11460
      ATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTGTGCCCTATTATGGCGCGGTGTATC

11461 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT 11520
      GTCTTGAAATTTTCACGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTA

11521 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 11580
      GAATGGCGACAACTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGACTAGAAGTCG

11581 ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA 11640
      TAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTCCGTTTTACGGCGTTT

11641 AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA 11700
      TTTCCCTTATTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAAT

                          amp启动子(11721,11749)<<<
                          |
11701 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA 11760
      AACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTT

11761 AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGA 11820
      TTTATTTGTTTATCCCCAAGGCGCGTGTAAAGGGGCTTTTCACGGTGGACTGCAGATTCT

11821 AACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCT 11880
      TTGGTAATAATAGTACTGTAATTGGATATTTTTATCCGCATAGTGCTCCGGGAAAGCAGA

11881 TCAAGAATTCAT 11892
      AGTTCTTAAGTA
```

图3D-16

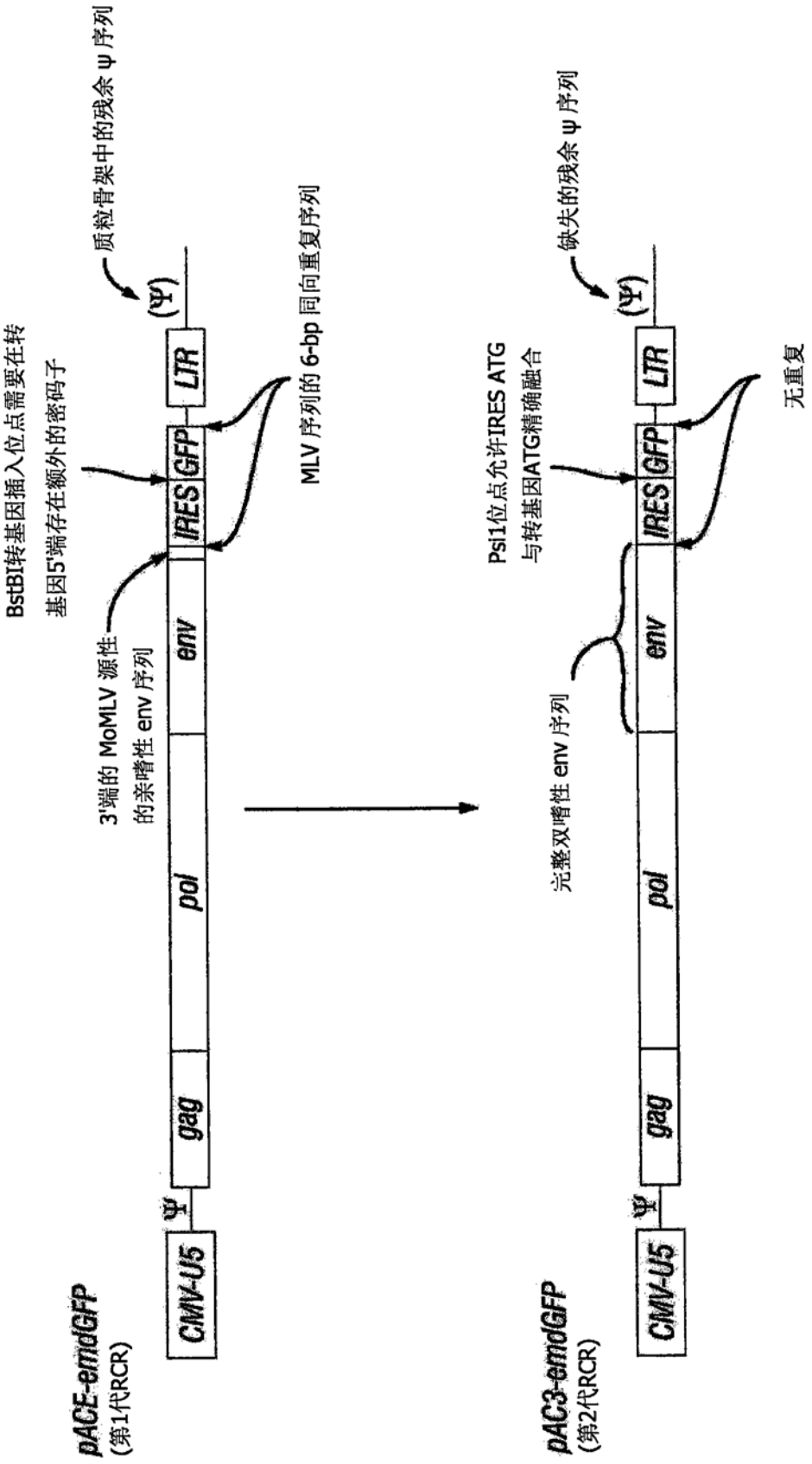

BstBI转基因插入位点需要在转
基因5'端存在额外的密码子
质粒骨架中的残余 ψ 序列
(Ψ)
3'端的 MoMLV 源性
的亲嗜性 env 序列
MLV 序列的 6-bp 同向重复序列
CMV-U5
Ψ
gag
pol
env
IRES
GFP
LTR
pAC2-emdGFP
(第1代RCR)
Psi1位点允许IRES ATG
与转基因ATG精确融合
缺失的残余 ψ 序列
完整双嗜性 env 序列
无重复
pAC3-emdGFP
(第2代RCR)

图3E

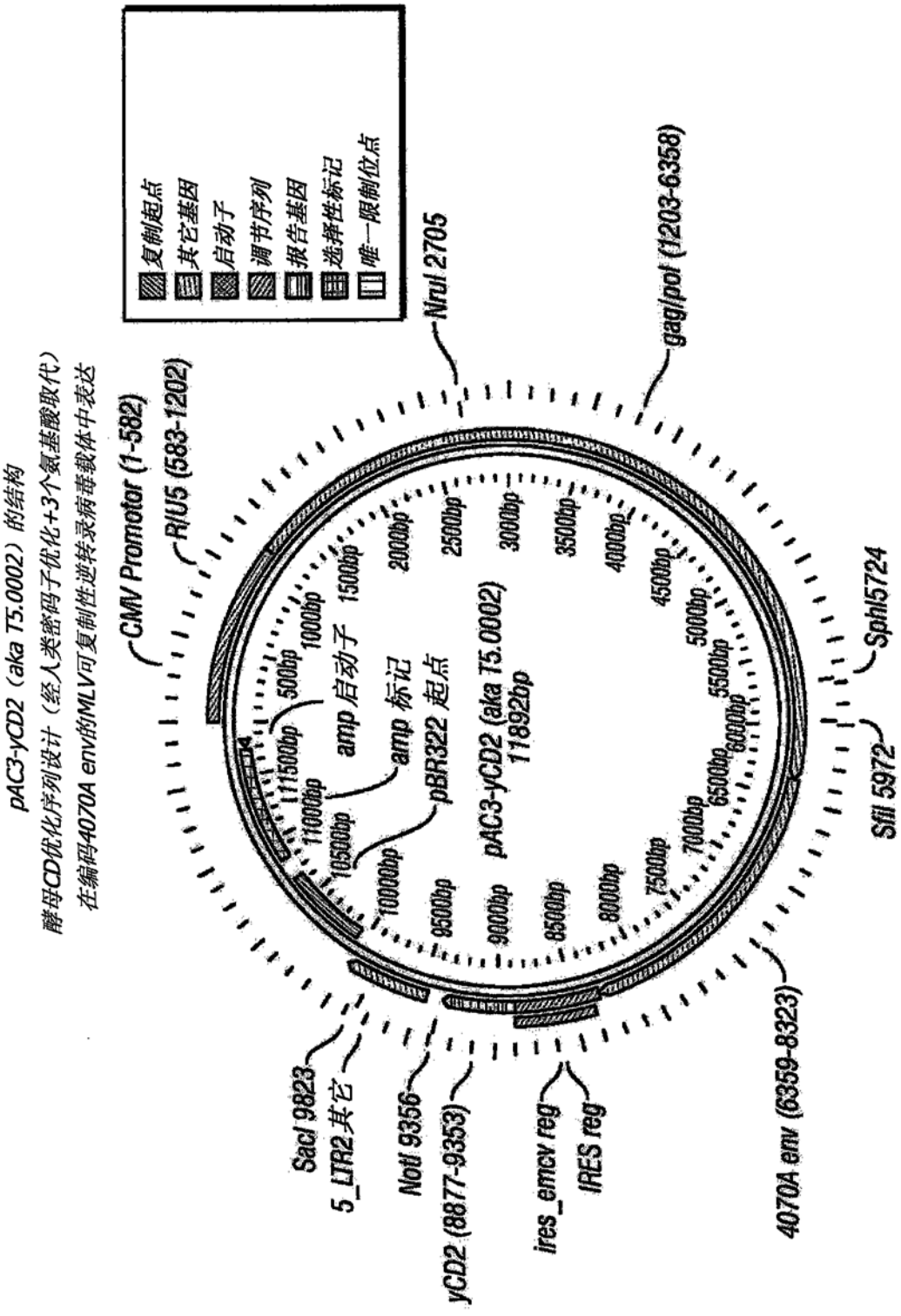

pAC3-yCD2（aka T5.0002）的结构
酵母CD优化序列设计（经人类密码子优化+3个氨基酸取代）
在编码4070A env的MLV可复制性逆转录病毒载体中表达
复制起点
其它基因
启动子
调节序列
报告基因
选择性标记
唯一限制位点
CMV Promotor (1-582)
R/U5 (583-1202)
NruI 2705
gag/pol (1203-6358)
SphI5724
SfiI 5972
4070A env (6359-8323)
ires_emcv reg
IRES reg
yCD2 (8877-9353)
NotI 9356
5_LTR2其它
SacI 9823
amp 启动子
amp 标记
pBR322 起点
pAC3-yCD2 (aka T5.0002)
11892bp

图3F-1

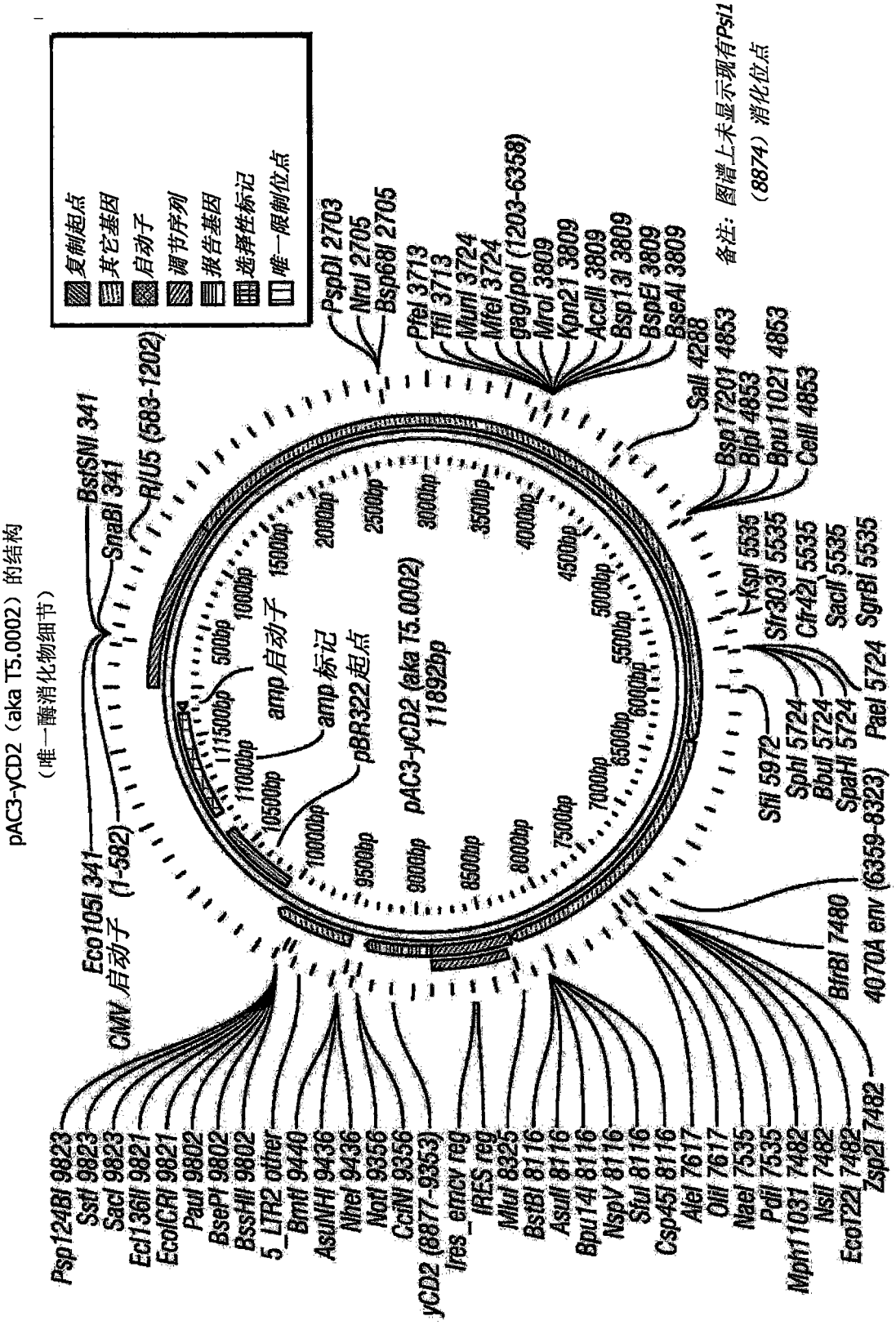

pAC3-yCD2（aka T5.0002）的结构
（唯一酶消化物细节）
复制起点
其它基因
启动子
调节序列
报告基因
选择性标记
唯一限制位点
pAC3-yCD2 (aka T5.0002)
11892bp
amp启动子
amp标记
pBR322起点
CMV 启动子 (1-582)
R/U5 (583-1202)
备注：图谱上未显示现有Psi1
（8874）消化位点
Psp124BI 9823
SstI 9823
SacI 9823
Ecl136II 9821
EcoICRI 9821
PauI 9802
BsePI 9802
BssHII 9802
5_LTR2 other
BmtI 9440
AsuNHI 9436
NheI 9436
NotI 9356
CciNI 9356
yCD2 (8877-9353)
Ires_emcv reg
IRES reg
MluI 8325
BstBI 8116
AsuII 8116
Bpu14I 8116
NspV 8116
SfuI 8116
Csp45I 8116
AleI 7617
OliI 7617
NaeI 7535
PdiI 7535
Mph1103I 7482
NsiI 7482
EcoT22I 7482
Zsp2I 7482
BtrBI 7480
4070A env (6359-8323)
PaeI 5724
SgrBI 5535
SacII 5535
Cfr42I 5535
Sfr303I 5535
KspI 5535
SalI 4288
Bsp1720I 4853
BlpI 4853
Bpu1102I 4853
CelII 4853
PspOI 2703
NruI 2705
Bsp68I 2705
PfeI 3713
TfiI 3713
MunI 3724
MfeI 3724
gagIpol (1203-6358)
MroI 3809
Kon2I 3809
AccIII 3809
Bsp13I 3809
BspEI 3809
BseAI 3809
BstSNI 341
SnaBI 341
Eco105I 341
SfiI 5972
SphI 5724
BbuI 5724
SpaHI 5724

图3F-2

pAC3-yCD2 的连续序列

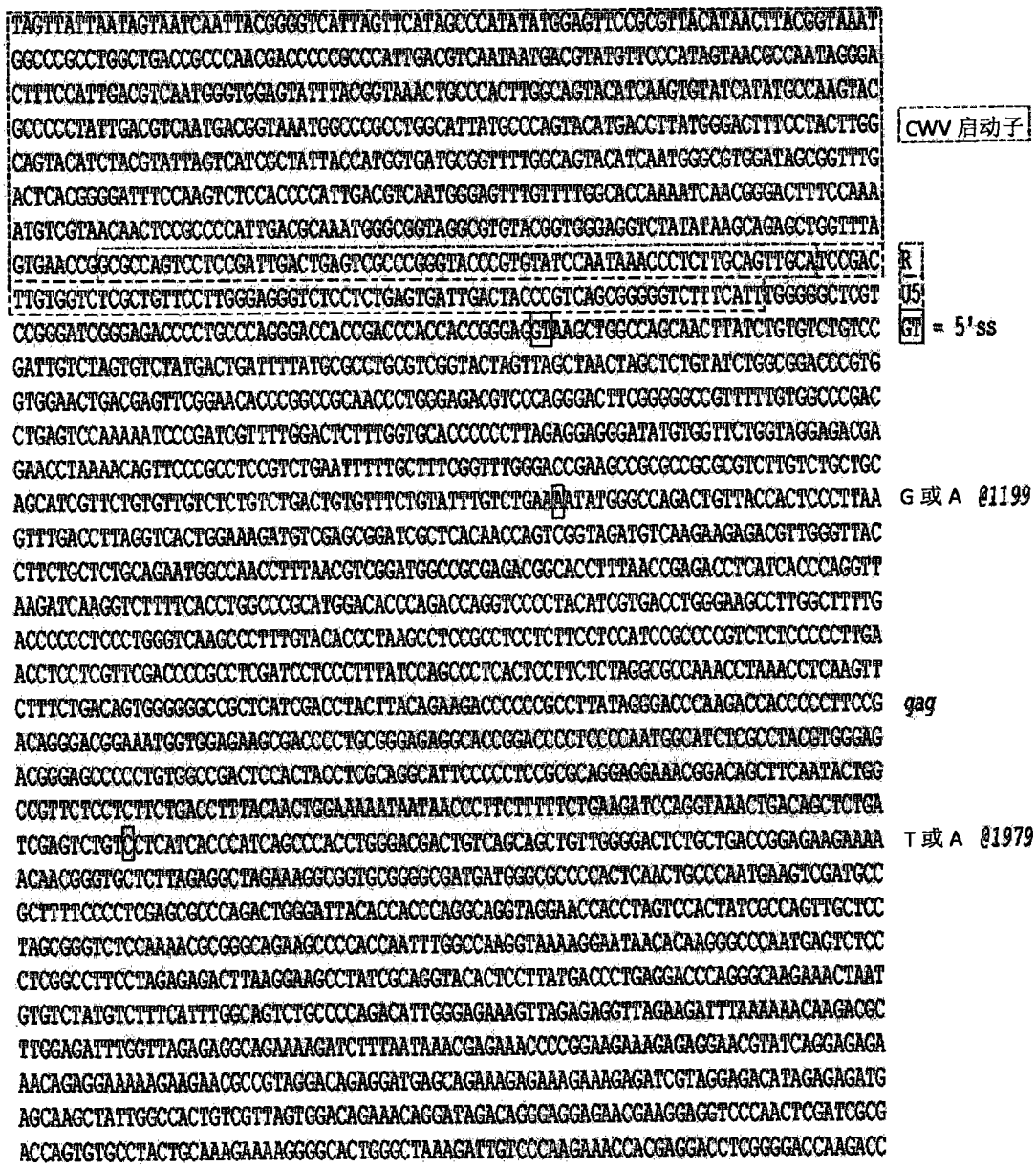


图3F-3

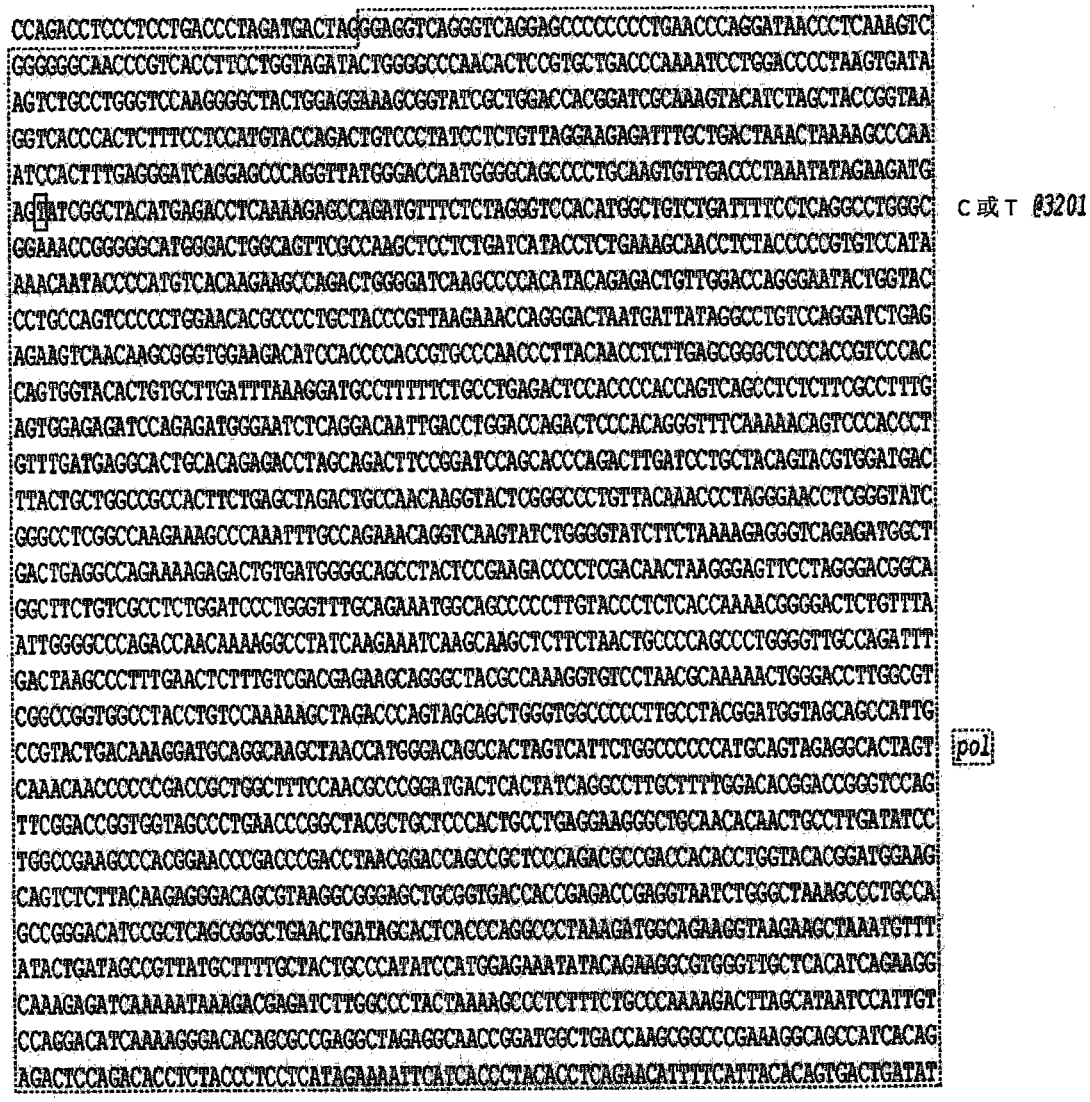


图3F-3(续)

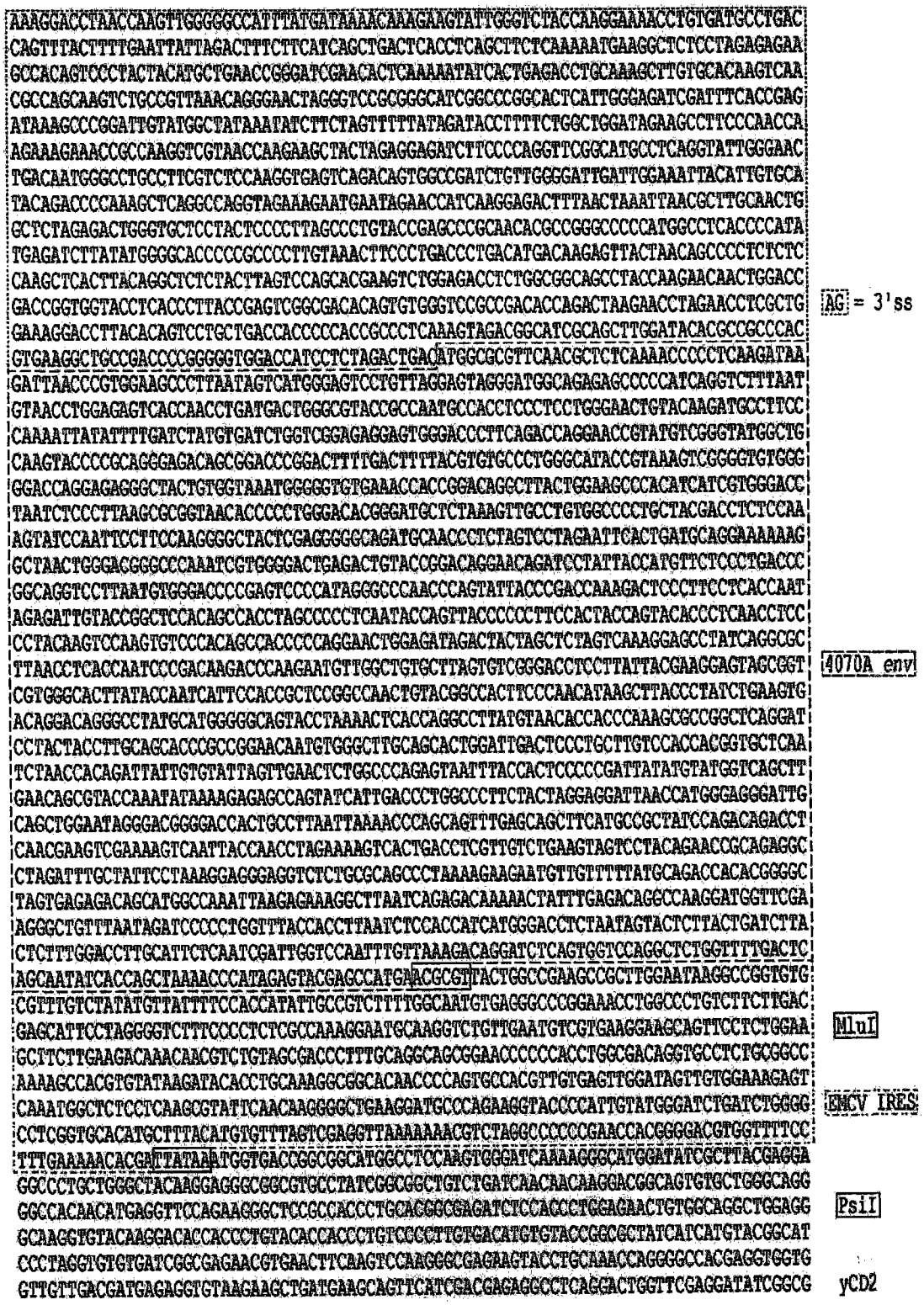


图3F-4

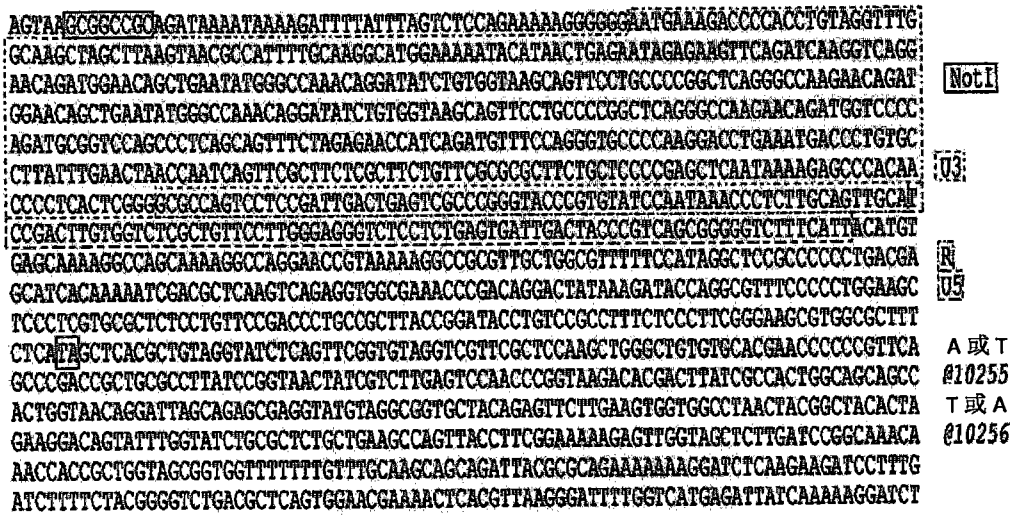


图3F-4(续)

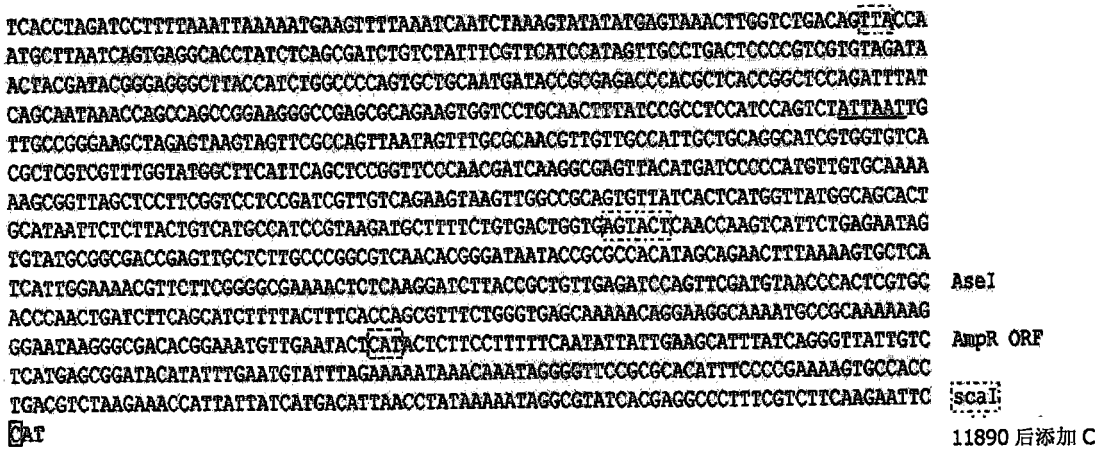


图3F-5

pAC3-yCD2 (唯一消化物序列)

CMV 启动子 (1-582)>>>

```
1    TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG 60
     ATCAATAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGTATATACCTCAAGGC

61   CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATT 120
     GCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGCGGGTAA

121  CACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCA 180
     CTGCAGTTATTACTGCATACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAACTGCAGT

181  ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC 240
     TACCCACCTCATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGG

241  AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA 300
     TTCATGCGGGGGATAACTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCAT
```

Eco105I
|
SnaBI
|
BstSNI
|

```
301  CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC 360
     GTACTGGAATACCCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGTAGCGATAATG

361  CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGG 420
     GTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGCCAAACTGAGTGCCCC

421  ATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG 480
     TAAAGGTTCAGAGGTGGGGTAACTGCAGTTACCCTCAAACAAAACCGTGGTTTTAGTTGC

481  GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT 540
     CCTGAAAGGTTTTACAGCATTGTTGAGGCGGGGTAACTGCGTTTACCCGCCATCCGCACA
```

R 区 (583-650)>>>
|

```
541  ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGGCGCCAGTCCTCCGATTG 600
     TGCCACCCTCCAGATATATTCGTCTCGACCAAATCACTTGGCCGCGGTCAGGAGGCTAAC
```

图3F-6

```
                                                          U5 区    (651-1202)>>>
                                                          |
601   ACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGT 660
      TGACTCAGCGGGCCCATGGGCACATAGGTTATTTGGGAGAACGTCAACGTAGGCTGAACA

661   GGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGTCAGCGGGGGTC 720
      CCAGAGCGACAAGGAACCCTCCCAGAGGAGACTCACTAACTGATGGGCAGTCGCCCCCAG

721   TTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCAGGGACCACCGACCCACCA 780
      AAAGTAAACCCCCGAGCAGGCCCTAGCCCTCTGGGGACGGGTCCCTGGTGGCTGGGTGGT

            5'SS (788)
            |
781   CCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGAC 840
      GGCCCTCCATTCGACCGGTCGTTGAATAGACACAGACAGGCTAACAGATCACAGATACTG

841   TGATTTTATGCGCCTGCGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCG 900
```

图3F-6(续)

```
      ACTAAAATACGCGGACGCAGCCATGATCAATCGATTGATCGAGACATAGACCGCCTGGGC

901   TGGTGGAACTGACGAGTTCGGAACACCCGGCCGCAACCCTGGGAGACGTCCCAGGGACTT 960
      ACCACCTTGACTGCTCAAGCCTTGTGGGCCGGCGTTGGGACCCTCTGCAGGGTCCCTGAA

961   CGGGGGCCGTTTTTGTGGCCCGACCTGAGTCCAAAAATCCCGATCGTTTTGGACTCTTTG 1020
      GCCCCCGGCAAAAACACCGGGCTGGACTCAGGTTTTTAGGGCTAGCAAAACCTGAGAAAC

1021  GTGCACCCCCCTTAGAGGAGGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAACAGT 1080
      CACGTGGGGGGAATCTCCTCCCTATACACCAAGACCATCCTCTGCTCTTGGATTTTGTCA

1081  TCCCGCCTCCGTCTGAATTTTTGCTTTCGGTTTGGGACCGAAGCCGCGCCGCGCGTCTTG 1140
      AGGGCGGAGGCAGACTTAAAAACGAAAGCCAAACCCTGGCTTCGGCGCGGCGCGCAGAAC

1141  TCTGCTGCAGCATCGTTCTGTGTTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAAA 1200
      AGACGACGTCGTAGCAAGACACAACAGAGACAGACTGACACAAAGACATAAACAGACTTT

        gag(1203,2819)>>>
        |
1201  ATATGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTCG 1260
      TATACCCGGTCTGACAATGGTGAGGGAATTCAAACTGGAATCCAGTGACCTTTCTACAGC

1261  AGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGTTGGGTTACCTTCTGCT 1320
      TCGCCTAGCGAGTGTTGGTCAGCCATCTACAGTTCTTCTCTGCAACCCAATGGAAGACGA

1321  CTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACCTTTAACCGAGACC 1380
      GACGTCTTACCGGTTGGAAATTGCAGCCTACCGGCGCTCTGCCGTGGAAATTGGCTCTGG

1381  TCATCACCCAGGTTAAGATCAAGGTCTTTTCACCTGGCCCGCATGGACACCCAGACCAGG 1440
      AGTAGTGGGTCCAATTCTAGTTCCAGAAAAGTGGACCGGGCGTACCTGTGGGTCTGGTCC

1441  TCCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCCTGGGTCAAGCCCT 1500
      AGGGGATGTAGCACTGGACCCTTCGGAACCGAAAACTGGGGGGAGGGACCCAGTTCGGGA

1501  TTGTACACCCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCTCTCCCCCTTGAAC 1560
      AACATGTGGGATTCGGAGGCGGAGGAGAAGGAGGTAGGCGGGGCAGAGAGGGGGAACTTG

1561  CTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGCG 1620
      GAGGAGCAAGCTGGGGCGGAGCTAGGAGGGAAATAGGTCGGGAGTGAGGAAGAGATCCGC

1621  CCAAACCTAAACCTCAAGTTCTTTCTGACAGTGGGGGGCCGCTCATCGACCTACTTACAG 1680
      GGTTTGGATTTGGAGTTCAAGAAAGACTGTCACCCCCCGGCGAGTAGCTGGATGAATGTC

1681  AAGACCCCCCGCCTTATAGGGACCCAAGACCACCCCCTTCCGACAGGGACGGAAATGGTG 1740
      TTCTGGGGGGCGGAATATCCCTGGGTTCTGGTGGGGGAAGGCTGTCCCTGCCTTTACCAC
```

图3F-7

```
1741  GAGAAGCGACCCCTGCGGGAGAGGCACCGGACCCCTCCCCAATGGCATCTCGCCTACGTG 1800
      CTCTTCGCTGGGGACGCCCTCTCCGTGGCCTGGGGAGGGGTTACCGTAGAGCGGATGCAC

1801  GGAGACGGGAGCCCCCTGTGGCCGACTCCACTACCTCGCAGGCATTCCCCCTCCGCGCAG 1860
      CCTCTGCCCTCGGGGGACACCGGCTGAGGTGATGGAGCGTCCGTAAGGGGGAGGCGCGTC

1861  GAGGAAACGGACAGCTTCAATACTGGCCGTTCTCCTCTTCTGACCTTTACAACTGGAAAA 1920
      CTCCTTTGCCTGTCGAAGTTATGACCGGCAAGAGGAGAAGACTGGAAATGTTGACCTTTT

1921  ATAATAACCCTTCTTTTTCTGAAGATCCAGGTAAACTGACAGCTCTGATCGAGTCTGTCC 1980
      TATTATTGGGAAGAAAAAGACTTCTAGGTCCATTTGACTGTCGAGACTAGCTCAGACAGG

1981  TCATCACCCATCAGCCCACCTGGGACGACTGTCAGCAGCTGTTGGGGACTCTGCTGACCG 2040
      AGTAGTGGGTAGTCGGGTGGACCCTGCTGACAGTCGTCGACAACCCCTGAGACGACTGGC
```

图3F-7(续)

```
2041  GAGAAGAAAAACAACGGGTGCTCTTAGAGGCTAGAAAGGCGGTGCGGGGCGATGATGGGC 2100
      CTCTTCTTTTTGTTGCCCACGAGAATCTCCGATCTTTCCGCCACGCCCCGCTACTACCCG

2101  GCCCCACTCAACTGCCCAATGAAGTCGATGCCGCTTTTCCCCTCGAGCGCCCAGACTGGG 2160
      CGGGGTGAGTTGACGGGTTACTTCAGCTACGGCGAAAAGGGGAGCTCGCGGGTCTGACCC

2161  ATTACACCACCCAGGCAGGTAGGAACCACCTAGTCCACTATCGCCAGTTGCTCCTAGCGG 2220
      TAATGTGGTGGGTCCGTCCATCCTTGGTGGATCAGGTGATAGCGGTCAACGAGGATCGCC

2221  GTCTCCAAAACGCGGGCAGAAGCCCCACCAATTTGGCCAAGGTAAAAGGAATAACACAAG 2280
      CAGAGGTTTTGCGCCCGTCTTCGGGGTGGTTAAACCGGTTCCATTTTCCTTATTGTGTTC

2281  GGCCCAATGAGTCTCCCTCGGCCTTCCTAGAGAGACTTAAGGAAGCCTATCGCAGGTACA 2340
      CCGGGTTACTCAGAGGGAGCCGGAAGGATCTCTCTGAATTCCTTCGGATAGCGTCCATGT

2341  CTCCTTATGACCCTGAGGACCCAGGGCAAGAAACTAATGTGTCTATGTCTTTCATTTGGC 2400
      GAGGAATACTGGGACTCCTGGGTCCCGTTCTTTGATTACACAGATACAGAAAGTAAACCG

2401  AGTCTGCCCCAGACATTGGGAGAAAGTTAGAGAGGTTAGAAGATTTAAAAAACAAGACGC 2460
      TCAGACGGGGTCTGTAACCCTCTTTCAATCTCTCCAATCTTCTAAATTTTTTGTTCTGCG

2461  TTGGAGATTTGGTTAGAGAGGCAGAAAAGATCTTTAATAAACGAGAAACCCCGGAAGAAA 2520
      AACCTCTAAACCAATCTCTCCGTCTTTTCTAGAAATTATTTGCTCTTTGGGGCCTTCTTT

2521  GAGAGGAACGTATCAGGAGAGAAACAGAGGAAAAAGAAGAACGCCGTAGGACAGAGGATG 2580
      CTCTCCTTGCATAGTCCTCTCTTTGTCTCCTTTTTCTTCTTGCGGCATCCTGTCTCCTAC

2581  AGCAGAAAGAGAAAGAAAGAGATCGTAGGAGACATAGAGAGATGAGCAAGCTATTGGCCA 2640
      TCGTCTTTCTCTTTCTTTCTCTAGCATCCTCTGTATCTCTCTACTCGTTCGATAACCGGT

2641  CTGTCGTTAGTGGACAGAAACAGGATAGACAGGGAGGAGAACGAAGGAGGTCCCAACTCG 2700
      GACAGCAATCACCTGTCTTTGTCCTATCTGTCCCTCCTCTTGCTTCCTCCAGGGTTGAGC

         NruI
          |
         Bsp68I
          |
       PspDI
       | |
2701  ATCGCGACCAGTGTGCCTACTGCAAAGAAAAGGGGCACTGGGCTAAAGATTGTCCCAAGA 2760
      TAGCGCTGGTCACACGGATGACGTTTCTTTTCCCCGTGACCCGATTTCTAACAGGGTTCT

                                                    pol(2820,6358)>>>
                                                       |
```

图3F-8

```
2761  AACCACGAGGACCTCGGGGACCAAGACCCCAGACCTCCCTCCTGACCCTAGATGACTAGG 2820
      TTGGTGCTCCTGGAGCCCCTGGTTCTGGGGTCTGGAGGGAGGACTGGGATCTACTGATCC

2821  GAGGTCAGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGCAAC 2880
      CTCCAGTCCCAGTCCTCGGGGGGGGACTTGGGTCCTATTGGGAGTTTCAGCCCCCCGTTG

2881  CCGTCACCTTCCTGGTAGATACTGGGGCCCAACACTCCGTGCTGACCCAAAATCCTGGAC 2940
      GGCAGTGGAAGGACCATCTATGACCCCGGGTTGTGAGGCACGACTGGGTTTTAGGACCTG

2941  CCCTAAGTGATAAGTCTGCCTGGGTCCAAGGGGCTACTGGAGGAAAGCGGTATCGCTGGA 3000
      GGGATTCACTATTCAGACGGACCCAGGTTCCCCGATGACCTCCTTTCGCCATAGCGACCT

3001  CCACGGATCGCAAAGTACATCTAGCTACCGGTAAGGTCACCCACTCTTTCCTCCATGTAC 3060
      GGTGCCTAGCGTTTCATGTAGATCGATGGCCATTCCAGTGGGTGAGAAAGGAGGTACATG

3061  CAGACTGTCCCTATCCTCTGTTAGGAAGAGATTTGCTGACTAAACTAAAAGCCCAAATCC 3120
      GTCTGACAGGGATAGGAGACAATCCTTCTCTAAACGACTGATTTGATTTTCGGGTTTAGG
```

图3F-8(续)

```
3121  ACTTTGAGGGATCAGGAGCCCAGGTTATGGGACCAATGGGGCAGCCCCTGCAAGTGTTGA 3180
      TGAAACTCCCTAGTCCTCGGGTCCAATACCCTGGTTACCCCGTCGGGGACGTTCACAACT

3181  CCCTAAATATAGAAGATGAGTATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTTCTC 3240
      GGGATTTATATCTTCTACTCATAGCCGATGTACTCTGGAGTTTTCTCGGTCTACAAAGAG

3241  TAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGCGGAAACCGGGGGCATGGGAC 3300
      ATCCCAGGTGTACCGACAGACTAAAAGGAGTCCGGACCCGCCTTTGGCCCCCGTACCCTG

                  3'SS (3314)
                  |
3301  TGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCTACCCCCGTGTCCA 3360
      ACCGTCAAGCGGTTCGAGGAGACTAGTATGGAGACTTTCGTTGGAGATGGGGGCACAGGT

3361  TAAAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGAC 3420
      ATTTTGTTATGGGGTACAGTGTTCTTCGGTCTGACCCCTAGTTCGGGGTGTATGTCTCTG

3421  TGTTGGACCAGGGAATACTGGTACCCTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCG 3480
      ACAACCTGGTCCCTTATGACCATGGGACGGTCAGGGGGACCTTGTGCGGGGACGATGGGC

3481  TTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAAGTCAACAAGC 3540
      AATTCTTTGGTCCCTGATTACTAATATCCGGACAGGTCCTAGACTCTCTTCAGTTGTTCG

3541  GGGTGGAAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTTGAGCGGGCTCCCAC 3600
      CCCACCTTCTGTAGGTGGGGTGGCACGGGTTGGGAATGTTGGAGAACTCGCCCGAGGGTG

3601  CGTCCCACCAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTTCTGCCTGAGACTCC 3660
      GCAGGGTGGTCACCATGTGACACGAACTAAATTTCCTACGGAAAAAGACGGACTCTGAGG

                                                          PfeI
                                                          |
                                                          TfiI
                                                          |
3661  ACCCCACCAGTCAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
      TGGGGTGGTCAGTCGGAGAGAAGCGGAAACTCACCTCTCTAGGTCTCTACCCTTAGAGTC

         MfeI
         |
         MunI
         |
```

图3F-9

```
3721  GACAATTGACCTGGACCAGACTCCCACAGGGTTTCAAAAACAGTCCCACCCTGTTTGATG 3780
      CTGTTAACTGGACCTGGTCTGAGGGTGTCCCAAAGTTTTTGTCAGGGTGGGACAAACTAC

                                  MroI
                                  |
                                  BseAI
                                  |
                                  Bsp13I
                                  |
                                  BspEI
                                  |
                                  Kpn2I
                                  |
                                  AccIII
                                  |
3781  AGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTAC 3840
      TCCGTGACGTGTCTCTGGATCGTCTGAAGGCCTAGGTCGTGGGTCTGAACTAGGACGATG

3841  AGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAAGGTACTC 3900
      TCATGCACCTACTGAATGACGACCGGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG
```

图3F-9(续)

```
3901  GGGCCCTGTTACAAACCCTAGGGAACCTCGGGTATCGGGCCTCGGCCAAGAAAGCCCAAA 3960
      CCCGGGACAATGTTTGGGATCCCTTGGAGCCCATAGCCCGGAGCCGGTTCTTTCGGGTTT

3961  TTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCTAAAAGAGGGTCAGAGATGGCTGA 4020
      AAACGGTCTTTGTCCAGTTCATAGACCCCATAGAAGATTTTCTCCCAGTCTCTACCGACT

4021  CTGAGGCCAGAAAAGAGACTGTGATGGGGCAGCCTACTCCGAAGACCCCTCGACAACTAA 4080
      GACTCCGGTCTTTTCTCTGACACTACCCCGTCGGATGAGGCTTCTGGGGAGCTGTTGATT

4081  GGGAGTTCCTAGGGACGGCAGGCTTCTGTCGCCTCTGGATCCCTGGGTTTGCAGAAATGG 4140
      CCCTCAAGGATCCCTGCCGTCCGAAGACAGCGGAGACCTAGGGACCCAAACGTCTTTACC

4141  CAGCCCCCTTGTACCCTCTCACCAAAACGGGGACTCTGTTTAATTGGGGCCCAGACCAAC 4200
      GTCGGGGGAACATGGGAGAGTGGTTTTGCCCCTGAGACAAATTAACCCCGGGTCTGGTTG

4201  AAAAGGCCTATCAAGAAATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAG 4260
      TTTTCCGGATAGTTCTTTAGTTCGTTCGAGAAGATTGACGGGGTCGGGACCCCAACGGTC

                                  SalI
                                   |
4261  ATTTGACTAAGCCCTTTGAACTCTTTGTCGACGAGAAGCAGGGCTACGCCAAAGGTGTCC 4320
      TAAACTGATTCGGGAAACTTGAGAAACAGCTGCTCTTCGTCCCGATGCGGTTTCCACAGG

4321  TAACGCAAAAACTGGGACCTTGGCGTCGGCCGGTGGCCTACCTGTCCAAAAAGCTAGACC 4380
      ATTGCGTTTTTGACCCTGGAACCGCAGCCGGCCACCGGATGGACAGGTTTTTCGATCTGG

4381  CAGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGGTAGCAGCCATTGCCGTACTGACAA 4440
      GTCATCGTCGACCCACCGGGGGAACGGATGCCTACCATCGTCGGTAACGGCATGACTGTT

4441  AGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTCTGGCCCCCCATGCAGTAG 4500
      TCCTACGTCCGTTCGATTGGTACCCTGTCGGTGATCAGTAAGACCGGGGGGGTACGTCATC

4501  AGGCACTAGTCAAACAACCCCCCGACCGCTGGCTTTCCAACGCCCGGATGACTCACTATC 4560
      TCCGTGATCAGTTTGTTGGGGGGCTGGCGACCGAAAGGTTGCGGGCCTACTGAGTGATAG

4561  AGGCCTTGCTTTTGGACACGGACCGGGTCCAGTTCGGACCGGTGGTAGCCCTGAACCCGG 4620
      TCCGGAACGAAAACCTGTGCCTGGCCCAGGTCAAGCCTGGCCACCATCGGGACTTGGGCC

4621  CTACGCTGCTCCCACTGCCTGAGGAAGGGCTGCAACACAACTGCCTTGATATCCTGGCCG 4680
      GATGCGACGAGGGTGACGGACTCCTTCCCGACGTTGTGTTGACGGAACTATAGGACCGGC

4681  AAGCCCACGGAACCCGACCCGACCTAACGGACCAGCCGCTCCCAGACGCCGACCACACCT 4740
      TTCGGGTGCCTTGGGCTGGGCTGGATTGCCTGGTCGGCGAGGGTCTGCGGCTGGTGTGGA

4741  GGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGA 4800
      CCATGTGCCTACCTTCGTCAGAGAATGTTCTCCCTGTCGCATTCCGCCCTCGACGCCACT
```

图3F-10

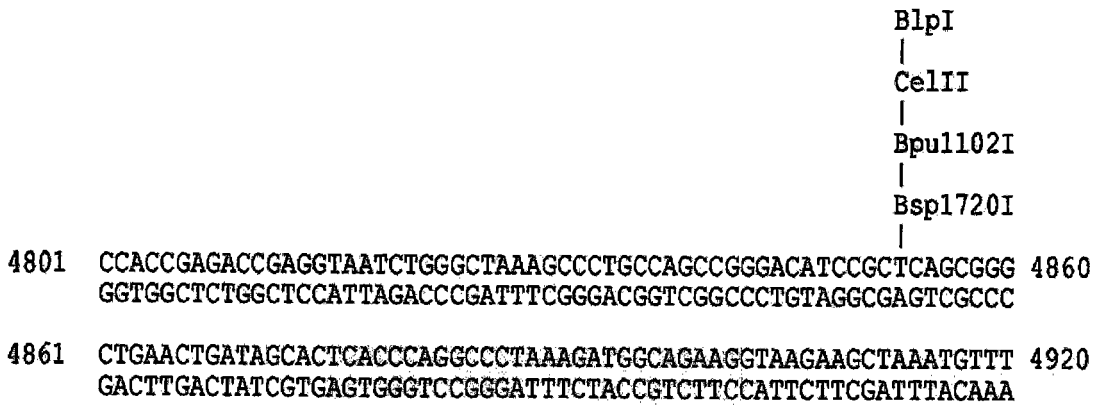

BlpI
CelII
Bpu1102I
Bsp1720I
4801 CCACCGAGACCGAGGTAATCTGGGCTAAAGCCCTGCCAGCCGGGACATCCGCTCAGCGGG 4860
GGTGGCTCTGGCTCCATTAGACCCGATTTCGGGACGGTCGGCCCTGTAGGCGAGTCGCCC
4861 CTGAACTGATAGCACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTT 4920
GACTTGACTATCGTGAGTGGGTCCGGGATTTCTACCGTCTTCCATTCTTCGATTTACAAA

图3F-10(续)

```
4921  ATACTGATAGCCGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGC 4980
      TATGACTATCGGCAATACGAAAACGATGACGGGTATAGGTACCTCTTTATATGTCTTCCG

4981  GTGGGTTGCTCACATCAGAAGGCAAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTAC 5040
      CACCCAACGAGTGTAGTCTTCCGTTTCTCTAGTTTTTATTTCTGCTCTAGAACCGGGATG

5041  TAAAAGCCCTCTTTCTGCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAAGG 5100
      ATTTTCGGGAGAAAGACGGGTTTTCTGAATCGTATTAGGTAACAGGTCCTGTAGTTTTCC

5101  GACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAAGGCAGCCA 5160
      CTGTGTCGCGGCTCCGATCTCCGTTGGCCTACCGACTGGTTCGCCGGGCTTTCCGTCGGT

5161  TCACAGAGACTCCAGACACCTCTACCCTCCTCATAGAAAATTCATCACCCTACACCTCAG 5220
      AGTGTCTCTGAGGTCTGTGGAGATGGGAGGAGTATCTTTTAAGTAGTGGGATGTGGAGTC

5221  AACATTTTCATTACACAGTGACTGATATAAAGGACCTAACCAAGTTGGGGGCCATTTATG 5280
      TTGTAAAAGTAATGTGTCACTGACTATATTTCCTGGATTGGTTCAACCCCCGGTAAATAC

5281  ATAAAACAAAGAAGTATTGGGTCTACCAAGGAAAACCTGTGATGCCTGACCAGTTTACTT 5340
      TATTTTGTTTCTTCATAACCCAGATGGTTCCTTTTGGACACTACGGACTGGTCAAATGAA

5341  TTGAATTATTAGACTTTCTTCATCAGCTGACTCACCTCAGCTTCTCAAAAATGAAGGCTC 5400
      AACTTAATAATCTGAAAGAAGTAGTCGACTGAGTGGAGTCGAAGAGTTTTTACTTCCGAG

5401  TCCTAGAGAGAAGCCACAGTCCCTACTACATGCTGAACCGGGATCGAACACTCAAAAATA 5460
      AGGATCTCTCTTCGGTGTCAGGGATGATGTACGACTTGGCCCTAGCTTGTGAGTTTTTAT

5461  TCACTGAGACCTGCAAAGCTTGTGCACAAGTCAACGCCAGCAAGTCTGCCGTTAAACAGG 5520
      AGTGACTCTGGACGTTTCGAACACGTGTTCAGTTGCGGTCGTTCAGACGGCAATTTGTCC
```

SacII
|
SgrBI
|
Cfr42I
|
Sfr303I
|
KspI
|

图3F-11

```
5521  GAACTAGGGTCCGCGGGCATCGGCCCGGCACTCATTGGGAGATCGATTTCACCGAGATAA 5580
      CTTGATCCCAGGCGCCCGTAGCCGGGCCGTGAGTAACCCTCTAGCTAAAGTGGCTCTATT

5581  AGCCCGGATTGTATGGCTATAAATATCTTCTAGTTTTTATAGATACCTTTTCTGGCTGGA 5640
      TCGGGCCTAACATACCGATATTTATAGAAGATCAAAAATATCTATGGAAAAGACCGACCT

5641  TAGAAGCCTTCCCAACCAAGAAAGAAACCGCCAAGGTCGTAACCAAGAAGCTACTAGAGG 5700
      ATCTTCGGAAGGGTTGGTTCTTTCTTTGGCGGTTCCAGCATTGGTTCTTCGATGATCTCC

                                  PaeI
                                  |
                                  BbuI
                                  |
                                  SpaHI
                                  |
                                  SphI
                                  |
5701  AGATCTTCCCCAGGTTCGGCATGCCTCAGGTATTGGGAACTGACAATGGGCCTGCCTTCG 5760
      TCTAGAAGGGGTCCAAGCCGTACGGAGTCCATAACCCTTGACTGTTACCCGGACGGAAGC
```

图3F-11(续)

```
5761  TCTCCAAGGTGAGTCAGACAGTGGCCGATCTGTTGGGGATTGATTGGAAATTACATTGTG 5820
      AGAGGTTCCACTCAGTCTGTCACCGGCTAGACAACCCCTAACTAACCTTTAATGTAACAC

5821  CATACAGACCCCAAAGCTCAGGCCAGGTAGAAAGAATGAATAGAACCATCAAGGAGACTT 5880
      GTATGTCTGGGGTTTCGAGTCCGGTCCATCTTTCTTACTTATCTTGGTAGTTCCTCTGAA

5881  TAACTAAATTAACGCTTGCAACTGGCTCTAGAGACTGGGTGCTCCTACTCCCCTTAGCCC 5940
      ATTGATTTAATTGCGAACGTTGACCGAGATCTCTGACCCACGAGGATGAGGGGAATCGGG

                                    SfiI
                                     |
5941  TGTACCGAGCCCGCAACACGCCGGGCCCCCATGGCCTCACCCCATATGAGATCTTATATG 6000
      ACATGGCTCGGGCGTTGTGCGGCCCGGGGGTACCGGAGTGGGGTATACTCTAGAATATAC

6001  GGGCACCCCCGCCCCTTGTAAACTTCCCTGACCCTGACATGACAAGAGTTACTAACAGCC 6060
      CCCGTGGGGGCGGGGAACATTTGAAGGGACTGGGACTGTACTGTTCTCAATGATTGTCGG

6061  CCTCTCTCCAAGCTCACTTACAGGCTCTCTACTTAGTCCAGCACGAAGTCTGGAGACCTC 6120
      GGAGAGAGGTTCGAGTGAATGTCCGAGAGATGAATCAGGTCGTGCTTCAGACCTCTGGAG

6121  TGGCGGCAGCCTACCAAGAACAACTGGACCGACCGGTGGTACCTCACCCTTACCGAGTCG 6180
      ACCGCCGTCGGATGGTTCTTGTTGACCTGGCTGGCCACCATGGAGTGGGAATGGCTCAGC

6181  GCGACACAGTGTGGGTCCGCCGACACCAGACTAAGAACCTAGAACCTCGCTGGAAAGGAC 6240
      CGCTGTGTCACACCCAGGCGGCTGTGGTCTGATTCTTGGATCTTGGAGCGACCTTTCCTG

6241  CTTACACAGTCCTGCTGACCACCCCCACCGCCCTCAAAGTAGACGGCATCGCAGCTTGGA 6300
      GAATGTGTCAGGACGACTGGTGGGGGTGGCGGGAGTTTCATCTGCCGTAGCGTCGAACCT

                                                    4070A env (6359-8323)>>>
                                                          |
6301  TACACGCCGCCCACGTGAAGGCTGCCGACCCCGGGGGTGGACCATCCTCTAGACTGACAT 6360
      ATGTGCGGCGGGTGCACTTCCGACGGCTGGGGCCCCCACCTGGTAGGAGATCTGACTGTA

6361  GGCGCGTTCAACGCTCTCAAAACCCCCTCAAGATAAGATTAACCCGTGGAAGCCCTTAAT 6420
      CCGCGCAAGTTGCGAGAGTTTTGGGGGAGTTCTATTCTAATTGGGCACCTTCGGGAATTA

6421  AGTCATGGGAGTCCTGTTAGGAGTAGGGATGGCAGAGAGCCCCCATCAGGTCTTTAATGT 6480
      TCAGTACCCTCAGGACAATCCTCATCCCTACCGTCTCTCGGGGGTAGTCCAGAAATTACA

6481  AACCTGGAGAGTCACCAACCTGATGACTGGGCGTACCGCCAATGCCACCTCCCTCCTGGG 6540
      TTGGACCTCTCAGTGGTTGGACTACTGACCCGCATGGCGGTTACGGTGGAGGGAGGACCC
```

图3F-12

```
6541  AACTGTACAAGATGCCTTCCCAAAATTATATTTTGATCTATGTGATCTGGTCGGAGAGGA 6600
      TTGACATGTTCTACGGAAGGGTTTTAATATAAAACTAGATACACTAGACCAGCCTCTCCT

6601  GTGGGACCCTTCAGACCAGGAACCGTATGTCGGGTATGGCTGCAAGTACCCCGCAGGGAG 6660
      CACCCTGGGAAGTCTGGTCCTTGGCATACAGCCCATACCGACGTTCATGGGGCGTCCCTC

6661  ACAGCGGACCCGGACTTTTGACTTTTACGTGTGCCCTGGGCATACCGTAAAGTCGGGGTG 6720
      TGTCGCCTGGGCCTGAAAACTGAAAATGCACACGGGACCCGTATGGCATTTCAGCCCCAC

6721  TGGGGGACCAGGAGAGGGCTACTGTGGTAAATGGGGGTGTGAAACCACCGGACAGGCTTA 6780
      ACCCCCTGGTCCTCTCCCGATGACACCATTTACCCCCACACTTTGGTGGCCTGTCCGAAT

6781  CTGGAAGCCCACATCATCGTGGGACCTAATCTCCCTTAAGCGCGGTAACACCCCCTGGGA 6840
      GACCTTCGGGTGTAGTAGCACCCTGGATTAGAGGGAATTCGCGCCATTGTGGGGGACCCT

6841  CACGGGATGCTCTAAAGTTGCCTGTGGCCCCTGCTACGACCTCTCCAAAGTATCCAATTC 6900
      GTGCCCTACGAGATTTCAACGGACACCGGGGACGATGCTGGAGAGGTTTCATAGGTTAAG
```

图3F-12(续)

```
6901  CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCTCTAGTCCTAGAATTCACTGATGC 6960
      GAAGGTTCCCCGATGAGCTCCCCCGTCTACGTTGGGAGATCAGGATCTTAAGTGACTACG

6961  AGGAAAAAAGGCTAACTGGGACGGGCCCAAATCGTGGGGACTGAGACTGTACCGGACAGG 7020
      TCCTTTTTTCCGATTGACCCTGCCCGGGTTTAGCACCCCTGACTCTGACATGGCCTGTCC

7021  AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCCTTAATGTGGGACCCCGAGT 7080
      TTGTCTAGGATAATGGTACAAGAGGGACTGGGCCGTCCAGGAATTACACCCTGGGGCTCA

7081  CCCCATAGGGCCCAACCCAGTATTACCCGACCAAAGACTCCCTTCCTCACCAATAGAGAT 7140
      GGGGTATCCCGGGTTGGGTCATAATGGGCTGGTTTCTGAGGGAAGGAGTGGTTATCTCTA

7141  TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACCAGTTACCCCCCTTCCACTACCAG 7200
      ACATGGCCGAGGTGTCGGTGGATCGGGGGAGTTATGGTCAATGGGGGGAAGGTGATGGTC

7201  TACACCCTCAACCTCCCCTACAAGTCCAAGTGTCCCACAGCCACCCCCAGGAACTGGAGA 7260
      ATGTGGGAGTTGGAGGGGATGTTCAGGTTCACAGGGTGTCGGTGGGGGTCCTTGACCTCT

7261  TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCGCTTAACCTCACCAATCCCGACAA 7320
      ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCGCGAATTGGAGTGGTTAGGGCTGTT

7321  GACCCAAGAATGTTGGCTGTGCTTAGTGTCGGGACCTCCTTATTACGAAGGAGTAGCGGT 7380
      CTGGGTTCTTACAACCGACACGAATCACAGCCCTGGAGGAATAATGCTTCCTCATCGCCA

7381  CGTGGGCACTTATACCAATCATTCCACCGCTCCGGCCAACTGTACGGCCACTTCCCAACA 7440
      GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTTGACATGCCGGTGAAGGGTTGT
```

EcoT22I
NsiI
Mph1103I
Zsp2I
BfrBI

图3F-13

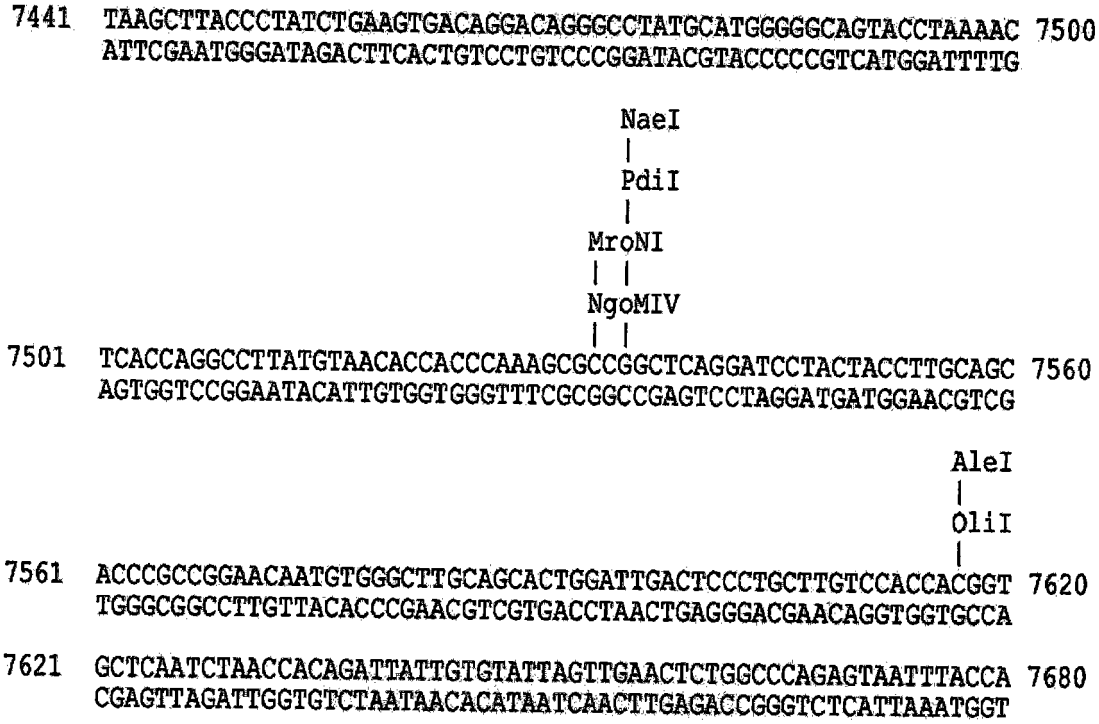

7441 TAAGCTTACCCTATCTGAAGTGACAGGACAGGGCCTATGCATGGGGGCAGTACCTAAAAC 7500
ATTCGAATGGGATAGACTTCACTGTCCTGTCCCGGATACGTACCCCCGTCATGGATTTTG
NaeI
PdiI
MroNI
NgoMIV
7501 TCACCAGGCCTTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTTGCAGC 7560
AGTGGTCCGGAATACATTGTGGTGGGTTTCGCGGCCGAGTCCTAGGATGATGGAACGTCG
AleI
OliI
7561 ACCCGCCGGAACAATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCACGGT 7620
TGGGCGGCCTTGTTACACCCGAACGTCGTGACCTAACTGAGGGACGAACAGGTGGTGCCA
7621 GCTCAATCTAACCACAGATTATTGTGTATTAGTTGAACTCTGGCCCAGAGTAATTTACCA 7680
CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAAATGGT

图3F-13(续)

```
7681  CTCCCCCGATTATATGTATGGTCAGCTTGAACAGCGTACCAAATATAAAAGAGAGCCAGT 7740
      GAGGGGGCTAATATACATACCAGTCGAACTTGTCGCATGGTTTATATTTTCTCTCGGTCA

7741  ATCATTGACCCTGGCCCTTCTACTAGGAGGATTAACCATGGGAGGGATTGCAGCTGGAAT 7800
      TAGTAACTGGGACCGGGAAGATGATCCTCCTAATTGGTACCCTCCCTAACGTCGACCTTA

7801  AGGGACGGGGACCACTGCCTTAATTAAAACCCAGCAGTTTGAGCAGCTTCATGCCGCTAT 7860
      TCCCTGCCCCTGGTGACGGAATTAATTTTGGGTCGTCAAACTCGTCGAAGTACGGCGATA

7861  CCAGACAGACCTCAACGAAGTCGAAAAGTCAATTACCAACCTAGAAAAGTCACTGACCTC 7920
      GGTCTGTCTGGAGTTGCTTCAGCTTTTCAGTTAATGGTTGGATCTTTTCAGTGACTGGAG

7921  GTTGTCTGAAGTAGTCCTACAGAACCGCAGAGGCCTAGATTTGCTATTCCTAAAGGAGGG 7980
      CAACAGACTTCATCAGGATGTCTTGGCGTCTCCGGATCTAAACGATAAGGATTTCCTCCC

7981  AGGTCTCTGCGCAGCCCTAAAAGAAGAATGTTGTTTTTATGCAGACCACACGGGGCTAGT 8040
      TCCAGAGACGCGTCGGGATTTTCTTCTTACAACAAAAATACGTCTGGTGTGCCCCGATCA

8041  GAGAGACAGCATGGCCAAATTAAGAGAAAGGCTTAATCAGAGACAAAAACTATTTGAGAC 8100
      CTCTCTGTCGTACCGGTTTAATTCTCTTTCCGAATTAGTCTCTGTTTTTGATAAACTCTG

                    NspV
                    |
                    BstBI
                    |
                    Bspl19I
                    |
                    AsuII
                    |
                    Csp45I
                    |
                    SfuI
                    |
                    Bpu14I
                    |
                    BspT104I
                    |
8101  AGGCCAAGGATGGTTCGAAGGGCTGTTTAATAGATCCCCCTGGTTTACCACCTTAATCTC 8160
      TCCGGTTCCTACCAAGCTTCCCGACAAATTATCTAGGGGGACCAAATGGTGGAATTAGAG

8161  CACCATCATGGGACCTCTAATAGTACTCTTACTGATCTTACTCTTTGGACCTTGCATTCT 8220
      GTGGTAGTACCCTGGAGATTATCATGAGAATGACTAGAATGAGAAACCTGGAACGTAAGA

8221  CAATCGATTGGTCCAATTTGTTAAAGACAGGATCTCAGTGGTCCAGGCTCTGGTTTTGAC 8280
      GTTAGCTAACCAGGTTAAACAATTTCTGTCCTAGAGTCACCAGGTCCGAGACCAAAACTG
```

图3F-14

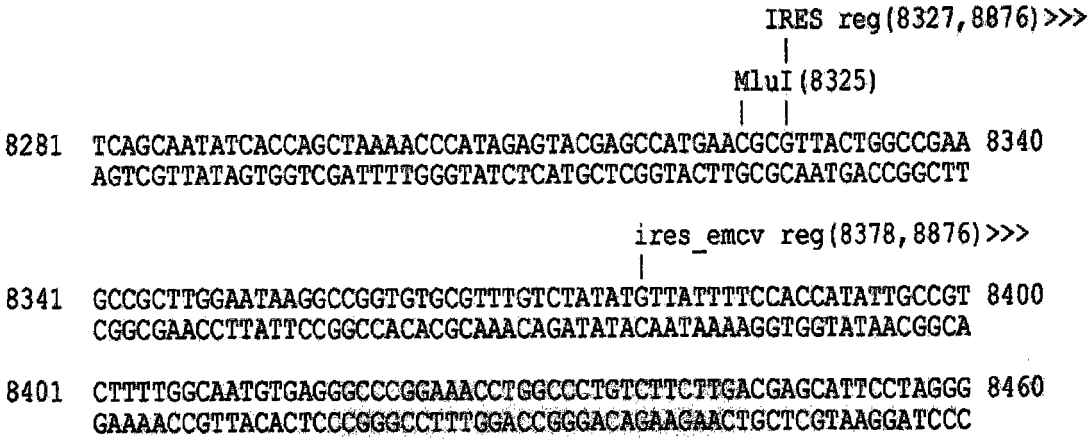

IRES reg(8327,8876)>>>
MluI(8325)
8281 TCAGCAATATCACCAGCTAAAACCCATAGAGTACGAGCCATGAACGCGTTACTGGCCGAA 8340
AGTCGTTATAGTGGTCGATTTTGGGTATCTCATGCTCGGTACTTGCGCAATGACCGGCTT
ires_emcv reg(8378,8876)>>>
8341 GCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGT 8400
CGGCGAACCTTATTCCGGCCACACGCAAACAGATATACAATAAAAGGTGGTATAACGGCA
8401 CTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGG 8460
GAAAACCGTTACACTCCCGGGCCTTTGGACCGGGACAGAAGAACTGCTCGTAAGGATCCC

图3F-14(续)

```
8461  GTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTC 8520
      CAGAAAGGGGAGAGCGGTTTCCTTACGTTCCAGACAACTTACAGCACTTCCTTCGTCAAG

8521  CTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACC 8580
      GAGACCTTCGAAGAACTTCTGTTTGTTGCAGACATCGCTGGGAAACGTCCGTCGCCTTGG

8581  CCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAA 8640
      GGGGTGGACCGCTGTCCACGGAGACGCCGGTTTTCGGTGCACATATTCTATGTGGACGTT

8641  AGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGC 8700
      TCCGCCGTGTTGGGGTCACGGTGCAACACTCAACCTATCAACACCTTTCTCAGTTTACCG

8701  TCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGG 8760
      AGAGGAGTTCGCATAAGTTGTTCCCCGACTTCCTACGGGTCTTCCATGGGGTAACATACC

8761  GATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAAC 8820
      CTAGACTAGACCCCGGAGCCACGTGTACGAAATGTACACAAATCAGCTCCAATTTTTTTG

                                                    yCD2 (8877,9353)>>>
                                                          |
                                                    PsiI(8874)
                                                      |   |
8821  GTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATTATAAATGG 8880
      CAGATCCGGGGGGGCTTGGTGCCCCTGCACCAAAAGGAAACTTTTTGTGCTAATATTTACC

8881  TGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTTACGAGGAGG 8940
      ACTGGCCGCCGTACCGGAGGTTCACCCTAGTTTTCCCGTACCTATAGCGAATGCTCCTCC

8941  CCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAGG 9000
      GGGACGACCCGATGTTCCTCCCGCCGCACGGATAGCCGCCGACAGACTAGTTGTTGTTCC

9001  ACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGC 9060
      TGCCGTCACACGACCCGTCCCCGGTGTTGTACTCCAAGGTCTTCCCGAGGCGGTGGGACG

9061  ACGGCGAGATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACA 9120
      TGCCGCTCTAGAGGTGGGACCTCTTGACACCGTCCGACCTCCCGTTCCACATGTTCCTGT

9121  CCACCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACG 9180
      GGTGGGACATGTGGTGGGACAGGGGAACACTGTACACATGGCCGCGATAGTAGTACATGC
```

图3F-15

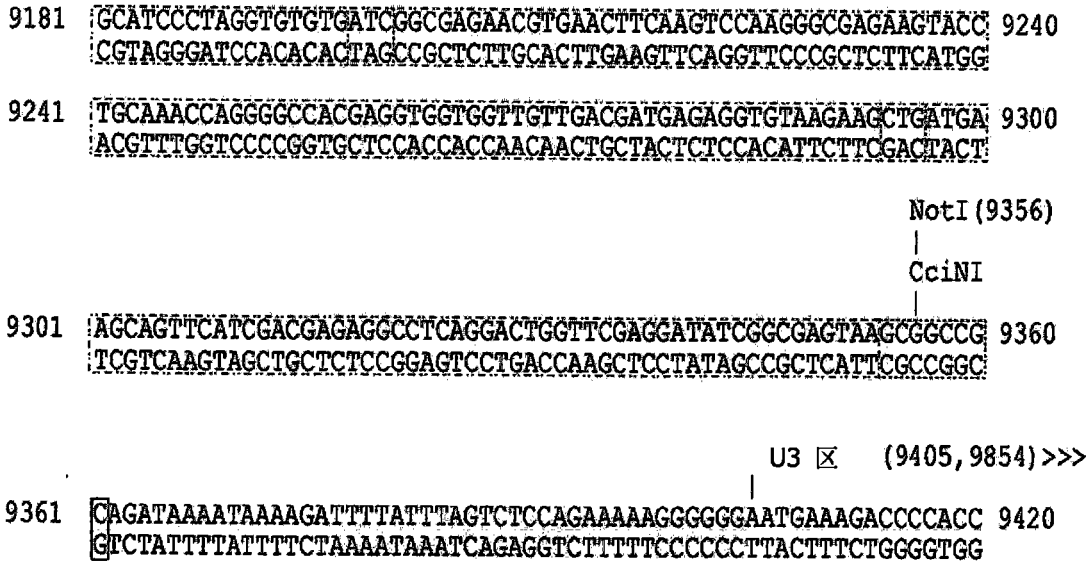

9181
GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCCAAGGGCGAGAAGTACC 9240
CGTAGGGATCCACACACTAGCCGCTCTTGCACTTGAAGTTCAGGTTCCCGCTCTTCATGG
9241
TGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTGATGA 9300
ACGTTTGGTCCCCGGTGCTCCACCACCAACAACTGCTACTCTCCACATTCTTCGACTACT
NotI(9356)
CciNI
9301
AGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTAAGCGGCCG 9360
TCGTCAAGTAGCTGCTCTCCGGAGTCCTGACCAAGCTCCTATAGCCGCTCATTCGCCGGC
U3 区 (9405,9854)>>>
9361
CAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACC 9420
GTCTATTTTATTTTCTAAAATAAATCAGAGGTCTTTTTCCCCCCTTACTTTCTGGGGTGG

图3F-15(续)

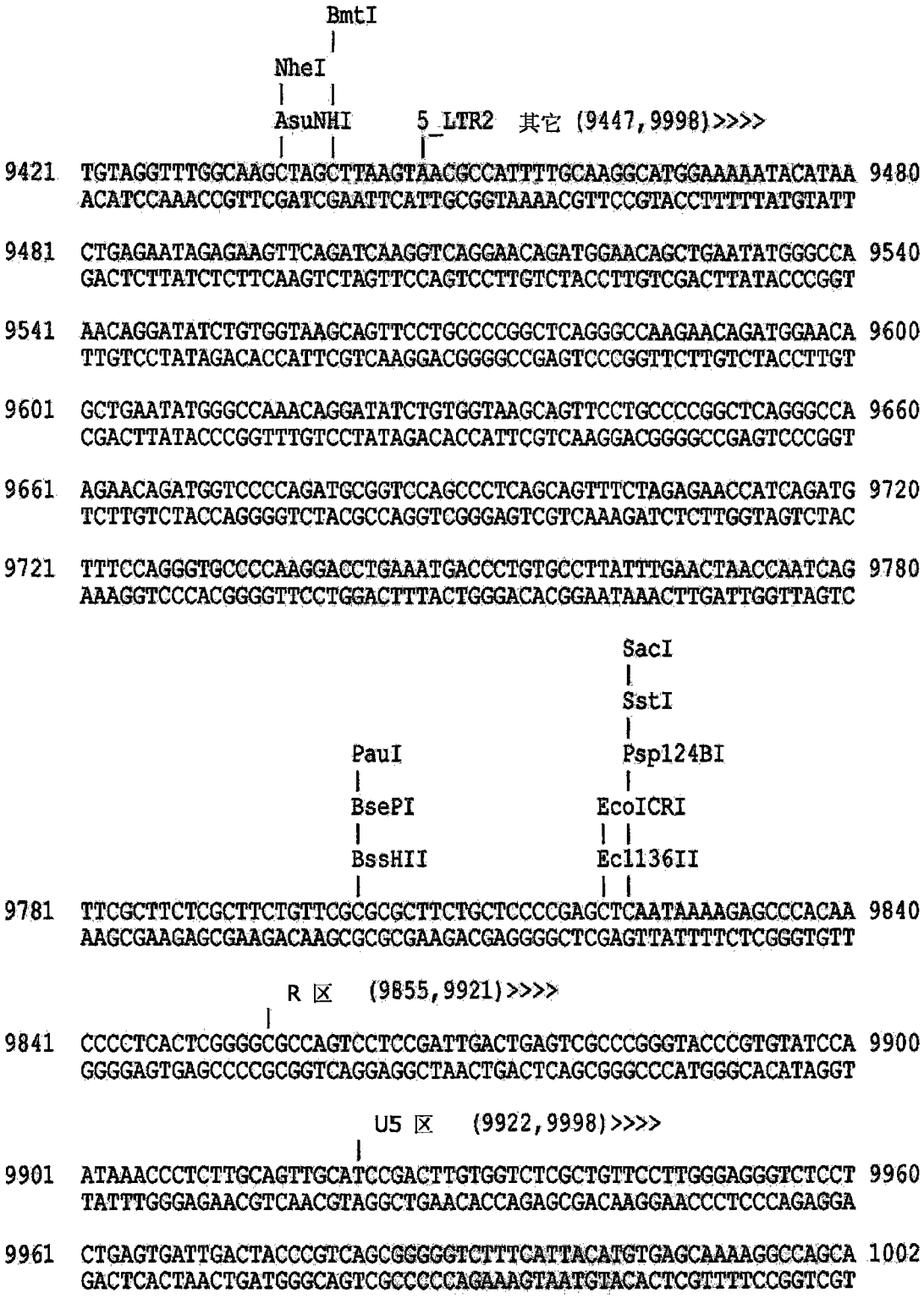

BmtI
NheI
AsuNHI
5_LTR2 其它 (9447,9998)>>>>
9421 TGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGGCATGGAAAAATACATAA 9480
ACATCCAAACCGTTCGATCGAATTCATTGCGGTAAAACGTTCCGTACCTTTTTATGTATT
9481 CTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCA 9540
GACTCTTATCTCTTCAAGTCTAGTTCCAGTCCTTGTCTACCTTGTCGACTTATACCCGGT
9541 AACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACA 9600
TTGTCCTATAGACACCATTCGTCAAGGACGGGGCCGAGTCCCGGTTCTTGTCTACCTTGT
9601 GCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCA 9660
CGACTTATACCCGGTTTGTCCTATAGACACCATTCGTCAAGGACGGGGCCGAGTCCCGGT
9661 AGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATG 9720
TCTTGTCTACCAGGGGTCTACGCCAGGTCGGGAGTCGTCAAAGATCTCTTGGTAGTCTAC
9721 TTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAG 9780
AAAGGTCCCACGGGGTTCCTGGACTTTACTGGGACACGGAATAAACTTGATTGGTTAGTC
SacI
SstI
PauI
Psp124BI
BsePI
EcoICRI
BssHII
Ec1136II
9781 TTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAA 9840
AAGCGAAGAGCGAAGACAAGCGCGCGAAGACGAGGGGCTCGAGTTATTTTCTCGGGTGTT
R 区 (9855,9921)>>>>
9841 CCCCTCACTCGGGGCGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCA 9900
GGGGAGTGAGCCCCGCGGTCAGGAGGCTAACTGACTCAGCGGGCCCATGGGCACATAGGT
U5 区 (9922,9998)>>>>
9901 ATAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCT 9960
TATTTGGGAGAACGTCAACGTAGGCTGAACACCAGAGCGACAAGGAACCCTCCCAGAGGA
9961 CTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTCATTACATGTGAGCAAAAGGCCAGCA 1002
GACTCACTAACTGATGGGCAGTCGCCCCCAGAAAGTAATGTACACTCGTTTTCCGGTCGT

图3F-16


pBR322 起点 (10045,10664)<<<
10021 AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC 10080
TTTCCGGTCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGG
10081 TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 10140
ACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGCTGTCCTGATAT
10141 AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC 10200
TTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGG

图3F-16(续)

```
10201 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC 10260
      CGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTATCGAG

10261 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA 10320
      TGCGACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCT

10321 ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC 10380
      TGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCATTGATAGCAGAACTCAGGTTGGG

10381 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 10440
      CCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCTAATCGTCTCGCTC

10441 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG 10500
      CATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTC

10501 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 10560
      CTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATC

10561 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA 10620
      GAGAACTAGGCCGTTTGTTTGGTGGCGACCATCGCCACCAAAAAAACAAACGTTCGTCGT

10621 GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA 10680
      CTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACTAGAAAAGATGCCCCAGACT

10681 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT 10740
      GCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTCCTA

10741 CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA 10800
      GAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACT

                          amp 标记 (10819,11679)<<<
                          |
10801 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG 10860
      CATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGAC

10861 TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA 10920
      AGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCT

10921 GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC 10980
      CCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGG

10981 AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC 11040
      TCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTG
```

图3F-17

```
11041 TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC 11100
      AAATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGG

11101 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTC 11160
      TCAATTATCAAACGCGTTGCAACAACGGTAACGACGTCCGTAGCACCACAGTGCGAGCAG

11161 GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC 11220
      CAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGG

11221 CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT 11280
      GTACAACACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTCATTCAA

11281 GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC 11340
      CCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAGAGAATGACAGTACGG

11341 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 11400
      TAGGCATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCAC
```

图3F-17(续)

```
11401 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAG 11460
      ATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTGTGCCCTATTATGGCGCGGTGTATC

11461 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT 11520
      GTCTTGAAATTTTCACGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTA

11521 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 11580
      GAATGGCGACAACTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGACTAGAAGTCG

11581 ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA 11640
      TAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTCCGTTTTACGGCGTTT

11641 AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA 11700
      TTTCCCTTATTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAAT

                          amp prom(11721,11749)<<<
                          |
11701 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA 11760
      AACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTT

11761 AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGA 11820
      TTTATTTGTTTATCCCCAAGGCGCGTGTAAAGGGGCTTTTCACGGTGGACTGCAGATTCT

11821 AACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCT 11880
      TTGGTAATAATAGTACTGTAATTGGATATTTTTATCCGCATAGTGCTCCGGGAAAGCAGA

11881 TCAAGAATTCCAT 11893
      AGTTCTTAAGGTA
```

图3F-18

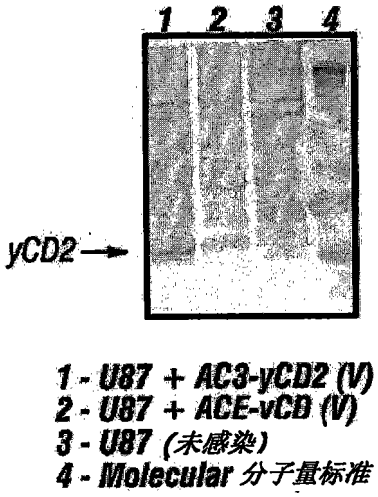

1 2 3 4
yCD2
1 - U87 + AC3-yCD2 (V)
2 - U87 + ACE-vCD (V)
3 - U87 (未感染)
4 - Molecular 分子量标准

图4

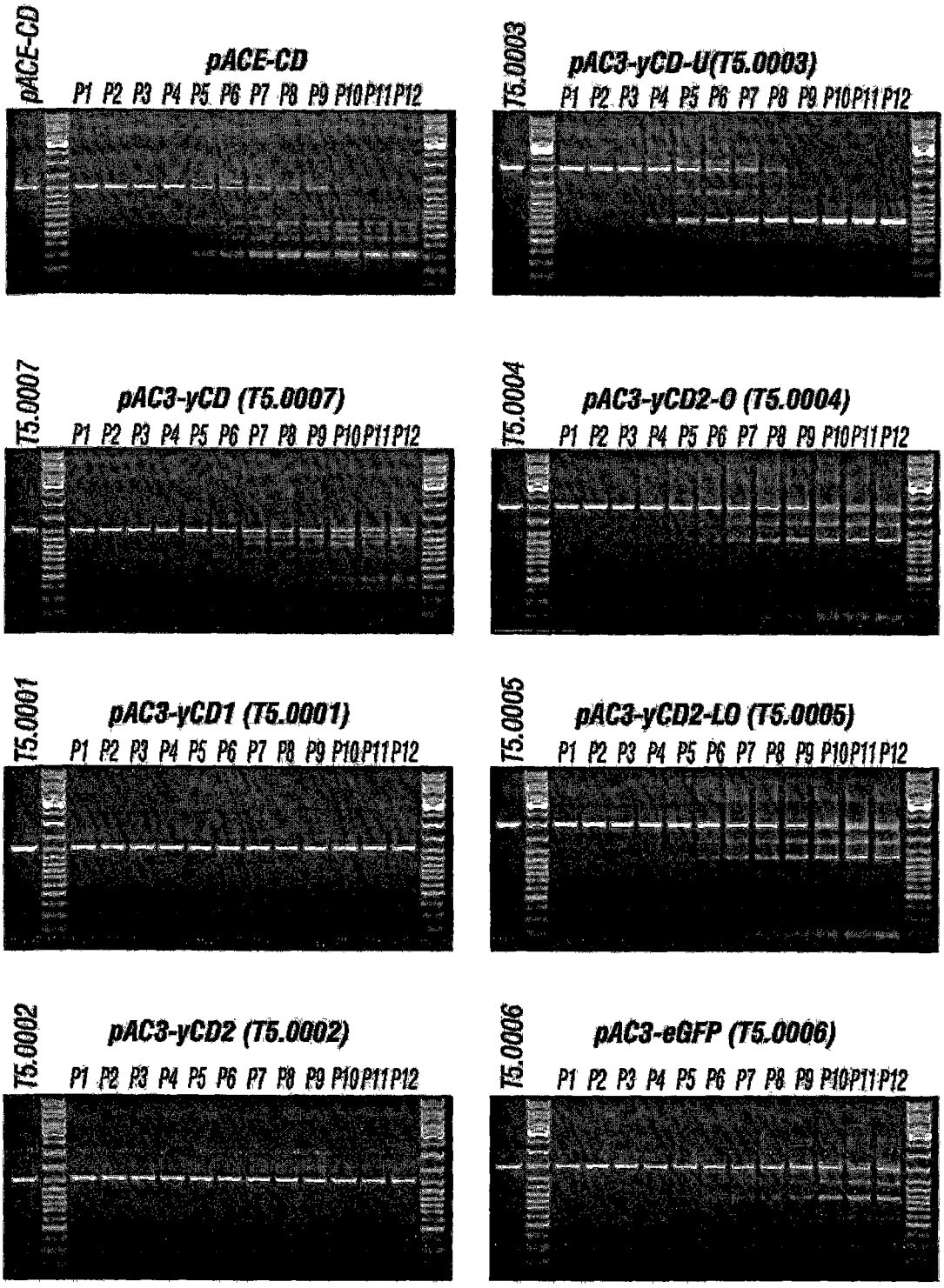

pACE-CD
pACE-CD
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0003
pAC3-yCD-U(T5.0003)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0007
pAC3-yCD (T5.0007)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0004
pAC3-yCD2-O (T5.0004)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0001
pAC3-yCD1 (T5.0001)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0005
pAC3-yCD2-LO (T5.0005)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0002
pAC3-yCD2 (T5.0002)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12
T5.0006
pAC3-eGFP (T5.0006)
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12

图5

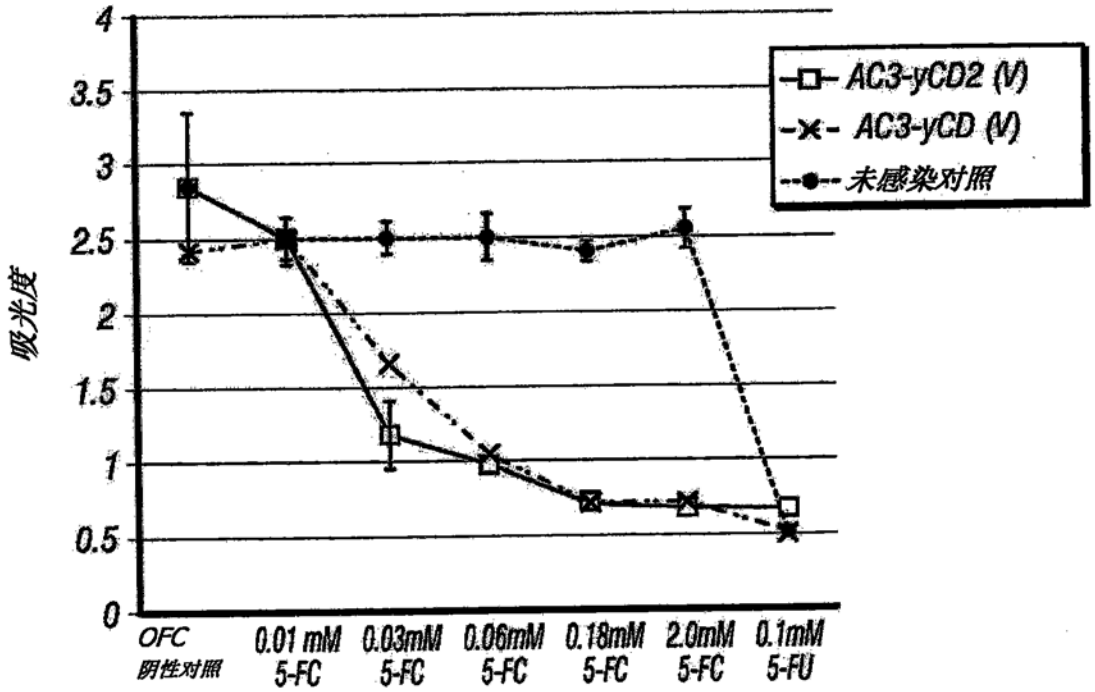

AC3-yCD2 (V)
AC3-yCD (V)
未感染对照
吸光度
4
3.5
3
2.5
2
1.5
1
0.5
0
OFC
阴性对照
0.01 mM 5-FC
0.03mM 5-FC
0.06mM 5-FC
0.18mM 5-FC
2.0mM 5-FC
0.1mM 5-FU

图6A

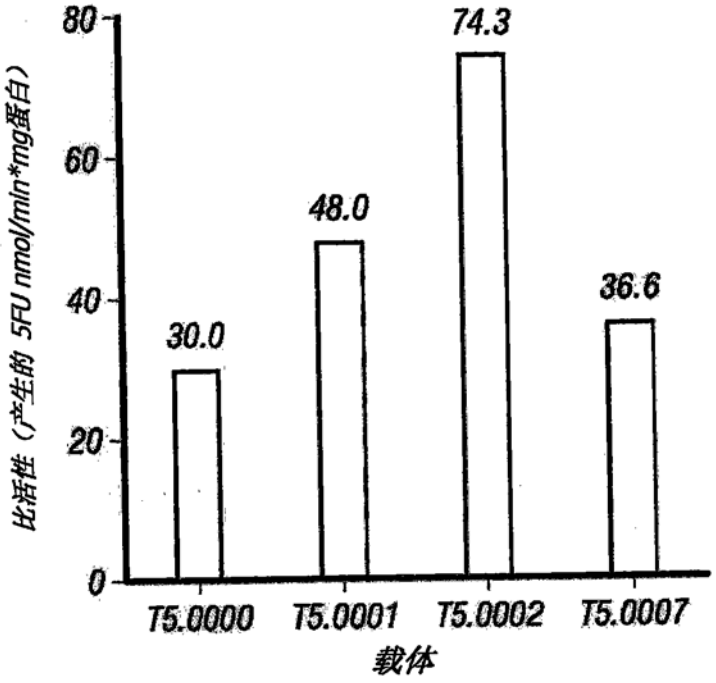

比活性（产生的 5FU nmol/min*mg蛋白）
80
60
40
20
0
30.0
48.0
74.3
36.6
T5.0000
T5.0001
T5.0002
T5.0007
载体

图6B

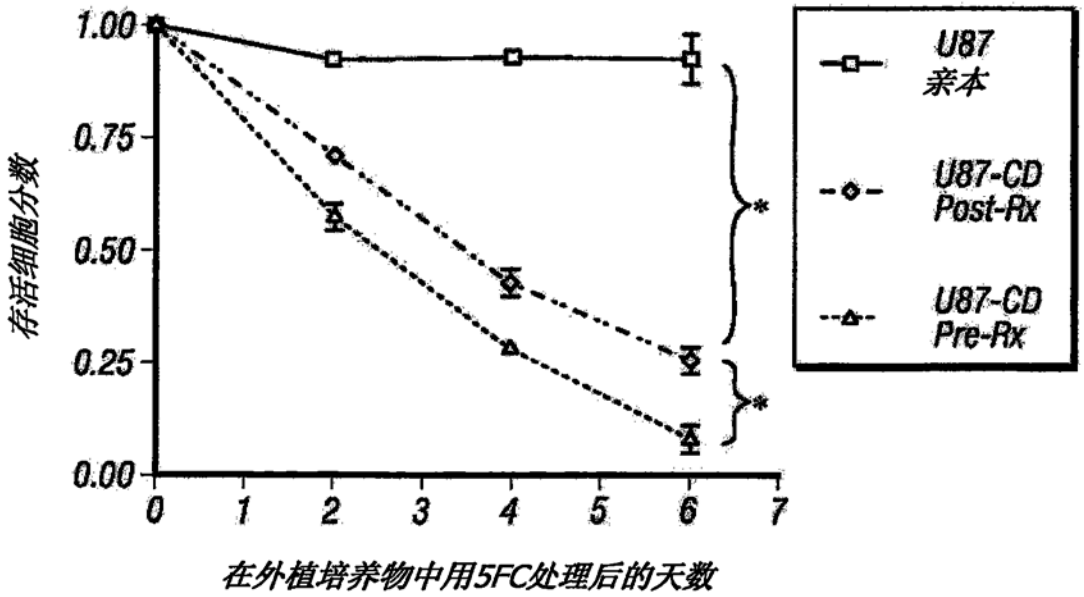

1.00
0.75
0.50
0.25
0.00
存活细胞分数
0 1 2 3 4 5 6 7
在外植培养物中用5FC处理后的天数
U87
亲本
U87-CD
Post-Rx
U87-CD
Pre-Rx
*
*

图7

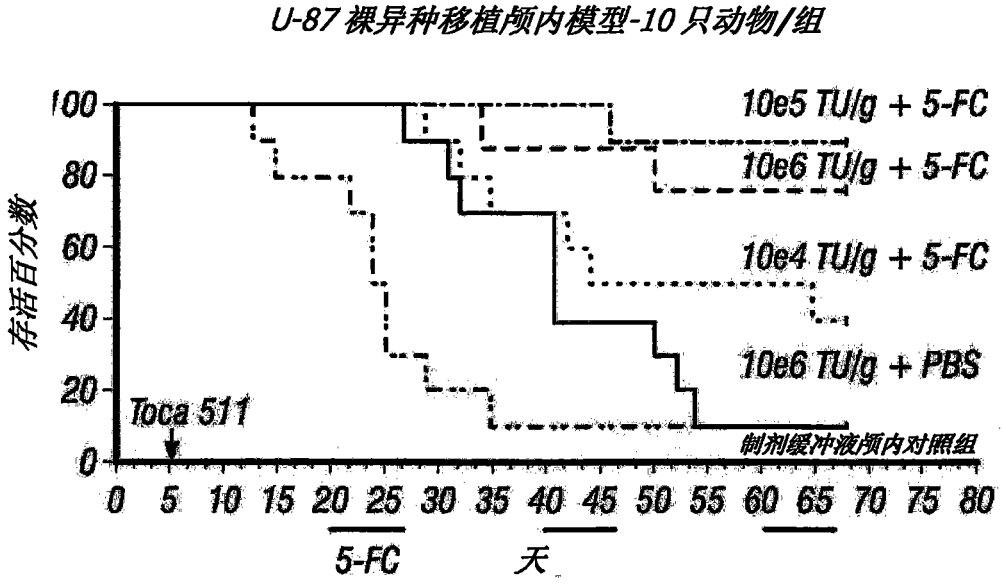

U-87 裸异种移植颅内模型-10 只动物/组
100
80
60
40
20
0
存活百分数
Toca 511
10e5 TU/g + 5-FC
10e6 TU/g + 5-FC
10e4 TU/g + 5-FC
10e6 TU/g + PBS
制剂缓冲液颅内对照组
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
5-FC
天

图8

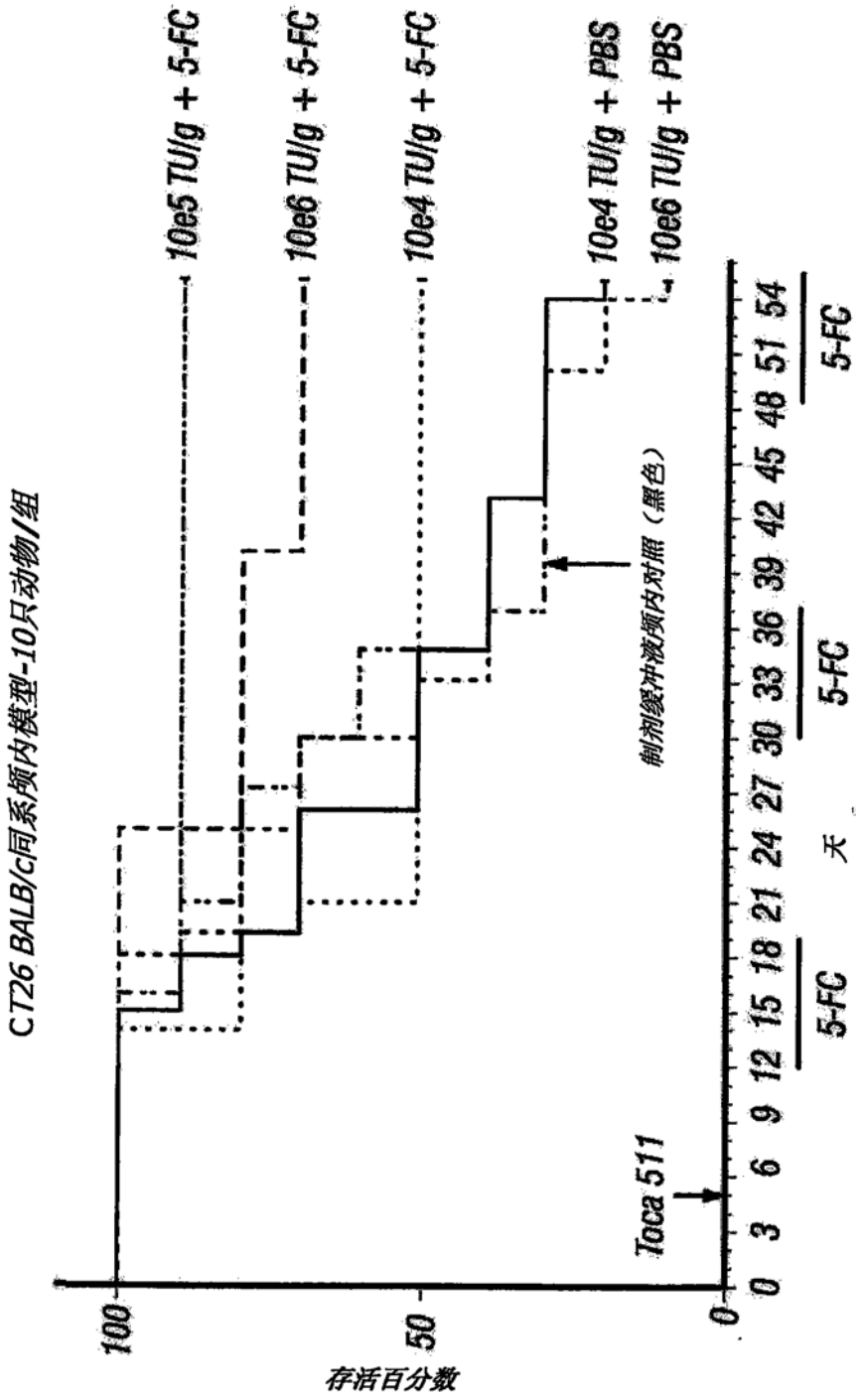
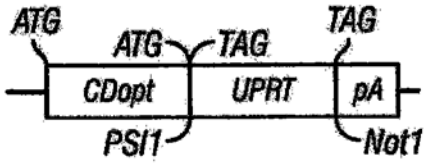

C726 BALB/c同系病内模型-10只动物/组
10e5 TU/g + 5-FC
10e6 TU/g + 5-FC
10e4 TU/g + 5-FC
10e4 TU/g + PBS
10e6 TU/g + PBS
制剂缓冲液病内对照（黑色）
Toca 511
5-FC
5-FC
5-FC
5-FC
0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54
天
100
50
0
存活百分数

图9

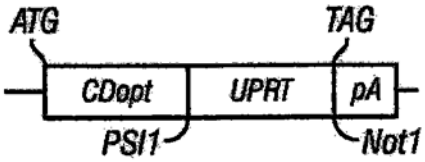
1.分开合成具有PSI1/Not1位点的CDopt和UPRT基因



2. 用PSI1消化并进行平端连接



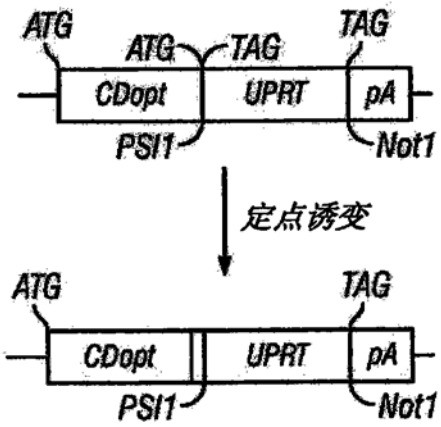
3. 通过定点诱变来缺失CDopt基因终止密码子和UPRT起始密码子



4. CDpot-UPRT基因大小：1.3kb

图10A

1.合成CDopt-连接子-UPRT：定点诱变插入连接子



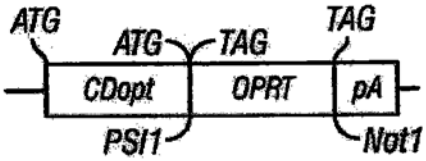
2. CDopt-连接子-UPRT基因大小:1.36kb

图10B

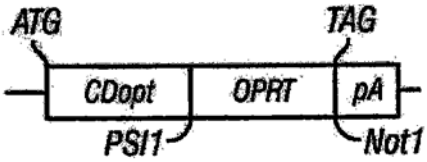
1. 分开合成具有PSI1/Not1位点的CDopt基因和OPRT域



2. 用PSI1消化并进行平端连接



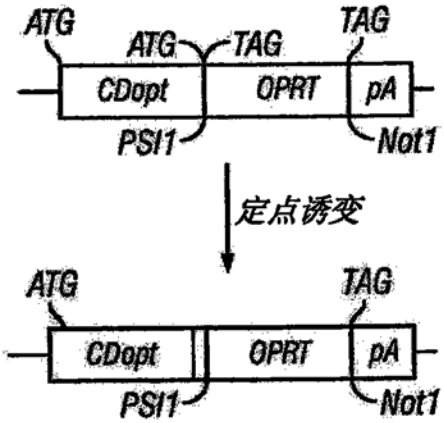
3. 通过定点诱变来缺失 CDopt 基因终止密码子和 UPRT 起始密码子



4. CDopt-OPRT基因大小：1.3kb

图10C

1. 合成CDopt-连接子-OPRT：定点诱变插入连接子



2. CD-OPRT基因大小：1.3kb

图10D

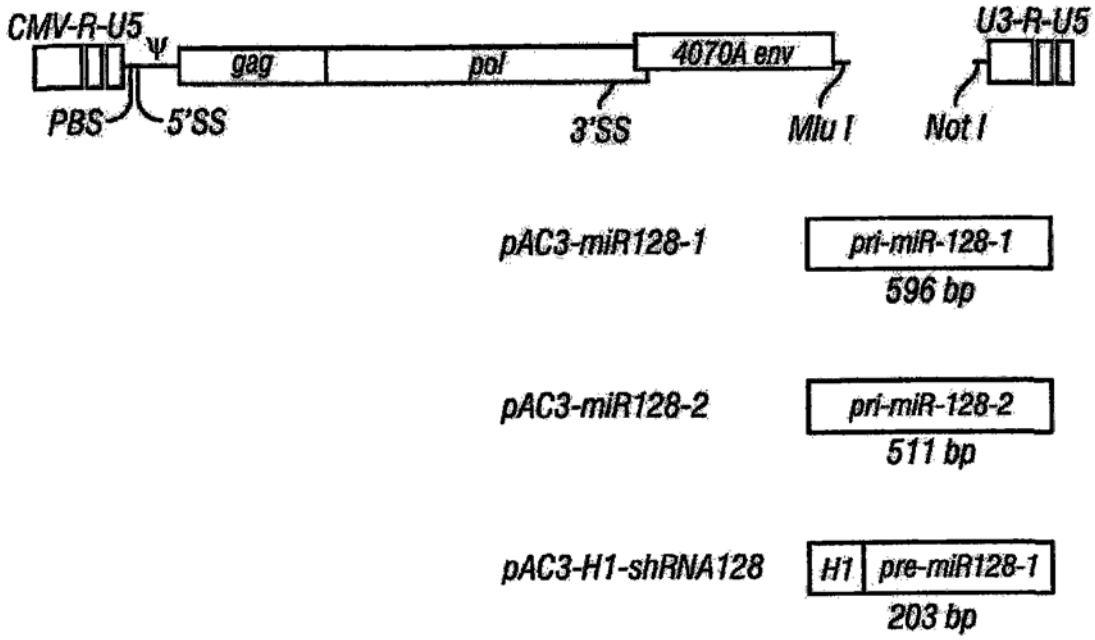

CMV-R-U5
Ψ
gag
pol
4070A env
U3-R-U5
PBS
5'SS
3'SS
Mlu I
Not I
pAC3-miR128-1
pri-miR-128-1
596 bp
pAC3-miR128-2
pri-miR-128-2
511 bp
pAC3-H1-shRNA128
H1
pre-miR128-1
203 bp

图11A

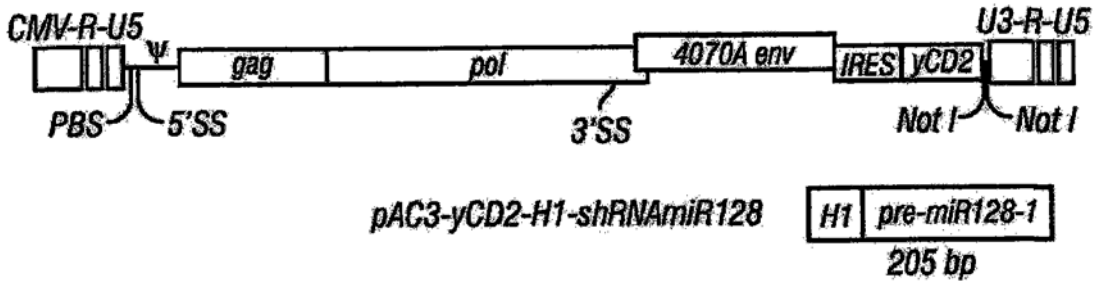

CMV-R-U5
Ψ
gag
pol
4070A env
IRES
yCD2
U3-R-U5
PBS
5'SS
3'SS
Not I
Not I
pAC3-yCD2-H1-shRNAmiR128
H1
pre-miR128-1
205 bp

图11B

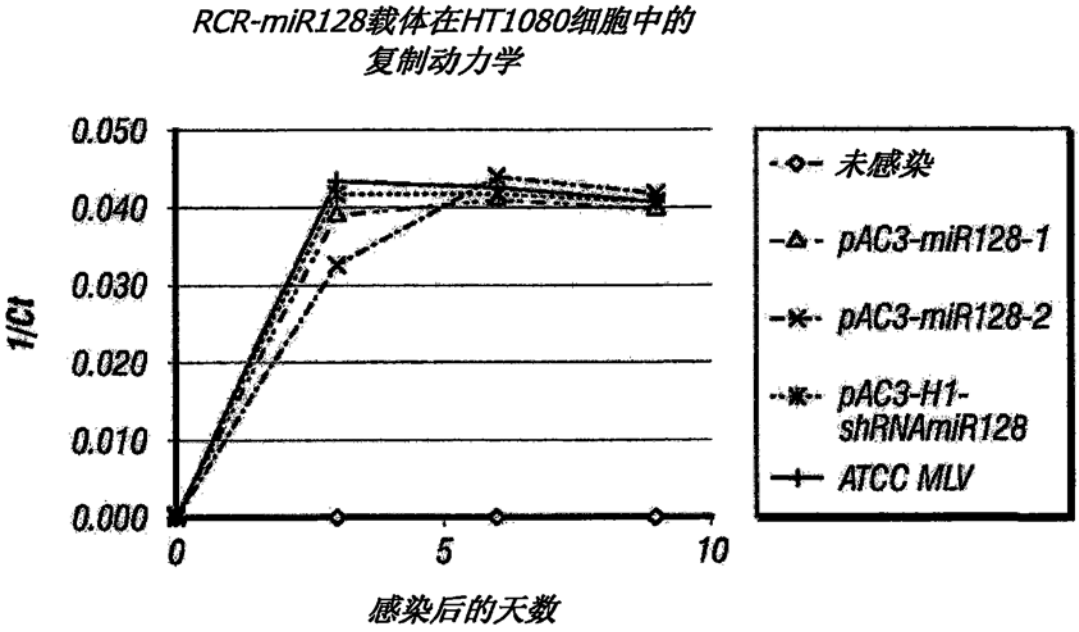

RCR-miR128载体在HT1080细胞中的
复制动力学
0.050
0.040
0.030
0.020
0.010
0.000
1/Ct
0
5
10
感染后的天数
未感染
pAC3-miR128-1
pAC3-miR128-2
pAC3-H1-shRNAmiR128
ATCC MLV

图12A

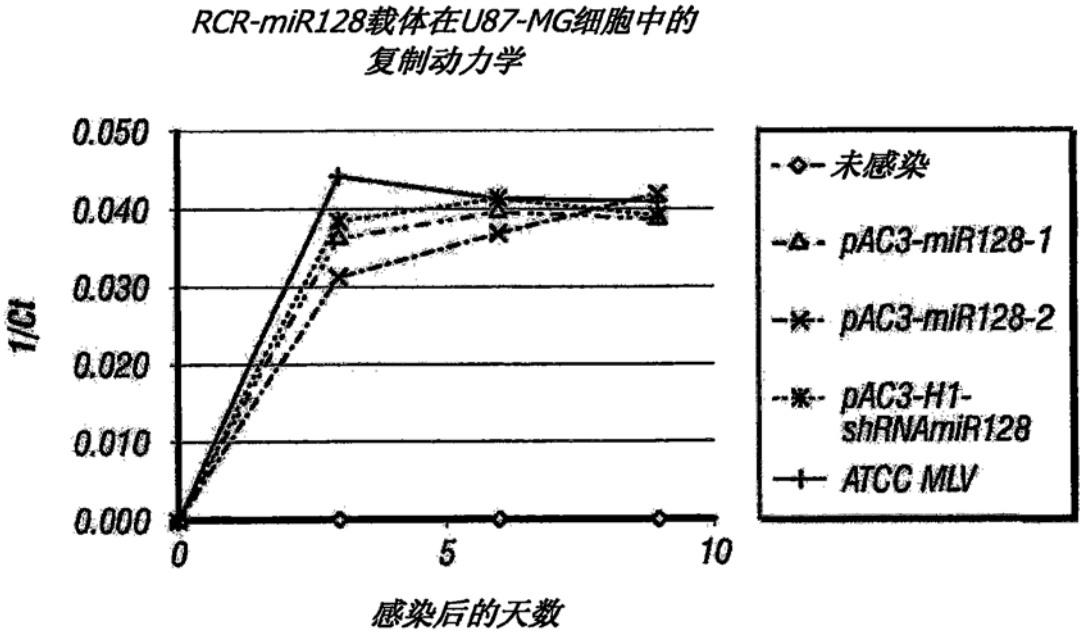

RCR-miR128载体在U87-MG细胞中的
复制动力学
0.050
0.040
0.030
0.020
0.010
0.000
1/Ct
0
5
10
感染后的天数
未感染
pAC3-miR128-1
pAC3-miR128-2
pAC3-H1-shRNAmiR128
ATCC MLV

图12B

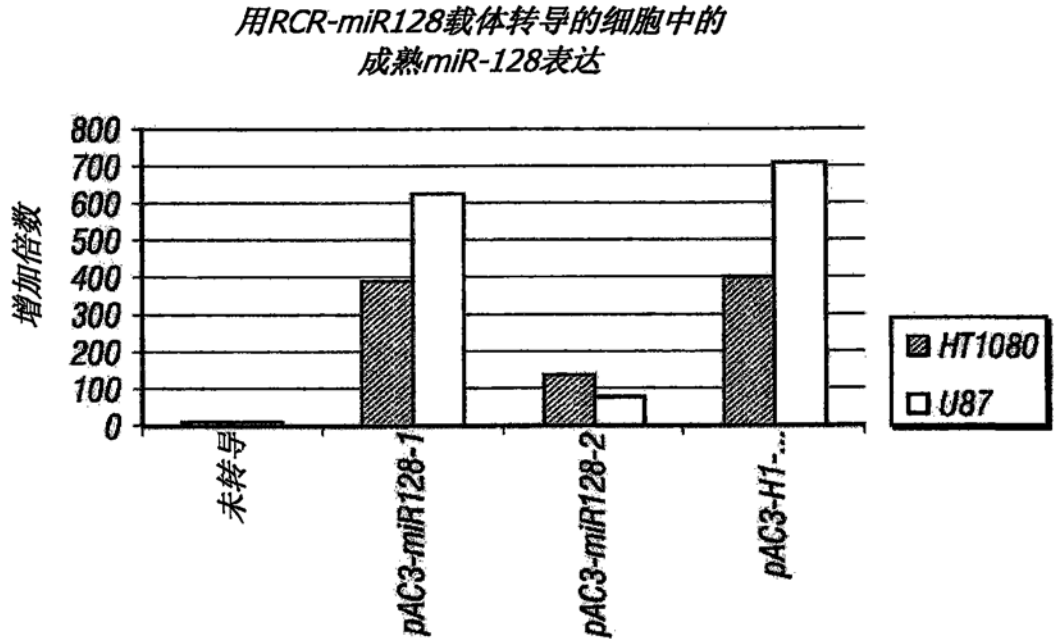

用RCR-miR128载体转导的细胞中的
成熟miR-128表达
增加倍数
800
700
600
500
400
300
200
100
0
未转导
pAC3-miR128-1
pAC3-miR128-2
pAC3-H1-...
HT1080
U87

图13

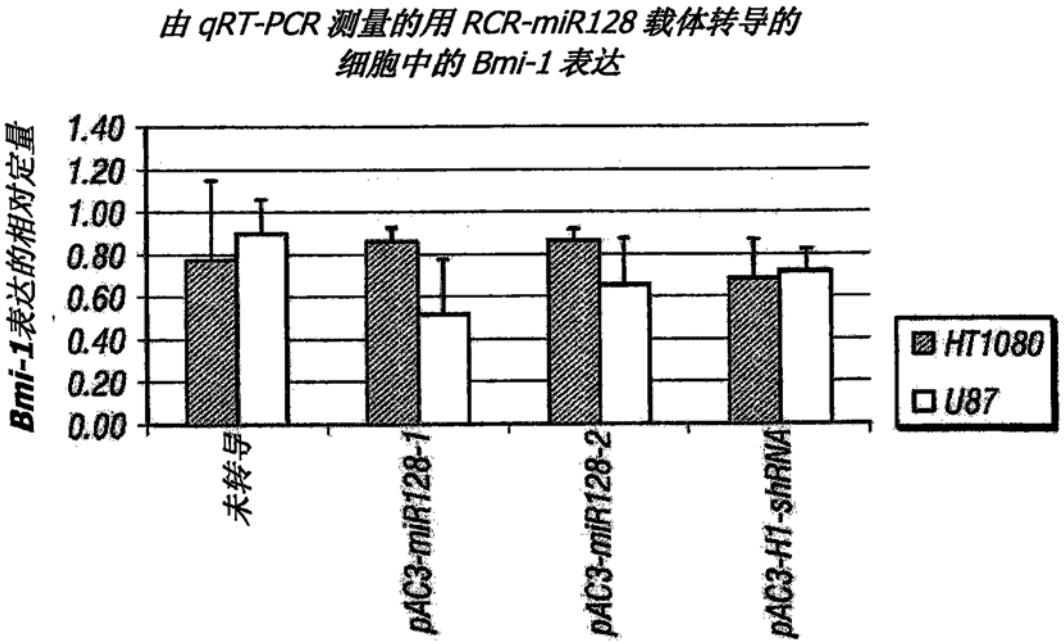

由 qRT-PCR 测量的用 RCR-miR128 载体转导的
细胞中的 Bmi-1 表达
Bmi-1表达的相对定量
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
未转导
pAC3-miR128-1
pAC3-miR128-2
pAC3-H1-shRNA
HT1080
U87

图14

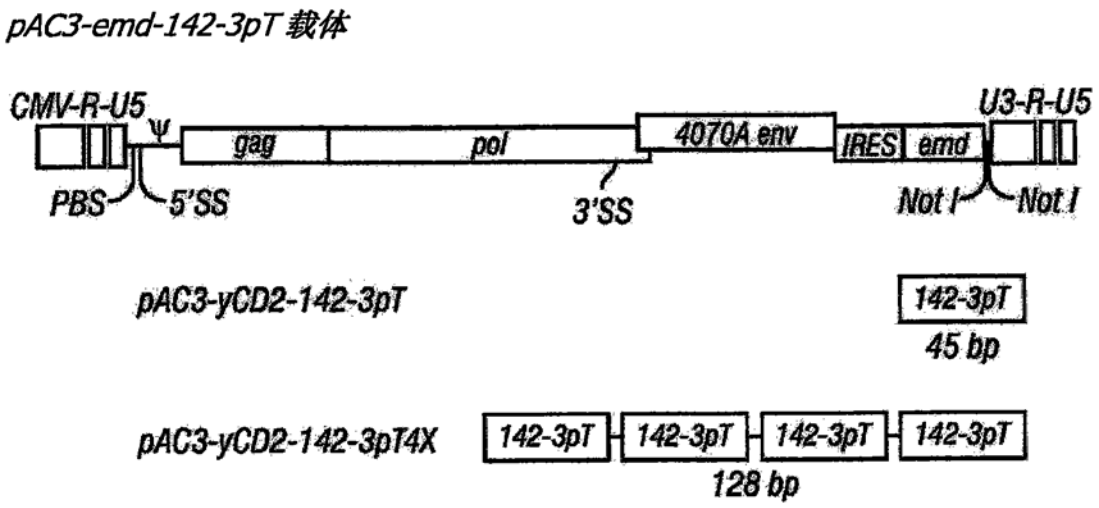

pAC3-emd-142-3pT 载体
CMV-R-U5
Ψ
gag
pol
4070A env
IRES
emd
U3-R-U5
PBS
5'SS
3'SS
Not I
Not I
pAC3-yCD2-142-3pT
142-3pT
45 bp
pAC3-yCD2-142-3pT4X
142-3pT
142-3pT
142-3pT
142-3pT
128 bp

图15

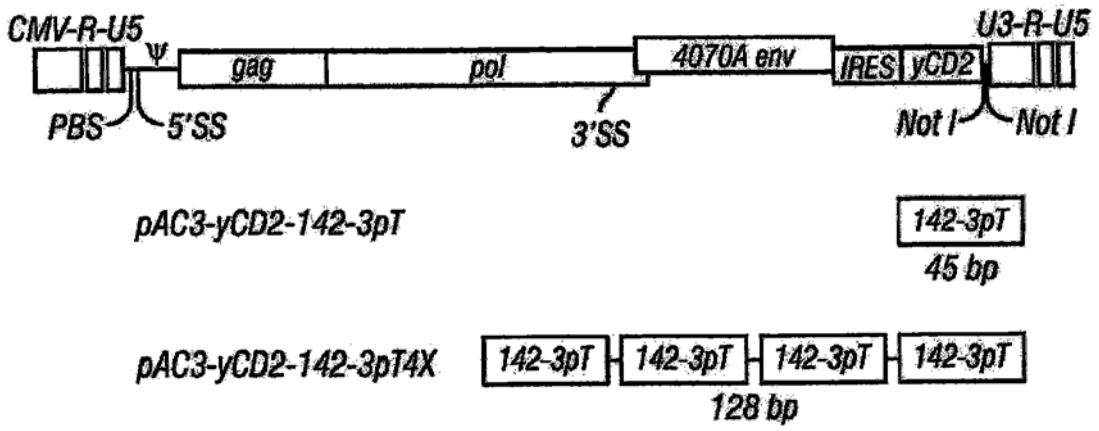

CMV-R-U5
Ψ
gag
pol
4070A env
IRES
yCD2
U3-R-U5
PBS
5'SS
3'SS
Not I
Not I
pAC3-yCD2-142-3pT
142-3pT
45 bp
pAC3-yCD2-142-3pT4X
142-3pT
142-3pT
142-3pT
142-3pT
128 bp

图16

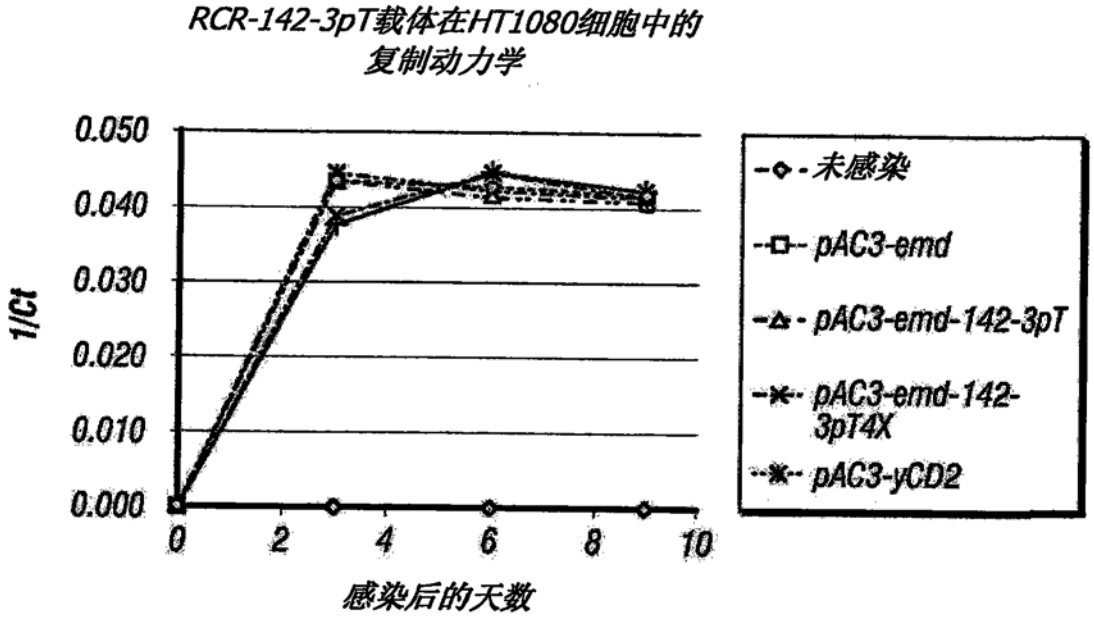

RCR-142-3pT载体在HT1080细胞中的
复制动力学
0.050
0.040
0.030
0.020
0.010
0.000
1/Ct
0
2
4
6
8
10
感染后的天数
未感染
pAC3-emd
pAC3-emd-142-3pT
pAC3-emd-142-3pT4X
pAC3-yCD2

图17A

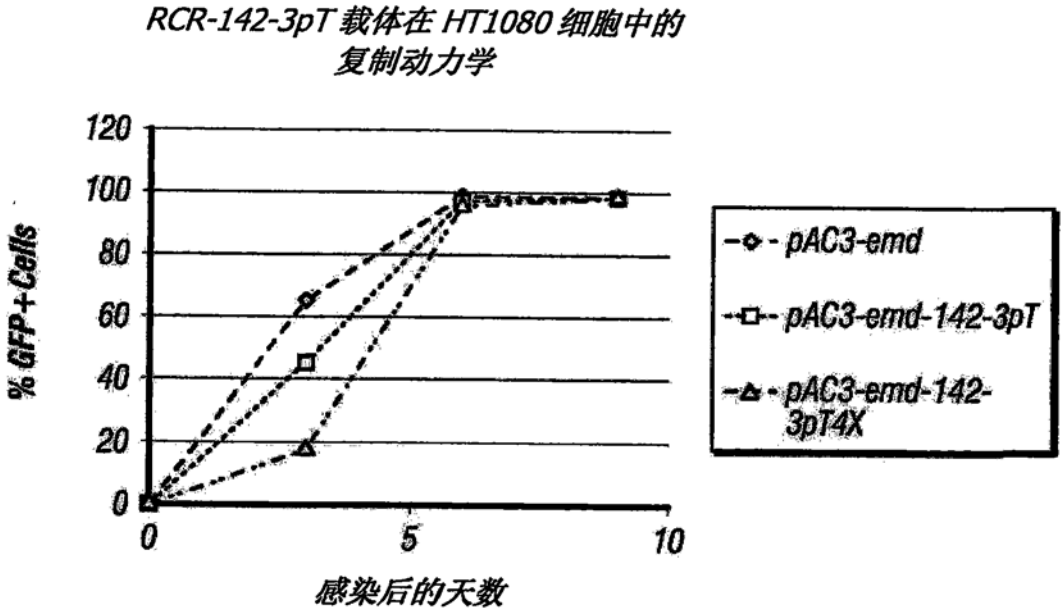

RCR-142-3pT 载体在 HT1080 细胞中的
复制动力学
120
100
80
60
40
20
0
% GFP+Cells
0
5
10
感染后的天数
pAC3-emd
pAC3-emd-142-3pT
pAC3-emd-142-3pT4X

图17B

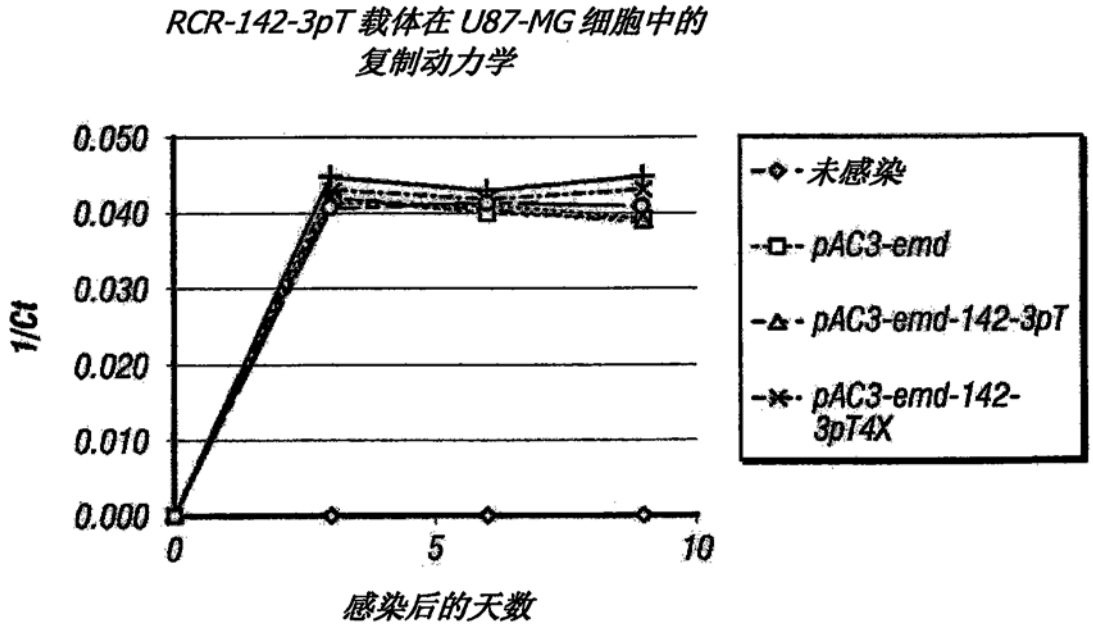

RCR-142-3pT 载体在 U87-MG 细胞中的
复制动力学
0.050
0.040
0.030
0.020
0.010
0.000
1/Ct
0
5
10
感染后的天数
未感染
pAC3-emd
pAC3-emd-142-3pT
pAC3-emd-142-3pT4X

图18A

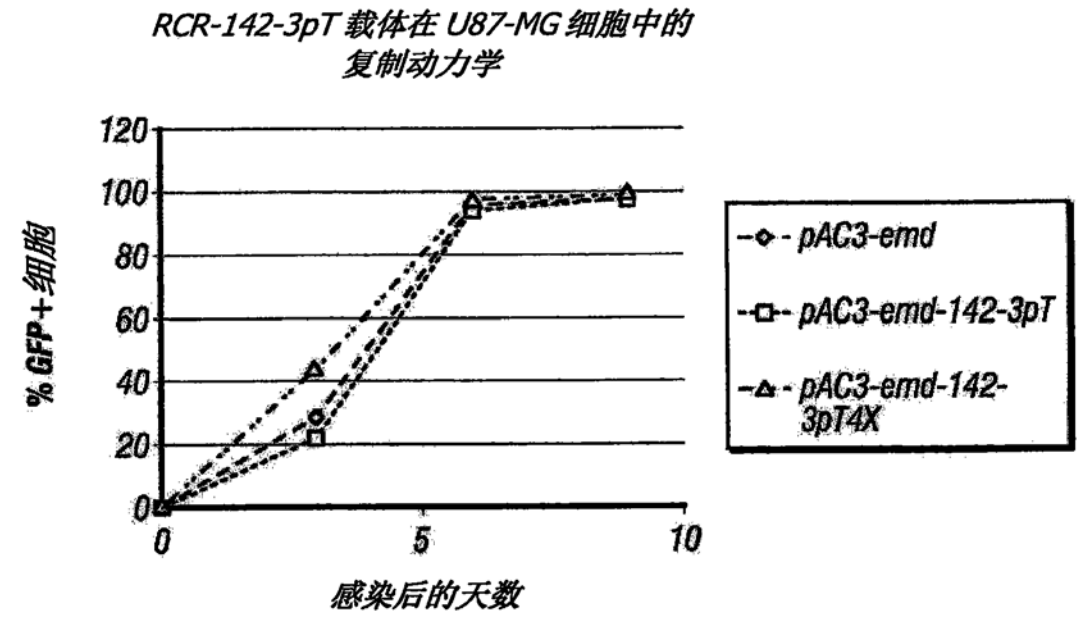

RCR-142-3pT 载体在 U87-MG 细胞中的
复制动力学
120
100
80
60
40
20
0
% GFP+细胞
0
5
10
感染后的天数
pAC3-emd
pAC3-emd-142-3pT
pAC3-emd-142-3pT4X

图18B

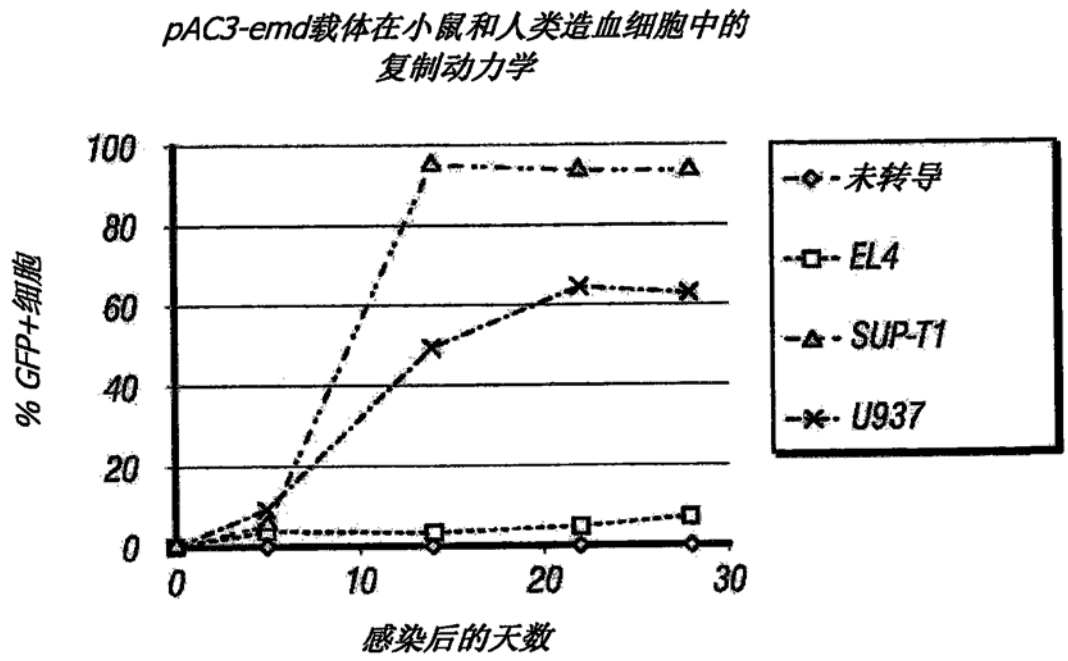

pAC3-emd载体在小鼠和人类造血细胞中的
复制动力学
% GFP+细胞
100
80
60
40
20
0
0
10
20
30
感染后的天数
未转导
EL4
SUP-T1
U937

图19

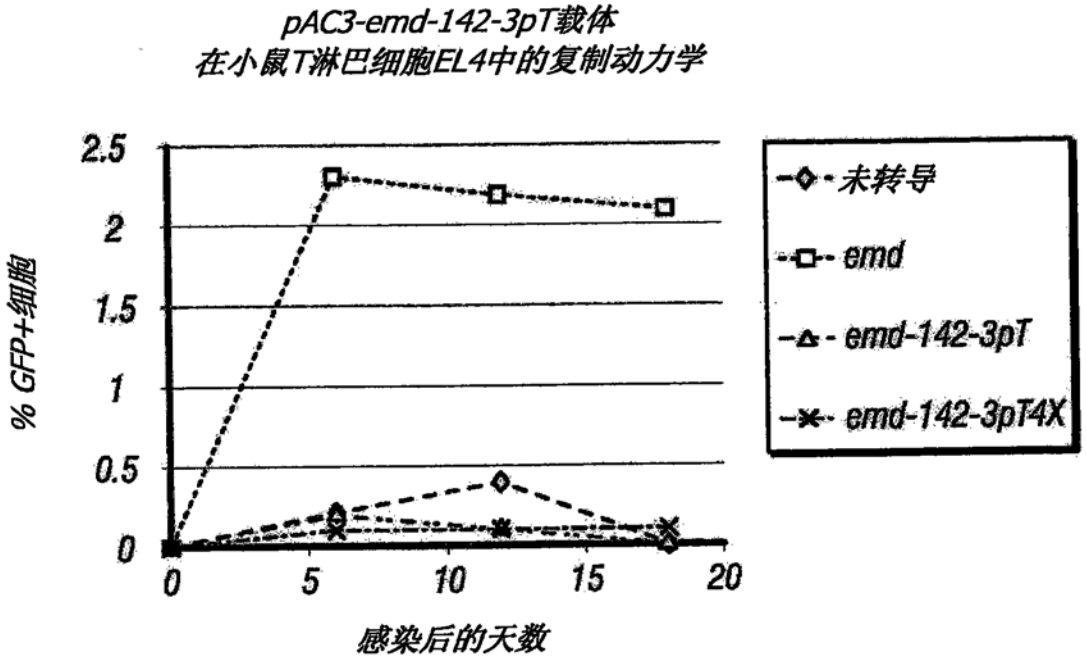

pAC3-emd-142-3pT载体
在小鼠T淋巴细胞EL4中的复制动力学
% GFP+细胞
2.5
2
1.5
1
0.5
0
0
5
10
15
20
感染后的天数
未转导
emd
emd-142-3pT
emd-142-3pT4X

图20A

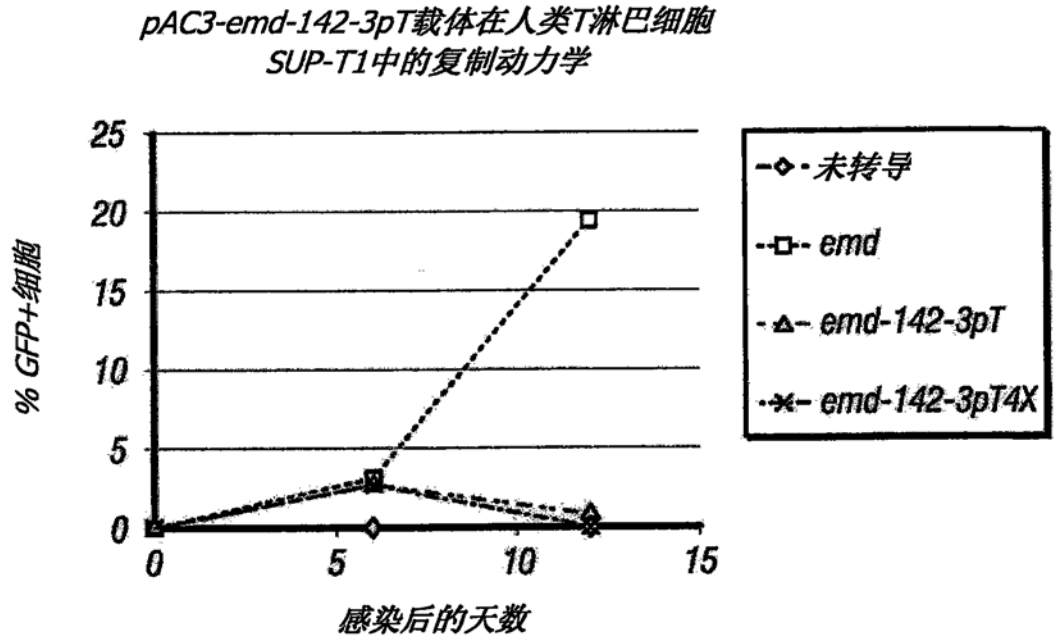

pAC3-emd-142-3pT载体在人类T淋巴细胞
SUP-T1中的复制动力学
% GFP+细胞
25
20
15
10
5
0
0
5
10
15
感染后的天数
未转导
emd
emd-142-3pT
emd-142-3pT4X

图20B

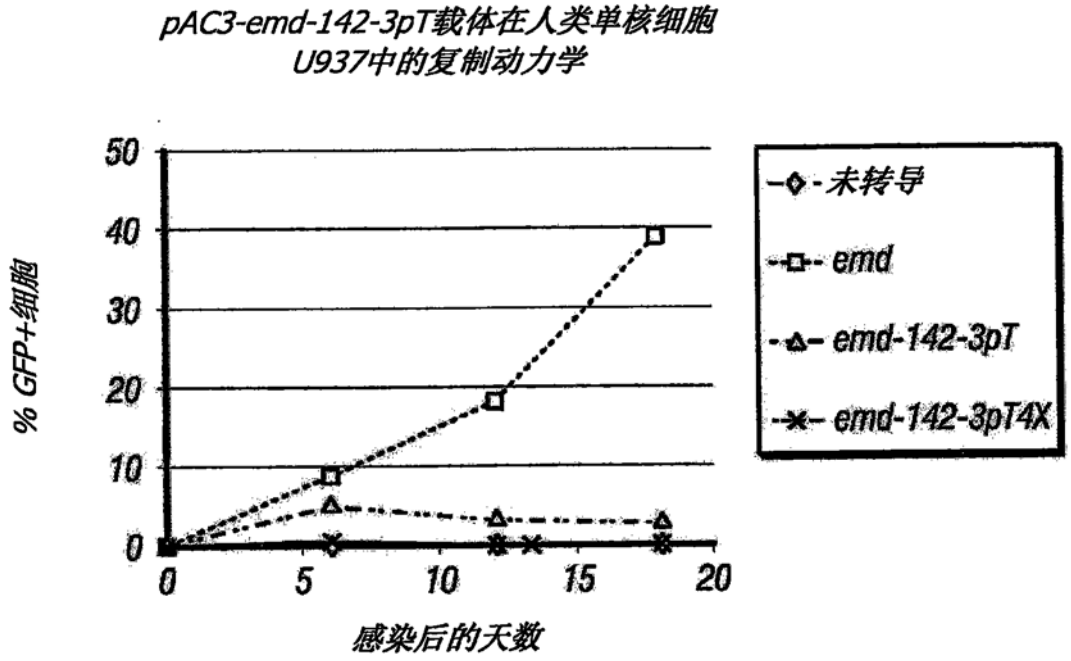

pAC3-emd-142-3pT载体在人类单核细胞
U937中的复制动力学
% GFP+细胞
50
40
30
20
10
0
0
5
10
15
20
感染后的天数
未转导
emd
emd-142-3pT
emd-142-3pT4X

图20C

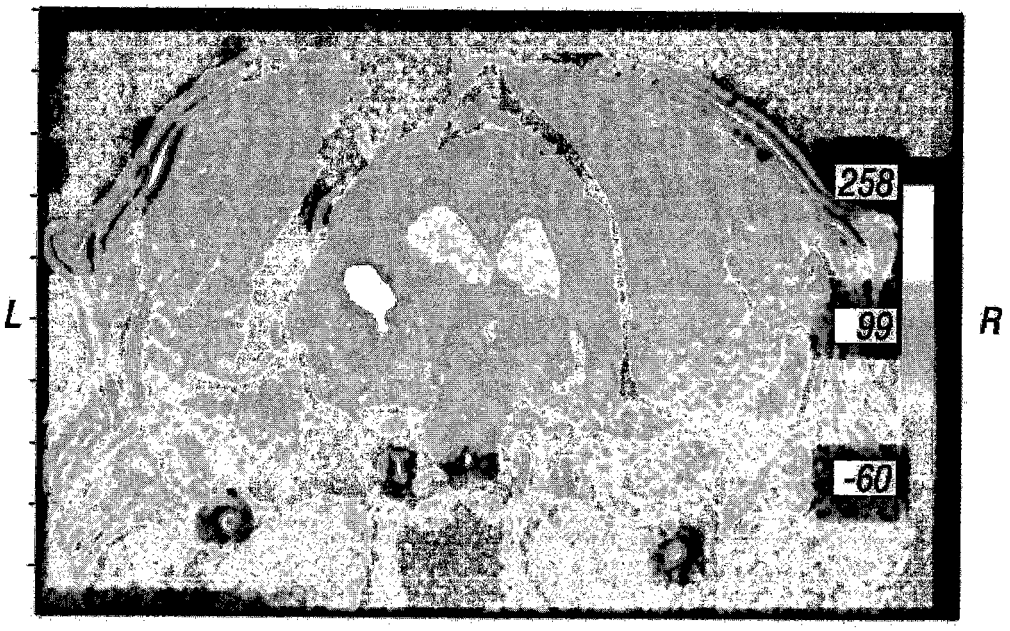


图21A

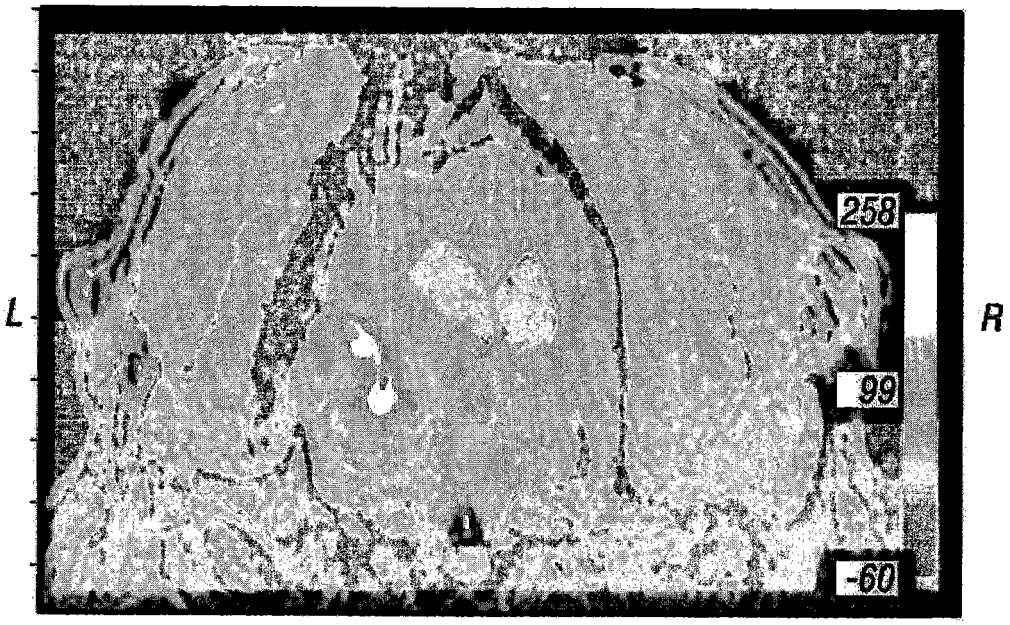


图21B

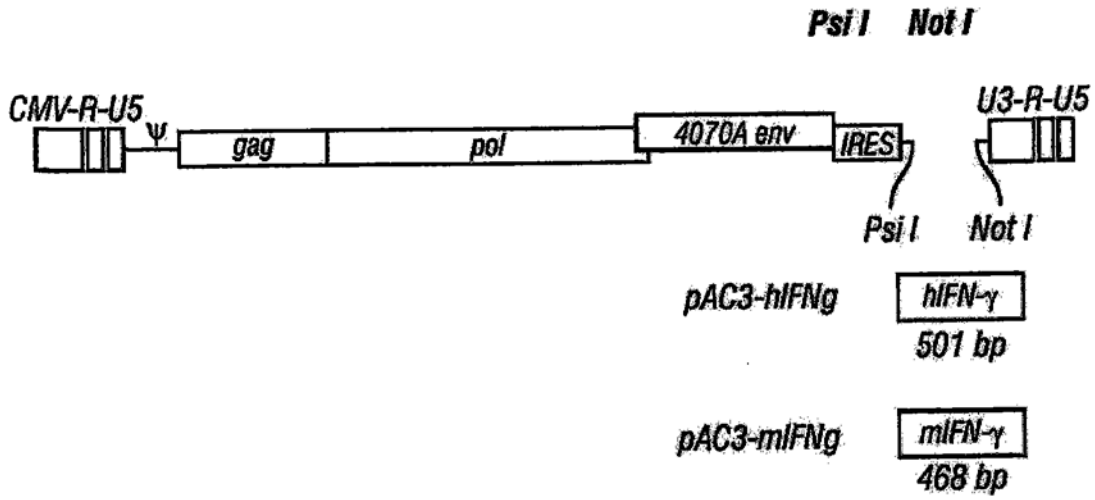

Psi I
Not I
CMV-R-U5
Ψ
gag
pol
4070A env
IRES
U3-R-U5
Psi I
Not I
pAC3-hIFNg
hIFN-γ
501 bp
pAC3-mIFNg
mIFN-γ
468 bp

图22

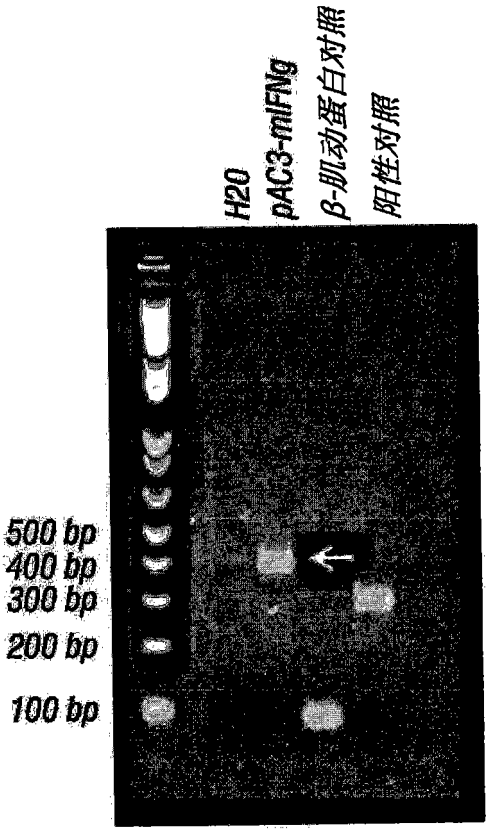

H20
pAC3-mIFNg
β-肌动蛋白对照
阳性对照
500 bp
400 bp
300 bp
200 bp
100 bp

图23

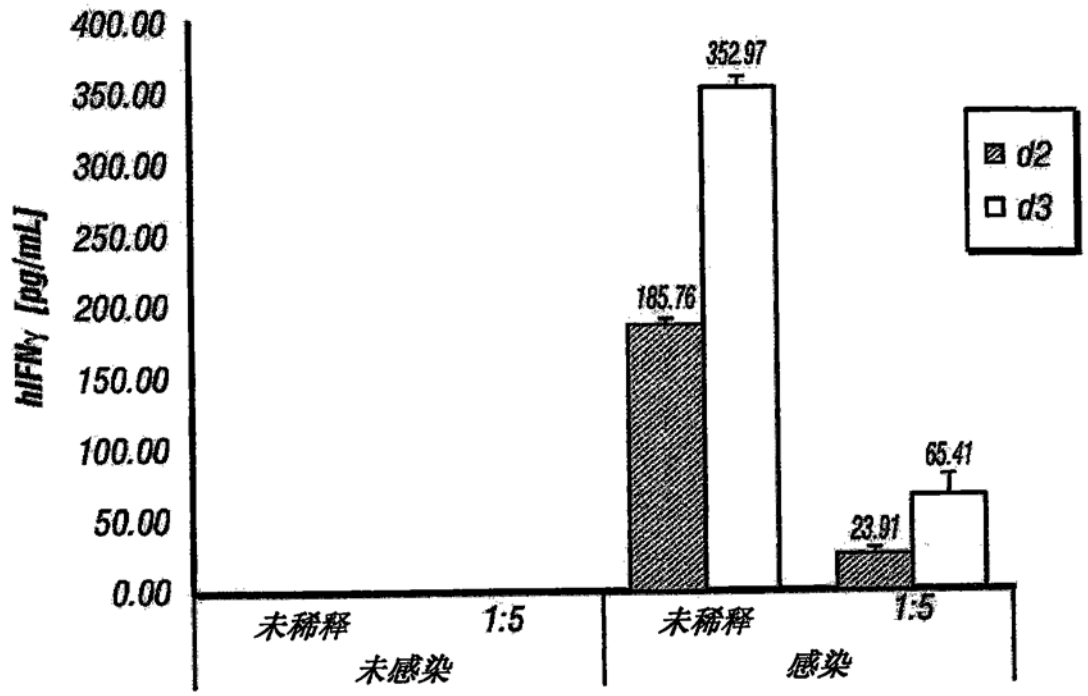

400.00
350.00
300.00
250.00
200.00
150.00
100.00
50.00
0.00
hIFNγ [pg/mL]
352.97
185.76
65.41
23.91
d2
d3
未稀释
1:5
未稀释
1:5
未感染
感染

图24

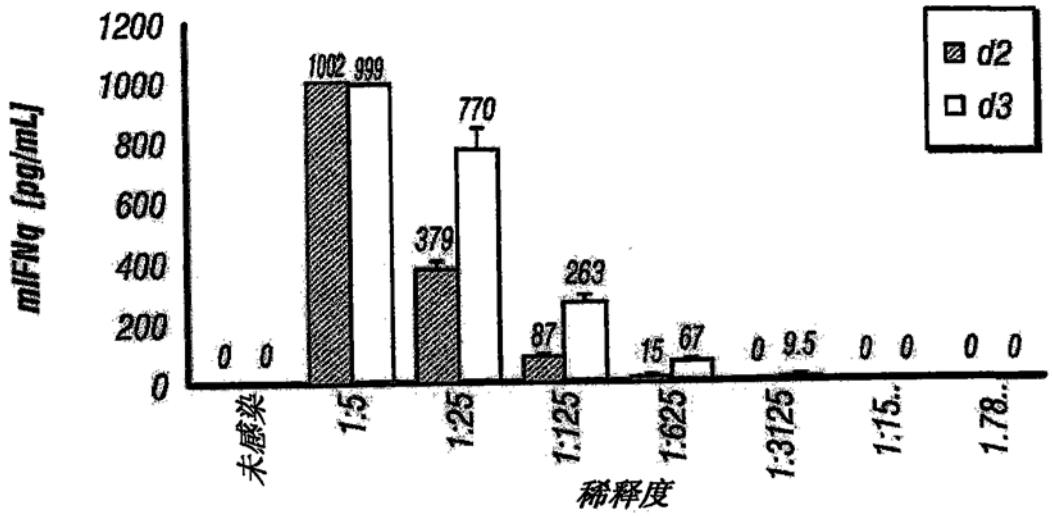

1200
1000
800
600
400
200
0
mIFNα [pg/mL]
1002
999
770
379
263
87
15
67
0
9.5
0
0
0
0
0
0
d2
d3
未感染
1:5
1:25
1:125
1:625
1:3125
1:15..
1.78..
稀释度

图25

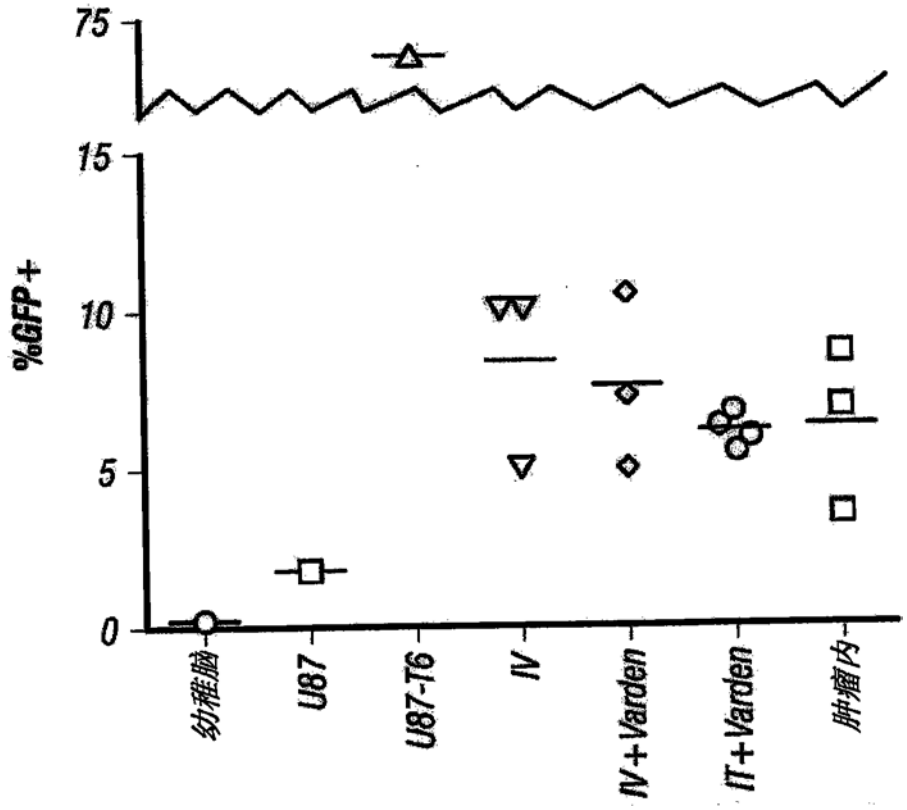


图26

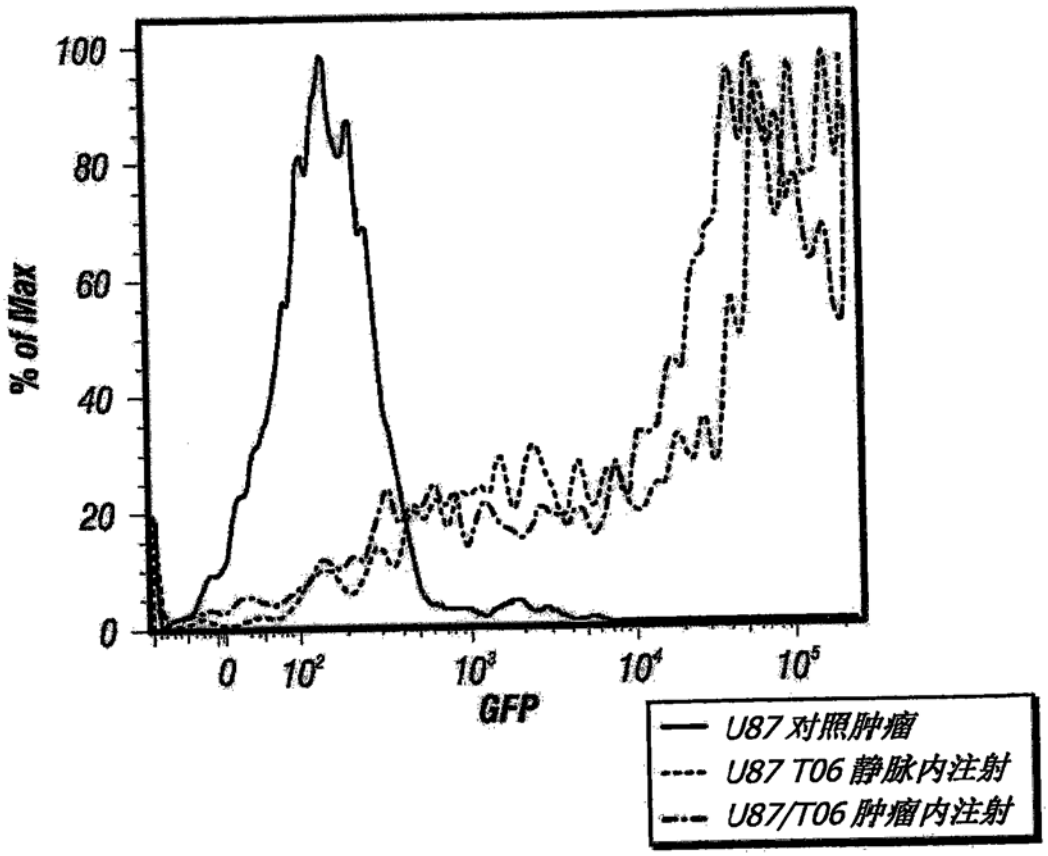

100
80
60
40
20
0
% of Max
0
10²
10³
10⁴
10⁵
GFP
U87 对照肿瘤
U87 T06 静脉内注射
U87/T06 肿瘤内注射

图27

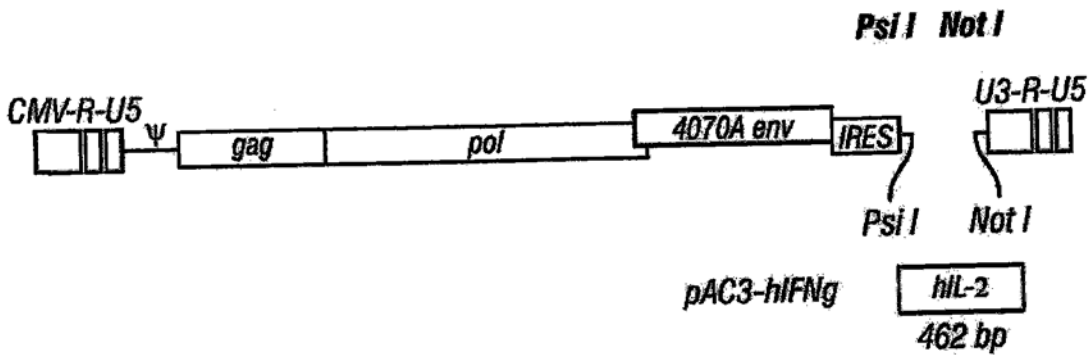

Psi I
Not I
CMV-R-U5
Ψ
gag
pol
4070A env
IRES
U3-R-U5
Psi I
Not I
pAC3-hIFNg
hIL-2
462 bp

图28