



(12) 发明专利申请

(10) 申请公布号 CN 104619832 A

(43) 申请公布日 2015. 05. 13

(21) 申请号 201380035802. X

C12N 15/00(2006. 01)

(22) 申请日 2013. 03. 14

C12N 5/00(2006. 01)

(30) 优先权数据

61/668839 2012. 07. 06 US

(85) PCT国际申请进入国家阶段日

2015. 01. 05

(86) PCT国际申请的申请数据

PCT/US2013/031644 2013. 03. 14

(87) PCT国际申请的公布数据

W02014/007858 EN 2014. 01. 09

(71) 申请人 衣阿华大学研究基金会

地址 美国衣阿华州

(72) 发明人 B. L. 戴维森 M. 希尔 R. 布德罗

(74) 专利代理机构 中国专利代理(香港)有限公司

72001

代理人 李慧惠 彭昶

(51) Int. Cl.

C12N 7/00(2006. 01)

权利要求书1页 说明书11页

序列表15页 附图29页

(54) 发明名称

经修饰的腺伴随病毒载体组合物

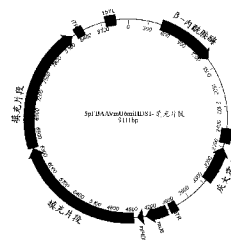
(57) 摘要

公开了包含长度为 3300-4200 个核苷酸的核酸的腺伴随病毒填充物组分。

5pFBAAVmU6mIHDS1-荧光质粒

图谱特征

494 - 1354	ori-ori质粒
1248 - 2645	TetR
2700 - 3242	质粒骨架
3810 - 3928	AAV ITR (130bp)
3927 - 4149	多克隆位点
4342 - 4417	mIHDS1
4461 - 5138	1.8kb荧光蛋白1
5293 - 5413	AAV ITR (130bp)
5761 - 5929	TetR



1. 腺伴随病毒(AAV)填充物组分,其包含与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 90% 同一性的长度为 3300-4200 个核苷酸的核酸。
2. 腺伴随病毒(AAV)填充物组分,其由与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 90% 同一性的长度为 3300-4200 个核苷酸的核酸组成。
3. 权利要求 1 或权利要求 2 的 AAV 填充物组分,其中所述核酸为 3500-4000 个核苷酸。
4. 权利要求 1 或权利要求 2 的 AAV 填充物组分,其中所述核酸为 3700-3850 个核苷酸。
5. 权利要求 1 至 4 中任一项的 AAV 填充物组分,其中所述核酸与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 95% 的同一性。
6. 腺伴随病毒(AAV)载体,其包含可操作地连接至表达盒的权利要求 1 至 5 中任一项的填充物组分。
7. 权利要求 6 的 AAV 载体,其中所述表达盒包含启动子。
8. 权利要求 7 的 AAV 载体,其中所述启动子是 pol III 启动子。
9. 权利要求 8 的 AAV 载体,其中所述启动子是 mU6 启动子。
10. 权利要求 6 至 9 中任一项的 AAV 载体,其进一步包含靶序列。
11. 权利要求 10 的 AAV 载体,其中所述靶序列是 RNAi 分子。

经修饰的腺伴随病毒载体组合物

[0001] 相关申请

本专利申请要求享有于 2012 年 7 月 6 日提交的美国申请系列号 61/668,839 的优先权权益,所述申请通过引用并入本文。

[0002] 序列表

本申请包含序列表,所述序列表已通过 EFS-Web 以 ASCII 格式递交,并且以其整体通过引用并入本文。所述 ASCII 副本,于 2013 年 3 月 14 日生成,命名为 17023.126W01_SL.txt,并且大小为 39,125 字节。

[0003] 背景

腺伴随病毒(AAV)是细小病毒科(parvoviridae family)的小的非致病病毒。AAV 通过其依靠辅助病毒进行复制而与该科其他成员区别。AAV 的约 5kb 基因组由一段正或负极性的单链 DNA 组成。基因组的末端是短的反向末端重复,其可以折叠为发夹结构并且作为病毒 DNA 复制起点。在物质上(Physically),细小病毒病毒体是无包膜的并且其二十面的衣壳直径约 20nm。

[0004] 迄今为止,已经鉴定很多种血清学上不同的 AAV,并且已经从人或灵长类中分离。Govindasamy 等,“Structurally Mapping the Diverse Phenotype of Adeno-Associated Virus Serotype 4,” *J. Vir.*, 80 (23):11556-11570 (2006)。例如,AAV2 的基因组长度为 4680 个核苷酸并且包含两个可读框(ORF)。左侧的 ORF 编码非结构性的 Rep 蛋白,Rep 40、Rep 52、Rep 68 和 Rep 78,它们除了涉及单链的后代基因组的产生,还涉及复制和转录的调节。Rep68/78 同样显示出具有 NTP 结合活性和 DNA 和 RNA 解旋酶活性。Rep 蛋白具有核定位信号和几个潜在的磷酸化位点。这些激酶位点之一的突变,导致复制活性的丢失。

[0005] 基因组的末端是短的反向末端重复(ITR),所述反向末端重复具有潜力以折叠为作为病毒 DNA 复制起点的 T 形发夹结构。在 ITR 区中已描述对于 ITR 功能重要的两个元件,GAGC 重复基序和末端解离位点(trs)。重复基序已显示出当 ITR 为线性或发夹构象时均结合 Rep。该结合作用于位置 Rep68/78 用于在 trs 切割,其以位点和链特异性的方式发生。

[0006] AAV 的以下特征使得其成为用于基因转移的有吸引力的载体。AAV 载体具有广泛的宿主范围;在体外和体内转导分裂的和分裂的细胞,并且维持转导基因高水平的表达。病毒颗粒是热稳定的、耐溶剂的、耐去污剂的、耐 pH 变化的、耐温度变化的,并且可以在 CsCl 梯度中浓缩。AAV 不与任何致病事件有关,并且使用 AAV 载体转导没有发现对细胞生长或分化诱导任何持续的负面影响。ITR 已显示为是允许完全去除病毒基因以产生载体系统的包装所需的唯一顺式(cis)元件。

[0007] 目前需要具有改进的包装特征的 AAV 载体。

[0008] 概述

在某些实施方案中,本发明提供腺伴随病毒(AAV)填充物(filler)组分(也称为“填充片段(stuffer)序列”),其包含与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 90% 同一性的长度为 3300-4200 个核苷酸的核酸。

[0009] 在某些实施方案中,本发明提供腺伴随病毒(AAV)填充物组分,其由与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 90% 同一性的长度为 3300-4200 个核苷酸的核酸组成。

[0010] 在某些实施方案中,本发明提供 AAV 载体,其包含上述填充物组分。

[0011] 附图和表格简述

图 1 是 5pFBAAVmU6miHDS1 填充片段(9110bp)的质粒图谱。

[0012] 图 2 是 5pFBAAVmU6miHDS1 填充片段(填充片段 #1) (SEQ ID NO:3) 的序列。

[0013] 图 3 提供 5pFBAAVmU6miHDS1 填充片段(SEQ ID NO:1, 4-11) 的各种单独组分的序列。

[0014] 图 4 是显示相对 Htt 表达的图。

[0015] 图 5 是 5pFBAAVmU6miHDS1- 填充片段的质粒图谱。

[0016] 图 6 是 5pFBAAVmU6miHDS1- 填充片段(SEQ ID NO:12) 的质粒序列。

[0017] 图 7 提供填充片段序列(填充片段 #2) (SEQ ID NO:2)。

[0018] 图 8. 完整病毒体相对于空病毒体的 EM 评估。两个空病毒体的实例通过箭头突出显示。该制备物只具有 ~4% 的空病毒体,这是非常低的。

[0019] 图 9. 银染色以检测纯化的病毒体的衣壳完整性。几种不同的 miRNA- 表达构建体与内含子 I/II 填充片段一同工程化入穿梭载体以生成接近野生型基因组大小。纯化的病毒显示出最佳的 VP1、VP2 和 VP3 蛋白比例。

[0020] 表 1. miR- 内含子 I/II 病毒体的包装效率%和污染物%。

[0021] 详述

AAV载体和表达盒

本发明的病毒载体利用 AAV 载体。“AAV”载体指腺伴随病毒,并且可以用来指天然存在的野生型病毒本身或其衍生物。除非另有要求,该术语涵盖全部亚型、血清型和假型,以及天然存在和重组形式。如本文使用的,术语“血清型”指这样的 AAV,所述 AAV 通过基于衣壳蛋白与确定的抗血清(例如,存在 8 种已知的灵长类 AAV 血清型,AAV-1 至 AAV-8) 的反应性进行鉴定并区别于其他 AAV。例如,血清型 AAV-2 用于指这样的 AAV,所述 AAV 包含从 AAV-2 的 cap 基因编码的衣壳蛋白和含有来自相同 AAV-2 血清型的 5' 和 3' ITR 序列的基因组。

[0022] 假型 AAV 指包含来自一种血清型的衣壳蛋白和包括第二种血清型的 5' -3' ITR 的病毒基因组的 AAV。假型 rAAV 期望具有衣壳血清型的细胞表面结合特性和与 ITR 血清型一致的遗传特性。假型 rAAV 使用本领域描述的标准技术生产。如本文所用的,例如,rAAV1 可以用于指具来自相同血清型的衣壳蛋白和 5' -3' ITR 的 AAV,或其可以指具有来自血清型 1 的衣壳蛋白和来自不同 AAV 血清型(例如 AAV 血清型 2) 的 5' -3' ITR 的 AAV。

[0023] 缩写“rAAV”指重组腺伴随病毒,同样指重组 AAV 载体(或“rAAV 载体”)。在一个实施方案中,AAV 表达载体使用已知技术构建,以至少提供在转录方向上的可操作地连接的组分,包括转录起始区、目标 DNA 和转录终止区的控制元件。选择在哺乳动物细胞中有功能的控制元件。产生的包含可操作地连接的组分的构建体侧接(5' 和 3') 有功能的 AAV ITR 序列。

[0024] “腺伴随病毒反向末端重复”或“AAV ITR”意指在 AAV 基因组每一端发现的本领域公知的区域,其以顺式一同作为 DNA 复制起点并且作为病毒的包装信号。

[0025] 已知 AAV ITR 区的核苷酸序列。如本文使用的,“AAV ITR”不需要具有描述的野

生型核苷酸序列,但是可以被改变,例如,通过核苷酸的插入、删除或取代。此外,AAV ITR 可以来自几种 AAV 血清型中的任一种,包括但不限于,AAV-1、AAV-2、AAV-3、AAV-4、AAV-5、AAV7 等。此外,在 AAV 载体中侧接所选择的核苷酸序列的 5' 和 3' ITR 不需要一定相同或源自相同的 AAV 血清型或分离物,只要它们的功能如预期的,即,允许删除和拯救来自宿主细胞基因组或载体的目的序列。

[0026] AAV ITR 可以从包含其的 AAV 载体质粒中切除,并使用标准连接技术融合存在于另一个载体中的所选核酸构建体的 5' 和 3', 例如 Sambrook 和 Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY (2001) 中描述的那些。例如,连接可以在 20 mM Tris-Cl pH 7.5、10 mM MgCl₂、10 mM DTT、33 μg/ml BSA、10 mM-50 mM NaCl 中与 40 μM ATP、0.01-0.02 (Weiss) 单位 T4 DNA 连接酶在 0°C (用于“粘性末端”连接) 或与 1 mM ATP、0.3-0.6 (Weiss) 单位 T4 DNA 连接酶在 14°C (用于“平末端”连接) 完成。分子间的“粘性末端”连接通常在总 DNA 浓度 30-100 μg/ml (总终浓度 5-100 nM) 实施。包含 ITR 的 AAV 载体在例如美国专利号 5,139,941 中描述。特别地,在其中描述了几种可从美国典型培养物保藏中心(“ATCC”) 在检索号 53222、53223、53224、53225 和 53226 下得到的 AAV 载体。

[0027] 腺伴随病毒优选地包装全长基因组,即近似天然基因组相同大小并且不太大或太小的基因组。很多靶核酸序列或编码靶核酸序列的表达盒非常小。为避免包装片段化的基因组,本发明人设计并测试了当连接至表达盒时在 ITR 之间产生长度上大小接近于正常的基因组的核酸序列。起始序列是哺乳动物起点,但是显著地修饰以确保该“填充物组分”(也称为“填充片段序列”) 除其他之外缺乏增强子、启动子、剪接调节子、非编码 RNAs 或反义序列。也就是说,填充片段序列是“沉默的”并且对表达盒不赋予活性。

[0028] 在本发明中,用于 AAV 载体的合适的 DNA 分子将包括,例如,填充片段序列和编码本发明 siRNA 分子的表达盒。很多表达盒非常小,例如那些表达抑制 RNAs (siRNAs 和 shRNAs) 的表达盒。因此,有必要添加序列至盒,从而构成全长或接近全长的 AAV 基因组。如果只有小基因组用于 AAV 生产,则重组病毒体会是异质的并且包含各种大小基因组。这是因为病毒喜欢包装全长基因组,所以它会挑取其他 DNA 片段以填充空间。填充片段不能太大,因为高于野生型基因组大小的 105% 的 AAV 基因组将通常不能包装。

[0029] 在某些实施方案中,本发明提供腺伴随病毒(AAV) 填充物组分(也称为“填充片段序列”),其包含与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 90% 同一性的长度为 3300-4200 个核苷酸的核酸。

(SEQ ID NO:1)

GAATTCGGGCTATCCCAGGTTGCCTTGGTTCATGGCAAATGGGACGTTAAGAGGGCAGAGAGAAT
ATGAACAGAACTGTTCTAATATTGGTCATTTAATGTGTAAGTATTGTTCTTTTTAAACCTCCTTC
ATTTTTTTCCAGGAATTGCTGGACACAGTGGCTTGGTGTGTGTCTGAGGACTGTAGGCCATGGCC
CTAGGTTGTGGTTTTAGGTCTCAGGTGCTCTTCCCTGGCTGTCTCCTTGCTTCTTCCCATGTCCTCTT
CTTTGTTCCAGCCATTTCTCCCTTATGCTTAAGTTTGGTGCAGCAGGGTTTGGCTGCTCTCAGATT
CCTGCTTCCTCAGATGCTGTAGTTGTCAGGCCAGCGGGCTGGCAGCGGGATCAGGATCTGGCTAG
GTTTGCTCTCACTGTGGCAGAGTAGGGGGAGGCGTGGGAGAGCACGTGTGACCCAGGCCAGCTG
TAGGGAGCATAGGCATGGTCACGTAGCCTCAGGTCCCTAGACTTTGTCTTCTCATGAGTATGGCTG
TGTGTGTATGGTGAAAACCTAGGTTCTACTTAGCCCAAGAAAATGGGCACATTTTGCATGTGGTTTC
TGTAGAGAAATGCACTGGGTATCTGACATAGCCTGGCAGCATGCCTCCCTCAGGTAGGTTAGTCTC
AGGCGGTGAAGCACGTGTGTCCAGCAAGAACTTCATATGTGGCATAAAGTCTCCGTTCTGTGAGGT
GCTGGCAAATCACCACCACCGTCAAGAGGCTGAAGTGATTTTTGTCTAGGGAGGCAGGAAAGGCT
TCCTGGAGTCAGCAGCCAGTAGGTGAAAGAGTAGATTGGAGACCTTCTTAATCATCACCGCCTCTT
GTCTCAAGGGGTGCCAGGAAGCTGTGGAGGCTGAACCCATCTTATGCTGCCAGAGAGTGGGACAC
CATGAGGGTCAGGTCAAGGGGTTGTACCTTGTTTGGTAGAGAATTAGGGGCTCTTGAAGACTTTGG
ATGTGGTCAGGGGAGTGTATCATTTAGGAAGAGTGACCCGGTGAGGACGTGGGGTAGAGGAGGAC
AGGTGGGAGGGAGTCCAGGTGGGAGTGAGTAGACCCAGCAGGAGTGCAGGGCCTCGAGCCAGGA
TGGTGGCAGGGCTGTGAGGAGAGGCAGCCACCTGTGTGTCTGCGGAAGCAGGGGCAAGAGGGAA
GAGGCCAGCAGCGTGCTGCCATCACCCAGCGACTGGCGTAGATTGTGAGAGACCATTCCCTGCTCT

TAGGAGGGGCTGAGTTTTAGTTTTCTCTGTGTTATACAATAAGCTTGGTATTTGTTTACAAAACATTT
GTAAAGCTAAATCAAGGTTTGATAAGGCTTCTAGTTTTATTTAAGAAGTAATGTTGAAAATAAATGT
TTGTCCAATTCGCTTTGCTCATTAAAGGACTTTCAGTACAAACTGCAACAACAGGATTAGGATTTA
AACGTTTTCTGAGATGTTTTACTCCTCAGAATTTCCAGAATGTGATCTGGTTTTGATTTTCAAGCT
TGCTGACCCAATAGGTTAACCCACAAGTTTTACGAAGACCATCTCAGTCCACTTACATCAACTGCC
CATGCCACGGTTAAAGAGATCATCGACTGATGTTTGGCACAGCTTCTCCCTCTTGGGTGGGCAAG
CATTTGGAAGAGAAGGCTCCTATGGGTGAGAGTGGGGCACCAAAGTCTTCCCTGTCCCATCCCCTA
GCTTGAGAAGCCCTTCTCTAATGTGGACTTTGTGCCGTTAGCATCGTTACTAGCTTGAAGTTGACCA
TCTGGACGTAATTTCTGGTTTAGCCTCACAAAGTGAGCAAGGAGGGTTGAGAGATGTGCTGTGAGGA
ATGTGGGGGCCAGCTGGCAGCAGGCTCTGGGTGAGGGGGCAGGGACCACGGGCATACCTGACA
GTGAGGAGGGTCTAGTAGGGGATCAGTCCCCTGTTGTTCTTTAGAATTTTCTGGATATTCTTCTTT
ATTGATTTTGGGATGTGAACAATAGAATCAACTTCTACTTGTAGATTGATTTAGGGAGAACTTATA
CCTCAGATGTTAAGTCACCCTGTCCAGAATGTGGGATGCTTTCCTATTTGTTGAGAACTTTTTAAAT
TACCTCAGAAGCACATGAAATTTAAAGGATTTTAAAAAAACTTAAAGATTATTTACATAGCTCT
TGCACATTTCTTGATAAATGAATCCTCAGGTATTCTCTGTTTTTGTACTAATAGTTACTTCTTATG
GGTTTTTTTTCCCCTGAAAATCATTATCAAACGTATGTGGCTTATTTTCTGAAGGATGTTTGATAA
TTTTGGAAGATATGAAAGTCTTCATATTTACAAGGTTTGAGGTCTCTTAAAGCTGCATGGTTCTCA
TGTCAGCTCCCAAAGCAGAAGACGGCATGTTGAAAAATGCCGTAGAGAAGATACTTCTTTTCCACC
TGTTTTCAACTCATATCATCTTGAATTTGAGGGCACCTTCCATGCTCCTAGTGCTTGCTATCTGTTT
ATTATTTTCTTCTGAATACCCTGAACTCCAGCATGTTCTGCTGTAATTCTGGCCTCCCTGGCATE
TTGGACTCCTGTTTCTTTGCTCTGTCATCCCCGCGTCAGCTCCTGCTGCGCAGCTTCTCAGCTGA
AGTGCGTTTTGGAGTGCCTGGCGTGTCTTGTGGATCTTTGAGTATTGCCCTCTGGTTTCTTGGTTCC
TTCTGCTGAGTTGCTCAGCGTCTCCACTCCCCATTTCTTGTGTGGCCCTTCTGCACTCCTCTGATTC
CTTTTGTCTTCCCCTGGTTTCTTGTCTTGGTTTCGAGTCTCCACAGAACTTTTGCAGCTCTTCTGAAGA
CCTGGAAGCTTTTTCATCTTAATTCTCATCTCATGACCTCTTTTCCCCTTCTTTGAGAGCTAGAACTTC
CCATGGTGAACCTTCTTTCCAGAATCCATGCCTTCTTTCCCCTCCCCTTACCTGTTGTCCAGGA
GAGGTCAGATTGCTGTGCATATTGGAGGAGAACCCTTTCTTCCCCTGGGCTCTTCATCTCACATGAC
ATCACCACATCACCTCGTTCCTTGGACCCTCAGTGGTGTCACTGCTGGATTTTTCTTTCCCTTTGGCT
GGCCTTAGGGCACACCAGGTTGACTAGCGTAGTCATGGTATTTAGATCCACTCACATTTTCAGTT
TCTGTGTCTGTCTTTGCCTGCTTCTGACTTCGCCAGAGAAAGCTTCTCTTTCACAAGGGTTCTTA
GATTTATGTTCACTGAGCACCTTCTTTTCTGAGGCAGTGTTTACCAATATTTATTTTCTAGTCAGT
CTCGCCTTACCTTTCTTGTATGCATGTCTTTGGTCTGACCCATTCTCTGAGTCTGTAAAATAGAA
TTGCTGTATAATTTAATTACATGAAATCCTTTAGAATCTTAACACATCTTACACCTGATTTAATATT
TTATTGTATCCAAATTGAACCAACCCTATGTGAATTTGACAGTGATTTCTCCCAGGGATCCTAGTGT
ATAAGGAATAGGACTTAGTATTTTCTATTTTTTGATATACCACATACCAGATACTGATTATGATGG
ACATTTAACCCTTTTTTCTCATTATGAAAGAAAGTTAGGAATATTTCTTCCAGTAGCGCCAGTGTA
ACCTGAAAGCCTTTGAAAGAGTAGTTTTTGTATAGCTATCTGAAAGGAATTTCTTCCAAAATATTT
TTCCAGTGCTGACAACAAACACGCAGACACACCCTGCAAGGTGAGTGTACGGCG

(SEQ ID NO:2)

GGGCTATCCCAGGTTGCCTTGGTTCATGGCAAATGGGACGTTAAGAGGGCAGAGAGAATATGAAC
AGAACTGTTCTAATATTGGTCATTTAATGTGTAAGTATTGTICTTTTTTAAACCTCCTTCATTTTT
TTCCAGGAATTGCTGGACACAGTGGCTTGGTGTGTGTCTGAGGACTGTAGGCCATGGCCCTAGGTT
GTGGTTTTAGGTCTCAGGTGCTCTTCTGGCTGTCTCCTTGTCTCTTTCCCATGTCTCTCTTTGTTT
CCAGCCATTTCTCCCATTATGCTTAAGTTTGGTGCAGCAGGGTTTGGCTGCTCTCAGATTCCCTGCTTC
CTCAGATGCTGTAGTTGTCAGGCCAGCGGGCTGGCAGCGGGATCAGGATCTGGCTAGGTTTGCTC
TCACTGTGGCAGAGTAGGGGGAGGCGTGGGAGAGCACGTGTGACCCCAGGCCAGCTGTAGGGAG
CATAGGCATGGTCACGTAGCCTTCAGGTCCTAGACTTTGTCTTCTCATGAGTATGGCTGTGTGTGTA
TGGTGAAAAGTAGGTTCTACTTAGCCCAAGAAAATGGGCACATTTTGCATGTGGTTTCTGTAGAGA
AATGCACTGGGTATCTGACATAGCCTGGCAGCATGCCTCCCTCAGGTAGGTTAGTCTCAGGCCGGT
AAGCACGTGTGTCCAGCAAGAACTTCATATGTGGCATAAAGTCTCCGTTCTGTGAGGTGCTGGCAA
ATCACCACCACCGTCAAGAGGCTGAAGTGATTTTTGTCTAGGGAGGCAGGAAAGGCTTCCTGGAG
TCAGCAGCCAGTAGGTGAAAGAGTAGATTGGAGACCTTCTTAATCATCACCGCCTCTTGTCTCAAG
GGGTGCCAGGAAGCTGTGGAGGCTGAACCCATCTTATGCTGCCAGAGAGTGGGACACCATGAGGG
TCAGGTCAAGGGTTGTACCTTGTGGTAGAGAATTAGGGGCTTGAAGACTTTGGATGTGGTC
AGGGGAGTGTATCATTAGGAAGAGTGACCCGGTGAGGACGTGGGGTAGAGGAGGACAGGTGGG
AGGGAGTCCAGGTGGGAGTGAGTAGACCCAGCAGGAGTGCAGGGCCTCGAGCCAGGATGGTGGC
AGGGCTGTGAGGAGAGGCAGCCACCTGTGTGTCTGCGGAAGCAGGGGCAAGAGGGAAGAGGCCA
GCAGCGTGTGCCATACCCAGCGACTGGCGTAGATTGTGAGAGACCATTCCCTGCTCTTAGGAGG
GGCTGAGTTTTAGTTTTCTTGTATACAATAAGCTTGGTATTTGTTACAAAACATTTGTAAAGC
TAAATCAAGGTTTGATAAGGCTTCTAGTTTTATTTAAGAAGTAATGTTGAAATAAATGTTTGTCCA
ATTCGCTTTGCTCATTAAAGGACTTTCAGTACAACTGCAACAACAGGATTAGGATTTAAACGTTT
CTGAGATGTTTTACTCCTCAGAATTTCCAGAATGTGATCTGGTTTTGATTTTCAAGCTTGTGAC
CCAATAGGTTAACCCACAAGTTTTACGAAGACCATCTCAGTCCACTTACATCAACTGCCCATGCCA
CGGTTAAAGAGATCATCGACTGATGTTGGCACAGCTTCTCCCTCTTGGGTGGGCAAGCATTGG
AAGAGAAGGCTCCTATGGGTGAGAGTGGGGCACCAAAGTCTTCCCTGTCCATCCCCTAGCTTGAG
AAGCCCTTCTAATGTGGACTTTGTGCCGTTAGCATCGTACTAGCTTGAAGTTGACCATCTGGAC
GTACTTTCTGGTTTAGCCTCACAAGTGAGCAAGGAGGGTTGAGAGATGTGCTGTGAGGAATGTGG
GGCCCCAGCTGGCAGCAGGCTCTGGGTGAGGGGGGAGGGACCACGGGCATACCTGACAGTGAG
GAGGGTCTAGTAGGGGATCAGTCCCCTGTGTCTTTAGAATTTCTGGATATTCTCTTTATTG
ATTTGGGATGTGAACAATAGAATCAACTTCTACTGTAGATTGATTTAGGGAGAAGTTATACCTC
AGATGTTAAGTCACCCGTCCAGAATGTGGGATGCTTTCCTATTTGTTTCAAGACTTTTAAATTACC
TCAGAAGCACATGAAATTTAAAGGATTTTAAAAAACTTAAAGATTATTTACATAGCTCTTGCA
CATTTCTTGATAAATGAATCCTCAGGTATTCCTCTGTTTTTGTACTAATAGTTACTTCTTATGGGTT
TTTTTCCCCTGAAAATCATTATCAAACGTATGTGGCTTATTTCTGAAGGATGTTTGATAATTTT
GGAAGATATGAAAGTCTTCATATTTTACAAGGTTTGGGGTCTCTTAAAGCTGCATGGTTCTCATGTC
AGCTCCCAAAGCAGAAGACGGCATGTTGAAAAATGCCGTAGAGAAGATACTTCTTTCCACCTGTT
TTCAACTCATATCATCTTGAATTTAGGGCACCTTCCATGCTCCTAGTGCTTGCTATCTGTTATTA
TTTTCTTCTGAATACCCTGAACTCCAGCATGTTCTGCTGTAATTTCTGGCCTCCCTGGCATCTTGG
ACTCCTGTTTCTTGTCTGTATCCCCGGGTGAGCTCCTGCTGCGCAGCTTCTCAGCTGAAGTG
CGTTGGAGTGCTGGCGTGTCTGCTGGATCTTTGAGTATTGCCTCTGGTTTCTTGGTTCCTTCTG


```

CTGAGTTGCTCAGCGTCTCCACTCCCCATTTCTTGTGTGGCCCTTCTCTGCACTCCTCTGATTCCTTTT
GTCTTCCCTGGTTTTCTTGGTTTTGTTTTGAGTCTCCACAGAACTTTTGCAGCTCTTCTGAAGACCTG
GAAGCTTTTTTCATCTTAAATTCTCATCTCATGACCTCTTTTCCCTTCTTTGAGAGCTAGAACTTCCCAT
GGTGAACCTTCTCTTTCCAGAATTCATGCCTTCTTTTCCCTCCCACCTTACCTGTTGTCCAGGAGAGG
TCAGATTGCTGTGCATATTGGAGGAGAACCCTTTCTTCCCTGGGCTCTTCATCTCACATGACATCAC
CACATCACCTCGTTCCTTGGACCCTCAGTGGTGTCACTGCTGGATTTTTCTTTCCCTTGGCTGGCCTT
AGGGCACACCCAGGTTGACTAGCGTAGTCATGGTATTTAGATCCACTCACATTTTCAGTTTCTGTGT
CTGTCTCTTGCCTGCTTCTGACTTCGCCCAGAGAAAGCTTCTCTTTCACAAGGGTTCTTAGATTTAT
GTTCACTGAGCACCTTCTTTTCTGAGGCAGTGTTTTACCAATATTTATTTTCCCTAGTCAGTCTCGCCT
TACCTTTCTTGTATGCATGTCTTTGGTCTGACCCATTCTCTGAGTCTGTAAAATAGAATTGCTGT
ATAATTTAATTACATGAAATCCTTTAGAATCTTAACACATCTTACACCTGATTTAATATTTTATTGT
ATCCAAATTGAACCAACCCTATGTGAATTTGACAGTGATTTCTCCAGGGATCCTAGTGTATAAGG
AATAGGACTTAGTATTTTCTATTTTTTGTATACCACATACCAGATACTGATTATGATGGACATTTA
ACCTTTTTTCTCATTATGAAAGAAAGTTAGGAATTTTCTTCCAGTAGCGCCAGTGTAACCTGAA
AGCCTTTGAAAGAGTAGTTTTTGTATAGCTATCTGAAAGGAATTTCTTTCCAAAATATTTTTTCCAGT
GCTGACAACAAAACACGCAGACACACCCTGCAAGGTGAGTGTACGGCG

```

[0030] 在某些实施方案中，本发明提供腺伴随病毒(AAV)填充物组分，其由与 SEQ ID NO: 1 或 SEQ ID NO:2 具有至少 90% 同一性的长度为 3300-4200 个核苷酸的核酸组成。在某些实施方案中，填充物组分由与 SEQ ID NO: 1 或 SEQ ID NO:2 至少 90% 的同一性组成。在某些实施方案中，填充物组分与 SEQ ID NO: 1 或 SEQ ID NO:2 具有 95% 同一性、98% 同一性、99% 同一性或甚至 100% 同一性。在某些实施方案中，填充物组分具有长度约 3500-4000 个核苷酸，或约 3700-3850 个核苷酸。在本发明中，填充物组分在生物活性方面是“沉默的”，因为其缺乏增强子、启动子、剪接调节子、非编码 RNAs、反义序列或编码序列。

[0031] 术语“核酸”指以单链或双链形式的脱氧核糖核酸(DNA)或核糖核酸(RNA)和其聚合物，其由包含糖、磷酸和可以是嘌呤或嘧啶的碱基的单体(核苷酸)构成。除非明确地限定，该术语包括包含天然核苷酸的已知类似物的核酸，所述类似物具有与参考核酸相似的结合特性并且以与天然存在的核苷酸相似的方式代谢。除非另有说明，具体的核酸序列同样包括其经保守修饰的变体(例如简并密码子取代)和互补序列，以及明确说明的序列。特别地，简并密码子取代可以通过生成其中一个或多个选择的(或全部的)密码子的第三位取代有混合的碱基和 / 或脱氧肌苷残基的序列实现。“核酸片段”是给出的核酸分子的一部分。

[0032] “核苷酸序列”是 DNA 或 RNA 的聚合物，其可以是单链或双链的，任选地包含可以掺入 DNA 或 RNA 聚合物中的合成的、非天然的或经修饰的核苷酸碱基。术语“核酸”、“核酸分子”、“核酸片段”、“核酸序列或区段”或“多核苷酸”可互换使用，并且可以同样与基因、通过基因编码的 cDNA、DNA 和 RNA 互换使用。

[0033] 本发明包括分离的或实质上纯化的核酸组合物。在本发明的上下文中，“分离的”或“纯化的”DNA 分子或 RNA 分子是远离其天然环境存在的 DNA 分子或 RNA 分子并且因此不是天然产物。分离的 DNA 分子或 RNA 分子可以以纯化的形式存在，或者可以在非天然环境，例如在转基因宿主细胞中存在。例如，“分离的”或“纯化的”核酸分子或其生物活性的部分，

在通过重组技术生产时实质上不含其他细胞材料或培养基,或者当化学合成时实质上不含化学前体或其他化学品。在一个实施方案中,“分离的”核酸不含在所述核酸来源于其的生物的基因组 DNA 中天然地侧接所述核酸的序列(即位于所述核酸的 5' 和 3' 端的序列)。例如,在多个实施方案中,分离的核酸分子可以包含小于约 5kb、4kb、3kb、2kb、1kb、0.5kb 或 0.1kb 的在所述核酸来源于其的细胞的基因组 DNA 中天然地侧接所述核酸分子的核苷酸序列。公开的核酸的片段和变体也包括在本发明中。“片段”或“部分”意指全长的核苷酸序列或小于全长的核苷酸序列。

[0034] “天然存在的”、“天然的”或“野生型”用于描述可以在自然中发现而区别于人工生产的对象。例如,存在于生物(包括病毒)中的蛋白或核苷酸序列(其可以从自然中的来源分离并且没有通过实验室人员有意地修饰)是天然存在的。

[0035] “基因组”指生物的全部遗传材料。

[0036] “载体”定义为包括,尤其是任何病毒载体,以及任何质粒、粘粒、噬菌体或二元载体,其以双链或单链、线性或环状形式,可以是或可以不是可自主转移的或可移动的,并且其可以转化原核或真核宿主。

[0037] AAV ITR

“AAV 病毒”或“AAV 病毒颗粒”指由至少一个 AAV 衣壳蛋白(优选由野生型 AAV 的全部衣壳蛋白)和使衣壳化的多核苷酸构成的病毒颗粒。如果颗粒包含异源的多核苷酸(即,野生型 AAV 基因组以外的多核苷酸,例如递送至哺乳动物细胞的转基因),其通常称为“rAAV”。

[0038] 在一个实施方案中,AAV 表达载体使用已知技术构建,以至少提供在转录方向上的可操作地连接的组分,包括转录起始区、目的 DNA 和转录终止区的控制元件。选择在哺乳动物细胞中有功能的控制元件。产生的包含可操作地连接的组分的构建体侧接(5' 和 3')有功能的 AAV ITR 序列。

[0039] “腺伴随病毒反向末端重复”或“AAV ITR”意指在 AAV 基因组的每一端发现的本领域公知的区域,其以顺式一同作为 DNA 复制起点并且作为病毒的包装信号。AAV ITR,与 AAV rep 编码区一同,提供从表达它们的质粒中的有效删除。

[0040] 已知 AAV ITR 区的核苷酸序列。如本文使用的,“AAV ITR”不需要具有描述的野生型核苷酸序列,但可以被修饰,例如通过核苷酸的插入、删除或取代。此外,AAV ITR 可以源自几种 AAV 血清型的任一种,包括但不限于,AAV1、AAV2、AAV3、AAV4、AAV5、AAV7 等。此外,在 AAV 载体中侧接所选择的核苷酸序列的 5' 和 3' ITR 不需要一定相同或来源于相同的 AAV 血清型或分离物,只要它们的功能如预期的,即,允许从载体删除和拯救目的序列和将期望的基因组包装入 AAV 病毒体。

[0041] 在一个实施方案中,AAV ITR 可以源自几种 AAV 血清型的任一种,包括但不限于,AAV1、AAV2、AAV3、AAV4、AAV5、AAV7 等。此外,在 AAV 表达载体中侧接选择的核苷酸序列的 5' 和 3' ITR 不需要一定相同或来源于相同的 AAV 血清型或分离物,只要它们的功能如预期的,即,允许从载体删除和拯救目的序列和允许将期望的基因组包装入 AAV 病毒体。

[0042] 在某些实施方案中,本发明提供腺伴随病毒(AAV)载体,其包含可操作地连接至表达盒的如上文所述的填充物组分。在某些实施方案中,表达盒包含启动子。在某些实施方案中,启动子是 pol III 启动子。在某些实施方案中,启动子是 mU6 启动子。在某些实施方案中,AAV 载体进一步包含靶序列。在某些实施方案中,靶序列是 RNAi 分子。

[0043] 如本文使用的“表达盒”意指能够在合适的宿主细胞中指导特定核苷酸序列表达的核酸序列,其可以包括可操作地连接至目的核苷酸序列的启动子,所述目的核苷酸序列可以可操作地连接至终止信号。编码区通常编码目的功能 RNA,例如 RNAi 分子。包括目的核苷酸序列的表达盒可以是嵌合的。表达盒还可以是天然存在的但是已经以用于异源表达的重组形式获得的表达盒。

[0044] 双链 RNA (dsRNA) 可以通过称为 RNA 干扰 (RNAi) 的过程在很多生物中诱导序列特异性的转录后的基因沉默。RNA 片段是 RNAi 的序列特异性的介体。通过这些 RNA 干扰 (RNAi) 分子的基因表达干扰如今被公认为在很多生物的细胞中沉默基因的天然存在的策略。

[0045] 本发明某些实施方案提供编码分离的 RNAi 分子的载体。如本文使用的术语“由……编码”以广义使用,类似于在专利术语中的术语“包含”。RNAi 分子包括 siRNA、shRNA 和其他小 RNA,其可以或有能力调节靶基因的表达,例如通过 RNA 干扰。这样的小 RNA 包括但不限于,shRNA 和微 RNA (miroRNA, miRNA)。

[0046] “可操作地连接”指在单一核酸片段上核酸序列的连接,从而使得序列之一的功能受另一个所影响。例如,将调节 DNA 序列称为“可操作地连接至”编码 RNA 或多肽的 DNA 序列或“与其连接”,如果这两个序列处于这样的位置,从而使得调节 DNA 序列影响编码 DNA 序列的表达(即,编码序列或功能 RNA 在启动子的转录控制之下)。编码序列可以以有义或反义方向可操作地连接至调节序列。

[0047] 可操作地连接的核酸是位于与另一核酸序列的功能关系中的核酸。例如,启动子或增强子可操作地连接至编码序列,如果其影响所述序列的转录;或核糖体结合位点可操作地连接至编码序列,如果它这样放置从而使得促进翻译。通常,可操作地连接的 DNA 序列是连续连接的 DNA 序列。然而,增强子不必是连续的。连接通过在方便的限制性位点连接来完成。如果这样的位点不存在,则依照常规实践使用合成的寡核苷酸接头 (adaptor) 或接头 (linker)。

[0048] 本发明现将通过以下非限制性实施例来说明。

[0049] 实施例 1

生成包含 AAV2 ITR、mU6 启动子、miHDS1 靶序列、填充物组分填充片段和 AAV 骨架的质粒 FBAAVmU6miHDS1 填充片段(图 1)。5pFBAAVmU6miHDS1AAV 填充片段的序列在图 2 中提供,并且该质粒的单独组分的序列在图 3 中提供。全长的填充物组分(“填充片段序列”)由 3776 个核苷酸组成。

[0050] 实施例 2

比较表达 miHDS1 的载体的体内沉默效率。构建 4 种载体:(1)表达控制序列(miSAFE)并且包含控制序列(eGFP)的载体,(2)表达靶序列(miHDS1)并且包含控制序列(eGFP)的载体,(3)表达控制序列(miSAFE)并且包含实施例 1 中所述的填充片段序列的载体,和(4)表达靶序列(miHDS 1)并且包含实施例 1 中所述的填充片段序列的载体。

- (1) AAV2/1 mU6miSAFE – eGFP (4.81E12 µg/ml)
- (2) AAV2/1 mU6miHDS1 – eGFP (4.81E12 µg/ml)
- (3) AAV2/1 mU6miSAFE – 填充片段 (4.81E12 µg/ml)
- (4) AAV2/1 mU6miHDS1 – 填充片段 (4.81E12 µg/ml)

[0051] miSAFE 和 miHDS1 的序列在先前已有讨论(见,PCT/US2012/024904,其在此以其整体通过引用并入本文)。野生型小鼠在纹状体中注射 4 种载体。小鼠在 1 个月处死并且通过 QPCR 测定相对于 Actb 表达水平的 Htt 表达。图 4 显示在 misafe/eGFP 和 miHDS1 / eGFP 表达盒之间存在表达 20% 的降低,而在 misafe/ 填充片段和 miHDS1 / 填充片段表达盒之间存在表达 60% 的降低,即,当使用填充片段时,表达降低 60%。

[0052] 实施例 3

生成包含 AAV2 ITR、mU6 启动子、miHDS1 靶序列、填充物组分填充片段和 AAV 骨架的质粒 5pFBAAVmU6miHDS1 填充片段(图 5)。质粒 5pFBAAVmU6miHDS1 AAV- 填充片段的序列在图 6 中提供。填充片段(填充片段 #2)的序列在图 7 中提供。

[0053] 实施例 4

AAV 包装的考虑因素之一是维持最佳的基因组大小。当这发生时,将缺少基因组形式的病毒体的比例降到最低。进行测试新填充片段序列的包装效率的实验并且发现高效率的包装。例如,见表 1“平均空的(Average empty)”和图 8。如果进行包装的遗传材料包含非 miRNA :内含子填充片段序列,则也测量它。发现用于病毒产生的非计划中的(unintended)基因组材料的掺入是极低的(Cap/rAAV、Amp/rAAV、Gent/rAAV)。最终,病毒质量通过在聚丙烯酰胺凝胶电泳后的银染色进行分析,并且发现包含合适比例的各种衣壳蛋白(VP1、VP2 和 VP3 ;图 9)。总之,内含子 I/II 填充片段序列允许将期望的转基因包装入 AAV 衣壳的最佳包装。

[0054] 全部公开、专利和专利申请通过引用并入本文。虽然在前述说明书中本发明已经关于其某些优选的实施方案进行了描述,并且很多细节已经为说明目的而进行陈述,但对本领域技术人员将显而易见的是本发明允许额外的实施方案,并且本文描述的某些细节可以在不背离本发明基本原则的情况下进行相当大地改变。

[0055] 在描述本发明的情况下使用的术语“一个(a)”和“一个(an)”和“所述(the)”和相似的指示物,应解释为涵盖单数和复数,除非本文另有说明或在上下文中明显矛盾。术语“包含(comprising)”、“具有(having)”、“包括(including)”和“含有(containing)”应解释为开放式(open-ended)术语(即,意为“包括,但不限于”),除非另有注释。本文的值的范围的详述仅仅旨在作为单独地指每一独立的值落入所述范围内的速记方法,除非本文另有说明,并且每一独立的值如同其被本文单独地描述一样并入本说明书中。本文描述的所有方法可以以任何合适的顺序实施,除非本文另有说明或另外在上下文中明显矛盾。本文提供的任何和全部实例、或示例性语言(例如“例如(such as)”)的使用,仅仅旨在更好地说明本发明并且不构成对本发明范围的限制,除非另有声明。在说明书中的任何语言均不应解释为表示对于本发明的实践是必需的任何未声明的要素。

[0056] 本文描述了本发明的实施方案,其包括本发明人已知地实施本发明的最佳模式。那些实施方案的变化对于本领域普通技术人员而言基于阅读前文描述可以是显而易见的。

本发明人期望技术人员合适地采用此类变化,并且本发明人旨在本发明以与本文明确描述不同的其他方式来实践。因此,本发明包括由适用法律允许的在本文附带的权利要求中描述的主题的所有修改和等同方案。此外,上述要素在其所有可能变化中的任何组合均包括于本发明中,除非本文另有说明或另外在上下文中明显矛盾。

[0001]

序列表

<110> UNIVERSITY OF IOWA RESEARCH FOUNDATION

<120> 经修饰的腺伴随病毒载体组合物

<130> 17023, 126W01

<140>

<141>

<150> 61/668, 839

<151> 2012-07-06

<160> 12

<170> PatentIn版本3.5

<210> 1

<211> 3774

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 合成的多核苷酸

<400> 1

```

gaattcgggc tatcccaggc tgcccttggt catggcaaat gggacgttaa gagggcagag      60
agaatatgaa cagaaactgt tctaataatg gtcatttaat gtgtaagtat tgttcttttt      120
taacctctct tcattttttt tccaggaatt gctggacaca gtgcttgggt gtgtgtctga      180
ggactgtagg ccatggccct aggttgggtt tttagtctc aggtgctctt cctggctgfc      240
tccttgcttc ttcccattgt cctctctttt gtttccagcc atttctccct tatgcttaag      300
tttgggtcag cagggtttgg ctgctctcag attcctgctt cctcagatgc tgtagtgtc      360
aggcccagcg ggctggcagc gggatcagga tetggctagg ttgtctctca ctgtggcaga      420
gtagggggag gcgtgggaga gcacgtgtga ccccaggcca gctgtaggga gcataggeat      480
ggtcacgtag ccttcaggtc ctagaatttg tctctcatg agtatgctg tgtgtgtatg      540
gtgaaaacta ggttctactt agcccaagaa aatgggcaca ttttgcattg ggtttctgta      600
gagaaatgca ctgggtatct gacatagcct ggcagcatgc ctccctcagg taggttagtc      660
tcaggcgggt aagcacgtgt gtccagcaag aacttcatat gtggcataaa gtcctcgttc      720
tgtgagggtc tggcaaatca ccaccaccgt caagagctg aagtattttt tgtctaggga      780
ggcaggaaag gcttctctga gtcagcagcc agtaggtgaa agagtagatt ggagaccctc      840
ttaateatca cgcctctctg tetcaagggg tgccaggaag ctgtggagge tgaacccatc      900
ttatgctgac agagagtggg acaccatgag ggtcaggtea aggggttgta cettgttttg      960
tagagaaita gggctctctg aagactttgg atgtgtctag gggagtgtat caittaggaa      1020
gagtgacccg gtgaggacgt ggggtagagg aggacagggt ggagggagtc caggtgggag      1080
tgagttagcc cagcaggagt gcaggcctc gagccaggat ggtggcaggg ctgtgaggag      1140
aggcagccac ctgtgtgtct gcggaagcag ggcgaagagg gaagaggcca gcagcgtgct      1200
gccatcacc cagcagctggc gtagattgtg agagaaccatt ccttctcttt aggagggget      1260
gagtttttagt ttctctctgt tatacaataa gcttgggtatt tgtttacaaa acattttgta      1320
agctaaatca aggtttgata aggettctag ttttatttaa gaagtaatgt tgaataaat      1380
gtttgtccaa ttctctttgc tcafttaagg actttcagta caaacgcaa caacaggatt      1440
aggatttaaa cgtttctgag atgtttttac tctctcagaat ttcccagaat gtgatctggt      1500

```

[0002]

tttgatttfc aagcttfgctg acceaafagg tfaaccacaca agttttacga agaccatctc	1560
agtccactta catcaactgc ccatgccacg gttaaagaga tcatcgactg atgtttggca	1620
cagcttccctc cctcttgggt gggcaagcat ttggaagaga aggctcciat gggfagagat	1680
ggggcaceaa agtcttccct gtcceatecc ctagecttgag aagcccttct ctaatgtgga	1740
ctttgtgccc ttagcatcgt tactagcttg aagttgacca tctggacgta cttctgtgtt	1800
tagcctcaca agtgagcaag gagggttgag agatgtgctg tgaggaaatgt ggggccccag	1860
ctggcagcag gctctgggtc aggggggcag ggaccacggg catacctgac agtgaggagg	1920
gtctagtagg ggatcagttc cctgtgtgtt ctttagaatt ttctggatat tcttctttat	1980
tgattttggg atgtgaacaa tagaatcaac ttctacttgt agattgattt agggagaact	2040
tatacctcag atgttaagtc accctgtcca gaatgtggga tgctttccta tttgttcaga	2100
actttttaaa ttacctcaga agcacatgaa atttaagga ttttaaaaa aacttaaaga	2160
ttatttcaca tagctcttgc acatttcttg ataaaigaat cctcaggtat tctctgtttt	2220
ttgttactaa tagttacttc ttatgggttt ttttccct gaaaatcatt tatcaaacgt	2280
atgtgctta tttttgaag gatgtttgat aattttgaa gatatgaaag tcttctatit	2340
ttacaaggtt tgaggctctt ttaagetgca tggttctcat gtcagctccc aaagcagaag	2400
acggcatgtt gaaaaatgce gtagagaaga tacttctttt ccacctgttt tcaactcata	2460
tcatcttgaa tticagggca cctttccatg ctcttagtgc ttgetatctg tttattattt	2520
tccttctga ataccctgaa ctccagcatg ttctgctgta attctggcct ccttggcacc	2580
ttggactcct gtttctttg ctctgtcate cccgggtca gctctgctg cgcagcttct	2640
cagctgaagt gcgtttggag tgcctggcgt gtcttctgtg atctttgagt attgcctctg	2700
gtttcttgg ttccttctgc tgagttgctc agegtctcca ctcccattt cttgtgtggc	2760
cttctctgca ctctctgat tcttttgc ttcccgttt tcttgetttg gttctgagtc	2820
tccacagaac ttttgcagct ctctgaaga cctggaagct tttctatctt aattctcate	2880
tcatgaecte ttttccctt tttagagctt agaacttccc atggggaact tctctttcca	2940
gaattccatg ccttctttc ctctccactt acctgttctc caggagaggt cagattgctg	3000
tgeatattgg aggagaacce tttcttccct gggtcttcca tctcactga catcaccaca	3060
tcaccctgll ccttggacc ctagtgggtt cacctctgga ttttcttct ctttggctgg	3120
ccttagggca caccaggtt gactagcgtg gtcatggfat tttagatcca tcaatitctc	3180
agttttctgt tctgtctctt gctgtctctt gacttccccc agagaangct tctctttcac	3240
aagggttctt agatttatgt tcactgagca ccttctttc tgaggcagtg ttttaccat	3300
atttatttct ctagctcagtc tegecttacc ttcttctgta tgeatgtctt tggctctgac	3360
ccattctctg agteigtaaa atagaattgc tgtataattt aattacatga aatcttttag	3420
aactttaaca catcttacac ctgatttaat attttatgtt atccaaattg aaccaacct	3480
atgtgaattt gacagtgatt tctcccagg atcttagtgt ataaggaata ggacttagta	3540
tttctatatt tttagatata cacataccag atactgatta tgatggacat ttaacccttt	3600
ttctcatta tgaagaaaag ttaggaatta tttcttccag tgcgccagt gtaacctgaa	3660
agcctttgaa agagtatttt ttgtatagct atctgaaagg aatttcttct caaatattt	3720
ttccagtctt gacaacaaac acgcagacac acctgcaag gtagtgtac ggcg	3774

[0003]

<210> 2
 <211> 3769
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列描述：合成的多核苷酸

<400> 2
 gggetatccc aggttgcctt ggttcatgce aaatgggacg ttaagagggc agagagaata 60
 tgaacagaaa ctgttctaata attggtcatt taatgtgtaa gtattgttct tttttaaacc 120
 tccittcattt tttttccagg aattgctgga cacagtggct tgggtgtgtg ctgaggactg 180
 taggcatggt ccctaggttg tggttttagg tctcagggtc tcttctctggc tgtctctctg 240
 ctctctctcc atgtctctct ctttctctcc agccatttct cccttatgct taagtttggc 300
 gcagcagggt ttgctctctc tcagattcct gcttctctcag atgctgtagt tgtcaggccc 360
 agcgggctgg cagcgggacg aggatctggc taggtttgct ctactgttg cagagtaggg 420
 ggaggcgtgg gagagcacgt gfgaccccag gecagctgta gggagcatag gcatggctac 480
 gtagccttca ggtcttagac ttgtctctct catgagtatg gctgtgtgtg tatggtgaaa 540
 actaggttct acttagccca agaaaatggg cacattttgc atgtggttgc tgtagagaaa 600
 tgcactgggt atctgacata gcttggcagc atgctctcct caggtaggtt agtctcaggc 660
 ggtgaagcac gtgtgtccag caagaacttc atatgtggca taaagtctcc gttctgtgag 720
 gtgtgtgcaa atcaccacca ccgtcaagag gctgaagtga tttttgtcta gggaggcagg 780
 aaaggcttcc tggagtcagc agccagtagg tgaagagta gattggagac ctctttaatc 840
 ataccgcct cttgtctcaa ggggtgccag gaagctgtgg aggetgaacc catcttatgc 900
 tgcagagagag tgggacacca tgagggtcag gtcaagggtt tgfaccctgt ttggtagaga 960
 attaggggct cttgaagact ttggatgtgg tcaggggagt gttacattta ggaagagtga 1020
 cccggtgagg acgtgggcta gaggaggaca ggtgggaggg agtccaggtg ggagttagta 1080
 gaccagcag gagtgcaggg cctcagccca gtaggtggc agggctgtga ggagaggcag 1140
 ccacctgtgt gctctcgcaa gcagggcaca gagggaagag gccagcagcg tgcctccatc 1200
 acccagcgac tggcgtagat tgtgagagac cattcctctc tcttaggagg ggctgagttt 1260
 tagttttctc ttgttataca ataagcttgg tatttgttta caaacattt gtaaagctaa 1320
 atcaaggttt gataaggctt ctagtattat ttaagaagta atgttgaat aatgtttgt 1380
 ccaattcgtt ttgtctattt aaggacttcc agtacaact gcaacaacag gattaggtat 1440
 taaacgttcc tgagatgttt ttactctcca gaatttccca gaatgtgac tggttttgat 1500
 ttcaagctt gctgaccaa taggttaacc cacaagtttt acgaagacca tctcagttca 1560
 ettacatcaa ctgcccagc cacggttaaa gagatcatcg actgatgttt ggcacagctt 1620
 ctctctctct ggggtggcaa gcatttggaa gagaaggctc ctatgggtga gagtgggca 1680
 ccaaaagctt cctgtctcca tcccctagct tgagaagccc tctctaatg tggactttgt 1740
 gccgttagca tegttaactag cttgaagttg accatctgga cgtacttctt ggtttagcct 1800
 caaagttag caaggagggt tgagagatgt gctgtgagga atgtggggcc ccagctggca 1860
 gcaggetctg ggtcaggggg gcagggacca cgggcatacc tgacagttag gaggggtcta 1920
 gtaggggacg agttccctg ttgttcttta gaattttctg gatattcttc tttattgatt 1980

[0004]

ttgggaatg aacaatagaa tcaacttcta ctgtagatt gatttagga gaactfatac 2040
 ctacagatgt aagtcacctt gccagaatg tgggatgctt tcctatttgt tcagaacttt 2100
 ttaaattacc tcagaagcac atgaaattta aaggatttta aaaaaaactt aaagattatt 2160
 tcacataget ctgacacatt tcttgataaa tgaatectca ggtattecte tgtttttgtt 2220
 actaatagtt acttcatttg ggtttttttt cccctgaaaa tcatttatca aacgtatgtg 2280
 gcttatttct tgaaggatgt ttgataattt tggagaatat gaaagfcttc atattttaca 2340
 aggtttgggg tctctttaag ctgcatggtt ctcatgtcag ctcccaaagc agaagacggc 2400
 atgttgaaaa atgecttaga gaagatactt cttttccacc tgttttcaac tcatatcacc 2460
 ttgaatttca gggcaccctt ccatgctcct agtgcttctt atctgtttat taitttcctt 2520
 cctgaatacc ctgaactcca gcatgttctg ctgtaattct ggctcctctg gcactcttga 2580
 ctctgttttc ctttctctg tcafeccccc ggtcagctcc tgcctgcagc ctctcagctt 2640
 gaagtgcctt tggagtgctt ggcgtgtctt gctggatctt tgagtattgc ctctggtttc 2700
 cttggttctt tctgtgagt tgcctcagct ctccactccc cttttcttgt gtggcccttc 2760
 ctgcactcct ctgattcctt ttgtcttccc tggtttcttg ctttggtttc gagtctccac 2820
 agaacttttg cagctcttct gaagacctgg aagcttttcc atcttaattc tcatctcatg 2880
 accttttct ccttctttga gagctagaac ttcccatggt gaacttctct ttcagaattt 2940
 ccatgccttc ttttccctcc cacttacctg ttgtccagga gaggtcagat tgetgtgcat 3000
 attggaggag aaccttttct tcctgggctt ctctcatctc catgacatca ccacatcacc 3060
 tegtctctg gacctcagt ggtgtcactg ctggattttt ctttcttttg gctggcctta 3120
 gggcacaccc aggttgacta gcgtagtcat ggtattttaga tccactcaca tttcagttt 3180
 ctgtgtctgt ctcttgctg cttctgactt cgeccagaga aagcttctct ttcacaaggg 3240
 ttcttagatt tatgttcact gagcaecttc tttctgagg cagtgtttta ccaatattta 3300
 ttttcttagt cagtctgcc ttacctttct tgttatgcat gtttttggc ctgacctatt 3360
 ctctgagctt gtaaaataga atgtctgat aatttaatta catgaaatcc tttagaatct 3420
 taacacatct tacacctgat ttaatatitt atgtatcca aattgaacca acctatgtg 3480
 aatttgacag tgatttctcc caggatcctt agtgataag gaataggact tagtatttcc 3540
 ttttttttga taaccacat accagatact gattatgat gacatttaac cctttttctt 3600
 cattatgaaa gaaagttagg aattatttct tcagtagcg ccagtgtaac ctgaaagcct 3660
 ttgaanagat agtttttgta tagctatctg aangnaattt ctttceaaaa tatttttcca 3720
 gtgtgacaa caaacacgca gacacacctt gcaagtgag tgtacggcg 3769

<210> 3

<211> 9110

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 合成的多核苷酸

<400> 3

ttctctttct tcccttctt tctcggcag ttgcggctt ttccccgta agctctaat 60
 cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt 120
 gattagggtg atggttcacg tagtgggcca tcgcccgat agacggtttt tgcctcttg 180

[0005]

acgttggagt ccacgttctt aatagtggac tcttgttcca aactggaaca acactcaacc	240
ctatctcggg etatfctttt gatttataag ggattttgcc gatttcggcc fatttggtaa	300
aaaatgagct gattfaacaa aaatttaacg cgaattttaa caaaafatta acgettacaa	360
ittaggtggc acttttcggg gaaatgtgcg cggaaacctt atttgtttat ttttctaaat	420
acattcaaat atgtatccgc tcatgagaca ataacctga taaatgett aataaatatg	480
aaaaaggaag agtafagta ttcaacattt cegtgtcgcc ettattccct tttttcggc	540
attttgectt cctgtttttg ctcaccaga aacgttggg aaagtaaaag atgctgaaga	600
tcagttgggt gcacgagtgg gttacatcga actggatctc aacagegga agatccttga	660
gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc tgctatgtg	720
cgcggtatta tcccgattt acgcccggca agagcaactc ggtcgcgcga tacactattc	780
tcagaatgac ttggttgggt actcaccagt cacagaaaag catcttacgg atggcatgac	840
agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact	900
tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggata	960
tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg	1020
tgacaccacg atgctgttag caatggcaac aacgttgcgc aaactattaa ctggcgaact	1080
acttactcta gcttcccgc aacaattaat agactggatg gagcgggata aagtgtcagg	1140
accaattctg cgtcggccc ttccgcttg ctggtttatf gctgataaat ctggagccgg	1200
tgagcgtggg tctcgggta tcattgcagc actggggcca gatggttaag cctcccgtat	1260
cgtagttatc tacacgacgg ggagtcagc aactatggat gaacgaaata gacagatcgc	1320
tgagataggt gcctcactga ttaagcattg gtaactgca gaccaagttt actcatatat	1380
actttagatt gatttaaaac ttcattttta atttaaaagg atctagtgta agatectttt	1440
tgataatctc atgaccaaaa tcccttaacg tgagttttcg ttcactgag cgtcagaccc	1500
cgtagaaaag atcaaaggat ctctctgaga tctttttttt ctgcggttaa tctgetgtt	1560
gcaaacaaaa aaaccaccgc taccagcgtt ggtttgttg ccggatcaag agctaccaac	1620
tctttttccg aaggttaactg gcttcagcag agcgcagata ccaataactg ttcttctagt	1680
gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgectacat acctcctct	1740
gctaactctg ttaccagtgg ctgctgccag tggcgataag tctgtctta ccgggttgg	1800
ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac	1860
acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg	1920
agaaagcgc acgcttccc aagggagaaa ggcggacagg tatccggtaa gcggcagggt	1980
cggaacagga gagegcacga gggagcttc aggggaaac gectggtatc ttatagtcc	2040
tgctgggttt cgcacctct gactttagcg tcatlittg tgatgctcgt cagggggcg	2100
gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttcttggcct ttgtctggcc	2160
ttttgctcac atgtttttt ctgcgttate cctgattct gtggataacc gtattaccgc	2220
ctttgagtga gctgataacc ctgcgcgag ccgaacgacc gagegcagcg agtcagtgag	2280
cgaggaagcg gaagagcgc tgatgcggtt ttttctctt acgeatctgt ggggtattc	2340
acaccgcata gaccagcgc gtaacctggc aaaatcggtt acggttgagt aataaatgga	2400
tgccctcgtt aagcgggtgt gggcgacaa taaagtctta aactgaacaa aatagateta	2460

[0006]

aactatgaca ataaagtctt aaactagaca gaatagttgt aaactgaaat cagfccagtt	2520
atgctgtgaa aaagcatact ggacttttgt tatggetaaa gcaaactctt cattttctga	2580
agtgcaaatf gcccgtcgtt ttaaagaggg gcgtggccaa gggcatggta aagactatat	2640
tcgcggcgtt gtgacaattt accgaacaac tcgcggcggc ggaagecgat ctgcgcttga	2700
acgaattgtt aggtggcggg acttgggtcg atatcaaagt gcatcacttc ttcccgatg	2760
cccaactttg tatagagagc cactgcggga tcgtcaccgt aatctgcttg caogtagatc	2820
acataagcac caagcgcgtt ggccctcatgc ttgaggagat tgatgagcgc ggtggcaatg	2880
ccctgcctcc ggtgctcgcg ggagactcgc agatcataga tatagatctc actacgcggc	2940
tgctcaaaact tgggcagAAC gtaagccgcg agagcgcCAA eaaccgcttc ttgttgaag	3000
gcagcaagcg cgatgaatgt ctactacggg agcaagtcc cgaggtaatc ggagtccggc	3060
tgatgttggg agtaggtggc tacgtctccg aactcaagac cgaagagatc aagagcagcc	3120
cgcatggatt tgacttggtc agggccgagc ctacatgtgc gaatgatgcc catacttgag	3180
ccacctaaact ttgttttagg gcgactgccc tctctcgtaa catcgttctt gctgcgtaac	3240
atcgtttctg ctccataaca tcaaacatcg acccaagcgg taacgcgctt gctgcttga	3300
tgcccagggc atagactgta caaaaaaca gtcataaca gccatgaaaa ccgccaactgc	3360
gcgcttaccA ccgctgcgtt cggtaagggt tctggaccag ttgctgagc gcataceta	3420
cttgcattac agtttacgaa ccgaacagc ttatgtcaac tgggttcgtg ccttcatccg	3480
tttcaagggt gtgcctcacc cggcaacctt gggcagcagc gaagtcaggc cttttctgc	3540
ctgcttggcg aacgagcgcA aggtttcggc ctccacgcac cgtcaggcat tggcggcctt	3600
gctgtttctc tacggcaagg tctgtgcac ggatctgccc tggttccagg agatccgaa	3660
acctcggcgg tcgcggcgtt tgccgggtgt gctgacccc gatgaagtgg ttgcctact	3720
cggttttctg gaaggcagc atcgtttgtt cgcaccagac tctagctata gttctagtgg	3780
ttgctacagc cttgcatgcc tgcaggcagc tgcgcgctcg ctgctcact gaggcgcgcc	3840
ggcgtcggg cgacctttgg tcgcccgcc tcagtgagcg agcagcgcg cagagaggga	3900
gtggccaact ccatcactag gggttcctg tagttaatga ttaaccgcgc atgctactta	3960
tctacgtagc catgctctag tgaattcga cgcgcacatc ctaggeccgc gccgceccc	4020
tcgcacagac ttgtgggaga agctcggcta ctcccctgcc ccggttaatt tgcataata	4080
atttctagt aactatagag gcttaatgtg cgataaaaga cagataatct gttctttta	4140
atactagcta cattttcat gatagcttg gatttctata agagatacaa atactaaatt	4200
attattttaa aaaacagcac aaaaggaaac tcaccctaac tgtaaagtaa ttgtgtgtt	4260
tgagactata aatatccctt ggagaaaage cttgtttgcg tttagtgaac cgtcagatgg	4320
taccgtttaa actcgagtga gcgatgctgg ctgcgaiggt cgatactgta aagccacaga	4380
tgggtgtcga ccattcagc cagcaccgcc tactagagcg gccgccacag cggggagatc	4440
cagacatgat aagatacatt ttttgaatte ggctatccc aggttgcctt ggttcatggc	4500
aatgggacg ttaagaggc agagagaata tgaacagaaa ctgttctaatt attgctcatt	4560
taatgtgtaa gtattgtct tttttaaacc tcttctattt ttttccagg aattgctgga	4620
cacagtggct tgggtgtgt ctgaggactg taggcatgg ccctaggttg tggttttagg	4680
tctcagtgcc tcttctggc tgtctctctg ctctctccc atgtctctt ctttgttcc	4740

[0007]

```

agccatttct cccctatgct taagtttggg gcagcagggt ttggctgctc tcagattcct 4800
gcttccctcag atgctgtagt tgcaggecc agcgggctgg cagcgggatc aggatctggc 4860
taggtttgct ctcactgtgg cagagtaggg ggagcgggag gagagcacgt gtgacccag 4920
gccagctgta gggagcatag gcatggtcac gtagccttca ggtccctagac ttigtcttct 4980
catgagtatg gctgtgtgtg taigtgaaa actaggttct acttagccca agaaaaatggg 5040
cacattttgc atgtggttcc ttagagaaa tgcactgggt atctgacata gcctggcagc 5100
atgctccctc caggtagggt agtctcaggc ggtgaagcac gtgtgtccag caagaacttc 5160
atatgtggca taaagtctcc gttctgtgag gtgctggcaa atcaccacca ccgtcaagag 5220
gctgaagtga tttttgtcta gggaggcagg aaaggcttcc tggagtcagc agccagttag 5280
tgaagagta gattggagac cttcttaate atcaccgctt cttgtctcaa ggggtgccag 5340
gaagctgtgg aggctgaacc catcttatgc tgcagagag tgggacacca tgagggtcag 5400
gtcaagggtt tgtaccttgt ttggtagaga attaggggct cttgaagact ttgatgtgg 5460
tcaggggagt gtatcattta ggaagagtga cccggtgagg acgtggggtg gaggaggaca 5520
ggtgggaggg agtccagggt ggagttagta gaccagcag gactgcaggg cctcagacca 5580
ggatggtggc agggctgtga ggagaggcag ccacctgtgt gtctgcggaa gcaggggcaa 5640
gagggagag gccagcagc tctctccate acccagcagc tggcgtagat tgtgagagac 5700
cattccctgc tcttaggagg gctgagttt tagttttctc ttgttataca ataagcttgg 5760
tatttgttta caaaacattt gtaaagctaa atcaaggttt gataaggctt ctagttttat 5820
ttaagaagta atgttgaat aatgtttgt ccaattcget ttgctcattt aaggactttc 5880
agtacaaact gcaacaacag gattaggatt taaacgttcc tgagatgttt ttactcctca 5940
gaatttccca gaatgtgac tggttttgat ttcaagctt gctgaeccaa taggttaacc 6000
cacaagtttt acgaagacca tctcagtcca ctacatcaa ctgcccatgc cacggttaaa 6060
gagatcatcg actgatgttt ggcacagett cctccctctt ggggtgggcaa gcattttgaa 6120
gagaaggctc ctatgggtga gactgggcca ccaaagcttt cctgtccca tcccctagct 6180
tgagaagccc ttctctaatg tggactttgt gccgttagca tctgtactag cttgaagtgt 6240
accatctgga cgtactttct gttttagcct cacaagttag caaggagggt tgagagatgt 6300
gctgtgagga atgtggggcc ccagctggca gcaggetctg ggtcaggggg gcagggacca 6360
cgggcatacc tgacagttag gagggtctag taggggatca gttccctgtt tttcttttag 6420
aattttctgg atattcttct ttattgattt tgggatgtga acaatagaat caacttctac 6480
ttgtagattg atttaggag aacttatacc tcagatgtta agtcacctg tccagaatgt 6540
gggatgtttt cctatttgtt cagaactttt taaattacct cagaagcaca tgaattttaa 6600
aggattttaa aaaaaactta aagatttttt cacatagctc ttgcacattt cttagataaa 6660
gaatectcag gtattcctct gtttttgta ctatagttt cttcttatgg gtttttttc 6720
cctgaaaaat cttttatcaa acgtatgtgg cttattttct gaaggatgtt tgataatttt 6780
ggaagatatg aaagtcttca tttttacaa gttttgaggt ctcttfaagc tgcattgttc 6840
tcatgtcagc tcccaaagca gaagacggca tgttgaaaaa tgcctgtagag aagatacttc 6900
ttttccacct gttttcaact catatcatct tgaatttcag ggcaccttcc catgctctca 6960
gtgettgtca tctgtttatt attttcttc ctgaatacce tgaactccag catgttctgc 7020

```

[0008]

tgtaattctg gectccctgg catcttggac tectgtttcc ttgtctctgt catccccgcg	7080
gtcagctcct gctgcgcagc ttctcagctg aagtgcggtt ggagtgccctg gcgtgtcttg	7140
ctggatcttt gagtaattgcc tcitgtttcc ttggttccct ctgctgagtt gctcagcgtc	7200
tccactcccc atttcttggg tggcccttcc tgcactctcc tgattccctt tgtcttccct	7260
ggtttcttgc ttgtgtttcg agtctccaca gaacttttgc agctcttctg aagacctgga	7320
agctttttca tcttaattct catctcatga ectcttttcc cttctttgag agetagaact	7380
tcccattgtg aacttctctt tccagaatcc catgcttctt ttccctcccc acttaacctg	7440
tgtccaggag aggtcagatt gctgtgcata ttggaggaga accctttctt ccttgggctc	7500
ttcatctcac atgaacacac cacatcacct cgttcccttg accctcagtg gtgtcactgc	7560
tggatttttc ttctctttgg ctggccttag ggcacacca ggttgactag cgtagtcatg	7620
gtatttagat ccactcacaat tticagtttc tgtgtctgtc tcttgccctg ttctgacttc	7680
gcccagagaa agcttctctt tcacaagggt tcttagattt atgttctctg agcaacctct	7740
ttcttgaggc agtgttttac caatatttat ttccctagtc agtctgcctt tacctttctt	7800
gttaigtcatg tctttgttcc tgaccacatc tctgagctctg taaaatagaa ttgtctgata	7860
atttaattac atgaaatcct ttagaatctt aacacatctt acacctgatt taatatttta	7920
tigtatecaa attgaaccaa cccctatgtga atttgacagt gatttctccc agggatccta	7980
gtgtataagg aataggactt agtattttct attttttgat ataccacata ccagatactg	8040
attatgatgg acatttaacc cttttttctc attatgaaag aaagtttaga attatttctt	8100
ccagtagcgc cagtgtaacg tgaagccctt tgaagagta gtttttgtat agctatctga	8160
aaggaatttc ttcccaaat atttttccag tgcctgacaac aaacacgcag acacacctg	8220
caagtgtagt gtacggcgca cttagagcatg gctacgtaga taagtagcat ggccgggttaa	8280
tcattaacta caaggaaccc ctagtgatgg agttggccac tccctctctg cgcgctcgt	8340
cgctcactga ggccgggcga ccaaaggctc cccgacgccc gggctttgcc cggggcgct	8400
cagtgagcga gcgagcgcgc agctgcttgc aggtctgaga caataacctt gataaatgct	8460
tcaataatgt aagcttctgc agaagtacta gaggatcata atcagccata ccacattgt	8520
agaggtttta ctgtctttta aaaaectccc acacctcccc ctgaacctga aacataaaat	8580
gaatgcaatt gtgtgtgtta actgttttat tgcagcttat aatgggtaca aataaagcaa	8640
tagcateaca aatttcacaa ataaagcatt ttttcaactg cttcttagtt gtgtttgtc	8700
caactcactc aatgtatctt atcatgtctg gatctgatea ctgatctgc ctaggagatc	8760
cgaaccagat aagtgaatc tagttccaaa ctattttgtc atttttaatt ttctatttag	8820
cttacgacgc tacacccagt tccatctat ttgtcactc ttccctaaat aatecttaaa	8880
aactccattt ecaacctctc cagtcccaa ctattttgtc cgeccacagc ggggcatttt	8940
tcttctgtti atgtttttaa tcaaacatcc tgccaactcc atgtgacaaa ccgtcactt	9000
cgctactttt ttctctgtca cagaatgaaa atttttctgt catctcttctg ttattaatgt	9060
ttgtaattga ctgaatatac acgttattt gcagcctgaa tggcgaatgg	9110

<210> 4
 <211> 86
 <212> DNA
 <213> 人工序列

[0009]

<220>		
<223>	人工序列描述：合成的寡核苷酸	
<400>	4	
	ctcagatgag cgatgctgge tegcatggtc gatacigttaa agccacagat ggggtgtegac	60
	catgagagcc agcaaccgct actaga	86
<210>	5	
<211>	311	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：合成的多核苷酸	
<400>	5	
	cgacgccc gcc atctctaggc cgcgcccgc cccctcgcac agacttgtgg gagaagctcg	60
	gctactcccc tgccccggtt aatttgata taataattcc tagtaactat agaggcttaa	120
	tgtgcgataa aagacagata atcigtctt ttaatacta gctacatctt acatgatagg	180
	cttggatttc tataagagat acaatacta aattattatt ttaaaaaaca gcacaaaagg	240
	aaactcacc taactgtaa gtaattgtgt gttttgagac tataaatatc ccttgagaa	300
	aagccttggt t	311
<210>	6	
<211>	94	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：合成的寡核苷酸	
<400>	6	
	ctgcgctc gctcgtcac tgaggccgc cggcgctgg gcgaccttg gtcgcccgc	60
	ctcagtgagc gagcgagcg gcagagagg agtg	94
<210>	7	
<211>	128	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：合成的多核苷酸	
<400>	7	
	aaggaacccc tagtgatgga gttggcaact cctctctg gcgctcgtc gctcactgag	60
	gccggcgac caaaggctgc cgcagccc ggctttgcc ggcgccctc agtgagcgag	120
	cgagcgcg	128
<210>	8	
<211>	534	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：合成的多核苷酸	
<400>	8	
	ttagggtggc gtaactgggt cgatacaaa gtgcatcaact tcttccgta tgcccaactt	60
	tgtatagaga gccactgcgg gatcgtcac gtaatctgct tgcaactaga teacataagc	120
	accaagcgcg ttggcctcat gcttggagag attgatgagc gcggtggcaa tgccctgctt	180

[0010]

ccggtgctcg ccggagactg cgagatcata gatatagata tcactacgcg gctgctcaaa 240
 cttgggcaga acgtaagccg cgagagcgcc aacaaccgct tcttggtcga aggcagcaag 300
 cgcgatgaat gcttactac ggagcaagtt cccgaggtaa tcggagtcg gctgatgtg 360
 ggagtaggtg gctacgtctc gaaactcacg accgaaaaga tcaagagcag cccgcattga 420
 tttgacttgg tcagggccga gctacatgt gegaatgatg cccatacttg agccacctaa 480
 ctttgtttta gggcgactgc cctgctgcgt aacatcgttg ctgctgcgta acat 534

<210> 9
 <211> 861
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列描述: 合成的多核苷酸

<400> 9
 atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttctc 60
 gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca 120
 cgagtggtt acatcgaact ggaictcaac agcggtaaga tccitgagag ttttcgcccc 180
 gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc 240
 egiatftaag ccgggcaaga gcaactcggf egccgcatac actattctca gaatgacttg 300
 gttgagtact caccagtcac agaaaagcat ctacggatg gcatgacagt aagagaatfa 360
 tgcagtctcg ccataacat gagtgataac actgcggcca acttaactct gacaacgac 420
 ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgctt 480
 gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg 540
 cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcaactact tactctagct 600
 tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 660
 tcggcccttc cggetgctg gtttattgct gataaatctg gagccgtga gegtgggtct 720
 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 780
 acgacgggga gtcaggcaac tatggatgaa cgaatatagac agatcgtga gataggtgcc 840
 tcactgatta agcattggta a 861

<210> 10
 <211> 225
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列描述: 合成的多核苷酸

<400> 10
 tgtggcgga caataaagtc ttaaaactgaa caaaatagat ctaaactatg acaataaagt 60
 cttaaaactag acagaatagt tgtaaaactga aatcagtcga gttatgctgt gaaaaagcat 120
 actggacttt tgttatgget aaagcaaaact ettcattttc tgaagtgcaa attgcccgte 180
 gtattaaaga gggcggtgca caaggcatg gtaaagacta tattc 225

<210> 11
 <211> 166
 <212> DNA
 <213> 人工序列

[0011]

<220>						
<223>	人工序列描述: 合成的多核苷酸					
<400>	11					
aaccagataa	gtgaaafcta	gttccaaact	atfttgcab	ttttaatfff	cgtattaget	60
tacgacgeia	eacccagttc	ccatctatft	tgtaectctf	ccctaaataa	tccttaaaaa	120
ctccatttce	acccctccca	gttcccaact	atfttgcog	cccaca		166
<210>	12					
<211>	9111					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	人工序列描述: 合成的多核苷酸					
<400>	12					
ttcgctttct	tcccttccct	tctcgcacg	ttcgcccgt	ttccccgtca	agctctaaat	60
cggggcctcc	ctttagggtt	ccgatttagt	gctttacgce	acctcgacct	caaaaaactt	120
gattaggggtg	atggttcacg	tagtgggcca	tcgcccgat	agacggtttt	tcgccccttg	180
acgttggagt	ccacgttctt	aatagtgga	tcctgttcca	aactggaaca	acactcaacc	240
ctatctcggg	ctatctcttt	gatttataag	ggattttgcc	gatttcggcc	tattggttaa	300
aaaafgaget	gatitaacaa	aaatftaacg	cgatftttaa	caaaatatta	acgtttacaa	360
tttaggtggc	acttttcggg	gaatgtgce	cggaacccct	atttgtttat	tttctaaat	420
acattcaaat	atgtatccgc	tcatgagaca	ataacctga	taaatgttc	aataatattg	480
aaaaaggaag	agtatgagta	ttcaacattt	ccgtgtcgc	cttattccct	tttttcggc	540
atfttgcctt	ctgtttttg	ctcacccaga	aacgcgggtg	aaagtaaaag	atgctgaaga	600
tcagttgggt	gcacgagtgg	gttacatcga	actggatctc	aacagcggta	agatccttga	660
gagttttcgc	cccgaagaac	gttttccaat	gatgagcact	tttaaagttc	tgctatgtgg	720
cgcggtatta	tccegtattg	acgcccggca	agageaactc	ggtcgcccga	tacaactatc	780
tcagaatgac	ttggttgagt	actcaaccgt	cacagaaaag	catcttacgg	atggcatgac	840
agtaagagaa	ttatgcagtg	ctgcataac	catgagtgat	aactctcgg	ccaacttact	900
tctgacaacg	atcggaggac	cgaaggagct	aaccgctttt	ttgcacaaca	tgggggatca	960
tgtaactcgc	cttgatcgtt	gggaaccgga	gctgaatgaa	gccataccaa	acgacgagcg	1020
tgacaccacg	atgcctgtag	caatggcaac	aacgttgcgc	aaactaftaa	ctggcgaact	1080
acttactcta	gcttcccggc	aacaattaat	agactggatg	gagcgggata	aagttgcagg	1140
accaettctg	cgctcggccc	ttccggctgg	ctggtttatt	gctgataaat	ctggagccgg	1200
tgagegtggg	tctcgcggta	tcattgcage	actggggcca	gatggtaagc	ctctccgtat	1260
cglagtllac	tacacgacgg	ggagtcagge	aactatggat	gaacgaaata	gacagatcgc	1320
tgagataggt	gcctcactga	ttaagcattg	giaactgtca	gaccaagttt	actcatatat	1380
actttagatt	gatttanaac	ttcattttta	atttanaagg	atctaggtga	agatcctttt	1440
tgataatctc	atgaccaaaa	tccttaacg	tgagttttcg	ttccactgag	cgicagaccc	1500
cgtagaaaaag	atcaaaaggat	ctctctgaga	tccttttttt	ctgcgcgtaa	tctgctgctt	1560
gcaaacaaaa	aaaccacgcg	taccagcggg	ggtttgtttg	ccgatcaag	agctaaccaac	1620
tctttttccg	aaggtaactg	gcttcagcag	agcgcagata	ccaataactg	ttcttctagt	1680

[0012]

gtagccgtag	ttagccace	acttcaagaa	cictgtagca	ccgcctacat	acctcgctct	1740
gctaatectg	ttaccagtgg	ctgctgccag	tggcgataag	tcgtgtotta	ccgggttggg	1800
ctcaagacga	tagttaccgg	ataaggcgca	gcggtcgggc	tgaaccgggg	gttcgtgcac	1860
acagcccagc	ttggagcgaa	cgacctacac	cgaactgaga	tacctacagc	gtgagctatg	1920
agaaaagccc	acgcttcccg	aaggagagaa	ggcggacagg	tatccggtaa	gcggcagggt	1980
cggaacagga	gagcgcacga	gggagcttcc	agggggaaac	gectggatc	tttatagtcc	2040
tgtcgggttt	cgccacctct	gacttgagcg	tcgatitttg	tgatgctcgt	cagggggggc	2100
gagcctatgg	aaaaacgcca	gcaacgcggc	ctttttacgg	ttcctgcect	tttgctggcc	2160
ttttgctcac	atgttcttct	ctgcgttate	ccctgattct	gtggataacc	gtattaccgc	2220
ctttgagtga	gctgataccg	ctcgcgcag	ccgaacgacc	gagcgcagcg	agtcagtggg	2280
cgaggaagcg	gaagagcgcc	tgatgcggta	ttttctctct	acgcatctgt	gcggtatttc	2340
acaccgcata	gaccagccgc	gtaacctggc	aaaatcgggt	acggttgagt	aataaatgga	2400
tgcctcgcgt	aagcgggtgt	ggcggacaa	taaagtotta	aactgaacaa	aatagatcta	2460
aaactatgaca	ataaagtctt	aaactagaca	gaatagtigt	aaactgaaat	cagtcagttt	2520
atgctgtgaa	aaagcatact	ggacttttgt	tatggctaaa	gaaaactctt	cattttctga	2580
agtgcanaatt	gcccgtcgtg	ttaaagaggg	gcgtggccaa	gggeatggta	aagactatct	2640
tcgcggcggt	gtgacaaftt	accgaacaac	tccgcggccg	ggaagccgat	ctcgcttga	2700
acgaattgtt	aggtggcggt	acttgggtcg	atatacaagt	gcactacttc	ttcccgtatg	2760
cccaactttg	tatagagagc	cactgcggga	tcgtcaccgt	aactgtcttg	cacgtagatc	2820
acataageac	caagcgcggt	ggcctcatgc	ttgaggagat	tgatgagcgc	ggtggcaatg	2880
ccctgectcc	ggtgctcgcc	ggagactgcg	agatcataga	tatagatctc	actacgcggc	2940
tgctcaaaact	tgggcagaa	gtaagccgcg	agagcggcaa	caaccgcttc	ttggtcgaag	3000
gcagcaagcg	cgatgaatgt	cttaactaegg	agcaagtccc	cgaggtaatc	ggagtcgggc	3060
tgatgttggg	agtaggtggc	tacgtctccg	aaactcagac	cgaaaagatc	aagagcagcc	3120
cgcattgatt	tgacttggtc	agggccgagc	ctacatgtgc	gaatgatgcc	catacttggg	3180
ccacctaaact	ttgttttagg	gcgactgccc	tgetgcgtaa	cctcgttget	gctgcgtaac	3240
atcgttctct	ctccataaca	tcaaacatcg	acccaaggcg	taaccgcttt	gctccttggg	3300
tgcccgaggc	atagactgta	caaaaaaaca	gtcataacaa	gcaatgaaaa	ccgccaactgc	3360
gccgttacca	ccgctgctgt	cggtcaaggt	tctggaccag	ttgctgagc	gcatacgeta	3420
cctgcattac	agtttacgaa	ccgaacaggc	ttatgtcaac	tgggttcgtg	cttctatccg	3480
tttcaacggg	gtgctgctcc	cggaacactt	gggcagcagc	gaagtcgagg	catctctgtc	3540
ctggctggcg	aacgagcgca	aggtttcggg	ctccacgcat	cgctaggeat	tggcggcttt	3600
gctgttcttc	tacggcaagg	tgctgtgcac	ggatctgccc	tggcttcagg	agatcggaag	3660
acctcggccg	tcgcggcget	tcccgttggg	gctgaccccg	gatgaagtgg	ttcgcactct	3720
cgtttttctg	gaagcgcagc	atcgtttgtt	cggccaggac	tctagctata	gttctagtgg	3780
ttgctacag	cttgcattgc	tgcagcagc	tgcgcctctg	ctcctctact	gagccgccc	3840
ggcgcctggg	cgaactttgg	tcgcccggcc	tcagtgagcg	agcagcgcg	cagagagggg	3900
gtggccaact	ccactactag	gggttctctg	tagttaatga	ttaacccgcc	atgctactta	3960

[0013]

tctacgtagc catgctctag tgaattcgac gccgccatct ctaggcccgc gccggccccc 4020
 tccacacagac ttgtgggaga agctcggcta ctcccctgcc ccggttaatt tgcataataat 4080
 atttcttagt aactatagag gcctaatgtg cgataaaaaga cagataatct gttcttttta 4140
 ataetageca cattttacat gataggcttg gatttctata agagatacaa atactaaatt 4200
 attattttta aaaacagcac aaaaggaaa tcaccetaac tgtaaagtaa ttgtgtgttt 4260
 tgagactata aatatccctt ggagaaaagc cttgtttgcg tttagtgaac cgtcagatgg 4320
 taccgtttta actcagatga gcgatgctgg ctcgcatggt cgatactgta aagccacaga 4380
 tgggtgtcga ccatgcgagc cagcaccgcc tactagagcg gccgccacag cggggagatc 4440
 cagacatgat aagatacatt ttttgaattc gggctatccc aggttgcctt ggttcatggc 4500
 aaatgggacg ttaagagggc agagagaata tgaacagaaa ctgtttctaat attggtcatt 4560
 taatgtgtaa gtattgttct tttttaaacc tcttctcattt tttttccagg aattgctgga 4620
 cacagtggct tgggtgtgtg ctgaggactg taggcatgg ccctaggttg tggttttagg 4680
 tctcaggtgc tcttctgagc tcttctcttg cttcttctcc atgtctctct ctttgtttcc 4740
 agccatttct cctttatget taagtttggt gcagcagggt ttggtctctc tcagattcct 4800
 gcttctcag atgtcttagt tctcaggccc agcgggctgg cagcgggac aggatctggc 4860
 taggtttgct ctcactgtgg cagagtaggg ggaggcgtgg gagagcacgt gtgaccccag 4920
 gccagctgta gggagcatag gcatggtcac gtagecttca ggtcttagac tttgtctct 4980
 catgagtatg gctgtgtgtg tatggtgaaa actaggttct acttagccca agaaaatggg 5040
 cacattttgc atgtggttgc tgtagagaaa tgcactgggt atctgacata gccctggcagc 5100
 atgcctccct caggtaggtt agtctcaggc ggtgaagcac gtgtgtccag caagaacttc 5160
 atatgtggca taaagctcc gttctgtgag gtgtgtgcaa atcaccacca ccgtcaagag 5220
 gctgaagtga tttttgtcta gggaggcagg aaagcctcc tggagtcagc agccagtagg 5280
 tgaagagta gattggagac ctctttaatc atcaccgctt cttgtctcaa ggggtgcccag 5340
 gaagctgtgg aggcctgaacc catcttatgc tgcagagag tgggacacca tgagggtcag 5400
 gtcaagggtt tgtacctgtt ttggtagaga attagggctt cttgaagact ttggaatggt 5460
 tcaggggagt gtatcattta ggaagagtga cccggtgagg acgtggggta gaggaggaca 5520
 ggtgggaggg agtccagglg ggagtgagta gaccagcag gagtgcaggg cctcagacca 5580
 ggatggtggc agggctgtga ggagaggcag ccacctgtgt gtctcggaa gcaggggcaa 5640
 gagggaagag gccagcagcg tgcctccatc acccagcagc tggcgtagat tgtgagagac 5700
 cattccctgc tcttaggagg ggctgagttt tagtttctc ttgttataca ataagcttgg 5760
 tatttgttta caaacatttt gtaaagctaa atcaaggttt gataaggctt ctagttttat 5820
 ttaagaagta atgttgaaat aatgtttgt ccaattcget ttgctcattt aaggacttcc 5880
 agtacaaact gcaacaacag gattaggatt taaacgttcc tgagatgttt ttactctca 5940
 gaatttccca gaatgtgate tggttttgat tttcaagctt gctgacccaa taggttaacc 6000
 cacaagtttt acgaagacca tctcagtcca cttacatcaa ctgccatgc cacggttaaa 6060
 gagatcctcg actgatgttt ggcacagctt cctccctctt ggggtggcaa gcatttggaa 6120
 gagaaggctc ctatgggtga gagtggggca ccaaagcttt cctgtccca tcccctagct 6180
 tgagaagccc ttctctaatg tggactttgt gccgttagca tctgtactag cttgaagttg 6240

[0014]

accatctgga cgtactttct ggttiagcct cacaagtgag caaggagggt tgagagatgt 6300
 gctgtgagga atgtggggcc ccagctggca gcaggetctg ggtcaggggg gcagggacca 6360
 cgggcataacc tgacagttag gaggggtcta gtaggggatc agttcccttg ttgttcttta 6420
 gaattttctg gatattcttc ttatttgatt ttgggatgtg aacaatagaa tcaacttcta 6480
 cttgtagatt gatttaggga gaactatac ctcatagtt aagtcacct gtccagaatg 6540
 tgggatgctt tcctatttgt tcagaacttt ttaaattacc tcagaagcac atgaaattta 6600
 aaggatttta aaaaaaactt aaagattatt tcacatagct cttgcacatt tcttgataaa 6660
 tgaatcctca ggtattcctc tgttttgtt actaatagtt acttcttatg ggittttttt 6720
 cccctgaaaa tcatttatca aacgtatgtg gcttatttcc tgaaggatgt ttgataattt 6780
 tggagatat gaaagtcttc atattttaca aggtttgggg tctctttaag ctgcatggtt 6840
 ctcatgtcag ctcccaaagc agaagacggc atgttgaaaa atgccgtaga gaagatactt 6900
 ctttccacc tgttttcaac tcatatcacc ttgaatttca gggcaccttt ccatgctctt 6960
 agtgcttctt atctgtttat tattttcctt cctgaatacc ctgaactcca gatgttctg 7020
 ctgtaattct ggctccttg gcactttgga ctctgttctt ctttgcctg tcactccccc 7080
 ggtcagctcc tgcctgcag cttctcagct gaagtgcctt tggagtgcct ggcgtgtctt 7140
 gctgatctt tgagtattgc ctctgttctt cttgttctt tetgtgagt tgcctcagct 7200
 ctccactccc cattcttctt gtggcccttc ctgaactctt ctgattcctt ttgtctccc 7260
 tggttttctg ctttggtttc gactctccac agaacttttg cagetctctt gaagacctgg 7320
 aagcttttct atcttaatto tcactcatg acccttttcc ccttcttga gagctagaac 7380
 ttcccatggt gaacttctct ttccagaatt ccatgecttc ttttccctcc cacttaacctg 7440
 ttgtccagga gaggtcagat tgctgtgcat attggaggag aaccctttct tccctgggct 7500
 ctctactca catgacatca ccacatcacc tcgttctctg gaccctcagt gggtcactg 7560
 ctggattttt ctttcttttg gctggcttca gggcacaccc aggtttagta gcttagtcat 7620
 ggtattttag tccactcaca ttctcagtt ctgtgtctgt ctcttgcctg cttctgaact 7680
 cccccagaga aagcttctct ttccaaagg ttcttagatt tatgttcact gagcaccttc 7740
 tttctgagg cagtgtttta ccaatatita ttttctagt cagtctcgc ttaaccttct 7800
 tgttatgcat gcttttggtc ctgaccatt ctctgagct gtaaataga attgctgtat 7860
 aatttaatta catgaaatcc tttagaatct taacacatct tacacctgat ttaattttt 7920
 attgtatcca aattgaacca acctatgtg aatttgacag tgatttctcc caggatctct 7980
 agtgataag gaataggact tagtatttct tattttttga tataccacat accagatact 8040
 gattatgatg gacatttaac cttttttctt cattatgaaa gaaagttagg aattatttct 8100
 tccagttagc ccagtgtaac ctgaaagcct ttgaaagagt agtttttga tagctatctg 8160
 aaaggaattt ctttccaaaa tatttttcca gtctgacaa caaacacgca gacacacct 8220
 gcaaggttag tgtacggcgc actagagcat ggctacgtag ataagtagca tggcgggtta 8280
 atcattaact acaaggaacc cctagtgatg gagttgcca ctccctctct gcgcctcgc 8340
 tctctcactg agccggggcg accaaaggtc gcccaagcc cgggccttgc ccggcgggcc 8400
 tcagttagcg agcgagcgc cagctgcctg caggcttagg acaataacce tgataaatgc 8460
 ttcaataatg taagcttctc gagaagtaact agaggatcat aatcagccat accacattg 8520

[0015]

tagaggtttt acttgcttta aaaaacctcc cacacctccc cctgaacctg aaacataaaa	8580
tgaatgcaat tgttggtgtt aacttgttta ttgcagetta taatggttac aaataaagca	8640
atagcateac aaatttcaca aataaagcat ttttticact gcattctagt tgtggtttgt	8700
ccaaactcat caatgtatct tateatgtct ggatcigate actgatacgc cctaggagat	8760
ccgaaccaga taagtgaat ctagttccaa actatittgt catttttaat tttegtatta	8820
gcttacgacg ctacaccag ttccatcta ttttgcact ctccctaaa taatcettaa	8880
aaaciccatt tccaccctc ccagttccca actatittgt ccgccacag cggggcattt	8940
ttcttctgt tatgtttta atcaaacatc ctgccaactc catgtgacaa accgtcatct	9000
teggetactt tttctctgtc acagaatgaa aatittctg tcactcttc gttattaatg	9060
tttgaattg actgaatc aacgettatt tgcagcctga atggcgaatg g	9111

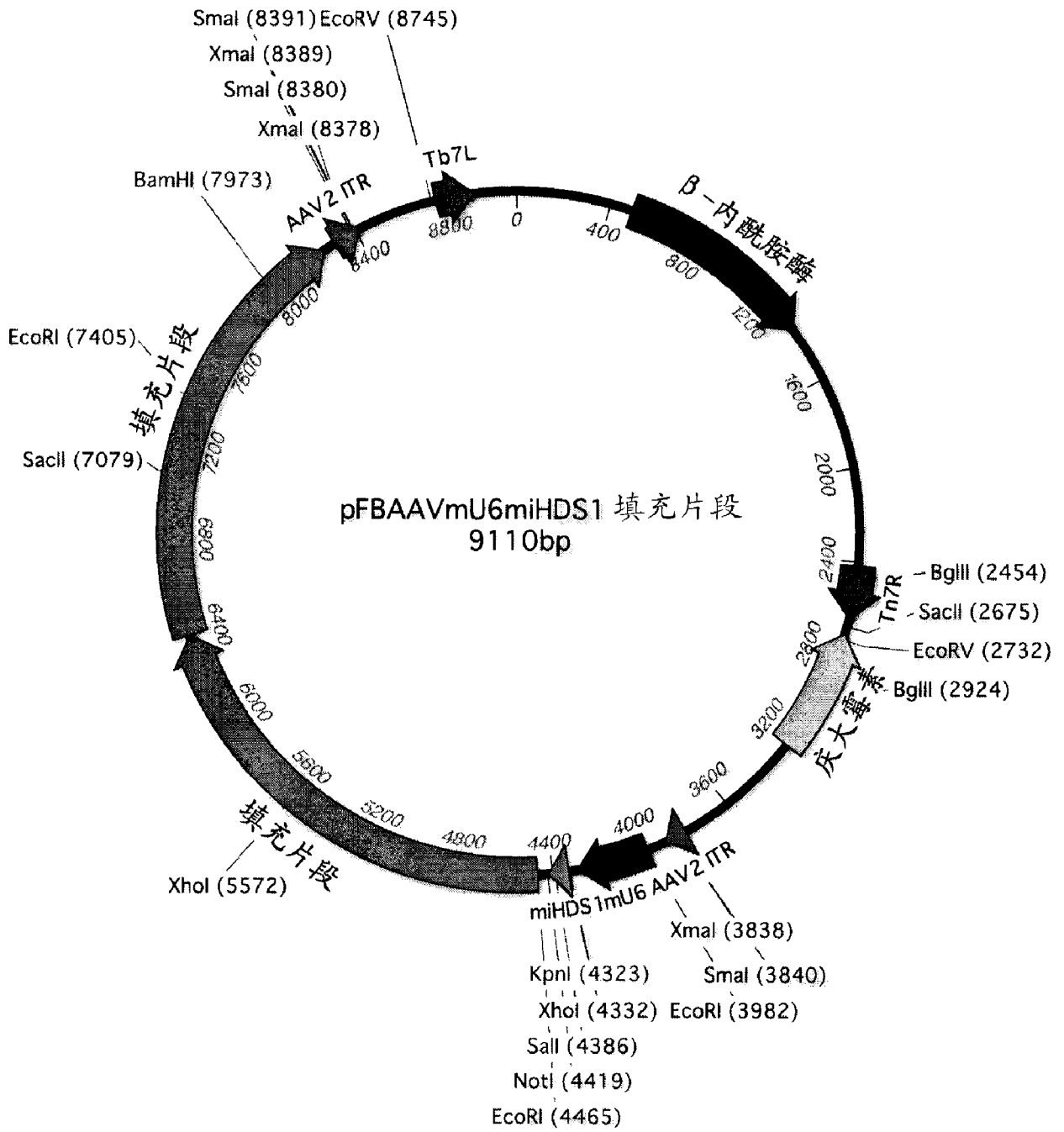


图 1

(SEQ ID NO:3)

序列: 5pFBAAVmU6miHDS1 填充片段装配范围: 1至9110

```

>5'_GTVG_G0202
|
      10      20      30      40      50
TTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCA
AAGCGAAAGAAGGGAAGGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCAGT

      60      70      80      90      100
AGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGGATTTAGTGCTTTACGGC
TCGAGATTTAGCCCCGAGGGAAATCCAAGGCTAAATCACGAAATGCCG

      110     120     130     140     150
ACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCA
TGGAGCTGGGGTTTTTTGAACTAATCCCACTACCAAGTGCATCACCCGGT

      160     170     180     190     200
TCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTT
AGCGGGACTATCTGCCAAAAGCGGGAAACTGCAACCTCAGGTGCAAGAA

      210     220     230     240     250
AATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
TTATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGAGCCA

      260     270     280     290     300
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAA
GATAAGAAAATAAATATTCCCTAAAACGGCTAAAGCCGGATAACCAATT

      310     320     330     340     350
AAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTA
TTTTACTCGACTAAATTGTTTTTAAATTTGCGCTTAAATTTGTTTTATAAT

      360     370     380     390     400
ACGCTTACAATTTAGGTGGCACTTTTTCGGGGAAATGTGCGCGGAACCCCT
TGCGAATGTTAAATCCACCGTGAAAAGCCCTTTACACGCGCCTTGGGGA

      410     420     430     440     450
ATTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA
TAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCGAGTACTCTGT

                                     > β-内酰胺酶
                                     |
      460     470     480     490     500
ATAACCCTGATAAATGCTTCAATAATATTGAAAAGGAAGAGTATGAGTA
TATTGGGACTATTTACGAAGTTATTATAACTTTTTTCCTTCTCATACTCAT

      510     520     530     540     550
TTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTTTCGGGCATTTTGCCTT
AAGTTGTAAAGGCACAGCGGGAATAAGGGAAAAACGCCGTAAAACGGAA

```

560 570 580 590 600
CCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGA
GGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTACGACTTCT

610 620 630 640 650
TCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTA
AGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTCGCCAT

660 670 680 690 700
AGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTTCCAATGATGAGCACT
TCTAGGAACTCTCAAAGCGGGCTTCTTGCAAAGGTTACTACTCGTGA

710 720 730 740 750
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AAATTTCAAGACGATACACCGCGCCATAATAGGGCATAACTGCGGCCCGT

760 770 780 790 800
AGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGT
TCTCGTTGAGCCAGCGGCGTATGTGATAAGAGTCTTACTGAACCAACTCA

810 820 830 840 850
ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TGAGTGGTCAGTGTCTTTTCGTAGAATGCCTACCGTACTGTCATTCTCTT

860 870 880 890 900
TTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACT
AATACGTACAGACGGTATTGGTACTCACTATTGTGACGCCGGTTGAATGA

910 920 930 940 950
TCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACA
AGACTGTTGCTAGCCTCCTGGCTTCTCTCGATTGGCGAAAAAACGTGTTGT

960 970 980 990 1000
TGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAA
ACCCCTAGTACATTGAGCGGAAC TAGCAACCCTTGGCCTCGACTTACTT

1010 1020 1030 1040 1050
GCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAAC
CGGTATGGTTTTGCTGCTCGCACTGTGGTGCTACGGACATCGTTACCGTTG

1060 1070 1080 1090 1100
AACGTTGCGCAAAC TATTA ACTGGCGAACTACTTACTCTAGCTTCCCGGC
TTGCAACGCGTTTTGATAATTGACCGCTTGATGAATGAGATCGAAGGGCCG

1110 1120 1130 1140 1150
AACAAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTG
TTGTTAATTATCTGACCTACCTCCGCCTATTTCAACGTCCTGGTGAAGAC

1160 1170 1180 1190 1200
CGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGG
GCGAGCCGGGAAGGCCGACCGACCAAATAACGACTATTTAGACCTCGGCC

1210 1220 1230 1240 1250
TGAGCGTGGGTCTCGCGGTATCATTCAGCACTGGGGCCAGATGGTAAGC

ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCG
1260 1270 1280 1290 1300
CCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGAT
GGAGGGCATAGCATCAATAGATGTGCTGCCCTCAGTCCGTTGATACCTA
1310 1320 1330 1340 1350
GAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTG
CTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAAC
1360 1370 1380 1390 1400
GTAAGTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
CATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
1410 1420 1430 1440 1450
TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTC
AAGTAAAAATTAATTTTCTAGATCCACTTCTAGGAAAAACTATTAGAG
1460 1470 1480 1490 1500
ATGACCAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCC
TACTGGTTTTAGGGAATTGCACTCAAAGCAAGGTGACTCGCAGTCTGGG
1510 1520 1530 1540 1550
CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAA
GCATCTTTTCTAGTTTTCTAGAAGAACTCTAGGAAAAAAGACGCGCATT
1560 1570 1580 1590 1600
TCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTG
AGACGACGAACGTTTGTTTTTTTGGTGGCGATGGTCGCCACCAAACAAAC
1610 1620 1630 1640 1650
CCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAG
GGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACCGAAGTCGTC
1660 1670 1680 1690 1700
AGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACC
TCGCGTCTATGGTTTATGACAAGAAGATCACATCGGCATCAATCCGGTGG
1710 1720 1730 1740 1750
ACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTG
TGAAGTTCTTGAGACATCGTGGCGGATGTATGGAGCGAGACGATTAGGAC
1760 1770 1780 1790 1800
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGA
AATGGTCACCGACGACGGTCACCGCTATTCAGCACAGAATGGCCCAACCT
1810 1820 1830 1840 1850
CTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGG
GAGTTCTGCTATCAATGGCCTATTCCGCGTCGCCAGCCCGACTTGCCCCC
1860 1870 1880 1890 1900
GTTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGA
CAAGCACGTGTGTGCGGTGGAACCTCGCTTGCTGGATGTGGCTTGACTCT

1910 1920 1930 1940 1950
TACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAA
ATGGATGTGCGACTCGATACTCTTTCGCGGTGCGAAGGGCTTCCCTCTTT

1960 1970 1980 1990 2000
GGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACAGGAGAGCGCACGA
CCGCCTGTCCATAGGCCATTCGCCGTCCAGCCTTGTCTCTCGCGTGCT

2010 2020 2030 2040 2050
GGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGTTT
CCCTCGAAGGTCCCCCTTTCGCGACCATAGAAATATCAGGACAGCCCAA

2060 2070 2080 2090 2100
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCG
CGGTTGGAGACTGAACTCGCAGCTAAAAACACTACGAGCAGTCCCCCGC

2110 2120 2130 2140 2150
GAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCT
CTCGGATACCTTTTTGCGGTCGTTGCGCCGGAAAAATGCCAAGGACCGGA

2160 2170 2180 2190 2200
TTTGCTGGCCTTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCT
AAACGACCGGAAAAACGAGTGTACAAGAAAGGACGCAATAGGGGACTAAGA

2210 2220 2230 2240 2250
GTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAG
CACCTATTGGCATAATGGCGGAACTCACTCGACTATGGCGAGCGGCGTC

2260 2270 2280 2290 2300
CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCC
GGCTTGCTGGCTCGCGTCGCTCAGTCACTCGCTCCTTCGCCTTCTCGCGG

2310 2320 2330 2340 2350
TGATGCGGTATTTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATA
ACTACGCCATAAAAGAGGAATGCGTAGACACGCCATAAAGTGTGGCGTAT

2360 2370 2380 2390 2400
GACCAGCCGCGTAACTGGCAAATCGGTTACGGTTGAGTAATAAATGGA
CTGGTCCGCGCATTGGACCGTTTTAGCCAATGCCAACTCATTATTTACCT

>Tn7R_8-10-11
|

2410 2420 2430 2440 2450
TGCCCTGCGTAAGCGGGTGTGGGCGGACAATAAAGTCTTAAACTGAACAA
ACGGGACGCATTCGCCACACCCGCCTGTTATTTTCAAGATTTGACTTGTT

2460 2470 2480 2490 2500
AATAGATCTAAACTATGACAATAAAGTCTTAAACTAGACAGAATAGTTGT
TTATCTAGATTTGATACTGTTATTTTCAAGATTTGATCTGTCTTATCAACA

2510 2520 2530 2540 2550
AAACTGAAATCAGTCCAGTTATGCTGTGAAAAAGCATACTGGACTTTTGT
TTTACTTTAGTCAGGTCAATACGACACTTTTTTCGTATGACCTGAAAACA

2560 2570 2580 2590 2600
TATGGCTAAAGCAAACCTCTTCATTTTCTGAAGTGCAAATTGCCCGTCGTA
ATACCGATTTCTGTTTGAGAAGTAAAAGACTTCACGTTTAAACGGGCAGCAT

2610 2620 2630 2640 2650
TTAAAGAGGGGCGTGGCCAAGGGCATGGTAAAGACTATATTCGCGGCGTT
AATTTCTCCCCGCACCGTTCCCGTACCATTTCTGATATAAGCGCCGCAA

2660 2670 2680 2690 2700
GTGACAATTTACCGAACAACTCCGCGGCCGGGAAGCCGATCTCGGCTTGA
CACTGTTAAATGGCTTGTTGAGGCGCCGGCCCTTCGGCTAGAGCCGAACT

2710 2720 2730 2740 2750
ACGAATTGTTAGGTGGCGGTACTTGGGTTCGATATCAAAGTGCATCACTTC
TGCTTAAACAATCCACCGCCATGAACCCAGCTATAGTTTCACGTAGTGAAG

2760 2770 2780 2790 2800
TTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGT
AAGGGCATAACGGGTTGAAACATATCTCTCGGTGACGCCCTAGCAGTGGCA

2810 2820 2830 2840 2850
AATCTGCTTGACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGC
TTAGACGAACGTGCATCTAGTGTATTCTGTTGCGCAACCCGAGTACG

2860 2870 2880 2890 2900
TTGAGGAGATTGATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGTCTGCC
AACTCCTCTAACTACTCGCGCCACCGTTACGGGACGGAGGCCACGAGCGG

2910 2920 2930 2940 2950
GGAGACTGCGGAGATCATAGATATAGATCTCACTACGCGGCTGCTCAAAC
CCTCTGACGCTCTAGTATCTATATCTAGAGTGATGCGCCGACGAGTTTGA

2960 2970 2980 2990 3000
TGGGCAGAACGTAAGCCGCGAGAGCGCCAACAACCGCTTCTTGGTCAAG
ACCCGTCTTGCAATTCGGCGCTCTCGCGGTTGTTGGCGAAGAACCAGCTTC

3010 3020 3030 3040 3050
GCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGGTAATC
CGTCGTTGCGGCTACTTACAGAATGATGCCTCGTTCAAGGGCTCCATTAG

3060 3070 3080 3090 3100
GGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCTCCGAACTCACGAC
CCTCAGGCCGACTACAACCCTCATCCACCGATGCAGAGGCTTGAGTGCTG

3110 3120 3130 3140 3150
CGAAAAGATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGC
GCTTTTCTAGTTCTCGTCGGGCGTACCTAAACTGAACCAGTCCCGGCTCG

3160 3170 3180 3190 3200
CTACATGTGCGAATGATGCCATACTTGAGCCACCTAACTTTGTTTTAGG
GATGTACACGCTTACTACGGGTATGAACTCGGTGGATTGAAACAAAATCC

<庆大霉素

3210 3220 3230 3240 3250
GCGACTGCCCTGCTGCGTAACATCGTTGCTGCTGCGTAACATCGTTGCTG
CGCTGACGGGACGACGCATTGTAGCAACGACGACGCATTGTAGCAACGAC

3260 3270 3280 3290 3300
CTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGA
GAGGTATTGTAGTTTGTAGCTGGGTGCCGCATTGCGCGAACGACGAACCT

3310 3320 3330 3340 3350
TGCCCGAGGCATAGACTGTACAAAAAACAGTCATAACAAGCCATGAAAA
ACGGGCTCCGTATCTGACATGTTTTTTTGTGTCAGTATTGTTCCGTACTTTT

3360 3370 3380 3390 3400
CCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAG
GGCGGTGACGCGGCAATGGTGGCGACGCAAGCCAGTTC AAGACCTGGTC

3410 3420 3430 3440 3450
TTGCGTGAGCGCATAACGCTACTTGCATTACAGTTTACGAACCGAACAGGC
AACGCACTCGCGTATGCGATGAACGTAATGTCAAATGCTTGGCTTGTCCG

3460 3470 3480 3490 3500
TTATGTCAACTGGGTTTCGTGCCTTCATCCGTTTCCACGGTGTGCGTCACC
AATACAGTTGACCCAAGCACGGAAGTAGGCAAAGGTGCCACACGCAGTGG

3510 3520 3530 3540 3550
CGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCCTGGCTGGCG
GCCGTTGGAACCCGTCGTGCTTCAGCTCCGTAAAGACAGGACCGACCGC

3560 3570 3580 3590 3600
AACGAGCGCAAGGTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCTT
TTGCTCGCGTTCCAAAGCCAGAGGTGCGTAGCAGTCCGTAAACCGCCGGAA

3610 3620 3630 3640 3650
GCTGTTCTTCTACGGCAAGGTGCTGTGCACGGATCTGCCCTGGCTTCAGG
CGACAAGAAGATGCCGTTCCACGACACGTGCCTAGACGGGACCGAAGTCC

3660 3670 3680 3690 3700
AGATCGGAAGACCTCGGCCGTCGCGGCGCTTGCCGGTGGTGCTGACCCCG
TCTAGCCTTCTGGAGCCGGCAGCGCCGCGAACGGCCACCACGACTGGGGC

3710 3720 3730 3740 3750
GATGAAGTGGTTCGCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTTGT
CTACTTCACCAAGCGTAGGAGCCAAAAGACCTTCCGCTCGTAGCAAACAA

3760 3770 3780 3790 3800
CGCCCAGGACTCTAGCTATAGTTCTAGTGGTTGGCTACAGCTTGCAATGCC
GCGGGTCTGAGATCGATATCAAGATCACCAACCGATGTCGAACGTACGG

```

>AAV_ITR_(94bp)
      |
    3810      3820      3830      3840      3850
TGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
ACGTCGGTCGACGCGCGAGCGAGCGAGTGACTCCGGCGGGCCCGCAGCCC

    3860      3870      3880      3890      3900
CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA
GCTGGAAACCAGCGGGCCGGAGTCACTCGCTCGCTCGCGCGTCTCTCCCT

    3910      3920      3930      3940      3950
GTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTAACCCGCC
CACCGGTTGAGGTAGTGATCCCCAAGGAACATCAATTACTAATTGGGCGG

```

```

                                >mU6启动子
                                |
    3960      3970      3980      3990      4000
ATGCTACTTATCTACGTAGCCATGCTCTAGTGAATTCGACGCCGCCATCT
TACGATGAATAGATGCATCGGTACGAGATCACTTAAGCTGCGGCGGTAGA

    4010      4020      4030      4040      4050
CTAGGCCCGCGCCGGCCCCCTCGCACAGACTTGTGGGAGAAGCTCGGCTA
GATCCGGGCGCGGCCGGGGGAGCGTGTCTGAACACCCTCTTCGAGCCGAT

    4060      4070      4080      4090      4100
CTCCCCTGCCCCGGTTAATTTGCATATAATATTTCTAGTAACTATAGAG
GAGGGGACGGGGCCAATTAAACGTATATTATAAAGGATCATTGATATCTC

    4110      4120      4130      4140      4150
GCTTAATGTGCGATAAAAAGACAGATAATCTGTTCTTTTTAATACTAGCTA
CGAATTACACGCTATTTTCTGTCTATTAGACAAGAAAAATTATGATCGAT

    4160      4170      4180      4190      4200
CATTTTACATGATAGGCTTGGATTTCTATAAGAGATACAAATACTAAATT
GTAAAATGTACTATCCGAACCTAAAGATATTCTCTATGTTTATGATTTAA

    4210      4220      4230      4240      4250
ATTATTTTAAAAAACAGCACAAAAGGAAACTCACCTAACTGTAAAGTAA
TAATAAAATTTTTTGTGCGTGTTTTCCTTTGAGTGGGATTGACATTTTCATT

    4260      4270      4280      4290      4300
TTGTGTGTTTTGAGACTATAAATATCCCTTGGAGAAAAGCCTTGTGTTGCG
AACACACAAAACCTCTGATATTTATAGGGAACCTCTTTTCGGAACAAACGC

```

```

                                >miHDS1
                                |
    4310      4320      4330      4340      4350
TTTAGTGAACCGTCAGATGGTACCGTTTAAACTCGAGTGAGCGATGCTGG
AAATCACTTGGCAGTCTACCATGGCAAATTTGAGCTCACTCGCTACGACC

    4360      4370      4380      4390      4400
CTCGCATGGTTCGATACTGTAAAGCCACAGATGGGTGTCGACCATGCGAGC
GAGCGTACCAGCTATGACATTTCCGGTGTCTACCCACAGCTGGTACGCTCG

```

4410 4420 4430 4440 4450
 CAGCACCGCCTACTAGAGCGGCCGCCACAGCGGGGAGATCCAGACATGAT
 GTCGTGGCGGATGATCTCGCCGGCGGTGTCGCCCTCTAGGTCTGTACTA

>填充片段

4460 4470 4480 4490 4500
 AAGATACATTTTTTTGAATTCGGGCTATCCCAGGTTGCCTTGGTTTCATGGC
 TTCTATGTAAAAAACTTAAGCCCGATAGGGTCCAACGGAACCAAGTACCG

4510 4520 4530 4540 4550
 AAATGGGACGTTAAGAGGGCAGAGAGAATATGAACAGAACTGTTCTAAT
 TTTACCCTGCAATTCTCCCGTCTCTCTTATACTTGTCTTTGACAAGATTA

4560 4570 4580 4590 4600
 ATTGGTCATTTAATGTGTAAGTATTGTTCTTTTTTAAACCTCCTTCATTT
 TAACCAGTAAATTACACATTCATAACAAGAAAAAATTTGGAGGAAGTAAA

4610 4620 4630 4640 4650
 TTTTTCCAGGAATTGCTGGACACAGTGGCTTGGTGTGTGTCTGAGGACTG
 AAAAAGGTCCTTAACGACCTGTGTCACCGAACCAACACACAGACTCCTGAC

4660 4670 4680 4690 4700
 TAGGCCATGGCCCTAGGTTGTGGTTTTAGGTCTCAGGTGCTCTTCCTGGC
 ATCCGGTACCGGGATCCAACACCAAATCCAGAGTCCACGAGAAGGACCG

4710 4720 4730 4740 4750
 TGTCTCCTTGCTTCTTTCCCATGTCCTCTTCTTTGTTTCCAGCCATTTCT
 ACAGAGGAACGAAGAAAGGGTACAGGAGAAGAAACAAAGGTCGGTAAAGA

4760 4770 4780 4790 4800
 CCCTTATGCTTAAGTTTGGTGCAGCAGGGTTTGGCTGCTCTCAGATTCCCT
 GGAATACGAATTCAAACCACGTCGTCCCAAACCGACGAGAGTCTAAGGA

4810 4820 4830 4840 4850
 GCTTCCTCAGATGCTGTAGTTGTCAGGCCAGCGGGCTGGCAGCGGGATC
 CGAAGGAGTCTACGACATCAACAGTCCGGGTGCGCCGACCGTCCGCCCTAG

4860 4870 4880 4890 4900
 AGGATCTGGCTAGGTTTGTCTCTCACTGTGGCAGAGTAGGGGGAGGCGTGG
 TCCTAGACCGATCAAACGAGAGTGACACCGTCTCATCCCCCTCCGCACC

4910 4920 4930 4940 4950
 GAGAGCACGTGTGACCCAGGCCAGCTGTAGGGAGCATAGGCATGGTCAC
 CTCTCGTGCACTGGGGTCCGGTGCACATCCCTCGTATCCGTACCAGTG

4960 4970 4980 4990 5000
 GTAGCCTTCAGGTCCTAGACTTTGTCTTCTCATGAGTATGGCTGTGTGTG
 CATCGGAAGTCCAGGATCTGAAACAGAAGAGTACTCATAACCGACACACAC

5010 5020 5030 5040 5050
TATGGTGAAAAGCTAGGTTCTACTTAGCCCAAGAAAATGGGCACATTTTGC
ATACCACTTTTGATCCAAGATGAATCGGGTTCCTTTTACCCGTGTAAAACG

5060 5070 5080 5090 5100
ATGTGGTTTCTGTAGAGAAATGCACTGGGTATCTGACATAGCCTGGCAGC
TACACCAAAGACATCTCTTTACGTGACCCATAGACTGTATCGGACCGTCC

5110 5120 5130 5140 5150
ATGCCTCCCTCAGGTAGGTTAGTCTCAGGCGGTGAAGCACGTGTGTCCAG
TACGGAGGGAGTCCATCCAATCAGAGTCCGCCACTTCGTGCACACAGGTC

5160 5170 5180 5190 5200
CAAGAACTTCATATGTGGCATAAAGTCTCCGTTCTGTGAGGTGCTGGCAA
GTTCTTGAAGTATACACCGTATTTTCAGAGGCAAGACACTCCACGACCGTT

5210 5220 5230 5240 5250
ATCACCACCACCGTCAAGAGGCTGAAGTGATTTTTGTCTAGGGAGGCAGG
TAGTGGTGGTGGCAGTTCTCCGACTTCACTAAAAACAGATCCCTCCGTC

5260 5270 5280 5290 5300
AAAGGCTTCCTGGAGTCAGCAGCCAGTAGGTGAAAGAGTAGATTGGAGAC
TTTCCGAAGGACCTCAGTCGTCCGTCATCCACTTTCTCATCTAACCTCTG

5310 5320 5330 5340 5350
CTTCTTAATCATCACCGCCTCTTGTCTCAAGGGGTGCCAGGAAGCTGTGG
GAAGAATTAGTAGTGGCGGAGAACAGAGTTCCCCACGGTCCCTTCGACACC

5360 5370 5380 5390 5400
AGGCTGAACCCATCTTATGCTGCCAGAGAGTGGGACACCATGAGGGTCAG
TCCGACTTGGGTAGAATACGACGGTCTCTCACCCCTGTGGTACTCCCAGTC

5410 5420 5430 5440 5450
GTCAAGGGGTTGTACCTTGTTTGGTAGAGAATTAGGGGCTCTTGAAGACT
CAGTTCCCCAACATGGAACAAACCATCTCTTAATCCCCGAGAACTTCTGA

5460 5470 5480 5490 5500
TTGGATGTGGTCAGGGGAGTGTATCATTTAGGAAGAGTGACCCGGTGAGG
AACCTACACCAGTCCCCTCACATAGTAAATCCTTCTCACTGGGCCACTCC

5510 5520 5530 5540 5550
ACGTGGGGTAGAGGAGGACAGGTGGGAGGGAGTCCAGGTGGGAGTGAGTA
TGCACCCCATCTCCTCCTGTCCACCCTCCCTCAGGTCCACCCTCACTCAT

5560 5570 5580 5590 5600
GACCCAGCAGGAGTGCAGGGCCTCGAGCCAGGATGGTGGCAGGGCTGTGA
CTGGGTCGTCCCTCACGTCCCGGAGCTCGGTCCCTACCACCGTCCCGACACT

5610 5620 5630 5640 5650
GGAGAGGCAGCCACCTGTGTGTCTGCGGAAGCAGGGGCAAGAGGGAAGAG
CCTCTCCGTCGGTGGACACACAGACGCCTTCGTCCCCGTTCTCCCTTCTC

5660 5670 5680 5690 5700
GCCAGCAGCGTGCTGCCATCACCCAGCGACTGGCGTAGATTGTGAGAGAC
CGGTCGTCGCACGACGGTAGTGGGTCGCTGACCGCATCTAACACTCTCTG

5710 5720 5730 5740 5750
CATTCCCTGCTCTTAGGAGGGGCTGAGTTTTAGTTTTCTCTTGTTATACA
GTAAGGGACGAGAATCCTCCCCGACTCAAAATCAAAAGAGAACAATATGT

5760 5770 5780 5790 5800
ATAAGCTTGGTATTTGTTTACAAAACATTTGTAAAGCTAAATCAAGGTTT
TATTCGAACCATAAACAAATGTTTTGTAAACATTTTCGATTTAGTTCCAAA

5810 5820 5830 5840 5850
GATAAGGCTTCTAGTTTTATTTAAGAAGTAATGTTGAAATAAATGTTTGT
CTATTCCGAAGATCAAAATAAATTCCTCATTACAACCTTTATTTACAAACA

5860 5870 5880 5890 5900
CCAATTCGCTTTGCTCATTTAAGGACTTTCAGTACAAACTGCAACAACAG
GGTTAAGCGAAACGAGTAAATTCCTGAAAGTCATGTTTGACGTTGTTGTC

5910 5920 5930 5940 5950
GATTAGGATTTAAACGTTTCTGAGATGTTTTTACTCCTCAGAATTTCCCA
CTAATCCTAAATTTGCAAAGACTCTACAAAATGAGGAGTCTTAAAGGGT

5960 5970 5980 5990 6000
GAATGTGATCTGGTTTTGATTTTCAAGCTTGCTGACCCAATAGGTTAACC
CTTACACTAGACCAAACTAAAAGTTTCGAACGACTGGGTTATCCAATTGG

6010 6020 6030 6040 6050
CACAAGTTTTACGAAGACCATCTCAGTCCACTTACATCAACTGCCCATGC
GTGTTCAAAATGCTTCTGGTAGAGTCAGGTGAATGTAGTTGACGGGTACG

6060 6070 6080 6090 6100
CACGGTTAAAGAGATCATCGACTGATGTTTGGCACAGCTTCCTCCCTCTT
GTGCCAATTTCTCTAGTAGCTGACTACAAACCGTGTCGAAGGAGGGAGAA

6110 6120 6130 6140 6150
GGGTGGGCAAGCATTGGAAGAGAAGGCTCCTATGGGTGAGAGTGGGGCA
CCCACCGTTCGTAAACCTTCTCTTCCGAGGATACCACTCTCACCCCGT

6160 6170 6180 6190 6200
CCAAAGTCTTCCCTGTCCCATCCCCTAGCTTGAGAAGCCCTTCTCTAATG
GGTTTCAGAAGGGACAGGGTAGGGGATCGAACTCTTCGGGAAGAGATTAC

6210 6220 6230 6240 6250
TGGACTTTGTGCCGTTAGCATCGTTACTAGCTTGAAGTTGACCATCTGGA
ACCTGAAACACGGCAATCGTAGCAATGATCGAACTTCAACTGGTAGACCT

6260 6270 6280 6290 6300
CGTACTTTCTGGTTTAGCCTCACAAGTGAGCAAGGAGGGTTGAGAGATGT
GCATGAAAGACCAAATCGGAGTGTTCACTCGTTCCTCCCACTCTCTACA

6310 6320 6330 6340 6350
GCTGTGAGGAATGTGGGGCCCCAGCTGGCAGCAGGCTCTGGGTGAGGGG

CGACACTCCTTACACCCCGGGGTCGACCGTCGTCCGAGACCCAGTCCCCC

6360 6370 6380 6390 6400
GCAGGGACCACGGGCATACCTGACAGTGAGGAGGGTCTAGTAGGGGATCA
CGTCCCTGGTGCCCGTATGGACTGTCACTCCTCCCAGATCATCCCCTAGT

6410 6420 6430 6440 6450
GTTCCCCTGTTGTTCTTTAGAAATTTTCTGGATATTCTTCTTTATTGATTT
CAAGGGGACAACAAGAAATCTTAAAAGACCTATAAGAAGAAATAACTAAA

6460 6470 6480 6490 6500
TGGGATGTGAACAATAGAATCAACTTCTACTTGTAGATTGATTTAGGGAG
ACCCTACACTTGTTATCTTAGTTGAAGATGAACATCTAACTAAATCCCTC

6510 6520 6530 6540 6550
AACTTATACCTCAGATGTTAAGTCACCCTGTCCAGAATGTGGGATGCTTT
TTGAATATGGAGTCTACAATTCAGTGGGACAGGTCTTACACCCTACGAAA

6560 6570 6580 6590 6600
CCTATTTGTTTCAGAACTTTTAAATTACCTCAGAAGCACATGAAATTTAA
GGATAAACAAGTCTTGAAAAATTTAATGGAGTCTTCGTGTACTTTAAATT

6610 6620 6630 6640 6650
AGGATTTTAAAAAACTTAAAGATTATTTACATAGCTCTTGACATTT
TCCTAAAATTTTTTTTGAATTTCTAATAAAGTGTATCGAGAACGTGTA

6660 6670 6680 6690 6700
CTTGATAAATGAATCCTCAGGTATTCCTCTGTTTTTGTACTAATAGTTA
GAACTATTTACTTAGGAGTCCATAAGGAGACAAAAACAATGATTATCAAT

6710 6720 6730 6740 6750
CTTCTTATGGGTTTTTTTTCCCTGAAAATCATTTATCAAACGTATGTGG
GAAGAATACCCAAAAAAGGGGACTTTTTAGTAAATAGTTTGCATACACC

6760 6770 6780 6790 6800
CTTATTTTCTGAAGGATGTTTGATAATTTTGGAAGATATGAAAGTCTTCA
GAATAAAAGACTTCCTACAACTATTTAAAACCTTCTATACTTTTCAGAAGT

6810 6820 6830 6840 6850
TATTTTACAAGGTTTGAGGTCTCTTTAAGCTGCATGGTTCTCATGTCAGC
ATAAAATGTTCCAACTCCAGAGAAATTCGACGTACCAAGAGTACAGTCG

6860 6870 6880 6890 6900
TCCCAAAGCAGAAGACGGCATGTTGAAAAATGCCGTAGAGAAGATACTTC
AGGGTTTCGTCTTCTGCCGTACAACTTTTTACGGCATCTCTTCTATGAAG

6910 6920 6930 6940 6950
TTTTCCACCTGTTTTCAACTCATATCATCTTGAATTTTCAGGGCACCTTTC
AAAAGGTGGACAAAAGTTGAGTATAGTAGAACTTAAAGTCCCCTGGAAAG

6960 6970 6980 6990 7000

CATGCTCCTAGTGCTTGCTATCTGTTTATTATTTTCCTTCCTGAATACCC
GTACGAGGATCACGAACGATAGACAAATAATAAAAGGAAGGACTTATGGG

7010 7020 7030 7040 7050
TGAACTCCAGCATGTTCTGCTGTAATTCTGGCCTCCCTGGCATCTTGAC
ACTTGAGGTCGTACAAGACGACATTAAGACCGGAGGGACCGTAGAACCTG

7060 7070 7080 7090 7100
TCCTGTTTCCCTTTGCTCTGTATCCCCGCGGTCAGCTCCTGCTGCGCAGC
AGGACAAAGGAAACGAGACAGTAGGGGCGCCAGTCGAGGACGACGCGTCG

7110 7120 7130 7140 7150
TTCTCAGCTGAAGTGCCTTTGGAGTGCCTGGCGTGTCTTGCTGGATCTTT
AAGAGTCGACTTCACGCAAACCTCACGGACCGCACAGAACGACCTAGAAA

7160 7170 7180 7190 7200
GAGTATTGCCTCTGGTTTCCTTGTTTCCTTCTGCTGAGTTGCTCAGCGTC
CTCATAACGGAGACCAAAGGAACCAAGGAAGACGACTCAACGAGTCGCAG

7210 7220 7230 7240 7250
TCCACTCCCCATTTCTTGTGTGGCCCTTCCTGCACTCCTCTGATTCCTTT
AGGTGAGGGGTAAAGAACACACCGGGAAGGACGTGAGGAGACTAAGGAAA

7260 7270 7280 7290 7300
TGTCTTCCCTGGTTTCTTGCTTTGGTTTCGAGTCTCCACAGAACTTTTGC
ACAGAAGGGACCAAAGAACGAAACCAAAGCTCAGAGGTGTCTTGAAAACG

7310 7320 7330 7340 7350
AGCTCTTCTGAAGACCTGGAAGCTTTTTTCATCTTAATTCTCATCTCATGA
TCGAGAAGACTTCTGGACCTTCGAAAAGTAGAATTAAGAGTAGAGTACT

7360 7370 7380 7390 7400
CCTCTTTTCCCTTCTTTGAGAGCTAGAACTTCCCATGGTGAACCTTCTCTT
GGAGAAAAGGGAAGAACTCTCGATCTTGAAGGGTACCACTTGAAGAGAA

7410 7420 7430 7440 7450
TCCAGAATTCCATGCCTTCTTTCCCTCCCACTTACCTGTTGTCCAGGAG
AGGTCTTAAGGTACGGAAGAAAAGGGAGGGTGAATGGACAACAGGTCCTC

7460 7470 7480 7490 7500
AGGTCAGATTGCTGTGCATATTGGAGGAGAACCCTTTCTTCCCTGGGCTC
TCCAGTCTAACGACACGTATAACCTCCTCTTGGGAAAGAAGGGACCCGAG

7510 7520 7530 7540 7550
TTCATCTCACATGACATCACCACATCACCTCGTTCCTTGGACCCTCAGTG
AAGTAGAGTGTACTGTAGTGGTGTAGTGGAGCAAGGAACCTGGGAGTCAC

7560 7570 7580 7590 7600
GTGTCACTGCTGGATTTTTCTTTCCCTTTGGCTGGCCTTAGGGCACACCCA
CACAGTGACGACCTAAAAGAAAGGAAACCGACCGGAATCCCCTGTGGGT

7610 7620 7630 7640 7650
GGTTGACTAGCGTAGTCATGGTATTTAGATCCACTCACATTTTCAGTTTC
CCAACCTGATCGCATCAGTACCATAAATCTAGGTGAGTGTAAGTCAAAG

```

      7660      7670      7680      7690      7700
TGTGTCTGTCTCTTGCCTGCTTCTGACTTCGCCCAGAGAAAGCTTCTCTT
ACACAGACAGAGAACGGACGAAGACTGAAGCGGGTCTCTTTCGAAGAGAA

      7710      7720      7730      7740      7750
TCACAAGGGTTCTTAGATTTATGTTCACTGAGCACCTTCTTTTCTGAGGC
AGTGTTCCCAAGAATCTAAATACAAGTGAAGTCTCGTGAAGAAAAGACTCCG

      7760      7770      7780      7790      7800
AGTGTTTTACCAATATTTATTTTCCTAGTCAGTCTCGCCTTACCTTTCTT
TCACAAAATGGTTATAAATAAAAAGGATCAGTCAGAGCGGAATGGAAAGAA

      7810      7820      7830      7840      7850
GTTATGCATGTCTTTGGTCCTGACCCATTCTCTGAGTCTGTAAAATAGAA
CAATACGTACAGAAACCAGGACTGGGTAAGAGACTCAGACATTTTATCTT

      7860      7870      7880      7890      7900
TTGCTGTATAATTTAATTACATGAAATCCTTTAGAATCTTAACACATCTT
AACGACATATTAATTAATGTACTTTAGGAAATCTTAGAATTGTGTAGAA

      7910      7920      7930      7940      7950
ACACCTGATTTAATATTTTATTGTATCCAAATTGAACCAACCCTATGTGA
TGTGGACTAAATTATAAAATAACATAGGTTTAACTTGGTTGGGATACT

      7960      7970      7980      7990      8000
ATTTGACAGTGATTTCTCCAGGGATCCTAGTGTATAAGGAATAGGACTT
TAAACTGTCACTAAAGAGGGTCCCTAGGATCACATATTCCTTATCCTGAA

      8010      8020      8030      8040      8050
AGTATTTTCTATTTTTTTGATATAACCACATAACCAGATACTGATTATGATGG
TCATAAAAGATAAAAAAATATATGGTGTATGGTCTATGACTAATACTACC

      8060      8070      8080      8090      8100
ACATTTAACCCTTTTTTCTCATTATGAAAGAAAGTTAGGAATTATTTCTT
TGTAATTTGGGAAAAAAGAGTAATACTTTCTTTCAATCCTTAATAAAGAA

      8110      8120      8130      8140      8150
CCAGTAGCGCCAGTGTAACCTGAAAGCCTTTGAAAGAGTAGTTTTTTGTAT
GGTCATCGCGGTCACATTGGACTTTCGGAACTTTCTCATCAAAAACATA

      8160      8170      8180      8190      8200
AGCTATCTGAAAGGAATTTCTTTCCAAAATATTTTTCCAGTGCTGACAAC
TCGATAGACTTTCTTAAAGAAAGGTTTTATAAAAAGGTCACGACTGTTG

                                     >3'_GTVG_G0202
                                     |
      8210      8220      8230      8240      8250
AAACACGCAGACACACCCTGCAAGGTGAGTGTACGGCGCACTAGAGCATG
TTTTGTGCGTCTGTGTGGGACGTTCCAATCACATGCCGCGTGATCTCGTAC

                                     >AAV_ITR_(128bp)
                                     |
      8260      8270      8280      8290      8300

```

GCTACGTAGATAAGTAGCATGGCGGGTTAATCATTAACTACAAGGAACCC
CGATGCATCTATTCATCGTACCGCCCAATTAGTAATTGATGTTCTTGGG

8310 8320 8330 8340 8350
CTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGA
GATCACTACCTCAACCGGTGAGGGAGAGACGCGCGAGCGAGCGAGTGACT

8360 8370 8380 8390 8400
GGCCGGGCGACCAAAGGTCGCCCGACGCCGGGCTTTGCCCGGGCGGCCT
CCGGCCCGCTGGTTTCCAGCGGGCTGCGGGCCCGAAACGGGCCCGCGGA

8410 8420 8430 8440 8450
CAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGTCTGAGACAATAACCT
GTCACCTCGCTCGCTCGCGCGTTCGACGGACGTCCAGACTCTGTTATTGGGA

8460 8470 8480 8490 8500
GATAAATGCTTCAATAATGTAAGCTTGTTCGAGAAGTACTAGAGGATCATA
CTATTTACGAAGTTATTACATTCGAACAGCTCTTCATGATCTCCTAGTAT

8510 8520 8530 8540 8550
ATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAACCTCCC
TAGTCGGTATGGTGTAACATCTCCAAAATGAACGAAATTTTTTGGAGGG

8560 8570 8580 8590 8600
ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTA
TGTGGAGGGGGACTTGGACTTTGTATTTTACTTACGTTAACAACAACAAAT

8610 8620 8630 8640 8650
ACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACA
TGAACAAATAACGTCGAATATTACCAATGTTTATTTTCGTTATCGTAGTGT

8660 8670 8680 8690 8700
AATTTTCAAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTG
TTAAAGTGTTTATTTTCGTAAAAAAGTGACGTAAGATCAACACCAAACAG

8710 8720 8730 8740 8750
CAAACCTCATCAATGTATCTTATCATGTCTGGATCTGATCACTGATATCGC
GTTTGAGTAGTTACATAGAATAGTACAGACCTAGACTAGTACTATAGCG

>Tb7L

8760 | 8770 8780 8790 8800
CTAGGAGATCCGAACCAGATAAGTGAAATCTAGTTCCAAACTATTTTGTG
GATCCTCTAGGCTTGGTCTATTCACTTTAGATCAAGGTTTGATAAAACAG

8810 8820 8830 8840 8850
ATTTTTAATTTTCGTATTAGCTTACGACGCTACACCCAGTTCCCATCTAT
TAAAAATTAAAAGCATAATCGAATGCTGCGATGTGGGTCAAGGGTAGATA

8860 8870 8880 8890 8900
TTTGTCACTCTTCCCTAAATAATCCTTAAAAACTCCATTTCCACCCCTCC
AAACAGTGAGAAGGGATTTATTAGGAATTTTTGAGGTAAAGGTGGGGAGG

8910 8920 8930 8940 8950

CAGTTCCCAACTATTTTGTCCGCCACAGCGGGGCATTTTCTTCCTGTT
GTCAAGGGTTGATAAAACAGGCGGGTGTGCCCCGTAAAAAGAAGGACAA

8960 8970 8980 8990 9000
ATGTTTTTAATCAAACATCCTGCCAACTCCATGTGACAAACCGTCATCTT
TACAAAAATTAGTTTGTAGGACGGTTGAGGTACACTGTTTGGCAGTAGAA

9010 9020 9030 9040 9050
CGGCTACTTTTTCTCTGTCCACAGAATGAAAATTTTTCTGTCATCTCTTCG
GCCGATGAAAAGAGACAGTGTCTTACTTTTAAAAAGACAGTAGAGAAGC

9060 9070 9080 9090 9100
TTATTAATGTTTGTAATTGACTGAATATCAACGCTTATTTGCAGCCTGAA
ATAATTACAAACATTAAGTACTGACTTATAGTTGCGAATAAACGTCGGACTT

9110
TGGCGAATGG (SEQ ID NO:3)
ACCGTTACC

图 2

填充片段序列 (SEQ ID NO:1)

GAATTCGGGCTATCCCAGGTTGCCTTGGTTCATGGCAAATGGGACGTTAAGAGGGCAGA
GAGAATATGAACAGAAACTGTTCTAATATTGGTCATTTAATGTGTAAGTATTGTTCTTT
TTTAAACCTCCTTCATTTTTTTTTCCAGGAATTGCTGGACACAGTGGCTTGGTGTGTGTCT
GAGGACTGTAGGCCATGGCCCTAGGTTGTGGTTTTAGGTCTCAGGTGCTCTTCCTGGCTG
TCTCCTTGCTTCTTTCCCATGTCCTCTTCTTTGTTTCCAGCCATTTCTCCCTTATGCTTAA
GTTTGGTGCAGCAGGGTTTGGCTGCTCTCAGATTCCTGCTTCCTCAGATGCTGTAGTTGT
CAGGCCCAGCGGGCTGGCAGCGGGATCAGGATCTGGCTAGGTTTGTCTCACTGTGGCAG
AGTAGGGGGAGGCGTGGGAGAGCACGTGTGACCCCAGGCCAGCTGTAGGGAGCATAGGC
ATGGTCACGTAGCCTTCAGGTCCTAGACTTTGTCTTCTCATGAGTATGGCTGTGTGTGTA
TGGTGAAAAC TAGGTTCTACTTAGCCCAAGAAAATGGGCACATTTTGCATGTGGTTTTCT
GTAGAGAAATGCACTGGGTATCTGACATAGCCTGGCAGCATGCCTCCCTCAGGTAGGTTA
GTCTCAGGCGGTGAAGCACGTGTGTCCAGCAAGAACTTCATATGTGGCATAAAGTCTCCG
TTCTGTGAGGTGCTGGCAAATCACCACCACCGTCAAGAGGCTGAAGTGATTTTTGTCTAG
GGAGGCAGGAAAGGCTTCCTGGAGTCAGCAGCCAGTAGGTGAAAGAGTAGATTGGAGAC
CTTCTTAATCATCACCGCCTCTTGTCTCAAGGGGTGCCAGGAAGCTGTGGAGGCTGAACC
CATCTTATGCTGCCAGAGAGTGGGACACCATGAGGGTCAGGTCAAGGGGTTGTACCTTGT
TTGGTAGAGAATTAGGGGCTCTTGAAGACTTTGGATGTGGTCAGGGGAGTGTATCATTT
AGGAAGAGTGACCCGGTGAGGACGTGGGGTAGAGGAGGACAGGTGGGAGGGAGTCCAGG
TGGGAGTGAGTAGACCCAGCAGGAGTGCAGGGCCTCGAGCCAGGATGGTGGCAGGGCTG
TGAGGAGAGGCAGCCACCTGTGTGTCTGCGGAAGCAGGGGCAAGAGGGAAGAGGCCAGC
AGCGTGCTGCCATCACCCAGCGACTGGCGTAGATTGTGAGAGACCATTCCCTGCTCTTAG
GAGGGGCTGAGTTTTAGTTTTCTTGTATACAATAAGCTTGGTATTTGTTTACAAAA
CATTTGTAAAGCTAAATCAAGGTTTGATAAGGCTTCTAGTTTTATTTAAGAAGTAATGT
TGAAATAAATGTTTGTCCAATTCGCTTTGCTCATTTAAGGACTTTCAGTACAAACTGCAA
CAACAGGATTAGGATTTAAACGTTTCTGAGATGTTTTTACTCCTCAGAATTTCCCAGAAT
GTGATCTGGTTTTGATTTTCAAGCTTGCTGACCCAATAGGTTAACCACAAGTTTTACGA
AGACCATCTCAGTCCACTTACATCAACTGCCCATGCCACGGTTAAAGAGATCATCGACTG
ATGTTTGGCACAGCTTCCTCCCTCTTGGGTGGGCAAGCATTGGAAGAGAAGGCTCCTAT
GGGTGAGAGTGGGGCACCAAAGTCTTCCTGTCCCATCCCCTAGCTTGAGAAGCCCTTCT
CTAATGTGGACTTTGTGCCGTTAGCATCGTTACTAGCTTGAAGTTGACCATCTGGACGTA
CTTTCTGGTTTAGCCTCACAAGTGAGCAAGGAGGGTTGAGAGATGTGCTGTGAGGAATG
TGGGGCCCCAGCTGGCAGCAGGCTCTGGGTCAGGGGGCAGGGACCACGGGCATACCTGA
CAGTGAGGAGGGTCTAGTAGGGGATCAGTTCCCTGTTGTTCTTTAGAATTTTCTGGATA
TTCTTCTTTATTGATTTTGGGATGTGAACAATAGAATCAACTTCTACTTGTAGATTGAT
TTAGGGAGAACTTATACCTCAGATGTTAAGTCACCCTGTCCAGAATGTGGGATGCTTTCC
TATTTGTTCAGAACTTTTTAAATTACCTCAGAAGCACATGAAATTTAAAGGATTTTAAA
AAAAACTTAAAGATTATTTACATAGCTCTTGACATTTCTTGATAAATGAATCCTCAG
GTATTCCTCTGTTTTTGTACTAATAGTTACTTCTTATGGGTTTTTTTTCCCTGAAAAT
CATTTATCAAACGTATGTGGCTTATTTTCTGAAGGATGTTTGATAATTTTGAAGATAT
GAAAGTCTTCATATTTTACAAGGTTTGAAGTCTCTTTAAGCTGCATGGTTCTCATGTCAG
CTCCCAAAGCAGAAGACGGCATGTTGAAAATGCCGTAGAGAAGATACTTCTTTCCACC
TGTTTTCAACTCATATCATCTTGAATTTCAAGGGCACCTTCCATGCTCCTAGTGCTTGT
ATCTGTTTATTATTTTCTTCTCCTGAATACCCTGAACTCCAGCATGTTCTGCTGTAATTCT

GGCCTCCCTGGCATCTTGGACTCCTGTTTCCTTTGCTCTGTCATCCCCGCGGTCAGCTCCT
GCTGCGCAGCTTCTCAGCTGAAGTGCCTTTGGAGTGCCTGGCGTGTCTTGCTGGATCTTT
GAGTATTGCCTCTGGTTTCCTTGGTTTCCTTCTGCTGAGTTGCTCAGCGTCTCCACTCCCCA
TTTCTTGTGTGGCCCTTCCTGCACTCCTCTGATTCTTTTGTCTTCCCTGGTTTCTTGCTT
TGGTTTCGAGTCTCCACAGAACTTTTGCAGCTCTTCTGAAGACCTGGAAGCTTTTTCATC
TTAATTCTCATCTCATGACCTCTTTTCCCTTCTTTGAGAGCTAGAACTTCCCATGGTGAA
CTTCTCTTCCAGAATTCCATGCCTTCTTTTCCCTCCCACCTTACCTGTTGTCCAGGAGAGG
TCAGATTGCTGTGCATATTGGAGGAGAACCCTTTCTTCCCTGGGCTCTTCATCTCACATG
ACATCACCACATCACCTCGTTCCCTTGGACCCTCAGTGGTGTCACTGCTGGATTTTTCTTT
CCTTTGGCTGGCCTTAGGGCACACCCAGGTTGACTAGCGTAGTCATGGTATTTAGATCCA
CTCACATTTTCAGTTTCTGTGTCTGTCTTGCCTGCTTCTGACTTCGCCCAGAGAAAGC
TTCTCTTTCACAAGGGTTCTTAGATTTATGTTCACTGAGCACCTTCTTTTCTGAGGCAGT
GTTTTACCAATATTTATTTTCTAGTCAGTCTCGCCTTACCTTTCTTGTTATGCATGTCT
TTGGTCTGACCCATTCTCTGAGTCTGTAATAAGAAATTGCTGTATAATTTAATTACATG
AAATCCTTTAGAATCTTAACACATCTTACACCTGATTTAATATTTTATTGTATCCAAATT
GAACCAACCCTATGTGAATTTGACAGTGATTTCTCCAGGGATCCTAGTGTATAAGGAAT
AGGACTTAGTATTTTCTATTTTGTATATACCACATAACCAGATACTGATTATGATGGAC
ATTTAACCTTTTCTCATTATGAAAGAAAGTTAGGAATTATTTCTTCCAGTAGCGCCA
GTGTAACCTGAAAGCCTTTGAAAGAGTAGTTTTTGTATAGCTATCTGAAAGGAATTTCT
TTCCAAAATATTTTCCAGTGCTGACAACAAACACGCAGACACACCCTGCAAGGTGAGTG
TACGGCG

mHDS1 序列 (SEQ ID NO:4)

CTCGAGTGAGCGATGCTGGCTCGCATGGTTCGATACTGTAAAGCCACAGATGGGTGTGCGAC
CATGCGAGCCAGCACCGCCTACTAGA

mU6 启动子 (SEQ ID NO:5)

CGACGCCGCATCTCTAGGCCCGCGCCGGCCCCCTCGCACAGACTTGTGGGAGAAGCTCG
GCTACTCCCCTGCCCCGTTAATTTGCATATAATTTTCTAGTAACTATAGAGGCTTAA
TGTGCGATAAAAGACAGATAATCTGTTCTTTTAAATACTAGCTACATTTTACATGATAG
GCTTGGATTTCTATAAGAGATACAAATACTAAATTATTATTTTAAAAAACAGCACAAAA
GGAAACTCACCTAACTGTAAAGTAATTGTGTGTTTTGAGACTATAAATATCCCTTGA
GAAAAGCCTTGTTT

AAV2 ITR (94bp) (SEQ ID NO:6)

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCCGGGCGACCTTTGGTCCGCCGGCC
TCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG

AAV2 ITR (128bp) (SEQ ID NO:7)

AAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAG
GCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAG
CGAGCGCG

庆大霉素 (SEQ ID NO:8)

TTAGGTGGCGGTA CTTGGGTCGATATCAAAGTGCATCACTTCTTCCCGTATGCCCACTT
TGTATAGAGAGCCACTGCGGGATCGTCACCGTAATCTGCTTGCACGTAGATCACATAAGC
ACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTGATGAGCGCGGTGGCAATGCCCTGCCT
CCGGTGCTCGCCGGAGACTGCGAGATCATAGATATAGATCTCACTACGCGGCTGCTCAA
CTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAACCGCTTCTTGGTGAAGGCAGCAAG
CGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGGTAATCGGAGTCCGGCTGATGTTG
GGAGTAGGTGGCTACGTCTCCGA ACTCACGACCGAAAAGATCAAGAGCAGCCCCGCATGGA
TTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGATGCCATACTTGAGCCACCTAA
CTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTGCGTAACAT

β -内酰胺酶(氨基青霉素) (SEQ ID NO:9)

ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGGCGGCATTTTGCCTTCCT
GTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGC
ACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCC
CGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATC
CCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTT
GGTTGAGTACTCACCAAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT
TATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA
TCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTGCC
TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA
TGCTGTAGCAATGGCAACAACGTTGCGCAA ACTATTA ACTGGCGAACTACTTACTCTAG
CTTCCCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC
GCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT
ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT
GCCTCACTGATTAAGCATTGGTAA

Tn7R (转座元件) (SEQ ID NO:10)

TGTGGGCGGACAATAAAGTCTTAAACTGAACAAAATAGATCTAAACTATGACAATAAAG
TCTTAAACTAGACAGAATAGTTGTAAACTGAAATCAGTCCAGTTATGCTGTGAAAAAGC
ATACTGGACTTTTTGTTATGGCTAAAGCAA ACTCTTCATTTTCTGAAGTGCAAATTGCCCG
TCGTATTAAAGAGGGGCGTGGCCAAGGGCATGGTAAAGACTATATTC

Tb7L (转座元件) (SEQ ID NO:11)

AACCAGATAAGTGAAATCTAGTTCCAAACTATTTTGTCAATTTTAATTTTCGTATTAGCT
TACGACGCTACACCCAGTTCCCATCTATTTTGTCACTCTTCCCTAAATAATCCTTAAAAA
CTCCATTTCCACCCCTCCAGTTCCCAACTATTTTGTCCGCCACA

图 3

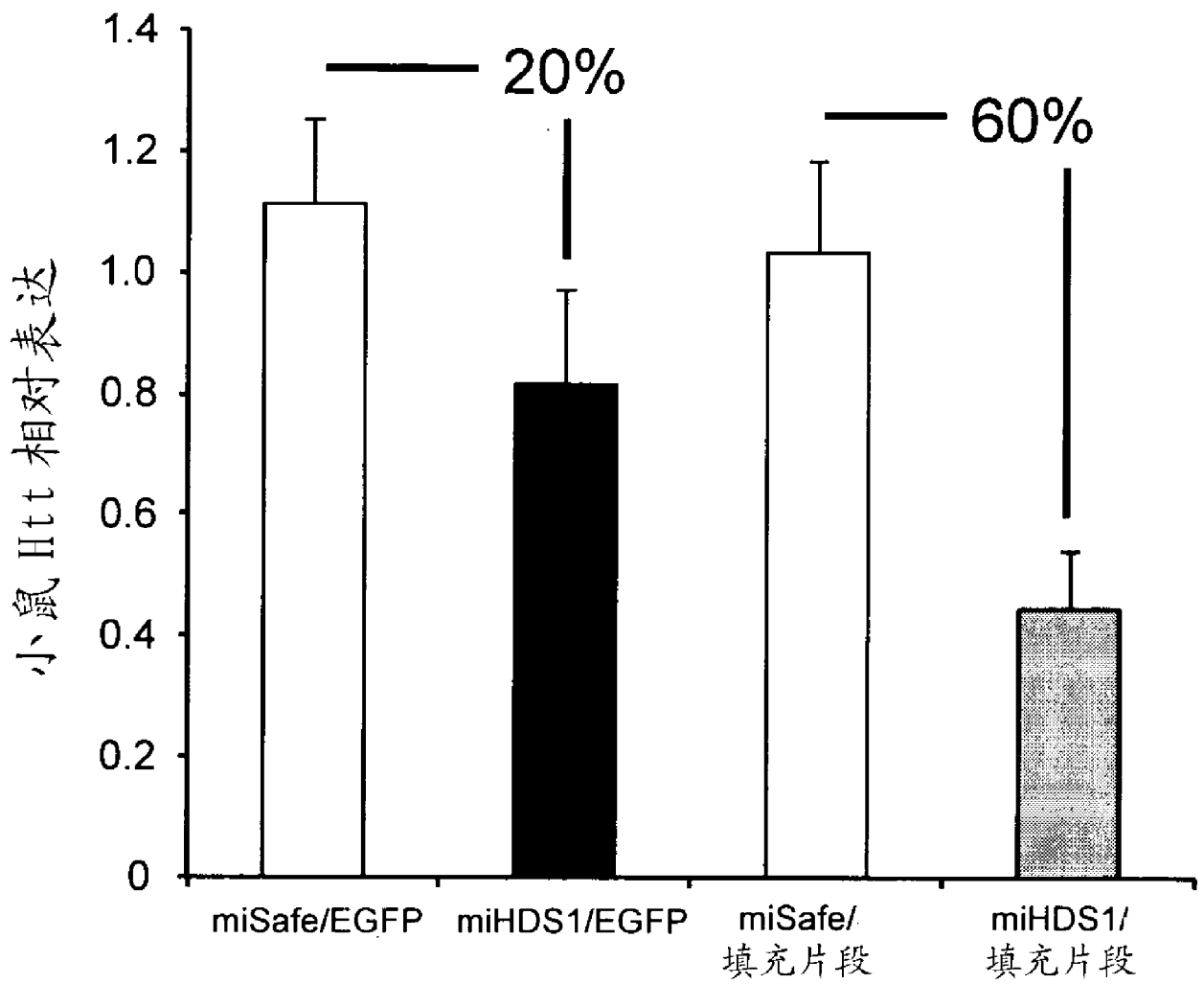


图 4

5pFBAAVmU6miHDS1-填充片段

图谱特征:

494 - 1354	β -内酰胺酶
2418 - 2642	Tn7R
27009 - 3242	庆大霉素
3810 - 3928	AAV ITR (119bp)
3967 - 4249	mU6启动子
4332 - 4417	miHDS1
4465 - 8239	填充片段序列
8293 - 8423	AAV ITR (130bp)
8764 - 8929	Tb7L

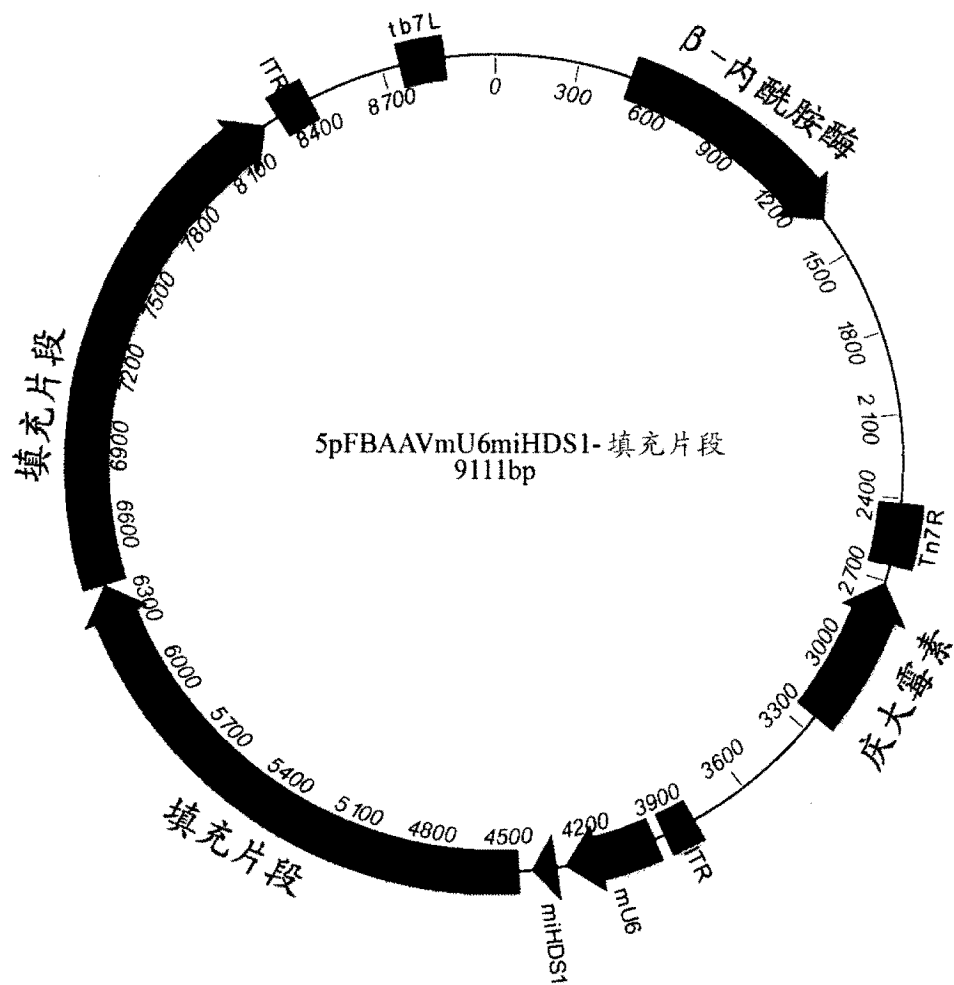


图 5

质粒序列 (SEQ ID NO:12):

TTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCCGCCGGCTTTCCCCGTCAAGCTCTAAATC
GGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTG
ATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGA
CGTTGGAGTCCACGTTCTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCC
TATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAA
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAA
TTTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAAT
ACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATT
GAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCCCTTATTCCCTTTTTTTCGGG
CATTTTGCCTTCCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAG
ATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTG
AGAGTTTTTCGCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTG
GCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCCCGCATACACTATT
CTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATG
ACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACAACACTGCGGCCAACTT
ACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGA
TCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGA
GCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGA
ACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGC
AGGACCCTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGC
CGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCG
TATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGA
TCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAAACCTTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATC
CTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCA
GACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGC
TGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTA
CCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTT
CTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATACCTC
GCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGG
TTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGGCTGAACGGGGGGTTC
GTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGA
GCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTT
TATAGTCCTGTCCGGTTTTCGCCACCTCTGACTTGAGCGTGCATTTTTGTGATGCTCGTCA
GGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTT
TGCTGGCCTTTTTGCTCACATGTTCTTCTGCGTTATCCCCTGATTCTGTGGATAACCGT
ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACCCGAACGACCGAGCGCAGCGAG
TCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTCTCCTTACGCATCTGTGC
GGTATTTACACCCGCATAGACCAGCCGCGTAACCTGGCAAAATCGGTTACGGTTGAGTAA

TAAATGGATGCCCTGCGTAAGCGGGTGTGGGCGGACAATAAAGTCTTAAACTGAACAAA
 ATAGATCTAAACTATGACAATAAAGTCTTAAACTAGACAGAATAGTTGTAAACTGAAAT
 CAGTCCAGTTATGCTGTGAAAAAGCATACTGGACTTTTGTATGGCTAAAGCAAACCTCT
 CATTTTCTGAAGTGCAAATTGCCCGTTCGTATTAAGAGGGGCGTGGCCAAGGGCATGGT
 AAAGACTATATTCGCGGCGTGTGACAATTTACCGAACAACTCCGCGGCCGGAAGCCGA
 TCTCGGCTTGAACGAATTGTTAGGTGGCGGTACTTGGGTCGATATCAAAGTGCATCACTT
 CTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTAATCTGCTT
 GCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTGATGAGCG
 CGGTGGCAATGCCCTGCCCTCCGGTGTCTCGCCGGAGACTGCGGAGATCATAGATATAGATCT
 CACTACGCGGCTGCTCAAACCTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAACCGCTT
 CTTGGTCAAGGCAGCAAGCGCGATGAATGCTTACTACGGAGCAAGTCCCGAGGTAAT
 CGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCTCCGAACTCACGACCGAAAAGAT
 CAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGATGC
 CCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGC
 TGCTGCGTAACATCGTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCT
 TGCTGCTTGATGCCCGAGGCATAGACTGTACAAAAAACAGTCATAACAAGCCATGAA
 AACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGA
 GCGCATACGCTACTTGCATTACAGTTTACGAACCGAACAGGCTTATGTCAACTGGGTTCCG
 TGCCTTCATCCGTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCCA
 GGCATTTCTGTCCTGGCTGGCGAACGAGCGCAAGGTTTCGGTCTCCACGCATCGTCAGGC
 ATTGGCGGCCTTGCTGTTCTTCTACGGCAAGGTGCTGTGCACGGATCTGCCCTGGCTTCA
 GGAGATCGGAAGACCTCGGCCGTCGCGGCGCTTGCCGGTGGTGTGACCCCGGATGAAGT
 GGTTCGCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTTGTTCGCCCAGGACTCTAGCTA
 TAGTTCTAGTGGTTGGCTACAGCTTGCATGCCTGCAGGCAGCTGCGCGCTCGCTCGCTCA
 CTGAGGCCGCCCCGGCGTCCGGGCGACCTTTGGTCCGCCGGCCTCAGTGAGCGAGCGAGCG
 CGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTTGTAGTTAATGATTAACCCG
 CCATGCTACTTATCTACGTAGCCATGCTCTAGTGAATTCGACGCCGCCATCTCTAGGCC
 GCGCCGGCCCCCTCGCACAGACTTGTGGGAGAAGCTCGGCTACTCCCCTGCCCGGTTAA
 TTTGCATATAAATTTTCTAGTAATAATAGAGGCTTAATGTGCGATAAAAGACAGATAA
 TCTGTTCTTTTTAATACTAGCTACATTTTACATGATAGGCTTGGATTTCTATAAGAGAT
 ACAATACTAAATTATTATTTTAAAAAACAGCACAAAAGGAAACTCACCTAACTGTAA
 AGTAATTGTGTGTTTTGAGACTATAAATATCCCTTGGAGAAAAGCCTTGTTTGCGTTTA
 GTGAACCGTCAGATGGTACCGTTTAAACTCGAGTGAGCGATGCTGGCTCGCATGGTCGAT
 ACTGTAAAGCCACAGATGGGTGTCGACCATGCGAGCCAGCACCGCCTACTAGAGCGGCCG
 CCACAGCGGGGAGATCCAGACATGATAAGATACATTTTTTTGAATTCGGGCTATCCCAGGT
 TGCCTTGGTTCATGGCAAATGGGACGTTAAGAGGGCAGAGAGAATATGAACAGAACTG
 TTCTAATAATTGGTCATTTAATGTGTAAGTATTGTTCTTTTTTAAACCTCCTTCATTTTTT
 TTCCAGGAATTGCTGGACACAGTGGCTTGGTGTGTGTCTGAGGACTGTAGGCCATGGCCC
 TAGGTTGTGGTTTTAGGTCTCAGGTGCTCTTCTGGCTGTCTCCTTGCTTCTTTCCCATG
 TCCCTTCTTTGTTTCCAGCCATTTCTCCCTTATGCTTAAGTTTGGTGCAGCAGGGTTTG
 GCTGCTCTCAGATTCCTGCTTCCCTCAGATGCTGTAGTTGTCAGGCCAGCGGGCTGGCAG
 CGGGATCAGGATCTGGCTAGGTTTGTCTCACTGTGGCAGAGTAGGGGGAGGCGTGGGA
 GAGCACGTGTGACCCAGGCCAGCTGTAGGGAGCATAGGCATGGTCACGTAGCCTTCAGG
 TCCTAGACTTTGTCTTCTCATGAGTATGGCTGTGTGTGTATGGTGAAACTAGGTTCTAC
 TTAGCCCAAGAAAATGGGCACATTTTGCATGTGGTTTCTGTAGAGAAATGCACTGGGTA
 TCTGACATAGCCTGGCAGCATGCCTCCCTCAGGTAGGTTAGTCTCAGGCGGTGAAGCACG

TGTGTCCAGCAAGAACTTCATATGTGGCATAAAGTCTCCGTTCTGTGAGGTGCTGGCAAA
 TCACCACCACCGTCAAGAGGCTGAAGTGATTTTTGTCTAGGGAGGCAGGAAAGGCTTCCT
 GGAGTCAGCAGCCAGTAGGTGAAAGAGTAGATTGGAGACCTTCTTAATCATCACCCGCTC
 TTGTCTCAAGGGGTGCCAGGAAGCTGTGGAGGCTGAACCCATCTTATGCTGCCAGAGAGT
 GGGACACCATGAGGGTCAGGTCAAGGGGTTGTACCTTGTTTGGTAGAGAATTAGGGGCT
 CTTGAAGACTTTGGATGTGGTCAGGGGAGTGTATCATTTAGGAAGAGTGACCCGGTGAG
 GACGTGGGGTAGAGGAGGACAGGTGGGAGGGAGTCCAGGTGGGAGTGAGTAGACCCAGC
 AGGAGTGCAGGGCCTCGAGCCAGGATGGTGGCAGGGCTGTGAGGAGAGGCAGCCACCTG
 TGTGTCTGCGGAAGCAGGGGCAAGAGGGGAAGAGGCCAGCAGCGTGCTGCCATCACCCAGC
 GACTGGCGTAGATTGTGAGAGACCATTCCCTGCTCTTAGGAGGGGCTGAGTTTTAGTTTT
 CTCTTGTTATAACAATAAGCTTGGTATTTGTTTACAAAACATTTGTAAGCTAAATCAAG
 GTTTGATAAGGCTTCTAGTTTTATTTAAGAAGTAATGTTGAAATAAATGTTTGTCCAAT
 TCGTTTTGCTCATTTAAGGACTTTCAGTACAAACTGCAACAACAGGATTAGGATTTAAAC
 GTTTCTGAGATGTTTTTACTCCTCAGAATTTCCCAGAATGTGATCTGGTTTTGATTTTCA
 AGCTTGCTGACCCAATAGGTTAACCACAAGTTTTACGAAGACCATCTCAGTCCACTTAC
 ATCAACTGCCCATGCCACGGTTAAAGAGATCATCGACTGATGTTGGCACAGCTTCCTCC
 CTCTTGGGTGGGCAAGCATTTGGAAGAGAAGGCTCCTATGGGTGAGAGTGGGGCACCAA
 AGTCTTCCCTGTCCCATCCCTAGCTTGAGAAGCCCTTCTCTAATGTGGACTTTGTGCCG
 TTAGCATCGTTACTAGCTTGAAGTTGACCATCTGGACGTACTTTCTGGTTTAGCCTCACA
 AGTGAGCAAGGAGGGTTGAGAGATGTGCTGTGAGGAATGTGGGGCCCCAGCTGGCAGCA
 GGCTCTGGGTGAGGGGGCAGGGACCAGGGCATACTGACAGTGAGGAGGGGTCTAGT
 AGGGGATCAGTTCCTGTTGTTCTTTAGAATTTCTGGATATTCTTCTTTATTGATTTT
 GGGATGTGAACAATAGAATCAACTTCTACTTGTAGATTGATTTAGGGAGAATTATACC
 TCAGATGTTAAGTCACCCGTCCAGAATGTGGGATGCTTTCCTATTTGTTTCAGAACTTTT
 TAAATTACCTCAGAAGCACATGAAATTTAAAGGATTTTAAAAAAACTTAAAGATTATT
 TCACATAGCTCTTGCACATTTCTTGATAAATGAATCCTCAGGTATTCCTCTGTTTTTGT
 ACTAATAGTTACTTCTTATGGGTTTTTTTTCCCTGAAAATCATTTATCAAACGTATGTG
 GCTTATTTTCTGAAGGATGTTTGATAATTTTGAAGATATGAAAGTCTTCATATTTTAC
 AAGTTTTGGGGTCTCTTTAAGCTGCATGGTTCTCATGTCAGCTCCCAAAGCAGAAGACGG
 CATGTTGAAAAATGCCGTAGAGAAGATACTTCTTTCCACCTGTTTTCAACTCATATCAT
 CTTGAATTTCAGGGCACCTTCCATGCTCCTAGTGCTTGCTATCTGTTTATTATTTTCT
 TCCTGAATAACCTGAACTCCAGCATGTTCTGCTGTAATTTCTGGCCTCCCTGGCATCTTGG
 ACTCCTGTTTCTTTGCTCTGTCATCCCCGCGGTGAGTCTCTGCTGCGCAGTTCTCAGCT
 GAAGTGCGTTTGGAGTGCCTGGCGTGTCTTGCTGGATCTTTGAGTATTGCCTCTGGTTTC
 CTTGGTTCTTCTGCTGAGTTGCTCAGCGTCTCCACTCCCATTTCTTGTGTGGCCCTTCC
 TGCACCTCTGATTCCTTTTGTCTTCCCTGGTTTTCTTGCTTTGGTTTTGAGTCTCCACAG
 AACTTTTTGCAGCTCTTCTGAAGACCTGGAAGCTTTTTTCATCTTAATTCTCATCTCATGAC
 CTCTTTTCCCTTCTTTGAGAGCTAGAACTTCCCATGGTGAAGTCTCTTTCCAGAATTCC
 ATGCCTTCTTTTCCCTCCACTTACCTGTTGTCCAGGAGAGGTCAGATTGCTGTGCATAT
 TGGAGGAGAACCCTTTCTTCCCTGGGCTCTTCATCTCACATGACATCACCATCACCTC
 GTTCCTTGGACCCTCAGTGGTGTCACTGCTGGATTTTTCTTTCTTTGGCTGGCCTTAGG
 GCACACCCAGGTTGACTAGCGTAGTCATGGTATTTAGATCCACTCACATTTTCAAGTTTCT
 GTGTCTGTCTCTTGCCTGCTTCTGACTTCGCCAGAGAAAAGCTTCTCTTTCACAAGGGTT
 CTTAGATTTATGTTCACTGAGCACCTTCTTTTCTGAGGCAGTGTTTTACCAATATTTATT
 TTCCTAGTCAGTCTCGCCTTACCTTTCTTGTTATGCATGTCTTTGGTCTGACCCATTCTC
 TGAGTCTGTAATAAGAATTGCTGTATAATTTAATTACATGAAATCCTTTAGAATCTTA

ACACATCTTACACCTGATTTAATATTTTATTGTATCCAAATTGAACCAACCCTATGTGAA
TTTGACAGTGATTTCTCCCAGGGATCCTAGTGTATAAGGAATAGGACTTAGTATTTTCT
ATTTTTTGATATAACCACATACCAGATACTGATTATGATGGACATTTAACCCTTTTTTCTC
ATTATGAAAGAAAGTTAGGAATTAATTTCTTCCAGTAGCGCCAGTGTAACCTGAAAGCCT
TTGAAAGAGTAGTTTTTGTATAGCTATCTGAAAGGAATTTCTTTCCAAAATATTTTTCC
AGTGCTGACAACAAACACGCAGACACACCCTGCAAGGTGAGTGTACGGCGCACTAGAGCA
TGGCTACGTAGATAAGTAGCATGGCGGGTTAATCATTAACTACAAGGAACCCCTAGTGA
TGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGG
TCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCC
TGCAGGTCTGAGACAATAACCCTGATAAATGCTTCAATAATGTAAGCTTGTGCGAGAAGT
ACTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAC
CTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAA
AGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCA
TGTCTGGATCTGATCACTGATATCGCCTAGGAGATCCGAACCAGATAAGTGAAATCTAG
TTCCAAACTATTTTGTCAATTTTAAATTTTCGTATTAGCTTACGACGCTACACCCAGTTCC
CATCTATTTTGTCACTCTTCCCTAAATAATCCTTAAAAACTCCATTTCCACCCCTCCCAG
TTCCCAACTATTTTGTCCGCCACAGCGGGGCATTTTTCTTCCCTGTTATGTTTTTAATCA
AACATCCTGCCAACTCCATGTGACAAACCGTCATCTTCGGCTACTTTTTCTCTGTACAG
AATGAAAATTTTTCTGTATCTCTTCGTTATTAATGTTTGTAATTGACTGAATATCAAC
GCTTATTTGCAGCCTGAATGGCGAATGG

图 6

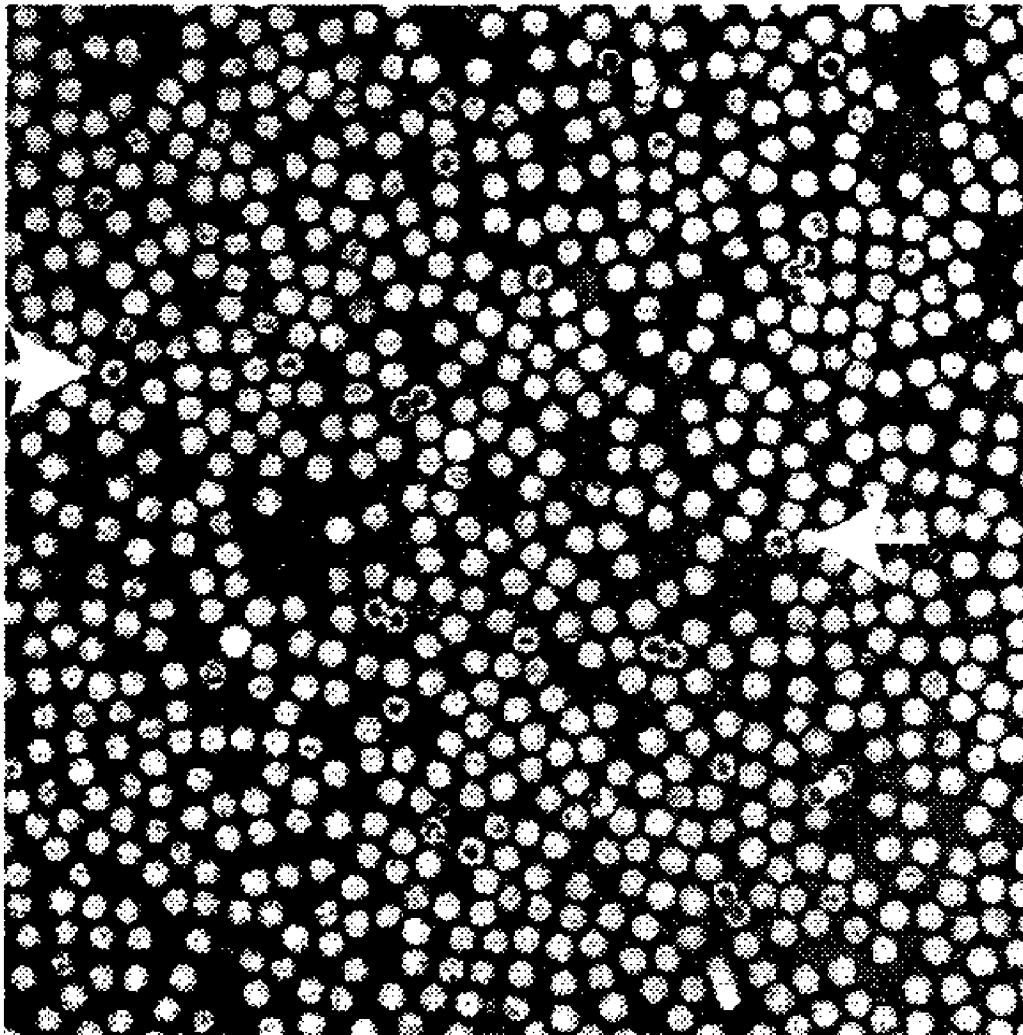
填充片段序列(填充片段#2) (SEQ ID NO: 2)

GGGCTATCCCAGGTTGCCTTGGTTCATGGCAAATGGGACGTTAAGAGGGCAGAGAGA
ATATGAACAGAACTGTTCTAATATTGGTCATTTAATGTGTAAGTATTGTTCTTTTT
TAAACCTCCTTCATTTTTTTTTCCAGGAATTGCTGGACACAGTGGCTTGGTGTGTGTCT
GAGGACTGTAGGCCATGGCCCTAGGTTGTGGTTTTAGGTCTCAGGTGCTCTTCCTGGC
TGTCTCCTTGCTTCTTTCCCATGTCTCTTCTTTGTTTTCCAGCCATTTCTCCCTTATGC
TTAAGTTTGGTGCAGCAGGGTTGGCTGCTCTCAGATTCCCTGCTTCCTCAGATGCTGT
AGTTGTCAGGCCAGCGGGCTGGCAGCGGGATCAGGATCTGGCTAGGTTTGGCTCTCAC
TGTGGCAGAGTAGGGGAGGCGTGGGAGAGCACGTGTGACCCAGGCCAGCTGTAGG
GAGCATAGGCATGGTCACGTAGCCTCAGGTCCTAGACTTTGTCTTCTCATGAGTAT
GGCTGTGTGTATGGTGAAGACTAGGTTCTACTTAGCCCAAGAAAATGGGCACATT
TTGCATGTGGTTTCTGTAGAGAAATGCACTGGGTATCTGACATAGCCTGGCAGCATG
CCTCCCTCAGGTAGGTTAGTCTCAGGCGGTGAAGCACGTGTGTCCAGCAAGAACTTCA
TATGTGGCATAAAGTCTCCGTTCTGTGAGGTGCTGGCAAATCACCACCACCGTCAAG
AGGCTGAAGTGATTTTTGTCTAGGGAGGCAGGAAAGGCTTCTGGAGTCAGCAGCCA
GTAGGTGAAAGAGTAGATTGGAGACCTTCTTAATCATCACCGCTCTTGTCTCAAGG
GGTGCCAGGAAGCTGTGGAGGCTGAACCCATCTTATGCTGCCAGAGAGTGGGACACC
ATGAGGGTCAGGTCAAGGGGTTGTACCTTGTTTGGTAGAGAATTAGGGGCTCTTGAA
GACTTTGGATGTGGTCAGGGGAGTGTATCATTTAGGAAGAGTGACCCGGTGAGGACG
TGGGGTAGAGGAGGACAGGTGGGAGGGAGTCCAGGTGGGAGTGAGTAGACCCAGCAG
GAGTGCAGGGCCTCGAGCCAGGATGGTGGCAGGGCTGTGAGGAGAGGCAGCCACCTG
TGTGTCTGCGAAGCAGGGGCAAGAGGGAAGAGGCCAGCAGCGTGTGCCATCACCC
AGCGACTGGCGTAGATTGTGAGAGACCATTCCCTGCTCTTAGGAGGGGCTGAGTTTT
AGTTTTCTCTTGTATACAATAAGCTTGGTATTTGTTTACAAAACATTTGTAAAGCT
AAATCAAGGTTTGATAAGGCTTCTAGTTTTATTTAAGAAGTAATGTTGAAATAAATG
TTTGTCCAATTTCGCTTTGCTCATTTAAGGACTTTTCAGTACAACTGCAACAACAGGA
TTAGGATTTAAACGTTTCTGAGATGTTTTACTCCTCAGAATTTCCAGAATGTGAT
CTGGTTTTGATTTTCAAGCTTGTGACCCAATAGGTTAACCACAAGTTTTACGAAG
ACCATCTCAGTCCACTTACATCAACTGCCCATGCCACGGTTAAAGAGATCATCGACTG
ATGTTTGGCACAGCTTCCCTCCCTCTTGGGTGGGCAAGCATTTGGAAGAGAAGGCTCCT
ATGGGTGAGAGTGGGGCACCAAAGTCTTCCCTGTCCCATCCCCTAGCTTGAGAAGCCC
TTCTCTAATGTGGACTTTGTGCCGTTAGCATCGTTACTAGCTTGAAGTTGACCATCTG
GACGTACTTTCTGGTTTAGCCTCACAAGTGAGCAAGGAGGGTTGAGAGATGTGCTGT
GAGGAATGTGGGGCCCCAGCTGGCAGCAGGCTCTGGGTGAGGGGGCAGGGACCACG
GGCATACTGACAGTGAGGAGGGGTCTAGTAGGGGATCAGTTCCCTGTTGTTCTTT
AGAATTTTCTGGATATTCTTCTTTATTGATTTTGGGATGTGAACAATAGAATCAACT
TCTACTTGTAGATTGATTTAGGGAGAACTTATACCTCAGATGTTAAGTCACCTGTG
CAGAATGTGGGATGCTTTCCTATTTGTTTCAAGACTTTTTAAATTACCTCAGAAGCAC
ATGAAATTTAAAGGATTTTAAAAAAACTTAAAGATTATTTACATAGCTCTTGAC
ATTTCTTGATAAATGAATCCTCAGGTATTCCTCTGTTTTTGTACTAATAGTTACTT
CTTATGGGTTTTTTTTCCCTGAAAATCATTTATCAAACGTATGTGGCTTATTTTCTG
AAGGATGTTTGATAATTTTGGGAAGATATGAAAGTCTTCATATTTTACAAGTTTGGG
GTCTCTTTAAGCTGCATGGTTCTCATGTGAGCTCCCAAAGCAGAAGACGGCATGTTG
AAAAATGCCGTAGAGAAGATACTTCTTTTCCACCTGTTTTCAACTCATATCATCTTG
AATTTCCAGGGCACCTTCCATGCTCCTAGTGCTTGTATCTGTTTATTATTTTCTTC
CTGAATACCCTGAACTCCAGCATGTTCTGCTGTAATTCGGCCTCCCTGGCATCTTGG

ACTCCTGTTTCCTTTGCTCTGTCATCCCCGCGGTCAGCTCCTGCTGCGCAGCTTCTCA
GCTGAAGTGCCTTTGGAGTGCCTGGCGTGTCTTGTGGATCTTTGAGTATTGCCTCTG
GTTTCCTTGGTTCCTTCTGCTGAGTTGCTCAGCGTCTCCACTCCCCATTTCTTGTGTG
GCCCTTCCTGCACTCCTCTGATTCCTTTTGTCTTCCCTGGTTCCTTGCCTTTGGTTTCG
AGTCTCCACAGAACTTTTGCAGCTCTTCTGAAGACCTGGAAGCTTTTTCATCTTAATT
CTCATCTCATGACCTCTTTTCCCTTCTTTGAGAGCTAGAACTTCCCATGGTGAACCTC
TCTTTCCAGAAATCCATGCCTTCTTTTCCCTCCCACCTTACCTGTTGTCCAGGAGAGGT
CAGATTGCTGTGCATATTGGAGGAGAACCCTTTCTTCCCTGGGCTCTTCATCTCACAT
GACATCACCACATCACCTCGTTCCTTGGACCCTCAGTGGTGTCACTGCTGGATTTTTC
TTTCCTTTGGCTGGCCTTAGGGCACACCCAGGTTGACTAGCGTAGTCATGGTATTTAG
ATCCACTCACATTTTTCAGTTTCTGTGTCTGTCTCTTGCCTGCTTCTGACTTCGCCCAG
AGAAAGCTTCTCTTTCACAAGGGTTCCTTAGATTTATGTTCACTGAGCACCTTCTTTTC
TGAGGCAGTGTTTTACCAATATTTATTTTCCCTAGTCAGTCTCGCCTTACCTTTCTTGT
TATGCATGTCTTTGGTCCTGACCCATTCTCTGAGTCTGTAAAATAGAATTGCTGTAT
AATTTAATTACATGAAATCCTTTAGAATCTTAACACATCTTACACCTGATTTAATAT
TTTATTGTATCCAAATTGAACCAACCCTATGTGAATTTGACAGTGATTTCTCCCAGG
GATCCTAGTGTATAAGGAATAGGACTTAGTATTTTCTATTTTTTTGATATAACCACATA
CCAGATACTGATTATGATGGACATTTAACCCTTTTTTCTCATTATGAAAGAAAGTTA
GGAATTATTTCTTCCAGTAGCGCCAGTGTAACCTGAAAGCCTTTGAAAGAGTAGTTT
TTGTATAGCTATCTGAAAGGAATTTCTTTCCAAAATATTTTTTCCAGTGCTGACAACA
AACACGCAGACACACCCTGCAAGGTGAGTGTACGGCG

图 7

AAV2/1 mU6miHDS25 内含子 I/II



96.6% 完整病毒体
1.07E+13 vg/mL

图 8

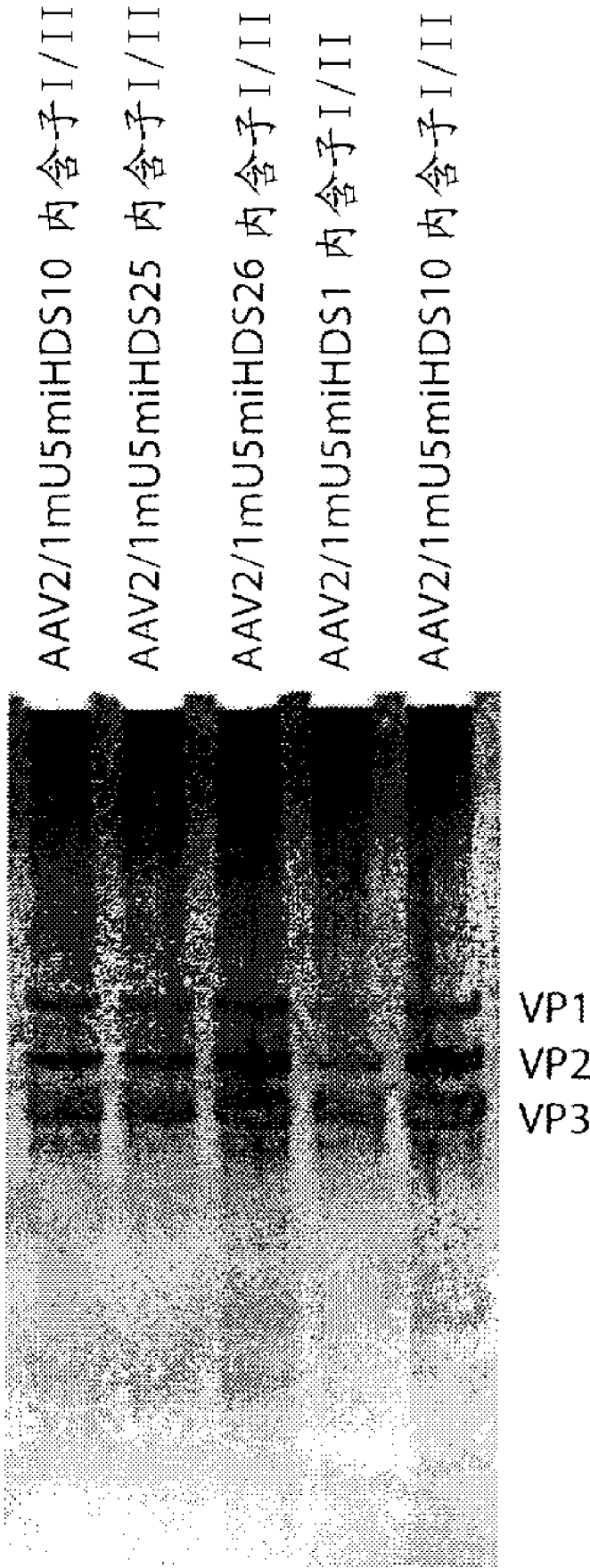


图 9

表1. miR-内含子1/II病毒体的包装效率%和污染物%。

	Cap/rAAV	Amp/rAAV	Gent/rAAV	平均 空的 %	QPCR 滴度 (vg/ml)	总vg/ml	以下的总# (pt/ml)
AAV2/1mU6miSafe 内含子I/II	0.00%	0.06%	0.15%	1.30%	2.75E+13	2.76E+13	2.79E+13
AAV2/1mU6miHDS26 内含子I/II	0.15%	1.81%	1.29%	2.00%	3.23E+12	3.34E+12	5.33E+12
AAV2/1mU6miHDS26 内含子I/II	0.80%	2.14%	7.87%	3.90%	1.09E+13	1.22E+13	1.27E+13
AAV2/1mU6miHDS25 内含子I/II	0.19%	1.34%	1.02%	0.90%	2.74E+12	2.81E+12	3.73E+12
AAV2/1mU6miHDS25 内含子I/II	0.08%	0.28%	1.98%	2.70%	1.07E+13	1.09E+13	1.12E+13
AAV2/1mU6miHDS10 内含子I/II	0.12%	1.40%	0.87%	5.60%	3.52E+12	3.60E+12	3.80E+12
AAV2/1mU6miHDS1 内含子I/II	0.01%	0.15%	0.15%	0.70%	1.81E+13	1.82E+13	2.08E+13