



(12) 发明专利

(10) 授权公告号 CN 110577972 B

(45) 授权公告日 2022. 10. 11

(21) 申请号 201910731412.5

C12N 15/113 (2010.01)

(22) 申请日 2019.08.08

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 110577972 A

US 2019161742 A1, 2019.05.30

US 2018312824 A1, 2018.11.01

WO 2019066490 A2, 2019.04.04

(43) 申请公布日 2019.12.17

CN 108795902 A, 2018.11.13

US 2018066242 A1, 2018.03.08

(73) 专利权人 复旦大学

WO 2021023307 A1, 2021.02.11

地址 200433 上海市杨浦区邯郸路220号

(72) 发明人 王永明 胡子英 王大奇 王帅

HU.Z.Y. 等. Discovery and engineering of small SlugCas9 with broad targeting range and high specificity and activity. 《bioRxiv》. 2020, 全文.

(74) 专利代理机构 上海正旦专利代理有限公司
31200

郑小梅等. CRISPR-Cas9介导的基因组编辑技术的研究进展. 《生物技术进展》. 2015, (第01期), 1-9页.

专利代理师 陆飞 陆尤

审查员 李有朝

(51) Int. Cl.

C12N 15/90 (2006.01)

C12N 15/864 (2006.01)

C12N 15/10 (2006.01)

C12N 9/22 (2006.01)

权利要求书2页 说明书7页

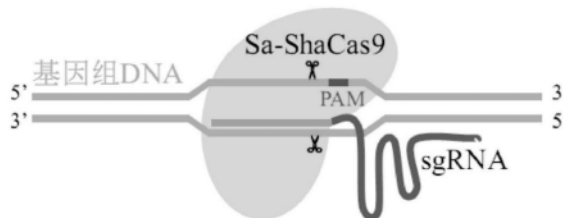
序列表8页 附图2页

(54) 发明名称

CRISPR/Sa-ShaCas9基因编辑系统及其应用

(57) 摘要

本发明属于基因编辑技术领域, 具体为一种CRISPR/Sa-ShaCas9基因编辑系统以及其应用。本发明基因编辑系统为Sa-ShaCas9蛋白与sgRNA形成的复合物, 能精确靶向DNA序列, 并产生切割, 使DNA发生双链断裂损伤; 所述基因编辑为在细胞中或体外进行基因编辑; Sa-ShaCas9为融合蛋白, 把SaCas9的PAM识别结构域(PAM interacting, PI) 替换为ShaCas9的PAM识别结构域(ShaCas9-PI), Sa-ShaCas9蛋白小, 为1055个氨基酸, 识别的PAM序列简单, 所述Sa-ShaCas9蛋白具有SEQ ID NO:1所示的氨基酸序列, 所述sgRNA具有SEQ ID NO:2所示的核苷酸序列。本发明在基因编辑领域中具有广泛的应用前景。



1. 一种CRISPR/Cas9基因编辑系统,基因编辑为在细胞中或体外进行基因编辑,其特征在于,所述CRISPR/Cas9系统为Sa-ShaCas9蛋白和sgRNA复合物,能精确定位靶向DNA序列,并产生切割,使DNA发生双链断裂损伤;所述Sa-ShaCas9蛋白为SEQ ID NO:1所示的氨基酸序列;所述sgRNA为SEQ ID NO:2所示的核苷酸序列。

2. 根据权利要求1所述的CRISPR/Cas9基因编辑系统,其特征在于,所述细胞包括真核细胞和原核细胞;所述真核生物细胞包括哺乳动物细胞和植物细胞;所述哺乳动物细胞包括中国仓鼠卵巢细胞、幼仓鼠肾细胞、小鼠Sertoli细胞、小鼠乳腺癌细胞、buffalo大鼠肝细胞、大鼠肝癌细胞、由SV40转化的猴肾CV1系、猴肾细胞、犬肾细胞、人宫颈癌细胞、人肺细胞、人肝细胞、HIH/3T3细胞、人U2-OS骨肉瘤细胞、人A549细胞、人K562细胞、人HEK293细胞、人HEK293T细胞、人HCT116细胞或人MCF-7细胞或TRI细胞。

3. 根据权利要求1所述的CRISPR/Cas9基因编辑系统,其特征在于,所述Sa-ShaCas9蛋白包括无切割活性或仅具有单链切割活性或具有双链切割活性的Sa-ShaCas9蛋白。

4. 根据权利要求1所述的CRISPR/Cas9基因编辑系统,其特征在于,所述精确定位DNA序列包括sgRNA中的5'端20bp或21bp序列与靶向DNA序列能形成碱基互补配对结构。

5. 根据权利要求1所述的CRISPR/Cas9基因编辑系统,其特征在于,所述精确定位靶向DNA序列包括Sa-ShaCas9蛋白和sgRNA复合物识别靶向DNA序列上的PAM序列。

6. 根据权利要求5所述的CRISPR/Cas9基因编辑系统,其特征在于,所述PAM序列为NGGRM,所述靶向DNA序列为SEQ ID NO:3所示。

7. 根据权利要求1所述的CRISPR/Cas9基因编辑系统,其特征在于,所述sgRNA可以进行磷酸化、硫化、甲基化或羟基化修饰。

8. 根据权利要求1所述的CRISPR/Cas9基因编辑系统,其特征在于,所述Sa-ShaCas9蛋白和sgRNA复合物能精确靶向DNA序列指Sa-ShaCas9蛋白和sgRNA复合物可以识别并结合特定DNA序列,或指将与Sa-ShaCas9蛋白融合的其他蛋白或特异性识别sgRNA的蛋白带至靶向DNA的位置。

9. 根据权利要求8所述的CRISPR/Cas9基因编辑系统,其特征在于,所述Sa-ShaCas9蛋白和sgRNA复合物或与Sa-ShaCas9蛋白融合的其他蛋白或特异性识别sgRNA的蛋白可以对靶向DNA区域进行修饰和调控,所述修饰和调控包括基因转录水平的调控、单碱基转换器或染色质成像追踪。

10. 根据权利要求9所述的CRISPR/Cas9基因编辑系统,其特征在于,所述单碱基转换器包括碱基腺嘌呤到鸟嘌呤、胞嘧啶到胸腺嘧啶、胞嘧啶到尿嘧啶或其它碱基之间的转换。

11. 如权利要求1-10之一所述的CRISPR/Cas9基因编辑系统在细胞中进行非治疗目的的基因编辑的方法,包括通过Sa-ShaCas9蛋白和sgRNA的复合物识别定位靶向DNA,对DNA进行编辑;最后检测编辑效率;具体步骤为:

(1) 合成人源化的ShaCas9-PI基因序列;并克隆到表达载体上,得到pAAV2_Sa-ShaCas9_ITR;

(2) 合成sgRNA对应的寡核苷酸单链DNA,即Oligo-F和Oligo-R序列,退火后连接至质粒pAAV2_Sa-ShaCas9_U6_BsaI的BsaI酶切位点,得到pAAV2_Sa-ShaCas9-hU6-sgRNA;

(3) 将表达Sa-ShaCas9蛋白、sgRNA的载体递送至含有靶位点的细胞中;

(4) 将编辑后的靶位点进行PCR扩增,T7EI酶切或者二代测序检测编辑效率。

12. 根据权利要求11所述的方法,其特征在于,所述pAAV2_Sa-ShaCas9-hU6-sgRNA为腺相关病毒骨架质粒,其包括AAV2 ITR、CMV增强子、CMV启动子、SV40 NLS、Sa-ShaCas9、nucleoplasmin NLS、3x HA、bGH poly(A)、人U6启动子、BsaI内切酶位点、sgRNA scaffold序列。

13. 根据权利要求11所述的方法,其特征在于,递送至细胞的CRISPR/Sa-ShaCas9系统包括表达Sa-ShaCas9蛋白或sgRNA的质粒、逆转录病毒、腺病毒、腺相关病毒载体或RNA或Sa-ShaCas9蛋白。

14. 根据权利要求11所述的方法,其特征在于,合成sgRNA对应的寡核苷酸单链DNA序列,即Oligo-F和Oligo-R序列为SEQ ID NO:4和SEQ ID NO:5所示;

步骤(3)中的细胞的靶位点具有SEQ ID NO:6所示的核苷酸序列;

步骤(4)中PCR的模板为编辑后的DNA;PCR扩增的引物序列为:SEQ ID NO:7、SEQ ID NO:8,SEQ ID NO:9、SEQ ID NO:10。

15. 如权利要求1-10之一所述CRISPR/Sa-ShaCas9基因编辑系统的试剂盒,该试剂盒包括Sa-ShaCas9蛋白或靶向DNA序列的sgRNA或靶向DNA。

16. 如权利要求1-10之一所述CRISPR/Sa-ShaCas9基因编辑系统的应用,包括基因敲除、定点碱基的改变、定点插入、基因转录水平的调控、单碱基转换器或染色质成像追踪。

CRISPR/Sa-ShaCas9基因编辑系统及其应用

技术领域

[0001] 本发明属于基因编辑技术领域,具体涉及一种能在细胞中进行基因编辑的CRISPR/Sa-ShaCas9系统以及其相关应用。

背景技术

[0002] CRISPR/Cas9是细菌和古细菌为抵御外源病毒或质粒入侵而进化的一种获得性免疫系统。在CRISPR/Cas9系统中,crRNA (CRISPR-derived RNA)、tracrRNA (trans-activating RNA) 以及Cas9蛋白形成复合体后,识别靶位点的PAM (Protospacer Adjacent Motif) 序列,crRNA 会与靶DNA序列形成互补结构,Cas9蛋白行使切割DNA的功能,使DNA发生断裂损伤。其中,tracrRNA和crRNA通过连接序列可以融合成为单链指导RNA (single guide RNA, sgRNA)。当DNA发生断裂损伤后,细胞内的两种主要DNA损伤修复机制负责修复:非同源末端连接 (Non-homologous end-joining, NHEJ) 和同源重组 (homologous recombination, HR)。NHEJ修复的结果会引起碱基的缺失或插入,可以进行基因敲除;在提供同源模板的情况下,利用HR修复可以进行基因的定点插入和碱基的精确替换。

[0003] 除了基础科研外,CRISPR/Cas9还具有广泛的临床应用前景。利用CRISPR/Cas9系统做基因治疗时,需要把Cas9和sgRNA导入到体内。目前做基因治疗最有效的递送载体是AAV 病毒。但是AAV病毒包装的DNA一般不超过4.5kb。SpCas9因为PAM序列简单(识别NGG)和活性高而得到广泛应用。但是SpCas9蛋白有1367个氨基酸,加上sgRNA和启动子,无法有效的包装到AAV病毒中,限制了其在临床中的应用。为了克服这个问题,几个小的Cas9 被发明出来了,包括SaCas9 (PAM序列为NNGRRT);St1Cas9 (PAM序列为NNAGAW); NmCas9 (PAM序列为NNNNGATT); Nme2Cas9 (PAM序列为NNNNCC); CjCas9 (PAM序列为NNNNRYAC),但是这些Cas9或者PAM序列复杂(在基因组中可靶向的DNA序列少),或者编辑效率低,难以广泛应用。寻找Cas9蛋白小,PAM序列简单CRISPR/Cas9系统是解决上述问题的希望所在。

发明内容

[0004] 针对上述问题,本发明目的在于提供一种编辑活性高、Cas9蛋白小、PAM序列简单的 CRISPR/Cas9新的基因编辑系统及其应用。

[0005] 本发明提供的CRISPR/Cas9基因编辑系统,为Sa-ShaCas9蛋白与sgRNA形成的复合体,记为CRISPR/Sa-ShaCas9基因编辑系统(即Sa-ShaCas9蛋白,它和single guide RNA (sgRNA) 共同作用实现基因编辑);能精确靶向DNA序列,并产生切割,使DNA发生双链断裂损伤;所述基因编辑为在细胞中或体外进行基因编辑;Sa-ShaCas9蛋白小,为1053个氨基酸,识别的PAM序列简单 (NNGRM),所述Sa-ShaCas9蛋白具有SEQ ID NO:1所示的氨基酸序列,或与SEQ ID NO:1所示的氨基酸序列至少80%相同的氨基酸序列;所述sgRNA具有SEQ ID NO:2所示的核苷酸序列,或者基于SEQ ID NO:2改造的sgRNA序列。

[0006] 本发明中,所述细胞包括真核细胞和原核细胞;所述真核生物细胞包括哺乳动物细胞和植物细胞。

[0007] 本发明中,所述哺乳动物细胞包括中国仓鼠卵巢细胞、幼仓鼠肾细胞、小鼠 Sertoli 细胞、小鼠乳腺癌细胞、buffalo 大鼠肝细胞、大鼠肝癌细胞、由 SV40 转化的猴肾 CV1 系、猴肾细胞、犬肾细胞、人宫颈癌细胞、人肺细胞、人肝细胞、HIH/3T3 细胞、人 U2-OS 骨肉瘤细胞、人 A549 细胞、人 K562 细胞、人 HEK293 细胞、人 HEK293T 细胞、人 HCT116 细胞或人 MCF-7 细胞或 TRI 细胞。

[0008] 本发明中,所述 CRISPR/Cas9 系统中 Sa-ShaCas9 是融合蛋白,是将 SaCas9 的 PI 结构域替换为 ShaCas9 的 PI 结构域,ShaCas9 为 *Staphylococcus haemolyticus* Cas9。Sa-ShaCas9 融合蛋白和 single guide RNA (sgRNA) 共同作用实现基因编辑。

[0009] 本发明中,所述 ShaCas9 蛋白属于溶血葡萄球菌属 (*Staphylococcus haemolyticus*),所述 ShaCas9 蛋白的 UniProt 的检索号为 A0A2T4SLN6。

[0010] 本发明中,所述 Sa-ShaCas9 蛋白包括无切割活性或仅具有单链切割活性或具有双链切割活性的 Sa-ShaCas9 蛋白。

[0011] 本发明中,所述精确定位 DNA 序列包括 sgRNA 中的 5' 端 20bp 或 21bp 序列与靶向 DNA 序列能形成碱基互补配对结构。

[0012] 本发明中,所述精确定位靶向 DNA 序列包括 Sa-ShaCas9 蛋白和 sgRNA 复合体识别靶向 DNA 序列上的 PAM 序列。

[0013] 本发明中,所述 PAM 序列为 NNGRM,所述靶向 DNA 序列为:

[0014] NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGRM (SEQ ID NO:3)。

[0015] 本发明中,所述 Sa-ShaCas9 蛋白和 sgRNA 复合体能精确靶向 DNA 序列指 Sa-ShaCas9 蛋白和 sgRNA 复合体可以识别并结合特定 DNA 序列,或指将与 Sa-ShaCas9 蛋白融合的其他蛋白或特异性识别 sgRNA 的蛋白带至靶向 DNA 的位置。

[0016] 本发明中,所述 Sa-ShaCas9 蛋白和 sgRNA 复合体或与 Sa-ShaCas9 蛋白融合的其他蛋白或特异性识别 sgRNA 的蛋白可以对靶向 DNA 区域进行修饰和调控,所述修饰和调控包括但不限于基因转录水平的调控、DNA 甲基化调控、DNA 乙酰化修饰、组蛋白乙酰化修饰、单碱基转换器或染色质成像追踪。

[0017] 本发明中,所述单碱基转换器包括但不限于碱基腺嘌呤到鸟嘌呤、或胞嘧啶到胸腺嘧啶、或胞嘧啶到尿嘧啶、或其它碱基之间的转换。

[0018] 本发明提供的基因编辑系统编辑活性高,与之前的 Cas9 相比具有明显的优势。

[0019] 本发明通过基因合成、分子克隆、细胞转染、PCR 产物深度测序、生物信息学分析等技术检测 CRISPR/Sa-ShaCas9 系统的编辑效率。

[0020] 本发明提供的 CRISPR/Sa-ShaCas9 基因编辑系统能够在细胞中进行基因编辑,包括通过 Sa-ShaCas9 蛋白和 sgRNA 的复合物识别定位靶向 DNA,对 DNA 进行编辑;最后检测编辑效率。具体步骤为:

[0021] (1) 合成人源化的 Sa-ShaCas9 基因序列;并克隆到表达载体上,得到 pAAV2_ Sa-ShaCas9_ITR;

[0022] (2) 合成 sgRNA 对应的寡核苷酸单链 DNA,即 Oligo-F 和 Oligo-R 序列,退火后连接至质粒 pAAV2_ Sa-ShaCas9_U6_BsaI 的 BsaI 酶切位点,得到 pAAV2_ Sa-ShaCas9-hU6-sgRNA;

[0023] (3) 将表达 Sa-ShaCas9 蛋白、sgRNA 的载体递送至含有靶位点的细胞中;

[0024] (4) 将编辑后的靶位点进行 PCR 扩增,T7EI 酶切或者二代测序检测编辑效率。

[0025] 本发明中,可以根据具体需要,针对待编辑的DNA序列设计任意靶向的sgRNA,并一定程度上对sgRNA进行本领域所熟知的修饰,所述修饰包括但不限于磷酸化、缩短、加长、硫化、甲基化、羟基化。

[0026] 本发明中,可以根据具体需要,递送至细胞的CRISPR/Sa-ShaCas9系统包括但不限于表达Sa-ShaCas9蛋白或sgRNA的质粒、逆转录病毒、腺病毒、腺相关病毒载体或RNA或Sa-ShaCas9蛋白。

[0027] 本领域技术人员可以理解的是,碱基N表示A、T、C或G四种碱基中的任何一种。

[0028] 进一步地,步骤(3)中,所述递送的手段包括但不限于脂质体、阳离子多聚物、纳米颗粒、多功能信封式纳米以及病毒载体。

[0029] 进一步地,步骤(3)中,所述细胞包括但不限于人细胞、动物细胞、植物细胞、细菌细胞、真菌细胞。

[0030] 进一步地,步骤(2)中,所述sgRNA具有SEQ ID NO:2所示的核苷酸序列,或与SEQ ID NO:2所示的核苷酸序列至少80%相同的核苷酸序列,或基于核苷酸序列进行的修饰,修饰包括但不限于磷酸化、缩短、加长、硫化或甲基化。

[0031] 更具体地,在一个具体实施例中,合成sgRNA对应的寡核苷酸单链DNA序列,即Oligo-F和Oligo-R,该寡核苷酸单链DNA序列如下:

[0032] Oligo-F CACCGCTCGGAGATCATCATTGCG, (SEQ ID NO:4)

[0033] Oligo-R AAACCGCAATGATGATCTCCGAGC. (SEQ ID NO:5)

[0034] 更具体地,在一个具体实施例中,本领域技术人员可以理解的是,Oligo-F和Oligo-R需要进行退火变为双链DNA,退火反应体系为1 μ L 100 μ M oligo-F,1 μ L 100 μ M oligo-R,28 μ L 水,震荡混匀后,放置于PCR仪中运行退火程序;退火程序如下:95 $^{\circ}$ C_5min, 85 $^{\circ}$ C_1min, 75 $^{\circ}$ C_1min,65 $^{\circ}$ C_1min,55 $^{\circ}$ C_1min,45 $^{\circ}$ C_1min,35 $^{\circ}$ C_1min,25 $^{\circ}$ C_1min,4 $^{\circ}$ C保存,降温速率0.3 $^{\circ}$ C/s。

[0035] 更具体地,在一个具体实施例中,所述质粒pAAV2_Sa-ShaCas9_ITR需要经过BsaI限制性内切酶(NEB)进行线性化处理。

[0036] 更具体地,在一个具体实施例中,所述Oligo-F和Oligo-R退火后的产物与线性化处理后的pAAV2_Sa-ShaCas9_ITR骨架载体通过DNA连接酶进行连接反应。

[0037] 更具体地,在一个具体实施例中,将连接产物转化至感受态细胞后,经Sanger测序验证正确的克隆,然后提取质粒备用。

[0038] 更具体地,在一个具体实施例中,步骤(3)中的细胞为HEK293T,其包含的靶位点具有SEQ ID NO:6所示的核苷酸序列。

[0039] 更具体地,在一个具体实施例中,步骤(3)中的递送手段为脂质体,包括Lipofectamine[®] 2000或PEI。

[0040] 更具体地,在本发明第一方面的一个具体实施例中,所述实验步骤(4)中PCR的模板为编辑后的HEK293T基因组DNA。

[0041] 更具体地,在一个具体实施例中,步骤(4)中PCR扩增的引物序列为:

[0042] F1-ACACTCTTTCCTACACGACGCTCTCCGATCTNNNNGCGAGAAAAGCCTTGTTT; (SEQ ID NO:7)

[0043] R1-ACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAACTTGTTGGCCGTTTAC; (SEQ ID NO:8)

[0044] F2-AATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGAC; (SEQ ID NO:9)

[0045] R2-CAAGCAGAAGACGGCATAACGATCACTGTGTGACTGGAGTTCAGACGTGTG; (SEQ ID NO:10)

[0046] 本发明还提供一种用于基因编辑的CRISPR/Sa-ShaCas9系统试剂盒,该试剂盒包括 Sa-ShaCas9蛋白或靶向DNA序列的sgRNA或靶向的DNA。

[0047] 本发明还提供CRISPR/Sa-ShaCas9基因编辑系统的应用,包括基因敲除、定点碱基的改变、定点插入、基因转录水平的调控、DNA甲基化调控、DNA乙酰化修饰、组蛋白乙酰化修饰、单碱基转换器或染色质成像追踪。

附图说明

[0048] 图1为CRISPR/Sa-ShaCas9基因编辑系统切割靶向DNA示意图。其中,灰色椭圆形表示Sa-ShaCas9蛋白,黑色弯曲状表示sgRNA序列,基因组上链中加深区域表示PAM序列NNGG。

[0049] 图2为质粒pAAV2_Sa-ShaCas9_U6_BsaI图谱示意图。其中,包括AAV2 ITR、CMV增强子、CMV启动子、SV40 NLS、Sa-ShaCas9、nucleoplasmin NLS、3x HA、bGH poly(A)、人 U6启动子(hU6)、BsaI内切酶位点、sgRNA scaffold序列等元件。

[0050] 图3为靶位点DNA序列被编辑后的部分二代测序结果。其中编辑结果有缺失、插入或错配,最后5bp表示PAM序列NNGRM。

[0051] 图4为内源位点的T7 Endonuclease I酶切结果。其中,箭头表示切开的片段大小。

具体实施方式

[0052] 以下将通过实施例进一步说明本发明,但实施例并不对本发明做任何形式的限定。

[0053] 除非特别说明,本发明采用的试剂、方法和设备为本技术领域常规试剂、方法和设备。除非特别说明,以下实施例所用试剂和材料均为市购。未注明具体条件的实验方法,通常按照常规条件,或制造厂商所建议条件实施。

[0054] 在一具体实施方式中,本发明提供的CRISPR/Sa-ShaCas9系统为一种新的基因编辑系统、方法、试剂盒及其应用。

[0055] 在本发明一具体实施例中,CRISPR/Sa-ShaCas9系统能够在细胞中进行基因编辑,所述方法包括如下步骤:

[0056] 1、构建质粒pAAV2_Sa-ShaCas9_ITR

[0057] 步骤(1),根据Sa-ShaCas9基因在UniProt上的检索号A0A2T4SLN6下载的ShaCas9蛋白序列加以改进,得到Sa-ShaCas9基因,其氨基酸序列为SEQ ID NO:1所示。

[0058] 步骤(2),将Sa-ShaCas9的氨基酸序列进行密码子优化,获得了Sa-ShaCas9在人细胞中高表达的编码序列,如SEQ ID NO:11所示。

[0059] 步骤(3),将获得基因Sa-ShaCas9的编码序列在公司进行基因合成,并构建至pAAV2_ITR骨架质粒上,得到质粒pAAV2_Sa-ShaCas9_ITR,如图2所示。

[0060] 2、制备线性化质粒pAAV2_Sa-ShaCas9_ITR

[0061] 步骤(1),用BsaI限制性内切酶将质粒pAAV2_Sa-ShaCas9_ITR进行酶切线性化,酶切体系:1μg质粒pAAV2_Sa-ShaCas9_ITR,5μL 10xCutSmart缓冲液,1μLBsaI内切酶,水补足

至50 μ l, 37 $^{\circ}$ C反应1小时。

[0062] 步骤(2), 将酶切后的产物在1%的琼脂糖凝胶上电泳, 120V电泳30分钟。

[0063] 步骤(3), 切除7430bp大小的DNA片段, 用胶回收试剂盒依据厂家提供的步骤进行回收, 最终用超纯水进行洗脱。

[0064] 步骤(4), 将回收的线性化质粒pAAV2_Sa-ShaCas9_ITR用NanoDrop测定DNA浓度, 备用或置于-20 $^{\circ}$ C进行长期保存。

[0065] 3、构建质粒pAAV2_Sa-ShaCas9-hU6-sgRNA

[0066] 步骤(1), 设计sgRNA序列。

[0067] 步骤(2), 在设计sgRNA序列对用的正义链和反义链上分别加上线性化质粒pAAV2_Sa-ShaCas9_ITR两侧对应的粘性末端序列, 并在公司合成两条寡核苷酸单链DNA, 具体序列为:

[0068] Oligo-F CACCGCTCGGAGATCATCATTGCG, (SEQ ID NO:4)

[0069] Oligo-R AAACCGCAATGATGATCTCCGAGC。(SEQ ID NO:5)

[0070] 步骤(3), 寡核苷酸单链DNA进行退火变为双链DNA, 退火反应体系: 1 μ L 100 μ Moligo-F, 1 μ L 100 μ Moligo-R, 28 μ L水, 震荡混匀后, 放置于PCR仪中运行退火程序: 95 $^{\circ}$ C_5min, 85 $^{\circ}$ C_1min, 75 $^{\circ}$ C_1min, 65 $^{\circ}$ C_1min, 55 $^{\circ}$ C_1min, 45 $^{\circ}$ C_1min, 35 $^{\circ}$ C_1min, 25 $^{\circ}$ C_1min, 4 $^{\circ}$ C保存, 降温速率0.3 $^{\circ}$ C/s。

[0071] 步骤(4), 将退火后的产物与线性化质粒pAAV2_Sa-ShaCas9_ITR在DNA连接酶的作用下依据产品提供的步骤进行连接。

[0072] 步骤(5), 取1 μ L连接产物进行化学感受态转化, 将生长的细菌克隆进行Sanger测序验证。

[0073] 步骤(6), 将测序验证连接正确的克隆摇菌, 提取质粒pAAV2_Sa-ShaCas9-hU6-sgRNA 备用。

[0074] 4、转染表达Sa-ShaCas9蛋白和sgRNA的质粒pAAV2_Sa-ShaCas9-hU6-sgRNA

[0075] 步骤(1), 第0天, 根据转染所需, 将含有sgRNA靶向位点的HEK293T细胞系在6孔板进行铺板, 细胞密度约30%左右, 靶位点序列如SEQ ID NO:6所示。

[0076] 步骤(2), 第1天, 进行转染, 转染体系如下,

[0077] i. 取2 μ g待转染质粒pAAV2_Sa-ShaCas9-hU6-sgRNA加入至100 μ l opti-MEM培养基中, 轻轻吹打混匀;

[0078] ii. 将Lipofectamine[®]2000轻弹混匀, 吸取5 μ l加入至100 μ l opti-MEM培养基中, 轻轻混匀, 室温静置5min;

[0079] iii. 将稀释的Lipofectamine[®]2000和稀释的质粒进行混合, 轻轻吹打混匀, 室温静置20min, 然后加入至待转染细胞的培养基中。

[0080] 步骤(3), 将细胞置于37 $^{\circ}$ C, 5%的CO₂培养箱中继续培养。

[0081] 5、制备二代测序文库

[0082] 步骤(1), 收集编辑3天后的HEK293T细胞, 用DNA试剂盒依据产品提供的步骤提取基因组DNA。

[0083] 步骤(2), 进行PCR建库第一轮PCR, 用2xQ5 Mastermix进行PCR反应, PCR引物如SEQ ID NO:7-SEQ ID NO:8所示, 反应体系如下:

	试剂	25 μ l 体系
[0084]	10 μ MF1	1.25 μ l
	10 μ MR1	1.25 μ l
	2X Q5 master mix	12.5 μ l
	基因组 DNA	5 μ g
	ddH2O	补足至 25 μ l

[0085] PCR运行程序如下:

	初始变性	98 °C	2 min
		95 °C	7 s
[0086]	35 个循环数	65 °C	20 s
		72 °C	10 s
	终延伸	72 °C	2 min
	保存	4 °C	∞

[0087] 步骤(3),进行PCR建库第二轮PCR,用2xQ5 Mastermix进行PCR反应,PCR引物如SEQ ID NO:9--SEQ ID NO:10所示,反应体系如下:

	试剂	25 μ l 体系
[0088]	10 μ MF2	1.25 μ l
	10 μ MR2	1.25 μ l
	2X Q5 master mix	12.5 μ l
	第一轮产物	3 μ l
	ddH2O	7 μ l

[0089] PCR运行程序如下:

	初始变性	98 °C	30 s
		95 °C	7 s
[0090]	15 个循环数	65 °C	20 s
		72 °C	10 s
[0091]	终延伸	72 °C	2 min
	保存	4 °C	∞

[0092] 步骤(4),将第二轮的PCR产物用胶回收试剂盒依据厂家提供的步骤,纯化366bp大小的DNA片段,二代测序文库制备完毕。

[0093] 6、二代测序结果分析

[0094] 步骤(1),将制备好的二代测序文库交由公司在HiSeqXTen上进行双端测序。

[0095] 步骤(2)生物信息学分析二代测序结果,部分编辑结果如图2和图3所示。

[0096] 7、内源位点验证

[0097] 步骤(1),将表达Sa-ShaCas9和sgRNA的质粒pAAV2_Sa-ShaCas9-hU6-sgRNA通过Lipofectamine® 2000依据厂家提供的步骤转染至HEK293T细胞中,其中,

[0098] sgRNA序列为:AGTGAGGGGAACAAAGTGGAC,(SEQ ID NO:12)

[0099] 靶位点的具体序列为:AGTGAGGGGAACAAAGTGGACATGGC;(SEQ ID NO:13)

[0100] 步骤(2),提取编辑5天后的细胞基因组DNA,通过2x Q5 Master mix,用引物Test-F 和Test-R将靶向DNA序列进行扩增;其中:

[0101] Test-F的具体序列为:GCTGCTTTCTGCTGTCTTC,(SEQ ID NO:14)

[0102] Test-R的具体序列为:AAGGAGGTCTCTGTCTGTGC;(SEQ ID NO:15)

[0103] 步骤(3),将PCR产物通过琼脂糖凝胶回收,纯化284bp大小的DNA片段;

[0104] 步骤(4),依照T7 Endonuclease I的说明书对纯化的DNA片段进行酶切,然后跑胶检测,结果如图4 所示,左侧为阴性对照组,转染时无sgRNA,T7 Endonuclease I切靶向序列后无切开的片段,说明未发生编辑;右侧为实验组,转染时有sgRNA,T7 Endonuclease I切靶向序列后出现切开的片段,说明发生了编辑。

[0001]	SEQUENCE LISTING																			
[0002]	<110> 复旦大学																			
[0003]	<120> CRISPR/Sa-ShaCas9基因编辑系统及其应用																			
[0004]	<130> 1115																			
[0005]	<160> 15																			
[0006]	<170> PatentIn version 3.5																			
[0007]	<210> 1																			
[0008]	<211> 1055																			
[0009]	<212> PRT																			
[0010]	<213> 人工序列(Artificial Sequence)																			
[0011]	<400> 1																			
[0012]	Met	Lys	Arg	Asn	Tyr	Ile	Leu	Gly	Leu	Asp	Ile	Gly	Ile	Thr	Ser	Val				
[0013]	1			5					10					15						
[0014]	Gly	Tyr	Gly	Ile	Ile	Asp	Tyr	Glu	Thr	Arg	Asp	Val	Ile	Asp	Ala	Gly				
[0015]				20					25					30						
[0016]	Val	Arg	Leu	Phe	Lys	Glu	Ala	Asn	Val	Glu	Asn	Asn	Glu	Gly	Arg	Arg				
[0017]				35				40					45							
[0018]	Ser	Lys	Arg	Gly	Ala	Arg	Arg	Leu	Lys	Arg	Arg	Arg	Arg	His	Arg	Ile				
[0019]		50						55					60							
[0020]	Gln	Arg	Val	Lys	Lys	Leu	Leu	Phe	Asp	Tyr	Asn	Leu	Leu	Thr	Asp	His				
[0021]		65				70					75				80					
[0022]	Ser	Glu	Leu	Ser	Gly	Ile	Asn	Pro	Tyr	Glu	Ala	Arg	Val	Lys	Gly	Leu				
[0023]					85					90				95						
[0024]	Ser	Gln	Lys	Leu	Ser	Glu	Glu	Glu	Phe	Ser	Ala	Ala	Leu	Leu	His	Leu				
[0025]				100						105				110						
[0026]	Ala	Lys	Arg	Arg	Gly	Val	His	Asn	Val	Asn	Glu	Val	Glu	Glu	Asp	Thr				
[0027]				115					120					125						
[0028]	Gly	Asn	Glu	Leu	Ser	Thr	Lys	Glu	Gln	Ile	Ser	Arg	Asn	Ser	Lys	Ala				
[0029]		130						135					140							
[0030]	Leu	Glu	Glu	Lys	Tyr	Val	Ala	Glu	Leu	Gln	Leu	Glu	Arg	Leu	Lys	Lys				
[0031]		145				150					155				160					
[0032]	Asp	Gly	Glu	Val	Arg	Gly	Ser	Ile	Asn	Arg	Phe	Lys	Thr	Ser	Asp	Tyr				
[0033]					165					170				175						
[0034]	Val	Lys	Glu	Ala	Lys	Gln	Leu	Leu	Lys	Val	Gln	Lys	Ala	Tyr	His	Gln				
[0035]				180						185				190						
[0036]	Leu	Asp	Gln	Ser	Phe	Ile	Asp	Thr	Tyr	Ile	Asp	Leu	Leu	Glu	Thr	Arg				
[0037]				195					200					205						
[0038]	Arg	Thr	Tyr	Tyr	Glu	Gly	Pro	Gly	Glu	Gly	Ser	Pro	Phe	Gly	Trp	Lys				
[0039]				210					215					220						
[0040]	Asp	Ile	Lys	Glu	Trp	Tyr	Glu	Met	Leu	Met	Gly	His	Cys	Thr	Tyr	Phe				
[0041]		225						230					235			240				

[0042]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
[0043]	245 250 255
[0044]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
[0045]	260 265 270
[0046]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
[0047]	275 280 285
[0048]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
[0049]	290 295 300
[0050]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
[0051]	305 310 315 320
[0052]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
[0053]	325 330 335
[0054]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
[0055]	340 345 350
[0056]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
[0057]	355 360 365
[0058]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
[0059]	370 375 380
[0060]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
[0061]	385 390 395 400
[0062]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
[0063]	405 410 415
[0064]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
[0065]	420 425 430
[0066]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
[0067]	435 440 445
[0068]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
[0069]	450 455 460
[0070]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
[0071]	465 470 475 480
[0072]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
[0073]	485 490 495
[0074]	Arg Asn Ala Ala Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
[0075]	500 505 510
[0076]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
[0077]	515 520 525
[0078]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
[0079]	530 535 540
[0080]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
[0081]	545 550 555 560
[0082]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
[0083]	565 570 575

[0084]	Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
[0085]	580 585 590
[0086]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
[0087]	595 600 605
[0088]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
[0089]	610 615 620
[0090]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
[0091]	625 630 635 640
[0092]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Ala Ala Leu
[0093]	645 650 655
[0094]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
[0095]	660 665 670
[0096]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
[0097]	675 680 685
[0098]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
[0099]	690 695 700
[0100]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
[0101]	705 710 715 720
[0102]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
[0103]	725 730 735
[0104]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
[0105]	740 745 750
[0106]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
[0107]	755 760 765
[0108]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
[0109]	770 775 780
[0110]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
[0111]	785 790 795 800
[0112]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
[0113]	805 810 815
[0114]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
[0115]	820 825 830
[0116]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
[0117]	835 840 845
[0118]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
[0119]	850 855 860
[0120]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
[0121]	865 870 875 880
[0122]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
[0123]	885 890 895
[0124]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Ser Tyr
[0125]	900 905 910

[0126] Arg Phe Asp Val Tyr Leu Thr Asp Lys Gly Tyr Lys Phe Val Ser Ile
 [0127] 915 920 925
 [0128] Thr Tyr Leu Asp Val Leu Lys Lys Glu Asn Tyr Tyr Tyr Ile Ser Glu
 [0129] 930 935 940
 [0130] Ala Lys Tyr Asp Lys Leu Lys Leu Asn Lys Gly Ile Asp Asp Lys Ala
 [0131] 945 950 955 960
 [0132] Lys Phe Ile Gly Ser Phe Tyr Tyr Asn Asp Leu Ile Glu Leu Asp Gly
 [0133] 965 970 975
 [0134] Glu Val Tyr Thr Val Ile Gly Val Asn Asn Asp Lys Asn Asn Val Ile
 [0135] 980 985 990
 [0136] Glu Leu Asn Leu Pro Glu Ile Arg Tyr Lys Glu Tyr Cys Glu Ile Asn
 [0137] 995 1000 1005
 [0138] Asn Ile Lys Gly Ser Gly Arg Leu Arg Ile Thr Ile Gly Lys Lys
 [0139] 1010 1015 1020
 [0140] Val Asn Ser Ile Arg Lys Leu Ser Thr Asp Val Leu Gly Asn Arg
 [0141] 1025 1030 1035
 [0142] Tyr Tyr Gln Ser Phe Ala Lys Lys Pro Gln Leu Val Phe Lys Lys
 [0143] 1040 1045 1050
 [0144] Gly Ile
 [0145] 1055
 [0146] <210> 2
 [0147] <211> 101
 [0148] <212> RNA
 [0149] <213> 人工序列(Artificial Sequence)
 [0150] <220>
 [0151] <221> misc_feature
 [0152] <222> (1)..(20)
 [0153] <223> n is a, c, g, or u
 [0154] <400> 2
 [0155] nnnnnnnnnn nnnnnnnnnn guuuuaguac ucuggaaaca gaaucuacua aaacaaggca 60
 [0156] aaauccgug uuuaucucgu caacuuguug gcgagauuuu u 101
 [0157] <210> 3
 [0158] <211> 26
 [0159] <212> DNA
 [0160] <213> 人工序列(Artificial Sequence)
 [0161] <220>
 [0162] <221> misc_feature
 [0163] <222> (1)..(23)
 [0164] <223> n is a, c, g, or t
 [0165] <400> 3
 [0166] nnnnnnnnnn nnnnnnnnnn nnngrm 26
 [0167] <210> 4

- [0168] <211> 24
[0169] <212> DNA
[0170] <213> 人工序列 (Artificial Sequence)
[0171] <400> 4
[0172] caccgctcgg agatcatcat tgcg 24
[0173] <210> 5
[0174] <211> 24
[0175] <212> DNA
[0176] <213> 人工序列 (Artificial Sequence)
[0177] <400> 5
[0178] aaaccgcaat gatgatctcc gagc 24
[0179] <210> 6
[0180] <211> 25
[0181] <212> DNA
[0182] <213> 人工序列 (Artificial Sequence)
[0183] <220>
[0184] <221> misc_feature
[0185] <222> (21) .. (25)
[0186] <223> n is a, c, g, or t
[0187] <400> 6
[0188] gctcggagat catcattgcg nnnnn 25
[0189] <210> 7
[0190] <211> 55
[0191] <212> DNA
[0192] <213> 人工序列 (Artificial Sequence)
[0193] <220>
[0194] <221> misc_feature
[0195] <222> (34) .. (37)
[0196] <223> n is a, c, g, or t
[0197] <400> 7
[0198] acactctttc cctacacgac gctcttccga tctnnnngcg agaaaagcct tgttt 55
[0199] <210> 8
[0200] <211> 51
[0201] <212> DNA
[0202] <213> 人工序列 (Artificial Sequence)
[0203] <400> 8
[0204] actggagtgc agacgtgtgc tcttccgatc tctgaacttg tggccgttta c 51
[0205] <210> 9
[0206] <211> 45
[0207] <212> DNA
[0208] <213> 人工序列 (Artificial Sequence)
[0209] <400> 9

[0210]	aatgatacgg cgaccaccga gatctacact ctttcctac acgac	45
[0211]	<210>	10
[0212]	<211>	52
[0213]	<212>	DNA
[0214]	<213>	人工序列(Artificial Sequence)
[0215]	<400>	10
[0216]	caagcagaag acggcatacg agatcaactgt gtgactggag ttcagacgtg tg	52
[0217]	<210>	11
[0218]	<211>	3165
[0219]	<212>	DNA
[0220]	<213>	人工序列(Artificial Sequence)
[0221]	<400>	11
[0222]	atgaagcgga actacatcct gggcctggac atcggcatca ccagcgtggg ctacggcatc	60
[0223]	atcgactacg agacacggga cgtgatcgat gccggcgtgc ggctgttcaa agaggccaac	120
[0224]	gtggaaaaca acgaggcag gcggagcaag agaggcgcca gaaggctgaa gcggcggagg	180
[0225]	cggcatagaa tccagagagt gaagaactg ctgttcgact acaacctget gaccgaccac	240
[0226]	agcgagctga gcggcatcaa cccctacgag gccagagtga agggcctgag ccagaagctg	300
[0227]	agcgaggaag agttctctgc cgccctgctg cacctggcca agagaagagg cgtgcacaac	360
[0228]	gtgaacgagg tggaagagga caccggcaac gagctgtcca ccaaagagca gatcagccgg	420
[0229]	aacagcaagg ccctggaaga gaaatactg gccgaactgc agctggaacg gctgaagaaa	480
[0230]	gacggcgaag tgcggggcag catcaacaga ttcaagacca gcgactacgt gaaagaagcc	540
[0231]	aaacagctgc tgaagtgca gaaggcctac caccagctgg accagagctt catcgacacc	600
[0232]	tacatcgacc tgctggaaac ccggcggacc tactatgagg gacctggcga gggcagcccc	660
[0233]	ttcggtgga aggacatcaa agaatggtac gagatgctga tgggccactg cacctacttc	720
[0234]	cccaggaac tgcggagcgt gaagtacgc tacaacgcc acctgtacaa cgccctgaac	780
[0235]	gacctgaaca atctcgtgat caccaggac gagaacgaga agctggaata ttacgagaag	840
[0236]	ttccagatca tcgagaacgt gttcaagcag aagaagaagc ccaccctgaa gcagatcgcc	900
[0237]	aaagaaatcc tcgtgaacga agaggatatt aagggtaca gaggtagaccag caccggcaag	960
[0238]	cccagttca ccaacctgaa ggtgtaccac gacatcaagg acattaccgc ccggaagag	1020
[0239]	attattgaga acgccagct gctggatcag attgccaaga tcctgaccat ctaccagagc	1080
[0240]	agcgaggaca tccaggaaga actgaccaat ctgaactccg agctgacca ggaagagatc	1140
[0241]	gagcagatct ctaatctgaa gggctatacc ggcaccaca acctgagcct gaaggccatc	1200
[0242]	aacctgatcc tggacgagct gtggcacacc aacgacaacc agatcgetat cttcaaccgg	1260
[0243]	ctgaagctgg tgcccaagaa ggtggacctg tcccagcaga aagagatccc caccacctg	1320
[0244]	gtggacgact tcattcctgag ccccgctcgt aagagaagct tcattccagag catcaaagt	1380
[0245]	atcaacgcca tcattcaagaa gtacggcctg cccaacgaca tcattatcga gctggcccgc	1440
[0246]	gagaagaact ccaaggacgc ccagaaaatg atcaacgaga tgcagaagcg gaacgccgc	1500
[0247]	accaacgagc ggatcgagga aatcatccgg accaccgca aagagaacgc caagtacctg	1560
[0248]	atcgagaaga tcaagctgca cgacatgcag gaaggcaagt gcctgtacag cctggaagcc	1620
[0249]	atccctctgg aagatctgct gaacaacccc ttcaactatg aggtggacca catcatcccc	1680
[0250]	agaagcgtgt ccttcgaaa cagcttcaac aacaaggtgc tcgtgaagca ggaagaaaac	1740
[0251]	agcaagaagg gcaaccggac ccattccag tacctgagca gcagcgacag caagatcagc	1800

[0252]	tacgaaacct tcaagaagca catcctgaat ctggccaagg gcaagggcag aatcagcaag	1860
[0253]	accaagaaag agtatctgct ggaagaacgg gacatcaaca ggttctccgt gcagaaagac	1920
[0254]	ttcatcaacc ggaacctggt ggataccaga tacgccaccg ccgccctgat gaacctgctg	1980
[0255]	cggagctact tcagagtga caacctggac gtgaaagtga agtccatcaa tggcggcttc	2040
[0256]	accagctttc tgcggcggaa gtggaagttt aagaaagagc ggaacaaggg gtacaagcac	2100
[0257]	cacgccgagg acgccctgat cattgccaac gccgatttca tcttcaaaga gtggaagaaa	2160
[0258]	ctggacaagg ccaaaaaagt gatggaaaac cagatgttcg aggaaaagca ggccgagagc	2220
[0259]	atgcccgaga tcgaaaccga gcaggagtac aaagagatct tcatcacccc ccaccagatc	2280
[0260]	aagcacatta aggacttcaa ggactacaag tacagccacc ggggtggacaa gaagccta	2340
[0261]	agagagctga ttaacgacac cctgtactcc acccggaagg acgacaaggg caacaccctg	2400
[0262]	atcgtgaaca atctgaacgg cctgtacgac aaggacaatg acaagctgaa aaagctgatc	2460
[0263]	aacaagagcc ccgaaaagct gctgatgtac caccacgacc cccagaccta ccagaaactg	2520
[0264]	aagctgatta tggaacagta cggcgacgag aagaatcccc tgtacaagta ctacgaggaa	2580
[0265]	accgggaact acctgaccaa gtactccaaa aaggacaacg gccccgtgat caagaagatt	2640
[0266]	aagtattacg gcaacaaact gaacgcccat ctggacatca ccgacgacta cccaacagc	2700
[0267]	agaaacaagg tcgtgaagct gtccctgaag agctaccgct tcgacgtgta cctgaccgac	2760
[0268]	aagggtaca agttcgtgag catcacctac ctggacgtgc tgaagaagga gaactactac	2820
[0269]	tacatcagcg aggccaagta cgacaagctg aagctgaaca agggcatcga cgacaaggcc	2880
[0270]	aagttcatcg gcagcttcta ctacaacgac ctgatagagc tggacggcga ggtgtacacc	2940
[0271]	gtgataggcg tgaacaacga caagaacaac gtgatagagc tgaacctgcc cgagatacgc	3000
[0272]	tacaaggagt actgcgagat aaacaacatc aagggcagcg gcaggctgcg catcaccatc	3060
[0273]	ggcaagaagg tgaacagcat ccgcaagctg agcaccgacg tgctgggcaa ccgctactac	3120
[0274]	cagagcttcg ccaagaagcc ccagctggtg ttcaagaagg gcac	3165
[0275]	<210> 12	
[0276]	<211> 21	
[0277]	<212> DNA	
[0278]	<213> 人工序列(Artificial Sequence)	
[0279]	<400> 12	
[0280]	agtgagggga acaaagtga c 21	
[0281]	<210> 13	
[0282]	<211> 26	
[0283]	<212> DNA	
[0284]	<213> 人工序列(Artificial Sequence)	
[0285]	<400> 13	
[0286]	agtgagggga acaaagtga catggc 26	
[0287]	<210> 14	
[0288]	<211> 20	
[0289]	<212> DNA	
[0290]	<213> 人工序列(Artificial Sequence)	
[0291]	<400> 14	
[0292]	gctgctttcc tgctgtcttc 20	
[0293]	<210> 15	

-
- [0294] <211> 20
[0295] <212> DNA
[0296] <213> 人工序列(Artificial Sequence)
[0297] <400> 15
[0298] aaggaggtct ctgtctgtgc 20

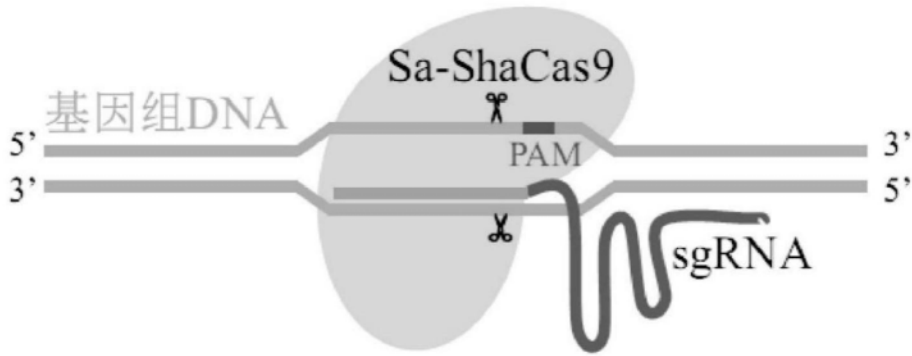


图1

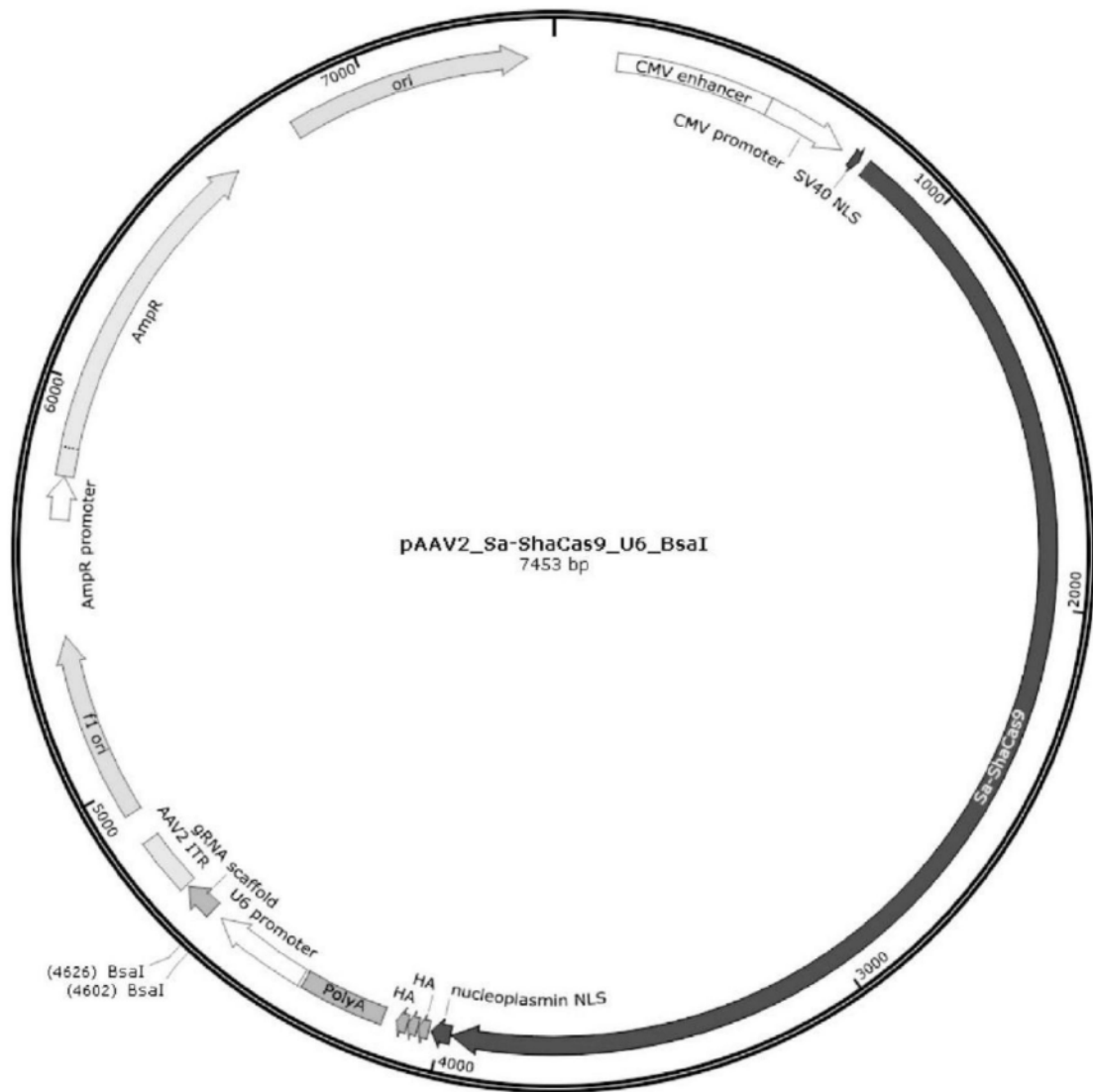


图2

ATGGAACGGCTCGGAGATCATCATTGCG NNGRM
ATGGAACGGCTCGGAGA.....GCG TGGGC
ATGGAACGGCTCGGAGATCACAC....GCG CGGGC
ATGGAACGGCTCGGAGATCATCA...GCG CGGAC
ATGGAACGGCTCGGAGATCT.....GCG CTGGA
ATGGAACGGCTCGGAGATCATCATTAGCG CCGGC
ATGGAACGGCTCGGAGATCATCATTGCG GAGGC

图3

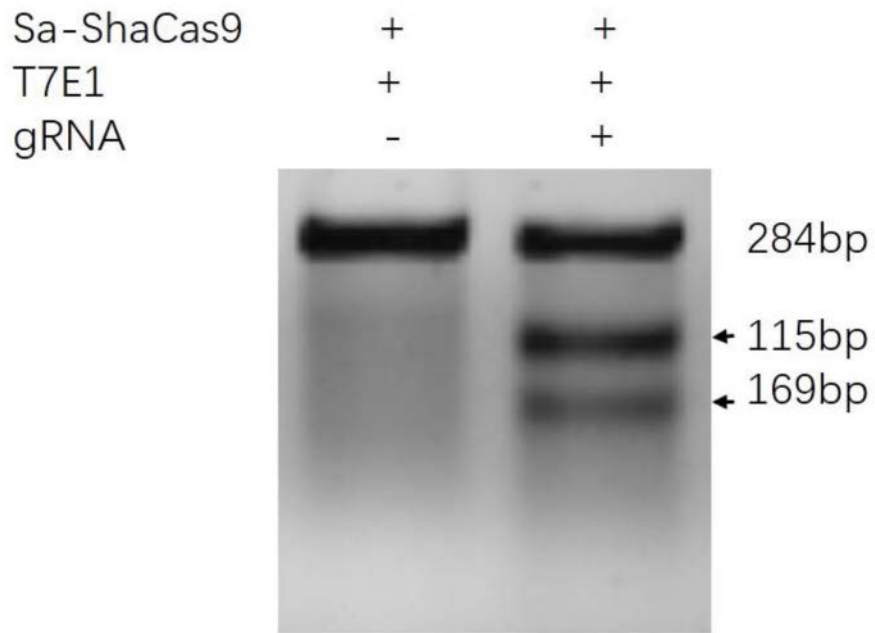


图4