



## (12) 发明专利

(10) 授权公告号 CN 105247066 B

(45) 授权公告日 2020.10.20

(21) 申请号 201480027950.1

A61K 38/43 (2006.01)

(22) 申请日 2014.03.14

A61K 38/47 (2006.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 105247066 A

(43) 申请公布日 2016.01.13

(30) 优先权数据

61/799,647 2013.03.15 US

61/838,178 2013.06.21 US

61/838,148 2013.06.21 US

61/921,007 2013.12.26 US

(85) PCT国际申请进入国家阶段日

2015.11.13

(86) PCT国际申请的申请数据

PCT/US2014/028630 2014.03.14

(87) PCT国际申请的公布数据

W02014/144288 EN 2014.09.18

(73) 专利权人 通用医疗公司

地址 美国马萨诸塞州

(72) 发明人 J.K.乔昂格 S.特赛

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 张文辉

(51) Int. Cl.

C12P 21/06 (2006.01)

C12P 19/34 (2006.01)

(56) 对比文件

WO 2012093833 A2, 2012.07.12

US 2010055793 A1, 2010.03.04

US 2015071899 A1, 2015.03.12

WO 2014127287 A1, 2014.08.21

WO 2013176772 A1, 2013.11.28

Yanfang Fu等.Targeted Genome Editing in Human Cells Using CRISPR/Cas Nucleases and Truncated Guide RNAs.《Methods Enzymol》.2014,第546卷第21-45页.

Yanfang Fu等.Targeted Genome Editing in Human Cells Using CRISPR/Cas Nucleases and Truncated Guide RNAs.《Methods Enzymol》.2014,第546卷第21-45页.

Lei S.Qi等.Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression.《Cell》.2013,第152卷(第5期),第1173-1183页.

Martin Jinek等.A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.《Science》.2012,第337卷(第6096期),第816-821页.

审查员 陈晋

权利要求书1页 说明书51页

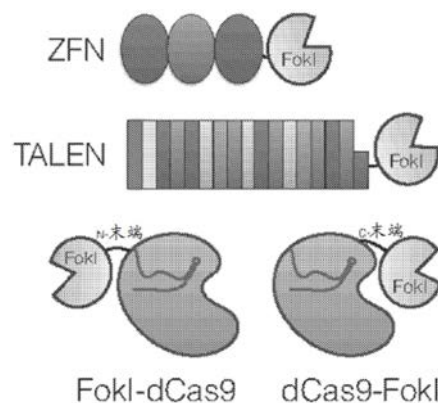
序列表126页 附图28页

(54) 发明名称

使用RNA引导的FokI核酸酶(RFN)提高RNA引导的基因组编辑的特异性

(57) 摘要

用于提高RNA引导的基因组编辑例如使用CRISPR/Cas9系统的编辑的特异性的方法。



1. 一种RNA引导的FokI核酸酶 (RFN) 融合蛋白, 其是无催化活性的CRISPR相关9 (dCas9) 与FokI催化结构域序列融合的融合蛋白, 其中所述dCas9通过间插接头在dCas9的氨基末端与所述FokI催化结构域序列融合, 且其中所述dCas9在D10、E762、H983或D986处; 和在H840或N863处包含突变, 其中所述dCas9来源于化脓链球菌。

2. 根据权利要求1所述的融合蛋白, 其中所述接头由2-30个氨基酸组成。

3. 根据权利要求2所述的融合蛋白, 其中所述接头包含Gly<sub>4</sub>Ser。

4. 根据权利要求1所述的融合蛋白, 其中所述FokI催化结构域包含SEQ ID NO:4的氨基酸388-583或408-583。

5. 根据权利要求1所述的融合蛋白, 其中所述dCas9在如下位置处包含突变:

(i) D10A或D10N; 和

(ii) H840A、H840Y或H840N。

6. 一种核酸, 其编码权利要求1-5中任一项所述的融合蛋白。

7. 一种载体, 其包含权利要求6的核酸。

8. 一种宿主细胞, 其表达权利要求1-5中任一项所述的融合蛋白。

9. 一种诱导细胞中的基因组序列中的序列特异性断裂的方法, 所述方法包括在所述细胞中表达权利要求1-5中任一项所述的RNA引导的FokI核酸酶 (RFN) 融合蛋白和引导RNA, 或将所述细胞与权利要求1-5中任一项所述的RNA引导的FokI核酸酶 (RFN) 融合蛋白和引导RNA接触, 所述引导RNA将所述RFN导向两个靶基因组序列, 其中所述两个靶基因组序列相隔10-20个核苷酸, 且其中所述两个靶基因组序列各自在3' 末端具有PAM序列。

10. 根据权利要求9所述的方法, 其中所述两个靶基因组序列相隔13-17个碱基对。

11. 根据权利要求9所述的方法, 其中所述引导RNA为:

(a) 两个单一引导RNA, 其中一个单一引导RNA靶向第一链, 并且另一个引导RNA靶向互补链, 并且FokI切割每一条链, 在相对的DNA链上导致一对切口, 从而产生双链断裂, 或

(b) tracrRNA和两个crRNA, 其中一个crRNA靶向第一链, 并且另一个crRNA靶向互补链, 并且FokI切割每一条链, 在相对的DNA链上导致一对切口, 从而产生双链断裂。

12. 根据权利要求9所述的方法, 其中所述两个引导RNA中的每一个包括与靶基因组序列的17-20个核苷酸互补的互补区。

13. 根据权利要求9-12中任一项所述的方法, 其中在所述两个靶序列之间诱导插入缺失突变。

14. 根据权利要求9-12中任一项所述的方法, 其中细胞中的RNA引导的基因组编辑的特异性相比于CRISPR/Cas9 RNA引导的基因组编辑得以提高。

15. 根据权利要求13所述的方法, 其中细胞中的RNA引导的基因组编辑的特异性相比于CRISPR/Cas9 RNA引导的基因组编辑得以提高。

16. 一种提高细胞中的RNA引导的基因组编辑的特异性的方法, 所述方法包括将所述细胞与权利要求1-5中任一项所述的RNA引导的FokI核酸酶 (RFN) 融合蛋白接触。

## 使用RNA引导的FokI核酸酶 (RFN) 提高RNA引导的基因组编辑的特异性

[0001] 优先权申明

[0002] 本申请依据35USC§119(e),要求2013年3月15日提交的美国专利申请系列第61/799,647号、2013年6月21日提交的美国专利申请系列第61/838,178号、2013年6月21日提交的美国专利申请系列第61/838,148号和2013年12月26日提交的美国专利申请系列第61/921,007号的优先权。前述专利申请系列号的完整内容在此通过引用并入。

[0003] 联邦资助的研究或开发

[0004] 本发明是在由美国国家卫生研究院授予的基金第DP1GM105378号下借助政府资助进行的。政府具有本发明的某些权利。

### 技术领域

[0005] 用于使用RNA引导的FokI核酸酶 (RFN) 例如FokI-dCas9融合蛋白提高RNA引导的基因组编辑,例如使用CRISPR/Cas9系统的编辑的特异性的方法。

[0006] 背景

[0007] 最近的工作已显示成簇规律间隔短回文重复 (CRISPR) /CRISPR-相关 (Cas) 系统 (Wiedenheft等,Nature 482,331-338 (2012);Horvath等,Science 327,167-170 (2010);Terns等,Curr Opin Microbiol 14,321-327 (2011)) 可用作在细菌、酵母和人细胞中以及在体内在完整生物体诸如果蝇、斑马鱼和小鼠中的基础基因组编辑 (Wang等,Cell 153,910-918 (2013);Shen等,Cell Res (2013);Dicarlo等,Nucleic Acids Res (2013);Jiang等,Nat Biotechnol 31,233-239 (2013);Jinek等,Elife 2,e00471 (2013);Hwang等,Nat Biotechnol 31,227-229 (2013);Cong等,Science 339,819-823 (2013);Mali等,Science 339,823-826 (2013c);Cho等,Nat Biotechnol 31,230-232 (2013);Gratz等,Genetics 194 (4):1029-35 (2013))。来自化脓性链球菌 (*S.pyogenes*) 的Cas9核酸酶 (在下文中简称为Cas9) 可通过工程化gRNA的前20个核苷酸与紧接前间区序列邻近基序 (PAM) 例如匹配序列NGG或NAG的PAM之后的目标靶基因组DNA序列的互补链之间的碱基对互补序列来引导 (Shen等,Cell Res (2013);Dicarlo等,Nucleic Acids Res (2013);Jiang等,Nat Biotechnol 31,233-239 (2013);Jinek等,Elife 2,e00471 (2013);Hwang等,Nat Biotechnol 31,227-229 (2013);Cong等,Science 339,819-823 (2013);Mali等,Science 339,823-826 (2013c);Cho等,Nat Biotechnol 31,230-232 (2013);Jinek等,Science 337,816-821 (2012))。在体外 (Jinek等,Science 337,816-821 (2012)) 于细菌中 (Jiang等,Nat Biotechnol 31,233-239 (2013)) 和人细胞中 (Cong等,Science 339,819-823 (2013)) 进行的先前研究已显示Cas9-介导的裂解在一些情况下可被gRNA/靶位点交界处上,特别地位于20个核苷酸 (nt) 的gRNA互补区的3'末端中的最后10-12个核苷酸 (nt) 中的单个错配消除。

[0008] 概述

[0009] 许多研究已显示CRISPR-Cas核酸酶可耐受多达5个错配并且仍然裂解;难以预测任何给定的单个错配或错配的组合对活性的影响。综上所述,这些核酸酶可显示显著的脱

靶效应,但预测这些位点可具有挑战性。本文中描述了用于使用CRISPR/Cas系统,例如使用RNA引导的FokI核酸酶(RFN),例如,FokI-Cas9或基于FokI-dCas9的融合蛋白提高基因组编辑的特异性的方法。

[0010] 在第一方面,本发明提供FokI-dCas9融合蛋白,其包含任选地通过间插接头,例如2-30个氨基酸,例如4-12个氨基酸的接头,例如Gly<sub>4</sub>Ser与dCas9的末端例如N末端融合的FokI催化结构域序列。在一些实施方案中,FokI催化结构域包含SEQ ID NO:4的氨基酸388-583或408-583。在一些实施方案中,dCas9在D10、E762、H983或D986;以及在H840或N863处包含突变;例如:(i) D10A或D10N;和(ii) H840A、H840Y或H840N。

[0011] 在其它方面,本发明提供编码这些融合蛋白的核酸、包含所述核酸的载体以及具有或表达所述核酸、载体或融合蛋白的宿主细胞。

[0012] 在另一个方面,本发明提供用于诱导细胞中的例如基因组序列中的双链DNA分子的序列特异性断裂的方法,所述方法包括在细胞中表达本文所述的FokI-dCas9融合蛋白,或将细胞与本文所述的FokI-dCas9融合蛋白接触,和:

[0013] (a) 两个单一引导RNA,其中两个单一引导RNA的每一个包括各自与靶序列的一条链互补序列,以使得使用两个引导RNA导致靶向两条链(即,一个单一引导RNA靶向第一链,并且另一个引导RNA靶向互补链),并且FokI切割每一条链,在相对的DNA链上产生一对切口,从而产生双链断裂,或

[0014] (b) tracrRNA和两个crRNA,其中两个crRNA的每一个包括与靶序列的一条链互补的序列,以使得使用两个crRNA导致靶向两条链(即,一个crRNA靶向第一链,并且另一个crRNA靶向互补链),并且FokI切割每一条链,在相对的DNA链上导致一对切口,从而产生双链断裂。

[0015] 在其它方面,本发明提供用于提高细胞中的RNA引导的基因组编辑的特异性的方法,所述方法包括将细胞与本文所述的RNA引导的FokI核酸酶(RFN)融合蛋白接触。

[0016] 所述方法还可包括在细胞中表达如下序列或将细胞与所述序列接触:(a) 两个单一引导RNA,其中两个单一引导RNA的每一个包括各自与靶序列的一条链互补序列,以使得使用两个引导RNA导致靶向两条链(即,一个单一引导RNA靶向第一链,并且另一个引导RNA靶向互补链),并且FokI切割每一条链,在相对的DNA链上产生一对切口,从而产生双链断裂,或

[0017] (b) tracrRNA和两个crRNA,其中两个crRNA的每一个包括与靶序列的一条链互补的序列,以使得使用两个crRNA都导致靶向两条链(即,一个crRNA靶向第一链,并且另一个crRNA靶向互补链),并且FokI切割每一条链,在相对的DNA链上导致一对切口,从而产生双链断裂。

[0018] 在一些实施方案中,两个靶基因组序列(即,与crRNA或单一引导RNA的靶互补区互补的序列)相隔10-20个碱基对,优选相隔13-17个碱基对。

[0019] 在一些实施方案中,在两个靶序列之间诱导插入缺失突变。

[0020] 在一些实施方案中,提高细胞中的RNA引导的基因组编辑的特异性。

[0021] 除非另有定义,否则本文所用的所有技术和/或科学术语具有与本发明所属领域的普通技术人员通常所理解的含义相同的含义。本文中描述了用于本发明的方法和材料;还可使用本领域中已知的其它合适的方法和材料。所述材料、方法和实施例仅为示例性的

并非旨在限制。本文中提及的所有出版物、专利申请、专利、序列、数据库条目和其它参考资料通过引用整体并入。如发生矛盾，则以本专利说明书包括定义为准。

[0022] 由以下具体实施方式和附图以及由权利要求可显而易见本发明的其它特征及优点。

[0023] 附图简述

[0024] 图1:举例说明结合于其靶DNA位点的gRNA/Cas9核酸酶复合物的示意图。剪刀表示Cas9核酸酶在基因组DNA靶位点上的大致裂解点。注意，引导RNA上的核苷酸的编号以相反方式从5'至3'进行。

[0025] 图2A:举例说明截短gRNA的5'互补区的基本原理的示意图。粗的灰色线=靶DNA位点，细的暗灰色线结构=gRNA，灰色椭圆=Cas9核酸酶，黑线表示gRNA与靶DNA位点之间的碱基配对。

[0026] 图2B:EGFP破坏测定的示意性概述。通过易错NHEJ介导的修复进行的单一整合的EGFP-PEST报告基因的靶向Cas9介导的双链断裂的修复导致破坏编码序列并且与荧光在细胞中的丢失相关的移码突变。

[0027] 图2C-F:在EGFP报告基因序列中的3个不同靶位点上测定的拥有具有(C)单个错配、(D)相邻的双错配、(E)可变间隔的双错配和(F)递增数目的相邻错配的sgRNA的RGN的活性。显示了针对完美地匹配的gRNA的活性标准化的平均复制活性(参见在线方法)。误差条指示平均值的标准误差。每一个gRNA中的错配位置在下面的网格中以灰色突出显示。3个EGFP靶位点的序列如下：

[0028] EGFP位点1GGGCACGGGCAGCTTGCCGGTGG (SEQ ID NO:1)

[0029] EGFP位点2GATGCCGTTCTTCTGCTTGTCGG (SEQ ID NO:2)

[0030] EGFP位点3GGTGGTGCAGATGAACTTCAGGG (SEQ ID NO:3)

[0031] 图2G:gRNA的5'末端上的错配产生比更多3'错配更敏感的CRISPR/Cas。除使用沃尔森-克里克颠换使其错配(即通过将gRNA在位置18及19上改变成其沃尔森-克里克伴侣来使EGFP位点#2M18-19错配)的用“m”指示的位置外的RNA与DNA之间的gRNA沃尔森-克里克碱基配对。虽然gRNA的5'附近的位置通常被极好地耐受，但当其它残基被错配时这些位置中的匹配对于核酸酶活性是非常重要的。当使所有四个位置错配时，核酸酶活性不再是可检测的。这进一步表明这些5'位置上的匹配可帮助补偿其它更多3'位置上的错配。注意利用可显示相较于密码子最优化形式较低绝对水平的核酸酶活性的Cas9的非密码子优化形式进行这些实验。

[0032] 图2H:由具有范围在15至25nt内的可变长度互补区的gRNA指导的Cas9核酸酶活性在基于人细胞的U2OS EGFP破坏测定的效率。gRNA从U6启动子的表达需要5' G的存在，从而只可能评价具有某些长度的与靶DNA位点的互补性(15、17、19、20、21、23和25nt)的gRNA。

[0033] 图3A:EGFP报告基因中的4个靶位点的由Cas9和全长或缩短的gRNA介导的人细胞中的EGFP破坏的效率。显示了互补区 and 对应靶DNA位点的长度。Ctrl=缺乏互补区的对照gRNA。

[0034] 图3B:通过匹配的标准和tru-RGN在7个不同的人内源基因靶上引入的靶向插入缺失突变的效率。显示了互补区 and 对应靶DNA位点的长度。通过T7EI测定测量插入缺失频率。Ctrl=缺乏互补区的对照gRNA。

[0035] 图3C:通过使用被靶向EMX1位点的tru-gRNA或匹配的全长gRNA的RGN诱导的插入缺失突变的DNA序列。与gRNA互补区相互作用的靶DNA位点的部分以灰色突出显示,PAM序列的第一碱基以小写字母显示。缺失由以灰色突出显示的虚线指示,插入由以灰色突出显示的斜体字母指示。缺失或插入的碱基的净数目和每一个序列的次数被分离并且示于右边。

[0036] 图3D:通过匹配的标准和tru-RGN在家个内源人基因上引入的精确HDR/ssODN介导的改变的效率。使用BamHI限制酶切消化测定(参见实施例2的实验程序)测量的HDR的百分比。对照gRNA=空U6启动子载体。

[0037] 图3E:利用可变量的全长gRNA表达质粒(顶)或tru-gRNA表达质粒(底)与固定量的Cas9表达质粒一起转染U20S.EGFP细胞,随后测定具有减少的EGFP表达的细胞的百分比。显示来自一式二份实验的平均值和所述平均值的标准误差。注意,利用tru-gRNA获得的数据与来自利用全长gRNA表达质粒而非tru-gRNA质粒针对这3个EGFP靶位点进行的实验的数据密切匹配。

[0038] 图3F:利用可变量的Cas9表达质粒与可变量的全长gRNA表达质粒(顶)或tru-gRNA表达质粒(底)一起转染U20S.EGFP细胞(针对来自图3E的实验的每一个tru-gRNA测定的量)。显示来自一式二份实验的平均值和所述平均值的标准误差。注意,利用tru-gRNA获得的数据与来自利用全长gRNA表达质粒而非tru-gRNA质粒针对这3个EGFP靶位点进行的实验的数据密切匹配。这些滴定的结果确定用于在实施例1和2中进行的EGFP破坏测定的质粒的浓度。

[0039] 图4A-C.基于RNA引导的FokI核酸酶和CRISPR/Cas亚型Ypest蛋白4(Csy4)的多重gRNA表达系统。

[0040] (a) RNA引导的FokI核酸酶的示意图概述。两个FokI-dCas9融合蛋白被两个不同的gRNA招募至相邻的靶位点以促进FokI二聚化和DNA裂解。

[0041] (b) 基于Csy4的多重gRNA表达系统的示意图概述。将两个gRNA(具有任何5'末端核苷酸)从U6启动子共表达在单个转录物中,每一个gRNA侧翼连接有Csy4识别位点。Csy4裂解并从转录物释放gRNA。Csy4识别位点保留在gRNA的3'末端,Csy4核酸酶结合于该位点。

[0042] (c) 基于Csy4的多重系统的验证。在人U20S.EGFP细胞中使用基于Csy4的系统与Csy4和Cas9核酸酶一起在单个RNA转录物中表达被靶向EGFP中的相邻位点的两个gRNA。显示了在这些细胞中诱导的插入缺失突变的序列。野生型序列示于顶部,两个靶位点以灰色突出显示,并且PAM序列以加以下划线的文本显示。缺失由灰色背景中的虚线指示,插入由灰色背景中的小写字母指示。在每一个序列的右边,指出了插入(+)或缺失( $\Delta$ )的大小。

[0043] 图5A-F.RNA引导的FokI核酸酶的设计和最优优化。

[0044] (a) ZFN、TALEN、FokI-dCas9融合物和dCas9-FokI融合物的示意图举例说明。

[0045] (b) 筛选FokI-dCas9融合物的EGFP破坏活性,gRNA对被以两个取向:PAM在内(左图)和PAM在外(右图)之一靶向半位点。半位点相隔具有在0至31bp的范围内的可变长度的间隔区序列。通过流式细胞术定量EGFP破坏, $n=1$ 。dCas9-FokI融合物和相同gRNA对的对应数据示于图5E中。

[0046] (c) 靶位点上的FokI-dCas9介导的EGFP破坏活性的另外的评估,所述靶位点具有以它们的PAM在外取向的并且具有在10至20bp的范围内的间隔区长度的半位点。EGFP破坏通过流式细胞术来定量。误差条表示平均值的标准误差(s.e.m.), $n=2$ 。

- [0047] (d) 按照间隔区长度分组的来自 (c) 的数据的平均EGFP破坏值。误差条表示s.e.m。
- [0048] (e) 这些曲线显示在利用60个具有0-31bp的间隔以及PAM在内和PAM在外取向的gRNA对在U2OS. EGFP细胞中进行的EGFP破坏测定中筛选dCas9-FokI活性的结果。
- [0049] (f) 显示了U2OS细胞中FokI-dCas9诱导的突变的序列。被Cas9或FokI-dCas9结合的23-nt靶序列标以灰色。前间区序列邻近基序或PAM序列以粗体和下划线标记。缺失以浅灰色背景中的虚线标记。插入以灰色突出显示。在紧邻序列右边的栏中标明插入或缺失的碱基的数目。
- [0050] 图6A-D. FokI-dCas9 RFN的二聚化是高效基因组编辑活性所需的。
- [0051] (a) 在正确靶向的gRNA对 (针对EGFP位点47和81) 和其中gRNA的一个或另一个已被另一个被靶向非EGFP序列 (在VEGFA基因中) 的gRNA替代的对存在的情况下, 评估的两个RFN对的EGFP破坏活性。EGFP破坏通过流式细胞术来定量。EGFP, 增强的绿色荧光蛋白; VEGFA, 血管内皮生长因子A。误差条表示平均值的标准误差 (s.e.m.),  $n=3$ 。
- [0052] (b) 通过利用来自用于 (a) 的EGFP破坏测定的相同细胞的基因组DNA进行的T7EI测定定量诱变频率。误差条表示s.e.m.,  $n=3$ 。
- [0053] (c) 被靶向APC、MLH1和VEGFA基因中的位点的RFN的活性。对于每一个靶, 我们共表达FokI-dCas9与一对同源gRNA, 仅一个gRNA用于“左”半位点, 或仅一个gRNA用于“右”半位点。诱变率通过T7EI测定来测量。APC, 腺瘤性结肠息肉病; MLH1, mutL同源物1; VEGFA, 血管内皮生长因子A。误差条表示s.e.m.,  $n=3$ 。
- [0054] (d) 被靶向用于靶向VEGFA位点1的gRNA之一的中靶位点上和5个先前已知的脱靶 (OT) 位点上的VEGFA位点1的RFN的诱变频率。突变频率通过深度测序来测定。从单步骤测序文库 (从由3个单独的转染实验混合的基因组DNA制备的) 测定每一个报导的值。针对中靶VEGFA位点1显示的值 (以星号标记的) 与下文中图4a中显示的值相同并且只在此处显示以便与该图中所示的值比较。
- [0055] 图7A-B. 与单一gRNA共表达的Cas9切口酶或FokI-dCas9的诱变活性。
- [0056] (a) 在被靶向6个不同人基因位点的两个gRNA之一存在的情况下由FokI-dCas9 (左条块) 或Cas9切口酶 (中条块) 诱导的插入缺失突变频率。对于每一个基因靶, 我们针对两个gRNA评估了插入缺失频率, 仅一个gRNA用于“左”半位点, 或仅另一个gRNA用于“右”半位点。通过深度测序测定突变频率。从单步骤测序文库 (从由3个单独的转染实验混合的基因组DNA制备的) 测定每一个报导的插入缺失频率值。VEGFA, 血管内皮生长因子A; DDB2, 损伤特异性DNA结合蛋白2; FANCF, 范科尼贫血互补群F; FES, 猫肉瘤癌基因; RUNX 1, Runt相关转录因子1。
- [0057] (b) 表示为插入缺失频率频率的倍数减少的来自 (a) 的数据, 该数据比较利用gRNA对对于每一个靶位点获得的值与通过每一个单一gRNA实验或对照实验 (无gRNA并且无Cas9切口酶或FokI-dCas9) 获得的值。计算FokI-dCas9 (每一对中的左条块, 更浅的灰色) 和Cas9切口酶 (每一对中的右条块, 更深的灰色) 的该倍数减少。
- [0058] 图8A-C: 单个Cas9切口酶可以高频率将点突变引入它们的靶位点。
- [0059] 在FokI-dCas9、Cas9切口酶或tdTomato对照存在的情况下, (a) VEGFA、(b) FANCF和 (c) RUNX1基因靶的被单一gRNA靶向的半位点中的每一个位置上发现的不同点突变的频率。突变频率通过深度测序来测定。从单步骤测序文库 (从由3个单独的转染实验混合的基因组

DNA制备的)测定每一个报导的点突变值。注意,从针对图7A-B中的插入缺失突变分析的相同细胞分离用于这些实验的基因组DNA。VEGFA,血管内皮生长因子A;FANCF,范科尼贫血互补群F;RUNX 1,Runt相关转录因子1。

[0060] 详述

[0061] CRISPR RNA引导的核酸酶(RGN)已快速形成用于基因组编辑的轻便高效的平台。虽然Marraffini及同事(Jiang等,Nat Biotechnol 31,233-239(2013))最近在细菌中进行了Cas9 RGN特异性的系统性研究,但RGN在人细胞中的特异性一直未被广泛地确定。如果这些核酸酶将被广泛用于研究和治疗性应用,理解RGN介导的脱靶效应在人和其它真核细胞中的范围将是非常必需的。本发明人已使用基于人细胞的报告基因测定来表征基于Cas9的RGN的脱靶裂解。取决于它们沿着引导RNA(gRNA)-DNA界面的位置,单和双错配被耐受至不同程度。由被靶向人细胞中的内源基因座的6个RGN中的4个诱导的脱靶改变可通过部分错配位点的检查来容易地检测。已鉴定的脱靶位点具有多至5个错配,并且许多以可与在期望的中靶位点上观察到的频率相当(或更高)的频率被诱变。因此,即使在人细胞中具有不完全匹配的RNA-DNA界面,RGN亦具有很高的活性,这是可能使它们在研究和治疗性应用中的用途受挫的发现。

[0062] 本文中描述的结果显示预测任何给定的RGN的特异性特征谱既不简单也不直接。EGFP报告基因测定实验显示单和双错配可对人细胞中的RGN活性具有可变的作用,所述作用不严格地取决于它们在靶位点中的位置。例如,与先前公开的报导一致,gRNA/DNA界面的3'半部分的改变通常具有比5'半部分的改变更大的效应(Jiang等,Nat Biotechnol 31,233-239(2013);Cong等,Science 339,819-823(2013);Jinek等,Science 337,816-821(2012));然而,3'末端中的单和双突变有时也表现被良好地耐受,然而5'末端中的双突变可极大地降低活性。此外,针对在任何给定的位置上的错配的这些效应的量级似乎是位点依赖性的。通过所有可能的核苷酸取代(涵盖用于我们的EGFP报告实验的沃尔森-克里克颠换)的测试进行的大系列RGN的全面图谱表征可帮助提供对潜在脱靶的范围的另外见解。在这一点上,最近描述的Marraffini及同事的基于细菌细胞的方法(Jiang等,Nat Biotechnol 31,233-239(2013))或由Liu及同事先前用于ZFN的在体外基于组合文库的裂解位点-选择法(Pattanayak等,Nat Methods 8,765-770(2011))可用于产生更大组的RGN特异性特征谱。

[0063] 尽管这些在全面预测RGN特异性中提出了挑战,但有可能通过检查一个亚组的与中靶位点相异在于1至5个错配的基因组位点来鉴定RGN的真正脱靶。值得注意的是,这些实验的条件下,RGN在这些脱靶位点的许多位点上诱导的突变的频率与在期望的中靶位点上观察到的所述频率相似(或更高),这使得能够使用T7EI测定(所述测定,如在我们实验室中进行的,具有~2%至5%的突变频率的可靠检测限)检测这些位点上的突变。因于这些突变率非常高,因此有可能避免使用先前需要用来检测低得多的频率的ZFN-和TALEN-诱导的脱靶改变的深度测序法(Pattanayak等,Nat Methods 8,765-770(2011);Perez等,Nat Biotechnol 26,808-816(2008);Gabriel等,Nat Biotechnol 29,816-823(2011);Hockemeyer等,Nat Biotechnol 29,731-734(2011))。人细胞中的RGN脱靶诱变的分析还确认了预测RGN特异性的困难-并非所有单和双错配脱靶位点显示突变的证据,虽然具有多至5个错配的一些位点也可显示改变。此外,鉴定的真正的脱靶位点不显示任何明显的相对于

期望的靶序列的朝向转换或颠换的偏倚性。

[0064] 虽然对于许多RGN看到脱靶位点,但这些位点的鉴定既非全面的也非全基因组规模的。对于6个研究的RGN,仅检查极小亚组的人基因组中的大多得的总数的潜在脱靶序列。虽然通过T7EI测定检查这样大数目的基因座的脱靶突变既然不现实也不是成本效益好的策略,但在将来的研究中使用高通量测序可能使得能够质询更大数目的候选脱靶位点并且提供用于检测真正的脱靶突变的更加灵敏的方法。例如,此类方法可能使得能够揭示我们对于其不能发现任何脱靶突变的两个RGN的另外的脱靶位点。另外,RGN特异性和可影响RGN在细胞中的活性的任何表观因子(例如,DNA甲基化和染色质状态)的更加深刻的理解可能也减少需要被检查的潜在位点的数目,从而使得RGN脱靶的全基因评估更可行和可承受得起。

[0065] 许多策略可用于使基因组脱靶突变的频率减小至最小。例如,RGN靶位点的特异性选择可被优化;鉴于在多至5个位置与期望的靶位点不同的脱靶位置可被RGN高效地突变,因此选择具有最少脱靶位点数目的靶位点(如通过错配计数判断的)似乎不可能是有效的;对于被靶向人基因组的序列的任何给定的RGN,通常存在在20bp RNA:DNA互补区内相异在于4或5个位置的数千潜在的脱靶位点。还可能的是,gRNA互补区的核苷酸含量可能影响潜在脱靶效应的范围。例如,已显示高GC含量稳定RNA:DNA杂交体(Sugimoto等,Biochemistry 34,11211-11216 (1995)),从而还可能预期使gRNA/基因组DNA杂交更稳定并且对错配更加耐受。需要利用更大数目的gRNA的另外的实验来评估这两个参数(基因组中的错配位点的数目和RNA:DNA杂交体的稳定性)是否影响和如何影响RGN的全基因组特异性。然而,重要地要指出,如果此类预测参数可被确定,则执行此类指导方针的作用将进一步限制RGN的靶向范围。

[0066] 用于减弱RGN诱导的脱靶效应的一个可能的一般策略可能是降低细胞中表达的gRNA和Cas9核酸酶的浓度。在U2OS.EGFP细胞中使用RGN针对VEGFA靶位点2和3测试该想法;转染较少的表达gRNA和Cas9的质粒减少中靶位点上的突变率但丝毫不改变相对脱靶突变率。与此一致,在两个其它人细胞类型(HEK293和K562细胞)中也观察到高水平的脱靶诱变率,即使绝对中靶诱变率比在U2OS.EGFP细胞中低。因此,降低细胞中的gRNA和Cas9的表达水平不可能提供减小脱靶效应的解决方法。此外,这些结果还表明在人细胞中观察到的高脱靶诱变率非由gRNA和/或Cas9的过表达引起。

[0067] 可在3个不同的人细胞类型中通过RGN诱导显著的脱靶诱变的发现对该基因组编辑平台的更广泛用途具有重要影响。对于研究应用,特别是对于牵涉及具有缓慢的世代时间的培养细胞或生物体的实验,将需要考虑高频率脱靶突变的潜在混淆作用,对于所述细胞或生物不期望的改变的异型杂交将具有挑战性。由于脱靶效应不是随机的而是与靶位点相关的,因此控制此类效应的一个方式可能是使用被靶向不同DNA序列的多个RGN来诱导基因组改变。然而,对于治疗性应用,这些发现明确地表明,如果这些核酸酶将在更长时期中被安全地用于治疗人疾病,则需要仔细地确定和/或提高RGN的特异性。

[0068] 用于提高特异性的方法

[0069] 如本文中所示,基于化脓性链球菌Cas9蛋白的CRISPR-Cas RNA引导的核酸酶可具有可与期望的中靶活性相当或比其更高的显著的脱靶诱变效应(实施例1)。此类脱靶效应对于研究,特别是对于潜在治疗性应用可以是有问题。因此,需要用于提高CRISPR-Cas RNA

引导的核酸酶 (RGN) 的特异性的方法。

[0070] 如实施例1中所述,Cas9 RGN可在人细胞中在脱靶位点诱导高频率的插入缺失突变(也参见Cradick等,2013;Fu等,2013;Hsu等,2013;Pattanayak等,2013)。这些不期望的改变可在基因序列中发生,其与期望的中靶位点相异在于多至5个错配(参见实施例1)。另外,尽管相较于3'末端的错配,gRNA互补区的5'末端上的错配通常被更好地耐受,但这些关系不是绝对的,并且显示位点-对-位点依赖性(参见实施例1和Fu等,2013;Hsu等,2013;Pattanayak等,2013)。因此,对于鉴定真正的脱靶位点,依赖于错配的数目和/或位置的计算方法目前具有有限的预测值。因此,如果RNA引导的核酸酶将被用于研究和治疗性应用,则用于减少脱靶突变的频率的方法仍然具有重要的优先性。

[0071] 二聚化是用于提高Cas9核酸酶的特异性的吸引人的潜在策略。这与不是真正的二聚体系统的成对Cas9切口酶法不同。成对切口酶通过将两个Cas9切口酶共定位在DNA的区段上来起作用,从而通过不明确的机制诱导高效率基因组编辑。因为二聚化不是酶促活性所需的,因此单个Cas9切口酶还可在某些位点(通过未知机制)高效地诱导插入缺失,从而可潜在地在基因组中引起不想要的脱靶突变。

[0072] 因此,提高RGN的特异性的一个策略是将FokI内切核酸酶结构域与具有D10A和H840A突变的Cas9的无催化活性的形式(也称为dCas9)融合。FokI核酸酶结构域用作二聚体,从而两个亚单位必须被招募至DNA的相同局部片段以诱导双链断裂。在该构型(图9A和实施例2)中,使用两个不同的gRNA以适当的构型招募两个FokI-dCas9融合物以产生双链断裂。因此,在该系统中,FokI-dCas9融合物可结合至为单个RGN的长度两倍长的位点,从而可预期该系统更具特异性。

[0073] 因此本文中提供了FokI-dCas9融合蛋白,其中FokI序列任选地通过间插接头,例如2-30个氨基酸,例如4-12氨基酸的接头,例如Gly<sub>4</sub>Ser (SEQ ID NO:23)或(Gly<sub>4</sub>Ser)<sub>3</sub>与dCas9(优选地与dCas9的氨基末端,而且还任选地与羧基末端)融合。在一些实施方案中,所述融合蛋白包括dCas9与FokI结构域之间的接头。可用于这些融合蛋白的接头(或串联结构中的融合蛋白之间)可包括不干扰融合蛋白的功能的任何序列。在优选实施方案中,所述接头是短的,例如2-20个氨基酸,并且通常是柔性的(即,包含具有高度自由的氨基酸诸如甘氨酸、丙氨酸和丝氨酸)。在一些实施方案中,所述接头包含一个或多个由GGGS (SEQ ID NO:22)或GGGG (SEQ ID NO:23)组成的单元,例如,2、3、4或更多个GGGS (SEQ ID NO:22)或GGGG (SEQ ID NO:23)单元的重复。还可使用其它接头序列。

[0074] 本文中还描述了RNA引导的FokI核酸酶平台,其中二聚化而非仅共定位是高效基因组编辑活性所需的。这些核酸酶可在人细胞中强劲地介导高效基因组编辑,并且可将脱靶突变降至不可检测的水平,如通过灵敏的深度测序法判断的。还描述了用于表达成对的具有任何5'末端核苷酸的gRNA,在RFN平台上赋予更宽的靶向范围的方法。最后,单体Cas9切口酶在单一gRNA存在的情况下通常引入比本文所述的核酸酶更多的不想要的插入缺失和点突变。这些结构确定了强健的用户友好的核酸酶平台,该平台具有良好表征的二聚化结构的特异性有利方面和相对于成对Cas9切口酶的改善的诱变特征谱、对于需要最高基因组编辑精确性的研究或治疗应用将很重要的特征。

[0075] 因此,本文中描述了用于在人细胞中进行强劲的高度特异性的基因组编辑的新的RNA引导的FokI核酸酶(RFN)平台。RFN需要两个gRNA来获得活性和以二聚体形式起作用。令

人惊讶地,活性RFN的工程化需要FokI核酸酶结构域与dCas9蛋白的氨基末端的融合(与其中FokI结构域与工程化锌指或转录激活剂样效应子重复阵列融合的来自ZFN和TALEN的不同的结构)。RFN还要求被每一个Fok-dCas9/gRNA复合物结合的半位点具有特定的相对取向(PAM在外),所述半位点相隔14至17bp的相对有限的间隔区长度(虽然以另外的间隔,活性也是可能的,但并不总是成功的)。

[0076] RFN的二聚体性质提供了相对于标准单体Cas9核酸酶的重要的特异性有利方面。在理想的二聚体系统中,对于单体(在半位点上)观察到很少或未观察到活性。本数据表明被单一gRNA引导的FokI-dCas9在RFN半位点上诱导非常少或未诱导诱变。通过共表达的FokI-dCas9测试12个单一gRNA(用于6个RFN靶位点),并且以极低频率(0.0045%至0.47%的范围),在一些情况下以低至在其中无gRNA或核酸酶表达的对照细胞中观察到的本底比率的水平观察到插入缺失。鉴于FokI核酸酶结构域以二聚体起作用,因此假定对于单一gRNA观察到的任何插入缺失可能归因于FokI-dCas9二聚体至DNA的招募。无论机制如何,鉴于当用单一gRNA在12个中靶半位点上测试FokI-dCas9时仅观察到极低水平的诱变,因此极不可能的是在部分错配的脱靶半位点上将诱导任何诱变。事实上,被靶向VEGFA的RFN在gRNA之一的已知脱靶位点上不诱导可检测的突变,如通过深度测序判断的。

[0077] 因为RFN是真正的二聚体系统,因此它们具有优于成对切口酶技术的许多重要有利方面,所述切口酶技术依赖于共定位但不需要二聚化。第一,本文中的直接比较显示单个Cas9切口酶通常以比被相同的单个gRNA指导的FokI-dCas9融合蛋白更高的效率诱导插入缺失突变。第二,单体Cas9切口酶还可在它们的靶半位点上以高效率诱导碱基配对取代,我们在本研究中发现的先前未知的诱变副作用。再次地,直接比较显示单体Cas9切口酶以比由相同的单一gRNA引导的FokI-dCas9融合物显著更高的比率诱导这些不想要的点突变。第三,成对Cas9切口酶在靶半位点的取向和间隔上显示比二聚体RFN更大的混乱,从而具有更大潜在范围的在其上可能诱导脱靶突变的位点。成对切口酶半位点可用它们的PAM在内或PAM在外来定向,并且间隔区序列的长度在0至1000bp的范围内(Ran等,Cell 154,1380-1389(2013);Mali等,Nat Biotechnol 31,833-838(2013);Cho等,Genome Res(2013))。该混乱存在,因为Cas9切口酶的基因组编辑活性不依赖于酶的二聚化而是仅依赖于两个切口的共定位。相反地,RFN在它们的特异性上要严格得多一半位点必须使它们的PAM在外,并且必需间隔14至17bp,这归因于需要两个适当地放置的FokI裂解结构域来进行高效裂解。

[0078] FokI

[0079] FokI是包括DNA识别结构域和催化(内切核酸酶)结构域的II型限制性内切核酸酶。本文所述的融合蛋白可包括所有FokI或仅仅催化内切核酸酶结构域,例如,基因库登录号AAA24927.1的氨基酸388-583或408-583,例如,如Li等,Nucleic Acids Res.39(1):359-372(2011);Cathomen和Joung,Mol. Ther.16:1200-1207(2008)中描述的,或如Miller等Nat Biotechnol 25:778-785(2007);Szczeppek等,Nat Biotechnol 25:786-793(2007);或Bitinaite等,Proc.Natl.Acad.Sci.USA.95:10570-10575(1998)中描述的FokI的突变形式。

[0080] FokI的示例性氨基酸序列如下:

10 20 30 40 50 60  
 MFLSMVSKIR TFGWVQNPGR FENLKRNVQV FDRNSKVHNE VKNIKIPTLV KESKIQKELV  
 70 80 90 100 110 120  
 AIMNQHDLIY TYKELVGTGT SIRSEAPCDA IIQATIADQG NKKGYIDNWS SDGFLRWHAH  
 130 140 150 160 170 180  
 LGFIEYINKS DSFVITDVGL AYSKSADGSA IEKEILIEAI SSYPPAIRIL TLLEDGQHLT  
 190 200 210 220 230 240  
 KFDLGKNLGF SGESGFTSLP EGILLDTLAN AMPKDKGEIR NNWEGSSDKY ARMIGGWLDK  
 250 260 270 280 290 300  
 LGLVKQKKKE FIIPTLGKPD NKEFISHAFK ITGEGCLKVL RAKGSTKFTV VPKRVYWEML  
 [0081] 310 320 330 340 350 360  
 ATNLTDKEYV RTRRALILEI LIKAGSLKIE QIQDNLKKLG FDEVETIEN DIKGLINTGI  
 370 380 390 400 410 420  
 FIEIKGRFYQ LKDHILQFVI PNRGVTKQLV KSELEEKKE LRHKLKYPH EYIELIEIAR  
 430 440 450 460 470 480  
 NSTQDRILEM KVMEFFMKVY GYRGKHLGGS RKPDAIYTV GSPIDYGVIV DTKAYSGGYN  
 490 500 510 520 530 540  
 LPIGQADEMQ RYVEENQTRN KHINPNEWK VYPSSVTEFK FLVSGHFKG NYKAQLTRLN  
 550 560 570 580  
 HITNCNGAVL SVEELLIGGE MIKAGTLTLE EVRRKFNNGE INF (SEQ ID NO:4)

[0082] 编码FoKI的示例性核酸序列如下：

[0083] ATGTTTTGAGTATGGTTTCTAAAATAAGAACTTTCGGTTGGGTTCAAAATCCAGGTAAATTTGAGAA  
 TTTAAACGAGTAGTTCAAGTATTTGATAGAAATTCTAAAGTACATAATGAAGTAAAAATATAAAGATACCAACC  
 CTAGTCAAAGAAAGTAAGATCCAAAAAGAACTAGTTGCTATTATGAATCAACATGATTTGATTTATACATATAAAG  
 AGTTAGTAGGAACAGGAACCTCAATACGTTCAAGAGCACCATGCGATGCAATTATTCAAGCAACAATAGCAGATCA  
 AGGAAATAAAAAAGGCTATATCGATAATTGGTCATCTGACGGTTTTTTCGGTTGGGCACATGCTTTAGGATTTATT  
 GAATATATAAATAAAAGTGATTCCTTTTGAATAACTGATGTTGGACTTGCTTACTCTAAATCAGCTGACGGCAGCG  
 CCATTGAAAAAGAGATTTTGATTGAAGCGATATCATCTTATCCTCCAGCGATTCGTATTTTAACTTTGCTAGAAGA  
 TGGACAACATTTGACAAAGTTTGATCTTGGCAAGAATTTAGGTTTTAGTGAGAAAGTGGATTTACTTCTCTACCG  
 GAAGGAATTCCTTTAGATACTCTAGCTAATGCTATGCCTAAAGATAAAGGCGAAATTCGTAATAATTGGGAAGGAT  
 CTTCAAGATAAGTACGCAAGAATGATAGGTGGTTGGCTGGATAAACTAGGATTAGTAAAGCAAGGAAAAAAGAATT  
 TATCATTCCTACTTTGGGTAAGCCGACAATAAAGAGTTTATATCCACGCTTTTAAATTAAGTGGAGAAGGTTTG  
 AAAGTACTGCGTCGAGCAAAAGGCTCTACAAAATTTACACGTGTACCTAAAAGAGTATATTGGGAAATGCTTGCTA  
 CAAACCTAACCGATAAAGAGTATGTAAGAACAAGAAGAGCTTTGATTTTAGAAATATTAATCAAAGCTGGATCATT  
 AAAAATAGAACAAATACAAGACAACCTGAAGAAATTAGGATTTGATGAAGTTATAGAACTATTGAAAATGATATC  
 AAAGGCTTAATTAACACAGGTATATTTATAGAAATCAAAGGGCGATTTTATCAATTGAAAGACCATATTCTTCAAT  
 TTGTAATACCTAATCGTGGTGTGACTAAGCAACTAGTCAAAAGTGAAGTGGAGGAGAAGAAATCTGAAGTTCGTCA  
 TAAATTGAAATATGTGCCTCATGAATATATTGAATTAATTGAAATTGCCAGAAATCCACTCAGGATAGAATTCTT  
 GAAATGAAGGTAATGGAATTTTTTATGAAAGTTTATGGATATAGAGGTAAACATTTGGGTGGATCAAGGAAACCGG  
 ACGGAGCAATTTATACTGTGCGATCTCCTATTGATTACGGTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTA  
 TAATCTGCCAATTGGCCAAGCAGATGAAATGCAACGATATGTGGAAGAAAATCAAACACGAAACAAACATATCAAC  
 CCTAATGAATGGTGGAAAGTCTATCCATCTTCTGTAACGGAATTTAAGTTTTTATTGTGAGTGGTCACTTTAAAG

GAAACTACAAAGCTCAGCTTACACGATTAAATCATATCACTAATTGTAATGGAGCTGTTCTTAGTGTAGAAGAGCTTTAATTGGTGGAGAAATGATTAAAGCCGGCACATTAACTTAGAGGAAGTGAGACGGAAATTTAATAACGGCGAGATAAACTTTTAA (SEQ ID NO:5)

[0084] 在一些实施方案中,本文所述的FokI核酸酶与SEQ ID NO:4,例如与SEQ ID NO:4的氨基酸388-583或408-583具有至少约50%的同一性。这些变异核酸酶必须保留裂解DNA的能力。在一些实施方案中,所述核苷酸序列与SEQ ID NO:4的氨基酸388-583或408-583具有约50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、99%或100%的同一性。在一些实施方案中,与SEQ ID NO:4的氨基酸388-583或408-583的任何差异在非保守区中。

[0085] 为了测定两个序列的百分比同一性,为了最佳比较目的,将所述序列比对(可根据需要在第一和第二氨基酸或核酸序列的一个或两个序列中引入缺口以进行最佳比对,并且为了比较目的可忽略非同源序列)。为了比较目的而比对的参照序列的长度为至少50%(在一些实施方案中,比对约50%、55%、60%、65%、70%、75%、85%、90%、95%或100%的参照序列的长度)。随后比较对应位置上的核苷酸或残基。当第一序列中的位置被与第二序列中的对应位置相同的核苷酸或残基占据时,则所述分子在该位置上是相同的。两个序列之间的百分比同一性是由序列共享的相同位置的数目的函数,该函数考虑了为了一两个序列的最佳比对而需要引入的缺口数目和每一个缺口的长度。

[0086] 序列的比较和两个序列之间的百分比同一性的测定可使用数学算法来实现。为了本申请的目的,两个氨基酸序列之间的百分比同一性使用Needleman和Wunsch((1970) J.Mol.Biol.48:444-453)算法,使用Blossum 62评分矩阵,利用为12的缺口罚分、为4的缺口延伸罚分和为5的移码缺口罚分来测定,该算法已被整合进GCG软件包中的GAP程序。

[0087] Cas9

[0088] 许多细菌表达Cas9蛋白变体。来自化脓性链球菌(*Streptococcus pyogenes*)的Cas9是目前最常使用的;一些另外的Cas9蛋白与化脓性链球菌Cas9具有高水平的序列同一性,并且使用相同的引导RNA。其它的更加多样,使用不同的gRNA,并且同样地识别不同的PAM序列(与由RNA指定的序列相邻的由蛋白质指定的2-5个核苷酸的序列)。Chylinski等将来自一大组细菌的Cas9蛋白分类(RNA Biology 10:5,1-12;2013),并且许多Cas9蛋白列于补充图1和其补充表1中,所述图表通过引用并入本文。另外的Cas9蛋白描述于Esvelt等, Nat Methods.2013年11月;10(11):1116-21和Fonfara等,“Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems.”Nucleic Acids Res.2013年11月22日中。[先于印刷的电子出版]doi:10.1093/nar/gkt1074。

[0089] 许多物种的Cas9分子可用于本文所述的方法和组合物。虽然化脓性链球菌和嗜热链球菌Cas9分子是本文中的许多公开内容的主题,但同样地可使用本文中所列的其它物种的Cas9分子、来源于或基于所述物种的Cas9蛋白的Cas9分子。换句话说,虽然本文中的许多描述使用化脓性链球菌和嗜热链球菌Cas9分子,但来自其它物种的Cas9分子可替代它们。此类物种包括下表中所示的那些物种,所述表是基于Chylinski等,2013的补充图1产生的。

[0090]

| 替代性 Cas9 蛋白 |                                                                   |
|-------------|-------------------------------------------------------------------|
| 基因库登录号      | 细菌                                                                |
| 303229466   | 非典型韦荣球菌( <i>Veillonella atypica</i> )ACS-134-V-Col 7a             |
| 34762592    | 具核梭杆菌奋森亚种( <i>Fusobacterium nucleatum</i> ssp. <i>vincentii</i> ) |
| 374307738   | <i>Filifactor alocis</i> ATCC 35896                               |
| 320528778   | <i>Solobacterium moorei</i> F0204                                 |
| 291520705   | 灵巧粪球菌( <i>Coprococcus catus</i> )GD-7                             |
| 42525843    | 齿垢密螺旋体( <i>Treponema denticola</i> )ATCC 35405                    |
| 304438954   | <i>Peptoniphilus duerdenii</i> ATCC BAA-1640                      |
| 224543312   | <i>Catenibacterium mitsuokai</i> DSM 15897                        |
| 24379809    | 变异链球菌( <i>Streptococcus mutans</i> )UA159                         |
| 15675041    | 酿脓链球菌( <i>Streptococcus pyogenes</i> )SF370                       |
| 16801805    | 无害李斯特菌( <i>Listeria innocua</i> Clip)11262                        |

[0091]

| 替代性 Cas9 蛋白 |                                                                          |
|-------------|--------------------------------------------------------------------------|
| 基因库登录号      | 细菌                                                                       |
| 116628213   | 嗜热链球菌( <i>Streptococcus thermophilus</i> )LMD-9                          |
| 323463801   | 伪中间型葡萄球菌( <i>Staphylococcus pseudintermedius</i> )ED99                   |
| 352684361   | <i>Acidaminococcus intestini</i> RyC-MR95                                |
| 302336020   | <i>Olsenella uli</i> DSM 7084                                            |
| 366983953   | <i>Oenococcus kitaharae</i> DSM 17330                                    |
| 310286728   | 两歧双歧杆菌( <i>Bifidobacterium bifidum</i> )S17                              |
| 258509199   | 鼠李糖乳杆菌( <i>Lactobacillus rhamnosus</i> )GG                               |
| 300361537   | 加氏乳杆菌( <i>Lactobacillus gasseri</i> )JV-V03                              |
| 169823755   | 大芬戈尔德菌( <i>Finegoldia magna</i> )ATCC 29328                              |
| 47458868    | 运动支原体( <i>Mycoplasma mobile</i> )163K                                    |
| 284931710   | 鸡败血性霉浆菌( <i>Mycoplasma gallisepticum</i> )菌株 F                           |
| 363542550   | 绵羊肺炎霉浆菌( <i>Mycoplasma ovipneumoniae</i> ) SC01                          |
| 384393286   | 犬支原体( <i>Mycoplasma canis</i> )PG 14                                     |
| 71894592    | 关节液支原体( <i>Mycoplasma synoviae</i> )53                                   |
| 238924075   | 直肠真杆菌( <i>Eubacterium rectale</i> )ATCC 33656                            |
| 116627542   | 嗜热链球菌( <i>Streptococcus thermophilus</i> )LMD-9                          |
| 315149830   | 粪肠球菌( <i>Enterococcus faecalis</i> )TX0012                               |
| 315659848   | 路邓葡萄球菌( <i>Staphylococcus lugdunensis</i> )M23590                        |
| 160915782   | 细长真杆菌( <i>Eubacterium dolichum</i> )DSM 3991                             |
| 336393381   | 棒状乳杆菌扭曲亚种( <i>Lactobacillus coryniformis</i> subsp. <i>torquens</i> )    |
| 310780384   | 营养泥杆菌( <i>Ilyobacter polytropus</i> )DSM 2926                            |
| 325677756   | 白色瘤胃球菌( <i>Ruminococcus albus</i> )8                                     |
| 187736489   | <i>Akkermansia muciniphila</i> ATCC BAA-835                              |
| 117929158   | 嗜酸纤维素分解菌( <i>Acidothermus cellulolyticus</i> )11B                        |
| 189440764   | 长双歧杆菌( <i>Bifidobacterium longum</i> )DJO10A                             |
| 283456135   | 齿双歧杆菌( <i>Bifidobacterium dentium</i> )Bd1                               |
| 38232678    | 白喉棒状杆菌( <i>Corynebacterium diphtheriae</i> )NCTC 13129                   |
| 187250660   | <i>Elusimicrobium minutum</i> Pei191                                     |
| 319957206   | <i>Nitratifractor salsuginis</i> DSM 16511                               |
| 325972003   | <i>Sphaerochaeta globus</i> str. Buddy                                   |
| 261414553   | 产琥珀酸丝状杆菌亚种( <i>Fibrobacter succinogenes</i> subsp. <i>succinogenes</i> ) |
| 60683389    | 脆弱拟杆菌( <i>Bacteroides fragilis</i> )NCTC 9343                            |
| 256819408   | 黄褐二氧化碳嗜纤维菌( <i>Capnocytophaga ochracea</i> )DSM 7271                     |
| 90425961    | 沼泽红假单胞菌( <i>Rhodopseudomonas palustris</i> )BisB                         |

[0092]

| 替代性 Cas9 蛋白 |                                                                   |
|-------------|-------------------------------------------------------------------|
| 基因库登录号      | 细菌                                                                |
|             | 18                                                                |
| 373501184   | 普氏菌藻( <i>Prevotella micans</i> )F0438                             |
| 294674019   | 反刍瘤胃亚菌( <i>Prevotella ruminicola</i> )23                          |
| 365959402   | 柱状黄杆菌( <i>Flavobacterium columnare</i> )ATCC 49512                |
| 312879015   | <i>Aminomonas paucivorans</i> DSM 12260                           |
| 83591793    | 深红螺菌( <i>Rhodospirillum rubrum</i> )ATCC 11170                    |
| 294086111   | <i>Candidatus Puniceispirillum marinum</i> IMCC1322               |
| 121608211   | <i>Verminephrobacter eiseniae</i> EF01-2                          |
| 344171927   | 蒲桃雷尔氏菌( <i>Ralstonia syzygii</i> )R24                             |
| 159042956   | <i>Dinoroseobacter shibae</i> DFL 12                              |
| 288957741   | 固氮螺菌属( <i>Azospirillum</i> )的一个种-B510                             |
| 92109262    | 汉氏硝化细菌( <i>Nitrobacter hamburgensis</i> )X14                      |
| 148255343   | 慢生根瘤菌属( <i>Bradyrhizobium</i> )的一个种--BTai1                        |
| 34557790    | 琥珀酸弧菌( <i>Wolinella succinogenes</i> )DSM 1740                    |
| 218563121   | <i>Campylobacter jejuni</i> subsp. <i>jejuni</i>                  |
| 291276265   | 鼯鼠螺杆菌( <i>Helicobacter mustelae</i> )12198                        |
| 229113166   | 蜡状芽孢杆菌( <i>Bacillus cereus</i> )Rock1-15                          |
| 222109285   | <i>Acidovorax ebreus</i> TPSY                                     |
| 189485225   | uncultured Termite group 1                                        |
| 182624245   | 产气荚膜梭菌 D 菌株( <i>Clostridium perfringens</i> D str.)               |
| 220930482   | 解纤维素梭菌( <i>Clostridium cellulolyticum</i> )H10                    |
| 154250555   | <i>Parvibaculum lavamentivorans</i> DS-1                          |
| 257413184   | <i>Roseburia intestinalis</i> LI-82                               |
| 218767588   | 奈瑟氏脑膜炎球菌( <i>Neisseria meningitidis</i> )Z2491                    |
| 15602992    | 多杀性巴氏杆菌亚种( <i>Pasteurella multocida</i> subsp. <i>multocida</i> ) |
| 319941583   | 华德萨特菌( <i>Sutterella wadsworthensis</i> )3 1                      |
| 254447899   | $\gamma$ -变形菌纲( <i>gamma proteobacterium</i> )HTCC5015            |
| 54296138    | 嗜肺军团菌巴黎菌株( <i>Legionella pneumophila</i> str. <i>Paris</i> )      |
| 331001027   | <i>Parasutterella excrementihominis</i> YIT 11859                 |
| 34557932    | 产琥珀酸沃廉菌( <i>Wolinella succinogenes</i> )DSM 1740                  |
| 118497352   | 新生儿弗朗西斯菌( <i>Francisella novicida</i> )U112                       |

[0093] 本文所述的构建体和方法包括任何那些Cas9蛋白以及它们对应的引导RNA或相容的其它引导RNA的使用。也已显示来自嗜热链球菌LMD-9CRISPR1系统的Cas9在Cong等(Science 339,819 (2013))中的人细胞中起作用。来自奈瑟氏脑膜炎球菌的Cas9直系同源物描述于Hou等,Proc Natl Acad Sci U S A.2013年9月24日;110 (39):15644-9和Esvelt等,Nat Methods.2013年11月;10 (11):1116-21中。另外,Jinek等人在体外显示来自嗜热链球菌和英诺克李斯特菌(*L. innocua*) (但非来自奈瑟氏脑膜炎球菌或空肠弯曲菌(*C. jejuni*)的Cas9直系同源物(其可使用不同的引导RNA)可被双重化脓性链球菌gRNA导向裂解靶质粒DNA,虽然效率略有降低。

[0094] 在一些实施方案中,本系统利用来自化脓性链球菌的Cas9蛋白(如在细菌中编码的或针对在哺乳动物细胞中的表达进行密码子最优化的),其在D10、E762、H983或D986和H840或N863上含有突变,例如D10A/D10N和H840A/H840N/H840Y,以使得蛋白的核酸酶部分催化失活;这些位置上的取代可以是丙氨酸(如它们在Nishimasu等,Cell 156,935-949 (2014)中一样)或它们可以是其它残基,例如谷氨酰胺、天冬酰胺、酪氨酸、丝氨酸或天冬氨酸,例如,E762Q、H983N、H983Y、D986N、N863D、N863S或N863H(图1C)。可用于本文所述的方法和组合物的自由化来活的化脓性链球菌Cas9的序列如下;D10A和H840A的示例性突变以粗体表示并加以下划线。

[0095]

|            |            |            |             |            |             |
|------------|------------|------------|-------------|------------|-------------|
| 10         | 20         | 30         | 40          | 50         | 60          |
| MDKKYSIGLA | IGTNSVGWAV | ITDEYKVPSK | KFKVLGNTDR  | HSIKKNLIGA | LLFDSGETAE  |
| 70         | 80         | 90         | 100         | 110        | 120         |
| ATRLKRTARR | RYTRRKNRIC | YLQEIFSNE  | AKVDDSFHR   | LEESFLVEED | KKHERHPIFG  |
| 130        | 140        | 150        | 160         | 170        | 180         |
| NIVDEVAYHE | KYPTIYHLRK | KLVDSTDKAD | LRLIYLALAH  | MIKFRGHFLI | EGDLNPDNSD  |
| 190        | 200        | 210        | 220         | 230        | 240         |
| VDKLFIQLVQ | TYNQLFEENP | INASGVDAKA | ILSARLSKSR  | RLENLIAQLP | GEKKNGLFGN  |
| 250        | 260        | 270        | 280         | 290        | 300         |
| LIALSLGLTP | NFKSNFDLAE | DAKLQLSKDT | YDDDLNLLA   | QIGDQYADLF | LAACKNLSDAI |
| 310        | 320        | 330        | 340         | 350        | 360         |
| LLSDILRVNT | EITKAPLSAS | MIKRYDEHHQ | DLTLLKALVR  | QQLPEKYKEI | FFDQSKNGYA  |
| 370        | 380        | 390        | 400         | 410        | 420         |
| GYIDGGASQE | EFYKFIKPIL | EKMDGTEELL | VKLNREDLLR  | KQRTFDNGSI | PHQIHLGELH  |
| 430        | 440        | 450        | 460         | 470        | 480         |
| AILRRQEDFY | PFLKDNREKI | EKILTFRIPY | YVGPLARGNS  | RFAWMTRKSE | ETITPWNFEE  |
| 490        | 500        | 510        | 520         | 530        | 540         |
| VVDKGASQAS | FIERMTNFDK | NLPNEKVLPK | HSLLEYFTV   | YNELTKVKYV | TEGMRKPAFL  |
| 550        | 560        | 570        | 580         | 590        | 600         |
| SGEQKKAIVD | LLFKTNRKVT | VKQLKEDYFK | KIECFDSVEI  | SGVEDRFNAS | LGTYHDLKI   |
| 610        | 620        | 630        | 640         | 650        | 660         |
| IKDKDFLDNE | ENEDILEDIV | LTLTLFEDRE | MIEERLKTYA  | HLFDDKVMKQ | LKRRRYTGWG  |
| 670        | 680        | 690        | 700         | 710        | 720         |
| RLSRKLINGI | RDQSGKTIL  | DFLKSDGFAN | RNFMQLIHDD  | SLTFKEDIQK | AQVSGQGDSL  |
| 730        | 740        | 750        | 760         | 770        | 780         |
| HEHIANLAGS | PAIKKGILQT | VKVVDLVKV  | MGRHKPENIV  | IEMARENQTT | QKGQKNSRER  |
| 790        | 800        | 810        | 820         | 830        | 840         |
| MKRIEEGIKE | LGSQILKEHP | VENTQLQNEK | LYLYYLQNGR  | DMYVDQELDI | NRLSDYDVDA  |
| 850        | 860        | 870        | 880         | 890        | 900         |
| IVPQSFLKDD | SIDNKVLTRS | DKNRGKSDNV | PSEEVKKMK   | NYWRQLLNAK | LITQRKFDNL  |
| 910        | 920        | 930        | 940         | 950        | 960         |
| TKAERGGLSE | LDKAGFIKRQ | LVETRQITKH | VAQILDSRMN  | TKYDENDKLI | REVKVITLKS  |
| 970        | 980        | 990        | 1000        | 1010       | 1020        |
| KLVSDFRKDF | QFYKVREINN | YHHAHDAYLN | AVVGITALIKK | YPKLESEFVY | GDYKVYDVRK  |
| 1030       | 1040       | 1050       | 1060        | 1070       | 1080        |
| MIKSEQEIG  | KATAKYFFYS | NIMNFFKTEI | TLANGEIRKR  | PLIETNGETG | EIVWDKGRDF  |
| 1090       | 1100       | 1110       | 1120        | 1130       | 1140        |
| ATVRKVLSP  | QVNIVKKTEV | QTGGFSKESI | LPKRNSDKLI  | ARKKDWDPPK | YGGFDSPTVA  |
| 1150       | 1160       | 1170       | 1180        | 1190       | 1200        |
| YSVLVVAKVE | KGSKKLKSV  | KELLGITIME | RSSFENPID   | FLEAKGYKEV | KKDLIIKLPK  |

|        |            |            |             |            |               |            |
|--------|------------|------------|-------------|------------|---------------|------------|
|        | 1210       | 1220       | 1230        | 1240       | 1250          | 1260       |
|        | YSLFELENGR | KRMLASAGEL | QKGNEALALPS | KYVNFLYLAS | HYEKLKGSPE    | DNEQKQLFVE |
| [0096] | 1270       | 1280       | 1290        | 1300       | 1310          | 1320       |
|        | QHKHYLDEII | EQISEFSKRV | ILADANLDKV  | LSAYNKHRDK | PIREQAENII    | HLFTLTNLGA |
|        | 1330       | 1340       | 1350        | 1360       | (SEQ ID NO:5) |            |
|        | PAAFKYFDTT | IDRKRYTSTK | EVLDATLIHQ  | SITGLYETRI | DLSQLGGD      |            |

[0097] 在一些实施方案中,本文中使用的Cas9核酸酶与化脓性链球菌Cas9的序列具有至少约50%的同一性,即与SEQ ID NO:5具有至少50%的同一性。在一些实施方案中,所述核苷酸序列与SEQ ID NO:5具有约50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、99%或100%的同一性。在一些实施方案中,与SEQ ID NO:5的任何差异在非保守区中,如通过Chylinski等, RNA Biology 10:5,1-12;2013(例如,在补充图1及其补充表1中);Esvelt等, Nat Methods.2013年11月;10(11):1116-21和Fonfara等, Nucl. Acids Res. (2014) 42(4):2577-2590中所示的序列的序列比对鉴定的。[2013年11月22日先于印刷的电子版] doi:10.1093/nar/gkt1074。如上文中所示测定同一性。

[0098] 引导RNA (gRNA)

[0099] 引导RNA一般而言出现在两个不同的系统中:系统1,其使用一起指导Cas9进行裂解的单独的crRNA和tracrRNA,和系统2,其使用在单个系统中组合两个单独的引导RNA的嵌合crRNA-tracrRNA杂交体(被称为单一引导RNA或sgRNA,也参见Jinek等, Science 2012; 337:816-821)。tracrRNA可被可变地截短,并且已显示许多长度在所述单独的系统(系统1)和嵌合gRNA系统(系统2)中都具有功能。例如,在一些实施方案中,可从其3'末端将tracrRNA截短至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35或40nt。在一些实施方案中,可将tracrRNA分子从其5'末端截短至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35或40nt。或者,可以从5'和3'末端截短tracrRNA分子,例如在5'末端截短至少1、2、3、4、5、6、7、8、9、10、15或20nt并且在3'末端截短至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35或40nt。参见,例如, Jinek等, Science 2012; 337:816-821; Mali等, Science. 2013年2月15日; 339(6121):823-6; Cong等, Science. 2013年2月15日; 339(6121):819-23; 以及Hwang和Fu等, Nat Biotechnol. 2013年3月; 31(3):227-9; Jinek等, Elife 2, e00471 (2013)。对于系统2,一般地更长长度的嵌合gRNA已显示更大的中靶活性,但不同长度的gRNA的相对特异性目前仍未确定,从而在某些情况下可能期望使用较短的gRNA。在一些实施方案中,gRNA与在转录起始位点上游约100-800bp内,例如在转录起始位点的上游约500bp内,包括转录起始位点,或在转录起始位点下游约100-800bp内,例如约500bp内的区域互补。在一些实施方案中,使用编码不止一个gRNA的载体(例如,质粒),例如编码导向靶基因的相同区域内的不同位点的2、3、4、5或更多个gRNA的质粒。

[0100] 可使用在其5'末端上具有与基因组DNA靶位点的互补链互补的17-20nt的引导RNA(例如单一gRNA或tracrRNA/crRNA)将Cas9核酸酶导向具有例如序列NGG的另外的邻近的前间区序列邻近基序(PAM)的特定17-20nt基因组靶。因此,本方法可包括单一引导RNA的使用,所述单一引导RNA包含与通常反式编码的tracrRNA融合的crRNA,例如Mali等, Science 2013年2月15日; 339(6121):823-6中描述的单个Cas9引导RNA,其在5'末端上具有25-17个,任选地20个或更少的核苷酸(nt)的与靶序列互补的序列,例如,紧接前间区序列邻近基序(PAM)例如NGG、NAG或NNGG的5'的靶序列的互补链的20、19、18或17nt,优选地17或18nt。在

一些实施方案中,所述单个Cas9引导RNA由如下序列组成:

[0101] (X<sub>17-20</sub>) GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCG (X<sub>N</sub>) (SEQ ID NO:6);

[0102] (X<sub>17-20</sub>) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X<sub>N</sub>) (SEQ ID NO:7);

[0103] (X<sub>17-20</sub>) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCG UUAUC (X<sub>N</sub>) (SEQ ID NO:8);

[0104] (X<sub>17-20</sub>) GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGC (X<sub>N</sub>) (SEQ ID NO:9);

[0105] (X<sub>17-20</sub>) GUUUUAGAGCUAGAAAUAGCAAGUUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGC (SEQ ID NO:10);

[0106] (X<sub>17-20</sub>) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAAUAAGGCUAGUCCGUUAUCAACU UGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:11);或

[0107] (X<sub>17-20</sub>) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAAUAAGGCUAGUCCGUUAUCAACU UGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:12);

[0108] 其中X<sub>17-20</sub>是与靶序列的17-20个连续核苷酸互补的核苷酸序列。先前已在文献 (Jinek等,Science.337 (6096):816-21 (2012) 和Jinek等,Elife.2:e00471 (2013)) 中描述了编码单一引导RNA的DNA。

[0109] 引导RNA可包括可以是不干扰核糖核酸对Cas9结合的任何序列的X<sub>N</sub>,其中N (在RNA中) 可以是0-200例如0-100、0-50或0-20。

[0110] 在一些实施方案中,所述引导RNA在3'末端上包括一个或多个腺嘌呤 (A) 或尿嘧啶 (U) 核苷酸。在一些实施方案中,作为用作终止RNA PolIII转录的终止信号的一个或多个T的任选的存在的结果,所述RNA在分子的3'末端包括一个或多个U,例如,1至8个或更多个U (例如,U,UU,UUU,UUUU,UUUUU,UUUUUU,UUUUUUU,UUUUUUUU)。

[0111] 虽然本文所述的一些实例利用单一gRNA,但也将所述方法与双重gRNA (例如在天然存在的系统中发现的crRNA和tracrRNA) 一起使用。在该情况下,可将单一tracrRNA与多个不同的使用本系统表达的crRNA结合使用,例如下列序列:

[0112] (X<sub>17-20</sub>) GUUUUAGAGCUA (SEQ ID NO:13);

[0113] (X<sub>17-20</sub>) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:14);或

[0114] (X<sub>17-20</sub>) GUUUUAGAGCUAUGCU (SEQ ID NO:15);和tracrRNA序列。在该情况下,将crRNA在本文中描述的方法和分子中用作引导RNA,并且可以从相同或不同DNA分子表达tracrRNA。在一些实施方案中,所述方法包括将细胞与tracrRNA接触,所述tracrRNA包含如下序列或由所述序列组成GGAACCAUUCAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACU UGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:16) 或其活性部分 (活性部分是保持与Cas9或dCas9形成复合物的能力的部分)。在一些实施方案中,可从其3'末端将tracrRNA分子截短至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35或40nt。在另一个实施方案中,可从其5'末端将tracrRNA截短至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35或40nt。或者,可从5'和3'末端将tracrRNA分子截短,例如,在5'末端截短至少1、2、3、4、5、6、7、8、9、10、15或20nt并且在3'末端截短至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35或40nt。除了SEQ ID NO:8以外,示例性tracrRNA还包括下列序列:UAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC

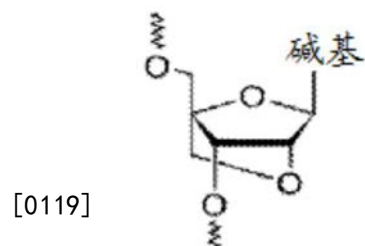
ACCGAGUCGGUGC (SEQ ID NO:17) 或其活性部分;或

[0115] AGCAUAGCAAGUUAUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:18) 或其活性部分。

[0116] 在其中将 (X<sub>17-20</sub>) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:14) 用作crRNA的一些实施方案中,使用下列tracrRNA:GGAACCAUUCACAAACAGCAUAGCAAGUUAUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:16) 或其活性部分。在其中将 (X<sub>17-20</sub>) GUUUUAGAGCUA (SEQ ID NO:13) 用作crRNA的一些实施方案中,使用下列tracrRNA:UAGCAAGUUAUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:17) 或其活性部分。在其中将 (X<sub>17-20</sub>) GUUUUAGAGCUAUGC (SEQ ID NO:15) 用作crRNA的一些实施方案中,使用下列tracrRNA:AGCAUAGCAAGUUAUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:18) 或其活性部分。

[0117] 在一些实施方案中,所述gRNA被靶向与基因组的其余部分中的任何序列相异至少3或更多个错配的位点以使脱靶效应减小至最小。

[0118] 经修饰的RNA寡核苷酸诸如锁核酸 (LNA) 已被证明通过以更有利的 (稳定的) 构象锁定经修饰的寡核苷酸来提高RNA-DNA杂交的特异性。例如,2'-O-甲基RNA是其中在2' 氧与4' 碳之间存在另外的共价键联的经修饰的碱基,当被掺入寡核苷酸中时其可提高总体热稳定性和选择性 (式I)。



### 式 I- 锁核酸

[0120] 因此,在一些实施方案中,本文中公开的tru-gRNA可包含一个或多个经修饰的RNA寡核苷酸。例如,本文中描述的截短的引导RNA分子可具有已被修饰的与靶序列互补的引导RNA的5' 区的17-18或17-19nt的一个或一些或全部被修饰,例如锁定的 (2'-O-4' -C亚甲基桥)、5'-甲基胞苷、2'-O-甲基-假尿苷,或其中磷酸核糖主链已被聚酰胺链 (肽核酸) 例如合成核糖核酸替代。

[0121] 在其它实施方案中,tru-gRNA序列的一个、一些或全部核苷酸可被修饰,例如锁定的 (2'-O-4' -C亚甲基桥)、5'-甲基胞苷、2'-O-甲基-假尿苷,或其中磷酸核糖主链已被聚酰胺链 (肽核酸) 例如合成核糖核酸替代。

[0122] 在一些实施方案中,所述单一引导RNA和/或crRNA和/或tracrRNA可在3' 末端上包括一个或多个腺嘌呤 (A) 或尿嘧啶 (U) 核苷酸。

[0123] 现有基于Cas9的RGN使用gRNA-DNA异源双链体形成来引导靶向目标基因组位点。然而,RNA-DNA异源双链体可形成比它们的DNA-DNA对应物更加混杂的范围的结构。实际上,DNA-DNA双链体对错配更加敏感,从而表明DNA引导的核酸酶可能不那么容易地结合于脱靶

序列,从而使得它们相较于RNA引导的核酸酶更具特异性。因此,可用于本文中描述的方法的引导RNA可以是杂交体,即,其中一个或多个脱氧核糖核苷酸例如短的DNA寡核苷酸替代gRNA的全部或部分,例如gRNA的互补区的全部或部分。该基于DNA的分子可替代单一gRNA系统中的gRNA的全部或部分或可选地可能替代双重crRNA/tracrRNA系统中的crRNA和/或tracrRNA的全部或部分。将DNA整合进互补区的此类系统应当因DNA-DNA双链体对错配的总不耐受性而相较于RNA-DNA双链体更容易靶向期望的基因组DNA序列。用于产生此类双链体的方法在本领域中是已知的,参见,例如,Barker等,BMC Genomics.2005年4月22日;6:57;和Sugimoto等,Biochemistry.2000年9月19日;39(37):11270-81。

[0124] 另外,在使用单独的crRNA和tracrRNA的系统中,其一种或两种可以是合成的并且包括一个或多个经修饰的(例如,锁定的)核苷酸或脱氧核糖核苷酸。

[0125] 在细胞背景中,Cas9与这些合成gRNA的复合物可用于提高CRISPR/Cas9核酸酶系统的全基因组特异性。

[0126] 所述方法可包括在细胞中表达本文中描述的Cas9 gRNA加融合蛋白,或将所述细胞与其接触。

[0127] 表达系统

[0128] 为了使用描述的融合蛋白,可能期望从编码它们的核酸表达它们。这可以以多种方式进行。例如,可将编码引导RNA的核酸克隆入用于转化进用于复制和/或表达的原核或真核细胞的中间载体。中间载体通常是原核载体,例如,质粒或穿梭载体或昆虫载体,其用于贮存或操纵编码融合蛋白的核酸以用于融合蛋白的产生。还可将编码融合蛋白的核酸克隆入表达载体,例如用于向植物细胞、动物细胞,优选地哺乳动物细胞或人细胞、真菌细胞、细菌细胞或原生动物细胞施用。

[0129] 为了获得表达,通常将编码融合蛋白的序列亚克隆入含有指导转录的启动子的表达载体。合适的细菌和真核启动子在本领域中是公知的,并且描述于例如Sambrook等,Molecular Cloning,A Laboratory Manual (2001年第3版);Kriegler,Gene Transfer and Expression:A Laboratory Manual (1990)和Current Protocols in Molecular Biology (Ausubel等,编辑,2010)中。用于表达工程化蛋白质的细菌表达系统可在例如大肠杆菌(E.coli)、芽孢杆菌属(Bacillus sp.)和沙门菌属(Salmonella) (Palva等,1983,Gene 22:229-235)中获得。此类表达系统的试剂盒是商购可得的。用于哺乳动物细胞、酵母和昆虫细胞的真核表达系统在本领域中是公知的并且也是商购可得的。

[0130] 用于指导核酸表达的启动子取决于具体应用。例如,强组成型启动子通常用于融合蛋白的表达和纯化。相反地,当将体内施用引导RNA以进行基因调控时,可使用组成型或诱导型启动子,这取决于引导RNA的具体用途。另外,用于施用引导RNA的优选启动子可以是弱启动子,诸如HSV TK或具有类似活性的启动子。启动子还可包括响应反式激活的元件,例如,缺氧应答元件、Gal4应答元件、lac阻遏应答元件和小分子控制系统诸如四环素调节的系统和RU-486系统(参见,例如,Gossen&Bujard,1992,Proc.Natl.Acad.Sci.USA,89:5547;Oligino等,1998,Gene Ther.,5:491-496;Wang等,1997,Gene Ther.,4:432-441;Neering等,1996,Blood,88:1147-55和Rendahl等,1998,Nat.Biotechnol.,16:757-761)。

[0131] 除了启动子以外,表达载体通常还含有包含核酸在宿主细胞(原核或真核的)中表达所需的所有另外的元件的转录单位或表达盒。常见表达盒从而含有可操作地连接于例如

编码gRNA的核酸序列的启动子和例如进行转录物的高效多腺苷酸化、转录终止、核糖体结合位点或翻译终止的所需的任何信号。表达盒的另外元件可包括例如增强子和异源剪接内含子信号。

[0132] 根据gRNA的期望用途(例如,在植物、动物、细菌、真菌、原生动物等中表达)选择用于将遗传信息转运至细胞的特定表达载体。标准细菌表达载体包括质粒诸如基于pBR322的质粒、pSKF、pET23D和商购可得的靶-融合表达系统诸如GST和LacZ。

[0133] 含有来自真核病毒的调控元件的表达载体通常用于真核表达载体,例如,SV40载体、乳头状瘤病毒载体和来源于爱泼斯坦-巴尔病毒的载体。其它示例性真核载体包括pMSG、pAV009/A+、pMT010/A+、pMAMneo-5、杆状病毒pDSVE和允许在如下启动子指导下表达蛋白质的任何其它载体:SV40早期启动子、SV40晚期启动子、金属硫蛋白启动子、鼠乳腺肿瘤病毒启动子、劳斯肉瘤病毒启动子、多角体蛋白启动子或经显示对于在真核细胞中的表达是有效的其它启动子。

[0134] 用于表达引导RNA的载体可包括驱动引导RNA表达的RNA Pol III启动子,例如H1、U6或7SK启动子。这些人启动子允许在质粒转染后在哺乳动物细胞中表达gRNA。或者,T7启动子可用于例如体外转录,并且所述RNA可被体外转录和纯化。可使用适合用于短的RNA例如siRNA、shRNA或其它小的RNA表达的载体。对于图4B中描述的基于Cys4的多重系统,可在单个转录物(由RNA Pol II或Pol III启动子驱动)中表达多个gRNA,随后从该更大的转录物切割出来。

[0135] 一些表达系统具有用于选择稳定地转染的细胞系的标志物诸如胸苷激酶、潮霉素B磷酸物转移酶和二氢叶酸还原酶。高产表达系统也是合适的,诸如在昆虫细胞中使用杆状病毒载体,利用在多角体蛋白启动子或其它强杆状病毒启动子的指导下的gRNA编码序列。

[0136] 通常被包括在表达载体中的元件还包括在大肠杆菌中起作用的复制子、编码抗生素抗性以允许选择具有重组质粒的细菌的基因和允许重组序列插入的质粒的非必需区中的独特限制性位点。

[0137] 标准转染法可用于产生表达大量蛋白质的细菌、哺乳动物、酵母或昆虫细胞系,随后使用标准技术(参见,例如,Colley等,1989,J.Biol.Chem.,264:17619-22;Guide to Protein Purification,于Methods in Enzymology,第182卷(Deutscher,编辑,1990)中)纯化所述蛋白质。真核和原核细胞的转化按照标准技术(参见,例如,Morrison,1977,J.Bacteriol.132:349-351;Clark-Curtiss&Curtiss,Methods in Enzymology 101:347-362(Wu等,编辑,1983)来进行。

[0138] 可使用用于将外来核苷酸序列引入宿主细胞的任何已知方法。这些方法包括使用磷酸钙转染、聚凝胺、原生质体融合、电穿孔、核转染、脂质体、显微注射、裸DNA、质粒载体、病毒载体(游离型和融合型)和用于将克隆的基因组DNA、cDNA、合成DNA或其它外来遗传物质引入宿主细胞的任何其它公知的方法(参见,例如,Sambrook等,同上)。唯一必需的是,使用的特定遗传工程方法能够成功地将至少一个基因引入能够表达gRNA的宿主细胞。

[0139] 本发明包括所述载体和包含所述载体的细胞。

## 实施例

[0140] 在下列实施例中进一步描述本发明,所述实施例不限制权利要求中描述的本发明

的范围。

[0141] 实施例1. 评估RNA引导的内切核酸酶的特异性

[0142] CRISPR RNA引导的核酸酶(RGN)已迅速形成用于基因组编辑的轻便高效的平台。本实施例描述基于人细胞的报告基因测定用于表征基于Cas9的RGN的脱靶裂解的用途。

[0143] 材料和方法

[0144] 下列材料和方法用于实施例1中。

[0145] 引导RNA的构建

[0146] 将具有用于Cas9靶向的可变20nt序列的DNA寡核苷酸退火以产生具有与至BsmBI-消化的质粒pMLM3636中的连接相容4bp悬突的短的双链DNA片段。这些退火的寡核苷酸的克隆产生编码在U6启动子的表达下的具有20个可变5'核苷酸的嵌合+103单链引导RNA的质粒(Hwang等, Nat Biotechnol 31, 227-229 (2013); Mali等, Science 339, 823-826 (2013))。用于本研究的pMLM3636和表达质粒pJDS246(编码Cas9的密码子最优化的形式)都可通过非营利质粒的分销服务Addgene (addgene.org/crispr-cas) 获得。

[0147] EGFP活性测定

[0148] 如先前所述(Reyon等, Nat Biotech 30, 460-465 (2012))培养具有单个整合的拷贝的EGFP-PEST融合基因的U2OS. EGFP细胞。为了进行转染, 按照制造商的方案, 使用SE细胞系4D-Nucleofector™X试剂盒(Lonza), 利用指定量的gRNA表达质粒和pJDS246与30ng的Td-tomato-编码质粒一起对200,000个细胞进行核转染。使用BD LSRII流式细胞仪在转染后2天分析细胞。以一式三份进行最优化gRNA/Cas9质粒浓度的转染, 并且以一式二份进行所有其它转染。

[0149] 内源人基因组位点的PCR扩增和序列验证

[0150] 使Phusion Hot Start II高保真DNA聚合酶(NEB)进行PCR反应。采用降落PCR(98℃, 10秒; 72-62℃; -1℃/循环, 15秒; 72℃, 30秒] 10个循环, [98℃, 10秒; 62℃, 15秒; 72℃, 30秒] 25个循环)成功地扩增大多数基因座。必要时, 在68℃或72℃的恒定退火温度下和3% DMSO或1M甜菜碱中进行用于其余靶的PCR, 进行35个循环。在QIAXCEL毛细管电泳系统上分析PCR产物以验证大小和纯度。利用ExoSap-IT(Affymetrix)处理验证的产物, 并通过Sanger法(MGH DNA Sequencing Core)对所述产物进行测序以验证每一个靶位点。

[0151] 人细胞中的RGN诱导的中靶和脱靶突变频率的测定

[0152] 对于U2OS. EGFP和K562细胞, 按照制造商的方案(Lonza), 使用4D Nucleofector系统, 用250ng的gRNA表达质粒或空U6启动子质粒(用于阴性对照)、750ng Cas9表达质粒和30ng的td-Tomato表达质粒转染 $2 \times 10^5$ 个细胞。对于HEK293细胞, 按照制造商的方案(Life Technologies), 使用脂质体LTX试剂, 利用125ng的gRNA表达质粒或空U6启动子质粒(用于阴性对照)、375ng的Cas9表达质粒和30ng的td-Tomato表达质粒转染 $1.65 \times 10^5$ 个细胞。按照制造商的说明书, 使用QIAamp DNA Blood Mini试剂盒(QIAGEN)从转染的U2OS. EGFP、HEK293或K562细胞细胞收获基因组DNA。为了产生足够的基因组DNA以扩增脱靶候选位点, 将来自3个核转染(对于U2OS. EGFP细胞)、2个核转染(对于K562细胞)或2个脂质体LTX转染的DNA混合在一起, 随后进行T7EI。对于每一个测试的条件, 该过程进行2次, 从而产生一式两份代表总共4或6个单独的转染的基因组DNA的混合物。随后如上所述使用这些基因组DNA作为模板进行PCR, 随后按照制造商的说明书使用Ampure XP珠(Agencourt)进行纯化。如先

前所述(Reyon等,2012,同上)进行T7EI测定。

[0153] NHEJ介导的插入缺失突变的DNA测序

[0154] 将用于T7EI测定的纯化的PCR产物克隆入Zero Blunt TOP0载体(Life Technologies),利用MGH DNA Automation Core,使用碱裂解小量制备法分离质粒DNA。随后利用Sanger法(MGH DNA Sequencing Core),使用M13正向引物(5'-GTAAAACGACGGCCAG-3'(SEQ ID NO:19)对质粒进行测序。

[0155] 实施例1a.单核苷酸错配

[0156] 为了开始确定人细胞中的RGN的特异性决定簇,进行大规模测试来评估多个gRNA/靶DNA界面内的系统性错配的不同位置的效应。为了进行该测定,使用先前描述的(参见上述方法和Reyon等,2012,同上)基于人细胞的定量提高绿色荧光蛋白(EGFP)破坏测定,该测定使得能够快速定量靶向核酸酶活性(图2B)。在该测定中,被靶向单一整合的EGFP报告基因的核酸酶的活性可通过评估通过灭活移码插入/缺失(插入缺失)突变(通过核酸酶诱导的双链断裂(DSB)的易错非同源末端连接(NHEJ)修复引起的)引起的人U2OS.EGFP细胞中的荧光信号的损失来定量(图2B)。为了进行此处描述的研究,使用如下被靶向EGFP内的不同序列的三个~100nt的单一gRNA(sgRNA):

[0157] EGFP位点1GGGCACGGGCAGCTTGCCGGTGG(SEQ ID NO:1)

[0158] EGFP位点2GATGCCGTTCTTCTGCTTGTCGG(SEQ ID NO:2)

[0159] EGFP位点3GGTGGTGCAGATGAACTTCAGGG(SEQ ID NO:3)

[0160] 这些sgRNA中的每一个可高效地指导Cas9介导的EGFP表达的破坏(参见实施例1e和2a,以及图3E(顶)和3F(顶))。

[0161] 在最初的实验中,测试3个EGFP-靶向sgRNA的互补靶向区中的20个核苷酸中的19个核苷酸上的单核苷酸错配的效应。为进行该测定,产生针对3个在位置1至19(以3'至5'方向编号1至20;参见图1)上具有沃尔森-克里克颠换错配的靶位点的每一个的变体sgRNA,并且测试这些不同的sgRNA在测试的人细胞中指导Cas9介导的EGFP破坏的能力(未产生在位置20上具有取代的变体sgRNA,因为该核苷酸为U6启动子序列的部分,从而必须保留鸟嘌呤来避免影响表达)。

[0162] 对EGFP靶位点#2,gRNA的位置1-10中的单个错配对相关Cas9活性具有显著作用(图2C,中图),与表明相较于3'末端上的错配,gRNA的5'末端上的错配被更好地被耐受的先前研究一致(Jiang等,Nat Biotechnol 31,233-239(2013);Cong等,Science 339,819-823(2013);Jinek等,Science 337,816-821(2012))。然而,对于EGFP靶位点#1和#3,gRNA中除少数位置外所有位置上的单个错配似乎被良好地耐受,即使在序列的3'末端内亦如此。此外,对错配敏感的特定位置对于两个靶是不同的(图2C,比较顶图与底图)-例如,靶位点#1对位置2上的错配特别敏感,而靶位点#3对于位置1和8上的错配是最敏感的。

[0163] 实施例1b.多个错配

[0164] 为了测试gRNA/DNA界面上的不止一个错配的效应,在相邻和分开的位置中产生一系列具有双沃尔森-克里克颠换错配的变体sgRNA,并使用EGFP破坏测定在人细胞中测试这些sgRNA指导Cas9核酸酶活性的能力。所有这三个靶位点通常显示更大的对双改变的敏感性,在所述双改变中一个或两个错配存于gRNA靶向区域的3'半部分内。然而,这些效应的量级显示位点特异性变化,靶位点#2显示最大的对这些双错配的敏感性,并且靶位点#1通常

显示最小的敏感性。为了测试可被耐受的相邻错配的数目,构建在gRNA靶向区的5'末端中具有在位置19至15的范围内的递增数目的错配位置的变体sgRNA(其中单和双错配似乎被更好地耐受)。

[0165] 这些递增错配的sgRNA的测试显示对于所有3个靶位点,3个或更多个相邻错配的引入导致RGN活性的显著丧失。当从5'末端的位置19开始,朝向3'末端添加更多错配时,对于3个不同的EGFP靶向gRNA,发生活性的突然下降。具体地,在位置19和19+18上含有错配的显示基本上完全的活性,然而在位置19+18+17、19+18+17+16和19+18+17+16+15上具有错配的那些gRNA显示相对于阴性对照基本上无差异(图2F)。(注意,我们在这些变体gRNA中在位置20上不错配,因为该位置需要保留来作为G,因为其为驱动gRNA的表达的U6启动子的部分)。

[0166] 在下列实验中获得缩短gRNA互补性可导致具有更大特异性的RGN的另一个证据:对于4个不同的EGFP靶向gRNA(图2H),在位置18和19上引入双错配不显著影响活性。然而,在位置10和11上引入另一个双错配然后进入这些gRNA导致活性几乎完全丧失。有趣地,仅10/11双错配的引入通常对活性没有同样巨大的影响。

[0167] 综上所述,人细胞中的这些结果确认了RGN的活性可对gRNA靶向序列的3'半部分中的错配更敏感。然而,所述数据还明确地显示RGN的特异性是复杂的并且是靶位点依赖性的,单和双错配通常被良好地耐受,即使当一个或多个错配在gRNA靶向区域的3'半部分中时。此外,这些数据还表明并非gRNA/DNA界面的5'半部分中的所有错配都一定被良好耐受。

[0168] 另外,这些结果强有力表明具有较短的互补性区域(具体地~17nt)的gRNA在它们的活性上具有更大的特异性。我们指出,与由PAM序列赋予的2nt的特异性组合的17nt的特异性导致19bp序列(在大的复杂基因组诸如在人细胞中发现的那些基因组中是唯一的足够长度之一)的特异性。

[0169] 实施例1c.脱靶突变

[0170] 为了确定针对被靶向内源人基因的RGN的脱靶突变是否可被鉴定,使用靶向VEGFA基因中的3个不同位点、EMX1基因中的一个位点、RNF2基因中的一个位点和FANCF基因中的一个位点的6个sgRNA。这6个sgRNA在人U2OS. EGFP细胞中在它们各自的内源基因座上高效地指导Cas9介导的插入缺失,如通过T7内切核酸酶I(T7EI)测定(上述方法)检测的。对这6个RGN的每一个,我们随后在U2OS. EGFP细胞中针对核酸酶诱导的NHEJ介导的插入缺失突变检查成打的潜在脱靶位点(数目在46至多至64的范围内)。被评估的基因座包括相异在于1或2个核苷酸的所有基因组位点,以及成亚组的差异在于3至6个核苷酸并且具有朝向在gRNA靶向序列的5'半部分中具有这些错配中的一个或多个的那些位点的偏向性的基因组位点。通过使用T7EI测定,容易地鉴定了针对VEGFA位点1的4个脱靶位点(所检查的53个候选位点中的)、针对VEGFA位点2的12个脱靶位点(所检查的46个中的)、针对VEGFA位点3的7个脱靶位点(所检查的64个中的)和针对EMX1位点的1个脱靶位点(所检查的46个中的)。在分别针对RNF2或FANCF基因检查的43个和50个潜在位点中未检测到脱靶突变。经验证的脱靶位点上的突变率非常高,在期望的靶位点上观察到在5.6%至125%(40%的平均值)的范围内的突变率。这些真正的脱靶包括在靶位点的3'末端具有错配并且具有多至总共5个错配的序列,大多数脱靶位点在蛋白质编码基因内。一个亚组的脱靶位点的DNA测序提供插入缺失突变在预期的RGN裂解位点上的另外的分子确认(图8A-C)。

[0171] 实施例1d. 其它细胞类型中的脱靶突变

[0172] 在已确定RGN可在U2OS. EGFP细胞中以高频率诱导脱靶突变后, 随后设法确定这些核酸酶是否在其它类型的人细胞中也具有这些效应。U2OS. EGFP细胞已被选择用于初始实验, 因为这些细胞先前已被用于评价TALEN<sup>15</sup>的活性, 但人HEK293和K562细胞已被更广泛地用于测试靶向核酸酶的活性。因此, 还在HEK293和K562细胞中评估被靶向VEGFA位点1、2和3以及EMX1位点的4个RGN的活性。在这两个另外的人细胞系中, 这4个RGN中的每一个在它们期望的中靶位点上高效地诱导NHEJ介导的插入缺失突变(如通过T7EI测定法测定的), 虽然具有比在U2OS. EGFP细胞中观察到的突变频率稍微更低的突变频率。最初在U2OS. EGFP细胞中鉴定的这4个RGN的24个脱靶位点的评估显示许多位点在HEK293和K562细胞中再次与在对应的中靶位点上的频率相似的频率被突变。如所预期的, 一个亚组的来自HEK293细胞的这些脱靶位点的DNA测序提供改变在预期的基因组基因座上发生的另外的分子证据。出于某种因不知道为什么在HEK293细胞中有4个在U2OS. EGFP细胞中鉴定的脱靶位点以及在K562细胞中有11个所述脱靶序列未显示可检测的突变。然而, 这些脱靶位点中的许多位点在U2OS. EGFP细胞中也显示相对较低的突变频率。因此, 这些位点在HEK293和K562细胞中的突变率可落在我们的T7EI测定的可靠检测限(~2-5%)下方, 因为在我们的实验中, RGN通常似乎在HEK293和K562细胞中相较于U2OS. EGFP细胞具有较低的活性。综上所述, 在HEK293和K562细胞中的结果提供了我们对于RNG观察到的高频脱靶突变将为在多个人细胞类型中看到的一般现象的证据。

[0173] 实施例1e. 用于EGFP破坏测定的表达gRNA和Cas9的质粒的量的滴定

[0174] 产生针对位于EGFP核苷酸502上游(在其上通过非同源末端连接引入移码突变可强烈地破坏EGFP的表达的位置)的3个不同序列序列(上文中显示的EGFP位点1-3)的单一引导RNA (sgRNA) (Maeder, M.L. 等, Mol Cell 31, 294-301 (2008); Reyon, D. 等, Nat Biotech 30, 460-465 (2012))。

[0175] 对于所述3个靶位点的每一个, 最初将一系列表达gRNA的质粒的量(12.5至250ng)与750ng表达Cas9核酸酶的密码子最优化形式的质粒一起转染进我们的U2OS. EGFP报告细胞, 所述细胞具有单拷贝的组成型表达的EGFP-PEST报告基因。所有3个RGN在gRNA质粒的最高浓度(250ng)上高效地破坏EGFP表达(图3E(顶))。然而, 当转染较少量的表达gRNA的质粒时, 针对靶位点#1和#3的RGN显示等同水平的破坏, 而当降低转染的表达gRNA的质粒的量时, RGN在靶位点#2上的活性立即下降(图3E(顶))。

[0176] 滴定在转染进我们的U2OS. EGFP报告细胞的编码Cas9的质粒的量(在50ng至750ng的范围内), 测定EGFP破坏。如图3F(顶)中显示的, 靶位点#1耐受转染的表达Cas9的质粒的量的3倍降低而无EGFP破坏活性的显著丧失。然而, RGN靶向靶位点#2和#3的活性随着转染的Cas9质粒的量的3倍降低而立即降低(图3F(顶))。基于这些结果, 分别将25ng/250ng、250ng/750ng和200ng/750ng的表达gRNA的质粒/表达Cas9的质粒用于EGFP靶位点#1、#2和#3, 以进行实施例1a-1d中描述的实验。

[0177] 一些gRNA/Cas9组合在破坏EGFP表达中表现比其它组合好的原因还不清楚, 这些组合中的一些组合对用于转染的质粒的量更敏感或不太敏感的原因也不清楚。虽然对于这3个sgRNA, 存在于基因组中的脱靶位点的范围有可能影响它们的每一个的活性, 但对于这些特定的靶位点的每一个, 在相异在于1至6bp的基因组位点的数目上未看到差异(表1), 所

述差异可解释3个sgRNA的差异行为。

[0178] 表1. 针对被靶向内源人基因的6个RGN和被靶向EGFP报告基因的3个RGN的人基因组中的脱靶位点的数目

[0179]

| 靶位点           | 与中靶位点错配的数目 |   |    |     |      |       |       |
|---------------|------------|---|----|-----|------|-------|-------|
|               | 0          | 1 | 2  | 3   | 4    | 5     | 6     |
| 靶1 (VEGFA位点1) | 1          | 1 | 4  | 32  | 280  | 2175  | 13873 |
| 靶2 (VEGFA位点2) | 1          | 0 | 2  | 35  | 443  | 3889  | 17398 |
| 靶3 (VEGFA位点3) | 1          | 1 | 17 | 377 | 6028 | 13398 | 35517 |
| 靶4 (EMX)      | 1          | 0 | 1  | 18  | 276  | 2309  | 15731 |
| 靶5 (RNF2)     | 1          | 0 | 0  | 6   | 116  | 976   | 7443  |
| 靶6 (FANCF)    | 1          | 0 | 1  | 18  | 271  | 1467  | 9551  |
| EGFP靶位点 #1    | 0          | 0 | 3  | 10  | 156  | 1365  | 9755  |
| EGFP靶位点 #2    | 0          | 0 | 0  | 11  | 96   | 974   | 7353  |
| EGFP靶位点 #3    | 0          | 0 | 1  | 14  | 165  | 1439  | 10361 |

[0180] 在人基因组序列构建GRCh37中鉴定了被靶向VEGFA、RNF2、FANCF和EMX1基因的6个RGN和被靶向EGFP靶位点#1、#2和#3的3个RGN的每一个的脱靶位点。仅允许在与gRNA退火的20nt区域中但不允许在PAM序列中存在错配。

[0181] 实施例2: 使用成对的引导RNA与FokI-dCas9融合蛋白

[0182] 单体CRISPR-Cas9核酸酶被广泛地用于靶向基因组编辑, 但可以以高频率诱导不想要的脱靶突变。本实施例描述新的二聚体RNA引导的FokI核酸酶 (RFN), 其识别延长的双倍长度的序列并且其裂解活性严格依赖于两个单一引导RNA (gRNA)。RFN可以以高效率强劲地编辑内源人基因的DNA序列。另外, 描述了用于表达具有任何5' 的gRNA的方法, 为二聚体RFN提供有用的靶向范围的关键进步。在直接比较中, 单体Cas9切口酶通常比由匹配的单一gRNA引起的RFN高的频率诱导不想要的插入缺失和预料之外的焦点突变。RFN组合基于CRISPR RNA的靶向的容易性与二聚化的特异性提高, 并且为需要高度精确的基因组编辑的研究和治疗性应用提供重要的新平台。

[0183] 材料和方法

[0184] 在实施例2中使用下列材料和方法。

[0185] 单一和多重gRNA表达质粒

[0186] 在退火的靶位点寡双链体(Integrated DNA Technologies)和恒定区寡双链体(对于多重gRNA)与BsmBI消化的侧翼连接Csy4的gRNA主链(pSQT1313; Addgene)的单步骤连接中装配编码单一或多重gRNA的质粒。

[0187] 通过将如下序列连接至利用BsmBI II型限制性酶消化的U6-Csy4位点-gRNA质粒主链中来构建多重gRNA编码质粒: 1) 编码第一靶位点的退火的寡核苷酸, 2) 编码crRNA、tracrRNA和Csy4结合位点的磷酸化的退火的寡核苷酸, 和3) 编码第二靶位点的退火的寡核苷酸。将Csy4RNA结合位点附接于gRNA序列的3' 和5' 末端, 并将其与Cas9一起在细胞中表达。将Csy4RNA结合位点序列'GUUCACUGCCGUUAUAGGCAGCUAAGAAA'(SEQ ID NO:20)与标准gRNA序列的5' 和3' 末端融合。

[0188] GUUCACUGCCGUUAUAGGCAGNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUAAAA  
AUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGUGCGUUCACUGCCGUUAUAGGCAGNNNNNNNN  
 NNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUAAAAUAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA

CCGAGUCGGUGCGUUCACUGCCGUAUAGGCAG (SEQ ID NO:21)

[0189] 该序列是侧翼连接有Csy4位点(加以下划线的)的多重gRNA序列。在功能上,在一个转录物上以多重方式编码这些应当具有与单独编码它们相同的结果。虽然在本文所述的实验中在多重背景中表达所有成对的侧翼连接有Csy4的sgRNA,但可在于一个转录物上编码的相隔Csy4位点的多重sgRNA中以及具有另外的Csy4序列的单个sgRNA中编码sgRNA。在该序列中,第一N20序列代表与靶基因组序列的一条链互补的序列,并且第二N20序列代表与靶基因组序列的另一条链互补的序列。

[0190] 用编码通过‘2A’肽键联间隔的Cas9和Csy4蛋白的质粒共转染编码含有Csy4识别位点的gRNA的质粒。结果显示通过使用先前描述的U20S-EGFP破坏测定,具有与5’和3’末端融合的Csy4的gRNA仍然能够在人细胞中指导Cas9介导的裂解。因此,可将Csy4RNA结合位点附接于gRNA序列的3’末端,并且这些含Csy4位点的gRNA与Cas9的复合物在细胞中仍然保持功能性。

[0191] 在一些实验中,使用编码Csy4-T2A-FokI-dCas9的构建体。FokI-dCas9融合物的序列示于下文中,并且包括FokI与dCas9之间的GGGGS (SEQ ID NO:23) 接头(加以下划线)和核定位序列。

[0192] FokI-dCas9氨基酸序列(FokI-G4S-dCas9-nls-3XFLAG)

[0193] MQLVKSELEEKKSELRHKLKYPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPD  
GAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWKVYPSSVTEFKFLFVSGHFKG  
NYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNGEINFGGGGSDKKYSIGLAIGTNSVGWA  
VITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD  
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIE  
GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLG  
LTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK  
RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL  
RKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITP  
WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL  
LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKLIKDKDFLDNEENEDI EDIVLTTLTFE  
DREMIEERLKYAHLFFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQM LIHDDSLT  
FKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIV IEMARENQTTQKGQKNSRE  
RMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK  
VLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV  
AQILDSRMNTKYDENDKLIREVKVITLKS LVSDFRKDFQFYK VREINNYHHAHDAYLNAVGTALIKKYPKLESE  
FVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV  
RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKS  
VKELLGITIMERSSFEKNPIDFLEAKGYKEVKDLIIKLPKYSLFELNGRKRMLASAGELQKGNELALPSKYVNF  
LYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKRVI LADANLDKVL SAYNKH RDKPIREQAENIIH  
LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGDGSPKKKRKVSSDYKDHDGD  
YKDHDIDYKDDDDK (SEQ ID NO:24)

[0194] FokI-dCas9核苷酸序列(FokI-G4S-dCas9-nls-3XFLAG)

[0195] ATGCAACTAGTCAAAAGTGAAGTGGAGGAGAAGAAATCTGAACCTTCGTCATAAATTGAAATATGTGCC  
TCATGAATATATTGAATTAATTGAAATTGCCAGAAATTCCTCAGGATAGAATTCTTGAATGAAGTAATGGAA  
TTTTTTATGAAAGTTTATGGATATAGAGGTAAACATTTGGGTGGATCAAGGAAACCGGACGGAGCAATTTATACTG  
TCGGATCTCCTATTGATTACGGTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTATAATCTGCCAATTGGCCA  
AGCAGATGAAATGCAACGATATGTCGAAGAAAATCAAACACGAAACAAACATATCAACCCTAATGAATGGTGGAAA  
GTCTATCCATCTTCTGTAACGGAATTTAAGTTTTTATTTGTGAGTGGTCACTTTAAAGGAACTACAAAGCTCAGC  
TTACACGATTAAATCATATCACTAATTGTAATGGAGCTGTTCTTAGTGTAGAAGAGCTTTTAATTGGTGGAGAAAT  
GATTAAGCCGGCACATTAACCTTAGAGGAAGTCAGACGGAAATTTAATAACGGCGAGATAAACTTTGGTGGCGGT  
GGATCCGATAAAAAGTATTCTATTGGTTTAGCCATCGGCACTAATTCGGTTGGATGGGCTGTCATAACCGATGAAT  
ACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATCGGTGC  
CCTCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGC  
AAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTCTTTTACCCTT  
TGGAAGAGTCCTTCTTGTGCAAGAGGACAAGAAACATGAACGGCACCCCATCTTGGAAACATAGTAGATGAGGT  
GGCATATCATGAAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTG  
AGGTAAATCTACTTGGCTCTTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGG  
ACAACTCGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAA  
TGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCA  
CAATTACCCGGAGAGAAGAAAAATGGGTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTA  
AGTCGAACCTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATCT  
ACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGCAATCCTCCTATCT  
GACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATC  
ACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCA  
GTGCAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATA  
TTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTT  
TCGACAACGGTAGCATTCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTAA  
TCCGTTCTCAAAGACAATCGTGAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTG  
GCCCAGGGAACTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAGGAAG  
TTGTGCATAAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAA  
AGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACT  
GAGGGCATGCGTAAACCCGCCCTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACC  
GCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCTGTGAGATCTCCGG  
GGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTCCTG  
GATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTG  
AGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGG  
CTGGGGACGATTGTGCGGAAACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTA  
AAGAGCGACGGCTTCGCCAATAGGAACCTTTATGCAGCTGATCCATGATGACTCTTAACCTTCAAAGAGGATATAC  
AAAAGGCACAGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAA  
AAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAACATT  
GTAATCGAGATGGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAGAATAG

AAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGAAAAATACCCAATTGCAGAACGAGAA  
ACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGAT  
TACGACGTCGATGCCATTGTACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGG  
ATAAGAACCGAGGGAAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAAATGAAGAACTATTGGCGGCAGCT  
CCTAAATGCGAAACTGATAACGCAAAGAAAGTTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTT  
GACAAGGCCGATTTATTAAACGTCAGCTCGTGGAAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATT  
CCCGAATGAATACGAAATACGACGAGAACGATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAT  
GGTGTGCGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAAATAACTACCACCATGCGCACGACGCT  
TATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATT  
ACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTT  
TTATTCTAACATTATGAATTTCTTTAAGACGGAATCACTCTGGCAAACGAGAGATACGCAAACGACCTTTAATT  
GAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCA  
TGCCCCAAGTCAACATAGTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAAGGAATCGATTCTTCAAAAAG  
GAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTT  
GCCTATTCTGTCTAGTAGTGGCAAAAGTTGAGAAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTGG  
GGATAACGATTATGGAGCGCTCGTCTTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGT  
AAAAAAGGATCTCATAATTAAGTACCAAAGTATAGTCTGTTTGAAGTTAGAAAATGGCCGAAAACGGATGTTGGCT  
AGCGCCGGAGAGCTTCAAAAGGGGAACGAACCTCGCACTACCGTCTAAATACGTGAATTCCTGTATTTAGCGTCCC  
ATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCT  
CGACGAAATCATAGAGCAAATTTTCGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTA  
AGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAAATATTATCCATTTGTTTACTCTTACCA  
ACCTCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGGT  
GCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAACTCGGATAGATTTGTCACAGCTTGGGGGT  
GACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACA  
TCGATTACAAGGATGACGATGACAAGTGA (SEQ ID NO:25)

[0196] 或者,使用构建体的人密码子最优化的形式,该形式含有N末端和C末端核定位信号,如下文中显示的。

[0197] N1s-FokI-dCas9-n1s氨基酸序列

[0198] MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA  
RRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD  
STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR  
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDYYDDLDNLLAQIGDQYADLFLAAKNL  
SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF  
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI  
PYVVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLEYFTVYNEL  
TKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI  
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG  
KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMG  
RHKPENIVIMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL

DINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE  
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVDREINN  
YHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE  
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG  
GFDSPTVAYSVLVVAKEVEKGSKKLKSVELLGITIMERSSEFKNPIDFLEAKGYKEVKDLIIKLPKYSLFELEN  
GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLAD  
ANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTL IHQSITGLYETRI  
DLSQLGGDGSPKKRKVSSDYKDHDGDYKDHDIDYKDDDDK (SEQ ID NO:26)

[0199] Nls-FokI-dCas9-nls核苷酸序列

[0200] ATGCCTAAGAAGAAGCGGAAGGTGAGCAGCCAACTTGTGAAGTCTGAACTCGAGGAGAAAAATCAGA  
GTTGAGACACAAGTTGAAGTACGTGCCACACGAATACATCGAGCTTATCGAGATCGCCAGAAACAGTACCCAGGAT  
AGGATCCTTGAGATGAAAGTCATGGAGTTCTTTATGAAGGTCTACGGTTATAGAGGAAAGCACCTTGGCGGTAGCA  
GAAAGCCCGATGGCGCCATCTATACTGTCGGATCTCCTATCGATTATGGGGTGATCGTGATACCAAAGCTTACTC  
AGGCGGGTACAACCTGCCCATAGGACAAGCCGACGAGATGCAGCGGTATGTGGAAGAGAACCAGACGCGCAACAAG  
CACATCAACCCCAATGAATGGTGGAAAGTGTAACCAAGTAGTGTGACTGAGTTCAAGTTCCTGTTTGTCTCCGGCC  
ACTTTAAGGGCAATTATAAAGCTCAGCTCACTAGACTCAATCACATCACAACTGCAACGGAGCTGTGTTGTCAGT  
GGAGGAGCTCCTGATTGGAGGCGAGATGATCAAAGCCGGCACCCCTTACACTGGAGGAGGTGCGGCGGAAGTTCAAC  
AATGGAGAGATCAACTTCGGTGGCGGTGGATCCGATAAAAAGTATTCTATTGGTTTAGCCATCGGCACTAATTCCG  
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCA  
TTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGA  
ACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCA  
AAGTTGACGATTCTTTCTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCC  
CATCTTTGGAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTA  
GTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATAAAGTTCCGTGGGCACT  
TTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCGACAAAACCTGTTTCATCCAGTTAGTACAAACCTATAA  
TCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAA  
TCCCGACGGCTAGAAAACCTGATCGCACAAATTACCCGAGAGAGAAGAAAAATGGGTTGTTCCGTAACCTTATAGCGC  
TCTCACTAGGCCTGACACCAAATTTTAAGTCGAAGTTGCACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGA  
CACGTACGATGACGATCTCGACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAA  
AACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTT  
CAATGATCAAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGA  
GAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAG  
GAATTCTACAAGTTTATCAAACCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAAGTCAATCGCG  
AAGATCTACTGCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGC  
TATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATTGAGAAAATCCTAACCTTT  
CGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAA  
CGATTACTCCATGGAATTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAA  
CTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAAT  
GAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGAAGAAAGCAA

TAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGA  
ATGCTTCGATTCTGTGCGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTA  
AAGATAATTAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTGACTCTTA  
CCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTGACGATAAGGTTATGAA  
ACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTGCGGAAACTTATCAACGGGATAAGAGACAAGCAA  
AGTGGTAAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGATG  
ACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGC  
GAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTC  
ATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAATCAAACGACTCAGAAGGGGCAAAAAA  
ACAGTCGAGAGCGGATGAAGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGT  
GGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAG  
GAACTGGACATAAACCGTTTATCTGATTACGACGTCGATGCCATTGTACCCCAATCCTTTTTGAAGGACGATTCAA  
TCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAA  
GAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAGAAAAGTTCGATAACTTAACTAAA  
GCTGAGAGGGGTGGCTTGCTGAACCTGACAAGGCCGATTTATTAAACGTCAGCTCGTGGAAACCCGCCAAATCA  
CAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACGAGAACGATAAGCTGATTGCGGAAGT  
CAAAGTAATCACTTTAAAGTCAAATTTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATA  
AATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTGCTAGGGACCGCACTCATTAGAAATACCCGAAGC  
TAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAGGAGAT  
AGGCAAGGCTACAGCCAAATACTTCTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAAC  
GGAGAGATACGCAAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACT  
TCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAAGTGAAGTGCAGACCGGAGGGTT  
TTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAAAAG  
TACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGCCAAAAGTTGAGAAGGGAAAATCCAAGA  
AACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAGAAGCCCATCGACTT  
CCTTGAGGCGAAAAGTTACAAGGAAGTAAAAAAGGATCTCATAATTAAGTACCAAAGTATAGTCTGTTTGAGTTA  
GAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGAGAGCTTCAAAGGGGAACGAACTCGCACTACCGTCTAAAT  
ACGTGAATTTCTGTATTTAGCGTCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACT  
TTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTA  
GCTGATGCCAATCTGGACAAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAA  
ATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGATTCAAGTATTTTGACACAACGATAGATCG  
CAAACGATACACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAACT  
CGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAAGTCTCGAGCGACTACAAAGACC  
ATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA (SEQ ID NO:27)

[0201] 组织培养和转染

[0202] 在HEK 293细胞、U2OS细胞中或具有稳定地整合的单拷贝去稳定化的EGFP基因的U2OS细胞(U2OS. EGFP细胞)中进行所有细胞培养实验。将细胞在37℃和5%CO<sub>2</sub>下于补充有10%FBS、2mM GlutaMax (Life Technologies) 和青霉素/链霉素的高级DMEM (Life Technologies) 进行培养。另外,在400μg/ml的G418存在的情况下培养U2OS. EGFP细胞。

[0203] 按照制造商的说明书,使用Lonza 4D-Nucleofector的DN-100程序转染U2OS细胞和U2OS.EGFP细胞。在最初的FokI-dCas9活性筛选和聚焦间隔区长度分析实验中,一起转染750ng的pCAG-Csy4-FokI-dCas9-nls核酸酶质粒和250ng的gRNA编码质粒,以50ng tdTomato表达质粒(Clontech)作为转染对照。在所有其它实验中,在U2OS和U2OS.EGFP细胞中,将975ng的人密码子最优化的pCAG-Csy4-T2A-nls-hFokI-dCas9-nls (SQT 1601) 或pCAG-Cas9-D 10A切口酶(NW3)与325ng的gRNA载体和10ng的Td tomato表达质粒一起转染,转染后3天进行分析。按照制造商的说明书使用脂质体(Life Technologies),利用750ng的核酸酶质粒、250ng的gRNA表达质粒和10ng的Td tomato转染HEK293细胞,并在转染后3天分析NHEJ介导的诱变。

[0204] 对于初始间隔区活性筛选,进行单次转染,并且对于聚焦间隔区长度分析进行一式二份转染。以一式三份进行所有其它转染。

[0205] EGFP破坏测定

[0206] 如先前所述(参见实施例1和Reyon等,Nat Biotech 30,460-465(2012)),使用U2OS.EGFP报告细胞进行EGFP破坏测定。使用BD Biosciences LSR II或Fortessa FACS分析仪测定细胞的EGFP和tdTomato表达。

[0207] 通过T7EI测定定量核酸酶或切口酶诱导的突变率

[0208] 如先前所述(Reyon等,Nat Biotech 30,460-465(2012))进行T7EI测定。简言之,转染后72小时,按照制造商的说明书,利用Sciclone G3液体处理工作站(Caliper),使用Agencourt DNAdvance基因组DNA分离试剂盒(Beckman Coulter Genomics)分离基因组DNA。使用Phusion Hot-start Flex DNA聚合酶(New England Biolabs)进行扩增基因组基因座的PCR反应。使用两步骤方案(98℃,30秒;(98℃,7秒;72℃,30秒)×35;72℃,5分钟)或降落PCR方案((98℃,10秒;72-62℃,-1℃/循环,15秒;72℃,30秒)×10个循环,(98℃,10秒;62℃,15秒;72℃,30秒)×25个循环)扩增样品。使200ng纯化的PCR变性、杂交,随后用T7内切核酸酶I(New England Biolabs)进行处理。如先前所述(Reyon等,Nat Biotech 30,460-465(2012))使用Qiaxcel毛细管电泳仪(Qiagen)定量突变频率。

[0209] 诱变的基因组DNA的Sanger测序

[0210] 对用于T7EI测定的相同的纯化的PCR产物进行Topo克隆(Life Technologies),随后分离单个克隆的质粒DNA,并使用M13反相引物(5'-GTAAAACGACGGCCAG-3';SEQ ID NO:19)对其进行测序。

[0211] Illumina文库的制备和分析

[0212] 使用Phusion Hot-start FLEX DNA聚合酶扩增短的200-350bp PCR产物。按照制造商的说明书,使用Ampure XP珠粒(Beckman Coulter Genomics)纯化PCR产物。在Sciclone G3液体处理工作站上使用高通量文库制备系统(Kapa Biosystems)制备双指数浮动型TruSeq Illumina深度测序文库。使用Qiaxcel毛细管电泳仪(Qiagen)定量最终的接头连接的文库。由Dana-Farber Cancer Institute Molecular Biology Core在Illumina MiSeq测序仪上进行150bp双端测序。

[0213] 使用bwa将MiSeq双端测序读数定位至人基因组参照GChr37。分析具有平均质量评分>30的读数的与期望的靶或候选脱靶核酸酶结合位点重叠的插入或缺失突变。使用基因组分析工具包(GATK)和Python进行突变分析。

[0214] 脱靶搜索算法:

[0215] 执行靶位点匹配算法,所述算法在整个人基因组范围内寻找在滑动窗口中具有少于指定数目的错配的匹配。

[0216] 实施例2a.用于设计二聚体RNA引导的核酸酶的基本原理

[0217] 假设可开发组合二聚化的特异性有利方面与Cas9靶向的容易性的单个平台。为此,将良好表征的二聚化依赖性FokI核酸酶结构域与RNA引导的无催化活性的Cas9 (dCas9) 蛋白融合。希望,与含FokI的ZFN和TALEN一样,当结合于由两个“半位点”(其间具有特定长度的“间隔区”序列)组成的靶位点时(图4A),这些融合物的二聚体可能介导序列特异性DNA裂解。有人假设,此类融合物具有提高的特异性,因为它们应当需要两个gRNA来具有活性(图4A)以及因为单个gRNA可假定地太无效以至不能或不能招募进行DNA裂解所需的两个含FokI融合蛋白。有人假设,此类嵌合系统可显示相对于标准单体Cas9核酸酶的提高特异性,并且还可潜在地具有优于其中单个切口酶仍然可产生不想要的诱变效应的成对切口酶系统的重要特异性有利方面。

[0218] 实施例2b.无5'末端核苷酸限制的gRNA的多重表达

[0219] 使用现有gRNA表达法,二聚体RNA引导的核酸酶的靶向范围可以是低的。两序列要求通常限制dCas9单体的靶向范围;对于由dCas9指定的5'-NGG的PAM序列的需要以及对gRNA的5'末端上的G核苷酸的需要可通过在大多数表达载体中使用U6启动子来施加。然而,如果可解除对gRNA中的5'G的需要,则靶向范围可拓宽16倍。

[0220] 为了开发允许表达具有任何5'核苷酸的gRNA的多重系统,构建可从其在从U6启动子转录的单一RNA内表达各自侧翼连接有Csy4核糖核酸酶的裂解位点的两个gRNA (Haurwitz等,Science 329,1355-1358 (2010)) (图4B)。预期Csy4加工该转录物从而释放2个gRNA。基于Csy4介导的裂解的已知机制(Haurwitz等,Science 329,1355-1358 (2010); Sternberg等,RNA 18,661-672 (2012)),每一个加工的gRNA应当在其3'末端保留Csy4识别位点,Csy4蛋白结合于该位点(图4B)。在该构型中,应当可能表达具有任何5'核苷酸的gRNA。通过使用其表达两个被靶向EGFP报告基因内的位点的gRNA来测试该系统。该转录物与Csy4和Cas9核酸酶一起在人细胞中的共表达导致两个EGFP靶位点上的插入缺失突变以及两个位点之间的序列的缺失的引入(图4C)。这些实验表明两个gRNA都从单个亲代RNA转录物加工并且两者都能够人在细胞中指导Cas9核酸酶活性。

[0221] 实施例2c.二聚体RNA引导的核酸酶的构建和最优化

[0222] 构建两个不同的具有FokI核酸酶结构域的杂交蛋白和dCas9蛋白:其中FokI核酸酶结构域与dCas9的羧基末端融合的一种(dCas9-FokI)和其中其与氨基末端融合的另一种(FokI-dCas9) (图5A)。dCas9-FokI蛋白在结构上与ZFN和TALEN类似(图5A)。为了确定这些融合物的一种或两种是否可介导DNA的位点特异性裂解,使用可快速且容易地定量NHEJ介导的插入缺失至EGFP报告基因内的引入的良好建立的基于人细胞的测定(实施例1中上述的EGFP破坏测定)。因为进行高效裂解所需的半位点的几何学是未知的,因此设计60对被靶向EGFP中的不同位点的gRNA。定向被这些gRNA对的每一个被向的两个半位点,以使它们的PAM序列都与间隔区序列紧密相邻(“PAM在内”取向)或都位于全长靶位点外界(“PAM在外”取向)(图5B)。另外,间隔区的长度也从0变化至31bp(图5B和表2)。

[0223]

表 2

| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+)'位点                     | SEQ ID<br>NO: | 序列(-)'位点                    | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
|-----------------------|------------|--------------|------------------------------|---------------|-----------------------------|---------------|------------------|-----|
| 1                     | EGFP 位点 1  | 74           | GAGCTGGACGGCGAC<br>GTAAACGG  | 28.           | CGCCGGACACGCTGAAC<br>TTGTGG | 29.           | 0                | 在内  |
| 2                     | EGFP 位点 2  | 174          | CCGGCAAGCTGCCCCG<br>TGCCCTGG | 30.           | GGTCAGGGTGGTCACGA<br>GGGTGG | 31.           | 1                | 在内  |
| 3                     | EGFP 位点 3  | 37           | CGAGGAGCTGTTTCAC<br>CGGGGTGG | 32.           | CCGTCCAGCTCGACCAGG<br>ATGGG | 33.           | 2                | 在内  |
| 4                     | EGFP 位点 4  | 37           | CGAGGAGCTGTTTCAC<br>CGGGGTGG | 34.           | GCCGTCCAGCTCGACCAG<br>GATGG | 35.           | 3                | 在内  |
| 5                     | EGFP 位点 5  | 174          | CCGGCAAGCTGCCCCG<br>TGCCCTGG | 36.           | GTAGGTCAGGGTGGTCAC<br>GAGGG | 37.           | 4                | 在内  |
| 6                     | EGFP 位点 6  | 34           | GGGCGAGGAGCTGTT<br>CACCGGG   | 38.           | CCGTCCAGCTCGACCAGG<br>ATGGG | 39.           | 5                | 在内  |
| 7                     | EGFP 位点 7  | 33           | AGGGCGAGGAGCTGT<br>TCACCGG   | 40.           | CCGTCCAGCTCGACCAGG<br>ATGGG | 41.           | 6                | 在内  |
| 8                     | EGFP 位点 8  | 32           | AAGGGCGAGGAGCTG<br>TTCACCG   | 42.           | CCGTCCAGCTCGACCAGG<br>ATGGG | 43.           | 7                | 在内  |
| 9                     | EGFP 位点 9  | 32           | AAGGGCGAGGAGCTG<br>TTCACCG   | 44.           | GCCGTCCAGCTCGACCAG<br>GATGG | 45.           | 8                | 在内  |
| 10                    | EGFP 位点 10 | 106          | CAGCGTGTCGGGCGA<br>GGCGAGG   | 46.           | CTTCAGGGTCAGCTTGCC<br>GTAGG | 47.           | 9                | 在内  |
| 11                    | EGFP 位点 11 | 34           | GGGCGAGGAGCTGTT<br>CACCGGG   | 48.           | CGTCGCCGTCCAGCTCGA<br>CCAGG | 49.           | 10               | 在内  |
| 12                    | EGFP 位点 12 | 33           | AGGGCGAGGAGCTGT<br>TCACCGG   | 50.           | CGTCGCCGTCCAGCTCGA<br>CCAGG | 51.           | 11               | 在内  |
| 13                    | EGFP 位点 13 | 32           | AAGGGCGAGGAGCTG<br>TTCACCG   | 52.           | CGTCGCCGTCCAGCTCGA<br>CCAGG | 53.           | 12               | 在内  |
| 14                    | EGFP 位点 14 | 155          | CTGAAGTTCATCTGCA<br>CCACCGG  | 54.           | GTGGTCACGAGGGTGGG<br>CCAGGG | 55.           | 13               | 在内  |

[0224]

表 2

| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+)'位点                     | SEQ ID<br>NO: | 序列(-)'位点                     | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
|-----------------------|------------|--------------|------------------------------|---------------|------------------------------|---------------|------------------|-----|
| 15                    | EGFP 位点 15 | 101          | AAGTTCAGCGTGTCCG<br>GCGAGGG  | 56.           | CTTCAGGGTCAGCTTGCC<br>GTAGG  | 57.           | 14               | 在内  |
| 16                    | EGFP 位点 16 | 100          | CAAGTTCAGCGTGTCC<br>GGCGAGG  | 58.           | CTTCAGGGTCAGCTTGCC<br>GTAGG  | 59.           | 15               | 在内  |
| 17                    | EGFP 位点 17 | 58           | GGTGCCCATCCTGGTC<br>GAGCTGG  | 60.           | CGCCGGACACGCTGAAC<br>TTGTGG  | 61.           | 16               | 在内  |
| 18                    | EGFP 位点 18 | 74           | GAGCTGGACGGCGAC<br>GTAAACGG  | 62.           | GGCATCGCCCTCGCCCTC<br>GCCGG  | 63.           | 17               | 在内  |
| 19                    | EGFP 位点 19 | 307          | GGAGCGCAACCATCTTC<br>TTCAAGG | 64.           | CTCGAACTTCACCTCGGC<br>GCCGG  | 65.           | 18               | 在内  |
| 20                    | EGFP 位点 20 | 155          | CTGAAGTTCATCTGCA<br>CCACCGG  | 66.           | GTCAGGGTGGTCACGAG<br>GGTGGG  | 67.           | 19               | 在内  |
| 21                    | EGFP 位点 21 | 95           | GGCCACAAAGTTCAGC<br>GTGTCCGG | 68.           | CTTCAGGGTCAGCTTGCC<br>GTAGG  | 69.           | 20               | 在内  |
| 22                    | EGFP 位点 22 | 203          | CTCGTGACCAACCCTGA<br>CCTACGG | 70.           | CGTGTGCTTTCATGTGGT<br>CGGGG  | 71.           | 21               | 在内  |
| 23                    | EGFP 位点 23 | 174          | CCGGCAAGCTGCCCCG<br>TGCCCTGG | 72.           | GCTGAAGCACTGCACGC<br>CGTAGG  | 73.           | 22               | 在内  |
| 24                    | EGFP 位点 24 | 107          | AGCGTGTCGGGCGAG<br>GGCGAGGG  | 74.           | GGTGGTGCAGATGAAC TT<br>CAGGG | 75.           | 23               | 在内  |
| 25                    | EGFP 位点 25 | 106          | CAGCGTGTCGGGCGA<br>GGGCGAGG  | 76.           | GGTGGTGCAGATGAAC TT<br>CAGGG | 77.           | 24               | 在内  |
| 26                    | EGFP 位点 26 | 49           | CACCGGGGTGGTGCC<br>CATCCTGG  | 78.           | CGCCGGACACGCTGAAC<br>TTGTGG  | 79.           | 25               | 在内  |
| 27                    | EGFP 位点 27 | 122          | GGCGAGGGCGATGCC<br>ACCTACGG  | 80.           | GGGCACGGGCAGCTTGC<br>CGGTGG  | 81.           | 26               | 在内  |
| 28                    | EGFP 位点 28 | 203          | CTCGTGACCAACCCTGA<br>CCTACGG | 82.           | AGAAAGTCGTGCTGTCA<br>TGTGG   | 83.           | 27               | 在内  |

[0225]

表 2

| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+)'位点                     | SEQ ID<br>NO: | 序列(-)'位点                     | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
|-----------------------|------------|--------------|------------------------------|---------------|------------------------------|---------------|------------------|-----|
| 29                    | EGFP 位点 29 | 337          | CAACTACAAGACCCG<br>CGCCGAGG  | 84.           | CGATGCCCTTCAGCTCGA<br>TGCGG  | 85.           | 28               | 在内  |
| 30                    | EGFP 位点 30 | 62           | CCCATCCTGTGTCGAGC<br>TGGACGG | 86.           | GGCATGCGCCCTCGCCCTC<br>GCCGG | 87.           | 29               | 在内  |
| 31                    | EGFP 位点 31 | 100          | CAAGTTCAGCGTGTCC<br>GGCGAGG  | 88.           | GGTGGTGCAGATGAAC TT<br>CAGGG | 89.           | 30               | 在内  |
| 32                    | EGFP 位点 32 | 74           | GAGCTGGACGGCGAC<br>GTAAACGG  | 90.           | GACCAGGATGGGCACCA<br>CCCCGG  | 91.           | 0                | 在外  |
| 33                    | EGFP 位点 33 | 314          | ACCATCTTCTTCAAGG<br>ACGACGG  | 92.           | CGCTCCTGGACGTAGCCT<br>TCGGG  | 93.           | 1                | 在外  |
| 34                    | EGFP 位点 34 | 122          | GGCGAGGGCGATGCC<br>ACCTACGG  | 94.           | CGCCGGACACGCTGAAC<br>TTGTGG  | 95.           | 2                | 在外  |
| 35                    | EGFP 位点 35 | 275          | TTCAAAGTCCGCCATGC<br>CCGAAGG | 96.           | GTCGTGCTGCTTTCATGTG<br>GTCGG | 97.           | 3                | 在外  |
| 36                    | EGFP 位点 36 | 275          | TTCAAAGTCCGCCATGC<br>CCGAAGG | 98.           | TCGTGCTGCTTTCATGTGG<br>TCGGG | 99.           | 4                | 在外  |
| 37                    | EGFP 位点 37 | 95           | GGCCACAAAGTTCAGC<br>GTGTCCGG | 100.          | CGTCGCCGTCCAGCTCGA<br>CCAGG  | 101.          | 5                | 在外  |
| 38                    | EGFP 位点 38 | 203          | CTCGTGACCCCTGA<br>CCTACGG    | 102.          | CCAGGGCACGGGCAGCT<br>TGCCGG  | 103.          | 6                | 在外  |
| 39                    | EGFP 位点 39 | 463          | CAGCCACAAACGTCTAT<br>ATCATGG | 104.          | TGTACTCCAGCTTGTGCC<br>CCAGG  | 105.          | 7                | 在外  |
| 40                    | EGFP 位点 40 | 95           | GGCCACAAAGTTCAGC<br>GTGTCCGG | 106.          | GCCGTCCAGCTCGACCAG<br>GATGG  | 107.          | 9                | 在外  |
| 41                    | EGFP 位点 41 | 95           | GGCCACAAAGTTCAGC<br>GTGTCCGG | 108.          | CCGTCCAGCTCGACCAGG<br>ATGGG  | 109.          | 10               | 在外  |
| 42                    | EGFP 位点 42 | 101          | AAGTTCAGCGTGTCCG<br>GCGAGGG  | 110.          | CGTCGCCGTCCAGCTCGA<br>CCAGG  | 111.          | 11               | 在外  |

[0226]

表 2

| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+ )位点                     | SEQ ID<br>NO: | 序列(-)位点                     | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
|-----------------------|------------|--------------|------------------------------|---------------|-----------------------------|---------------|------------------|-----|
| 43                    | EGFP 位点 43 | 350          | CGCGCCGAGGTGAAG<br>TTCGAGGG  | 112.          | GCCGTCGTCCTTGAAGAA<br>GATGG | 113.          | 12               | 在外  |
| 44                    | EGFP 位点 44 | 174          | CCGGCAAGCTGCCCCG<br>TGCCCTGG | 114.          | CTTCAGGGTCAGCTTGCC<br>GTAGG | 115.          | 13               | 在外  |
| 45                    | EGFP 位点 45 | 100          | CAAGTTCAGCGTGTCC<br>GGCGAGG  | 116.          | GCCGTCCAGCTCGACCAG<br>GATGG | 117.          | 14               | 在外  |
| 46                    | EGFP 位点 46 | 100          | CAAGTTCAGCGTGTCC<br>GGCGAGG  | 118.          | CCGTCCAGCTCGACCAGG<br>ATGGG | 119.          | 15               | 在外  |
| 47                    | EGFP 位点 47 | 101          | AAGTTCAGCGTGTCCG<br>GCGAGGG  | 120.          | CCGTCCAGCTCGACCAGG<br>ATGGG | 121.          | 16               | 在外  |
| 48                    | EGFP 位点 48 | 107          | AGCGTGTCGGCGGAG<br>GGCGAGGG  | 122.          | CGTCGCCGTCCAGCTCGA<br>CCAGG | 123.          | 17               | 在外  |
| 49                    | EGFP 位点 49 | 155          | CTGAAGTTCATCTGCA<br>CCACCCG  | 124.          | GGCATCGCCCTCGCCCTC<br>GCCGG | 125.          | 18               | 在外  |
| 50                    | EGFP 位点 50 | 106          | CAGCGTGTCGGCGCA<br>GGCGGAGG  | 126.          | GCCGTCCAGCTCGACCAG<br>GATGG | 127.          | 20               | 在外  |
| 51                    | EGFP 位点 51 | 95           | GGCCACAAGTTCAGC<br>GTGTCCGG  | 128.          | GACCAGGATGGCAACCA<br>CCCCGG | 129.          | 21               | 在外  |
| 52                    | EGFP 位点 52 | 107          | AGCGTGTCGGCGGAG<br>GGCGAGGG  | 130.          | CCGTCCAGCTCGACCAGG<br>ATGGG | 131.          | 22               | 在外  |
| 53                    | EGFP 位点 53 | 337          | CAACTACAAGACCCG<br>CGCCGAGG  | 132.          | GCGTCTCTGGACGTAGCC<br>TTCGG | 133.          | 23               | 在外  |
| 54                    | EGFP 位点 54 | 337          | CAACTACAAGACCCG<br>CGCCGAGG  | 134.          | CGTCTCTGGACGTAGCCT<br>TCGGG | 135.          | 24               | 在外  |
| 55                    | EGFP 位点 55 | 397          | GCTGAAGGGCATCGA<br>CTTCAAGG  | 136.          | CCTCGAACTTCACCTCGG<br>CGCGG | 137.          | 25               | 在外  |
| 56                    | EGFP 位点 56 | 100          | CAAGTTCAGCGTGTCC<br>GGCGAGG  | 138.          | GACCAGGATGGCAACCA<br>CCCCGG | 139.          | 26               | 在外  |

[0227]

表 2

| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+ )位点                     | SEQ ID<br>NO: | 序列(-)位点                      | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
|-----------------------|------------|--------------|------------------------------|---------------|------------------------------|---------------|------------------|-----|
| 57                    | EGFP 位点 57 | 101          | AAGTTCAGCGTGTCCG<br>GCGAGGG  | 140.          | GACCAGGATGGGCACCA<br>CCCCGG  | 141.          | 27               | 在外  |
| 58                    | EGFP 位点 58 | 400          | GAAGGGCATCGACTTC<br>AAGGAGG  | 142.          | CCTCGAACTTCACCTCGG<br>CGCGG  | 143.          | 28               | 在外  |
| 59                    | EGFP 位点 59 | 337          | CAACTACAAGACCCG<br>CGCCGAGG  | 144.          | CTGGACGTAGCCTTCGGG<br>CATGG  | 145.          | 29               | 在外  |
| 60                    | EGFP 位点 60 | 307          | GGAGCGCACCATCTTC<br>TTCAAGG  | 146.          | AGAAAGTCGTGCTGCTTCA<br>TGTGG | 147.          | 31               | 在外  |
| 61                    | EGFP 位点 61 | 100          | CAAGTTCAGCGTGTCC<br>GGCGAGG  | 148.          | CGTCGCCGTCCAGCTCGA<br>CCAGG  | 149.          | 10               | 在外  |
| 62                    | EGFP 位点 62 | 286          | CATGCCCCGAAGGCTAC<br>GTCCAGG | 150.          | AGAAGTCGTGCTGCTTCA<br>TGTGG  | 151.          | 10               | 在外  |
| 63                    | EGFP 位点 63 | 337          | CAACTACAAGACCCG<br>CGCCGAGG  | 152.          | TGAAGAAGATGGTGGCT<br>CCTGG   | 153.          | 10               | 在外  |
| 64                    | EGFP 位点 64 | 382          | GGTGAACCGCATCGA<br>GCTGAAGG  | 154.          | CCTCGAACTTCACCTCGG<br>CGCGG  | 155.          | 10               | 在外  |
| 65                    | EGFP 位点 65 | 275          | TTCAAGTCCGCCATGC<br>CCGAAGG  | 156.          | GCTTCATGTGTCGGGGT<br>AGCGG   | 157.          | 11               | 在外  |
| 66                    | EGFP 位点 66 | 349          | CCGCGCCGAGGTGAA<br>GTTCGAGG  | 158.          | GCCGTCGTCCTTGAAGAA<br>GATGG  | 159.          | 11               | 在外  |
| 67                    | EGFP 位点 67 | 382          | GGTGAACCGCATCGA<br>GCTGAAGG  | 160.          | CTCGAACTTCACCTCGGC<br>GCGGG  | 161.          | 11               | 在外  |
| 68                    | EGFP 位点 68 | 383          | GTGAACCGCATCGAG<br>CTGAAGGG  | 162.          | CCTCGAACTTCACCTCGG<br>CGCGG  | 163.          | 11               | 在外  |
| 69                    | EGFP 位点 69 | 520          | CAAGATCCGCCACAA<br>CATCGAGG  | 164.          | GATGCCGTTCTTCTGCTT<br>GTCGG  | 165.          | 11               | 在外  |
| 70                    | EGFP 位点 70 | 383          | GTGAACCGCATCGAG<br>CTGAAGGG  | 166.          | CTCGAACTTCACCTCGGC<br>GCGGG  | 167.          | 12               | 在外  |

[0228]

表 2

| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+)'位点                     | SEQ ID<br>NO: | 序列(-)'位点                     | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
|-----------------------|------------|--------------|------------------------------|---------------|------------------------------|---------------|------------------|-----|
| 71                    | EGFP 位点 71 | 415          | CAAGGAGGACGGCAA<br>CATCCTGG  | 168.          | TCAGCTCGATGCGGTTCA<br>CCAGG  | 169.          | 13               | 在外  |
| 72                    | EGFP 位点 72 | 286          | CATGCCCCGAAGGTAC<br>GTCCAGG  | 170.          | GTCGTGCTGCTTCATGTG<br>GTCGG  | 171.          | 14               | 在外  |
| 73                    | EGFP 位点 73 | 415          | CAAGGAGGACGGCAA<br>CATCCTGG  | 172.          | CAGCTCGATGCGGTTCA<br>CAGGG   | 173.          | 14               | 在外  |
| 74                    | EGFP 位点 74 | 416          | AAGGAGGACGGCAAC<br>ATCCTGGG  | 174.          | TCAGCTCGATGCGGTTCA<br>CCAGG  | 175.          | 14               | 在外  |
| 75                    | EGFP 位点 75 | 101          | AAGTTCAGCGTGTCCG<br>GCGAGGG  | 176.          | GCCGTCACAGCTCGACCAG<br>GATGG | 177.          | 15               | 在外  |
| 76                    | EGFP 位点 76 | 286          | CATGCCCCGAAGGTAC<br>GTCCAGG  | 178.          | TCGTGCTGCTTCATGTGG<br>TCGGG  | 179.          | 15               | 在外  |
| 77                    | EGFP 位点 77 | 416          | AAGGAGGACGGCAAC<br>ATCCTGGG  | 180.          | CAGCTCGATGCGGTTCA<br>CAGGG   | 181.          | 15               | 在外  |
| 78                    | EGFP 位点 78 | 417          | AGGAGGACGGCAACA<br>TCCTGGGG  | 182.          | TCAGCTCGATGCGGTTCA<br>CCAGG  | 183.          | 15               | 在外  |
| 79                    | EGFP 位点 79 | 524          | ATCCGCCACAAACATCG<br>AGGACGG | 184.          | GATGCCGTTCTTCTGTCTT<br>GTCGG | 185.          | 15               | 在外  |
| 80                    | EGFP 位点 80 | 106          | CAGCGTGTCCGGCGA<br>GGCGGAGG  | 186.          | CGTCGCCGTCCAGCTCGA<br>CCAGG  | 187.          | 16               | 在外  |
| 81                    | EGFP 位点 81 | 174          | CCGGCAAGCTGCCCCG<br>TGCCCTGG | 188.          | CAGGGTCAGCTTGCCGTA<br>GGTGG  | 189.          | 16               | 在外  |
| 82                    | EGFP 位点 82 | 286          | CATGCCCGAAGGCTAC<br>GTCCAGG  | 190.          | CGTGTGCTTCATGTGGT<br>CGGGG   | 191.          | 16               | 在外  |
| 83                    | EGFP 位点 83 | 417          | AGGAGGACGGCAACA<br>TCCTGGGG  | 192.          | CAGCTCGATGCGGTTCA<br>CAGGG   | 193.          | 16               | 在外  |
| 84                    | EGFP 位点 84 | 427          | CAACATCCTGGGGCAC<br>AAGCTGG  | 194.          | CGATGCCCTTCAGCTCGA<br>TGCGG  | 195.          | 16               | 在外  |

[0229]

| 表 2                   |            |              |                              |               |                             |               |                  |     |
|-----------------------|------------|--------------|------------------------------|---------------|-----------------------------|---------------|------------------|-----|
| FokI-dCas9<br>EGFP 对# | 名称         | 靶起始位<br>置(+) | 序列(+ )位点                     | SEQ ID<br>NO: | 序列(-)位点                     | SEQ ID<br>NO: | 边-至-边'间<br>隔区'距离 | PAM |
| 85                    | EGFP 位点 85 | 397          | GCTGAAGGGGCATCGA<br>CTTCAAGG | 196.          | GTCGCCCTCGAACTTCAC<br>CTCGG | 197.          | 20               | 在外  |

[0230] 令人惊讶地,当与60个gRNA对中的任一个在人U20S. EGFP细胞中共表达时,dCas9-FokI蛋白不显示可检测的EGFP破坏活性(图5E)。然而,利用相同的60个gRNA对筛选FokI-

dCas9蛋白在以PAM在外取向存在的并且间隔区长度为13至17bp和26bp(大致一圈超过13-17bp间隔区长度的DNA双螺旋)的半位点组成的靶位点上未显示EGFP破坏活性(图5B)。在另外25个具有在10至20bp的范围内的间隔区长度和具有以PAM在外取向存在的半位点的靶DNA位点上测试FokI-dCas9显示对具有13至18bp的间隔区长度的靶的高效裂解(图5C-D)。在这些实验中,测试每一个位点的17或18bp的间隔区长度,并且并非所有具有13bp间隔区长度的位点显示活性。通过T7EI分析和Sanger测序对一个亚组的被成功靶向的位点的分析进一步确认了插入缺失在期望的位置上的存在。因此,FokI-dCas9可被两个适当地放置的gRNA导向高效地裂解目标全长靶位点。为简便起见,两个FokI-dCas9融合物与两个gRNA的复合物在本文中被称为RNA引导的FokI核酸酶(RFN)。

[0231] 为了扩展利用EGFP报告基因的初始发现和确认RFN是否可用于进行内源人基因的常规基因组编辑,设计针对9个不同人基因中的12个不同靶位点的gRNA对(表2)。所测试的12个RFN中的11个以高效率(3至40%的范围)在人U2OS. EGFP细胞中在它们的期望的靶位点上诱导插入缺失,如通过T7EI测定判断的(表2)。利用这些相同的12个RFN对在HEK293细胞中获得类似结果(表2)。来自U2OS. EGFP细胞的被成功地靶向的等位基因的Sanger测序显示一系列插入缺失(主要是缺失)在预期的裂解位点的引入(图5F)。在两个不同的人细胞系中观察到的修饰的高成功率和高效率显示RFN用于修饰内源人基因的健壮性。

[0232] 实施例2d. RFN对于它们的裂解位点具有扩大的特异性

[0233] 为了测试RFN是否具有与二聚化相关的提高的识别特异性,检查这些核酸酶是否严格依赖于两个gRNA在对中的存在。在理想的二聚体系统中,单一gRNA不应当能够高效地指导FokI-dCas9诱导的插入缺失。为了进行初始测试,使用已显示在人U2OS. EGFP细胞中高效地将FokI-dCas9诱导的插入缺失导向它们的靶位点(EGFP位点47和81)的两对被导向EGFP中的两个靶位点的gRNA(图5C)。用被靶向VEGFA中的无关位点的gRNA替代这两对的每一对中的一个或另一个gRNA导致至不可检测的水平EGFP破坏活性的降低(图6A)和靶向突变的减少,如通过T7EI测定判断的(图6B)。类似地,使用高效地在人APC、MLH1和VEGFA基因(表2)中引入FokI-dCas9介导的插入缺失的对测试只使用每一两个gRNA的一个的效应,并且再次通过T7EI测定观察到可检测的RFN诱导的插入缺失的丢失(图6C)。这些结果表明,通过RFN进行的基因组编辑的高效引入需要两个与全长靶位点具有适当的互补性的gRNA。

[0234] 鉴于我们的RFN的活性依赖于两个gRNA的表达,因此希望它们对对中的单一gRNA之一的已知脱靶位点的诱变效应应当是可微不足道的。进行这些直接比较需要知道通过单一gRNA引导的单体Cas9核酸酶的脱靶位点,所述单一gRNA本身也可用作靶向二聚体RFN所需的两个gRNA之一。虽然已在文献中确定了极少的单体Cas9核酸酶脱靶位点,但针对我们用于靶向人VEGFA基因中的二聚体RFN位点的gRNA之一先前已鉴定了5个脱靶位点(实施例1)。深度测序用于确定这5个脱靶位点是否在其中已表达靶向VEGFA的RFN的细胞(这些细胞是我们用于图6C中显示的T7EI测定的相同细胞)中显示突变的证据。所有5个脱靶位点上的插入缺失突变的频率与本底无区别(图6D和表3)。这些结果表明,RFN的使用可基本上消除原来由Cas9核酸酶和单一gRNA诱导的脱靶效应,并且与我们的观察(与FokI-dCas9一起表达的单一gRNA不高效地诱导插入缺失)一致。虽然目前不可能在另外的位点上进行这些直接比较-但此类实验将必须等待比也可靶向二聚体RFN的半位点的更多的单一gRNA位点的脱靶位点的鉴定,得出二聚体RFN具有相对于标准单体Cas9核酸酶的提高的特异性。

[0235]

表 3

| 基因名称           | 左靶序列                        | SEQ ID No. | 右靶序列                         | SEQ ID No. | U2OS或293细胞中的RFN靶位点的内源序列(间隔区序列以小写字母表示)                              | SEQ ID No. |
|----------------|-----------------------------|------------|------------------------------|------------|--------------------------------------------------------------------|------------|
| APC1           | CCAGAACTACGAGCGCCGC<br>CCGG | 198.       | TGGCAGTGAGTGAGGCT<br>GCAGG   | 199.       | CCGGGCGCGCTCGTACTTCTGGccactg9g9cgag<br>cgtcTGGCAGGTGAGTGAGGCTGCAGG | 200.       |
| BRCA1          | GAATACCATCTGTACGCT<br>TCGG  | 201.       | GGCGAACCTGAGAGGCG<br>TAAGG   | 202.       | CCGAAGCTACAGATGGGTATTCTtttgac99999g<br>tagggGCGGAACCTGAGAGGCGTAAGG | 203.       |
| DDB2           | AATATTCAAGCAGCAGGCA<br>CAGG | 204.       | CTCGCGCAGGAGGCTGCA<br>GCGGG  | 205.       | CCTGTGCTGCTGCTTGAATATTtccgccttttag<br>9gt9CTCGCGCAGGAGGCTGCAGCGGG  | 206.       |
| EMX1           | CCCAAGCCTGGCCAGGGA<br>GTGG  | 207.       | GCCCCACAGGCTTGAAG<br>CCCCG   | 208.       | CCACTCCCTGGCCAGGCTTGGGgag9g9cct9gagt<br>cat9GCCCCACAGGCTTGAAGCCCCG | 209.       |
| FANCF -<br>位点1 | CCCTACTTCGCTTTCACC<br>TTGG  | 210.       | GGATCCCTTCTGCAGCA<br>CCTGG   | 211.       | CCAAGGTGAAGCGGAAGTAGGccttc9gcaccc<br>tcatGGAATCCCTTCTGCAGCACCTGG   | 212.       |
| FANCF -<br>位点2 | CGCTCCAGAGCCGTGCGAA<br>TGGG | 213.       | TGGAGCAAGAGGCGGC<br>TTTGG    | 214.       | CCCATTGCGACGGCTCTGGAGCGcg9gctgcacaa<br>ccagTGGAGGCAAGAGGCGGCTTTGG  | 215.       |
| FES            | CGAGGAGACTGGGACTGT<br>AGGG  | 216.       | CCAGCTGCTGCCCTTGCCT<br>CCAGG | 217.       | CCCTACAGTCCCCAGCTCTCGtcccatgcctcc<br>gtctCCAGCTGCTGCTTGCCTCCAGG    | 218.       |
| GLI1           | CATAGCTACTGATTGGTGG<br>TGGG | 219.       | COGGCCCTCCCCAGTCA<br>GGGGG   | 220.       | CCCACCAATCAGTAGCTATGcgag9cctgct<br>gtctCCGGCCCCCTCCCAAGTCAGGGG     | 221.       |
| MLH1           | GGAAACGTCTAGATGCTCA<br>ACGG | 222.       | CAAAATGTCGTTCTGTGGC<br>AGTGG | 223.       | CCGTTGAGCATCTAGACGTTTCcttg9ctctctg<br>gc9CCAAATGTCGTTCTGTGGCAGGGG  | 224.       |
| RARA1          | CTGTTGCTGGCCATGCCAA<br>GCGG | 225.       | CCTGGGGCGGCACCTC<br>AATGG    | 226.       | CCGCTTGGCATGGCCAGCAACAGcagctcctgccc<br>gacaCCTGGGGCGGCACCTCAATGG   | 227.       |
| RUNX           | TTCGGAGCGAAACCAAGA<br>CAGG  | 228.       | GAGTCCCCCGCCTTCAGA<br>AGAGG  | 229.       | CCTGCTTTGGTTTTCGCTCCGAAGgtataa9aaat<br>cattGAGTCCCCCGCTTCAGAAGAGG  | 230.       |
| SS18           | GGCCCGTCCGACTCCGGGC<br>CCGG | 231.       | TGCTGGGAATCAGCAGTG<br>TTTGG  | 232.       | CCGGGCCCGGAGTCACCGGGCCgag9cgag9cg<br>ggcctGCTGGGAATCAGCAGTGTGG     | 233.       |
| VEGFA -<br>位点1 | GGGTGGGGGAGTTTGCTC<br>CTGG  | 234.       | TCCCTCTTAGCCAGAGC<br>CGGGG   | 235.       | CCAGGAGCAAACTCCCCACCCctttccaaagc<br>ccatTCCCTCTTAGCCAGAGCGGGG      | 236.       |
| VEGFA -<br>位点2 | GCCCGCGCCGGGAGGAG<br>GTGG   | 237.       | GGCGACCGCGGCGAGGG<br>GCCGG   | 238.       | CCACTCTCTCCCCGGCGGCGg9acagt9gacg<br>c99CGCGAGCCCGGGCAGGGCCCG       | 239.       |
| VEGFA -<br>位点3 | CCGTCTGCACACCCCGGCT<br>CTGG | 240.       | CTCGGCCACACAGGGAA<br>GCTGG   | 241.       | CCAGAGCCGGGTGTGACAGCGcagtcactag9g<br>ggcgCTCGGCCACACAGGGAGCTGG     | 242.       |

[0236]

| 表 3          |                                          |            |                               |            |              |        |            |                     |
|--------------|------------------------------------------|------------|-------------------------------|------------|--------------|--------|------------|---------------------|
| 基因名称         | 用于 T7E1 测定的引物 1                          | SEQ ID NO: | 用于 T7E1 测定的引物 2               | SEQ ID NO: | 加 DMSO (是/否) | 热循环仪方案 | 扩增子大小 (bp) | T7E1 中的裂解产物的大小 (bp) |
| APC1         | GGCTGTGGGAAGCCAGCAA<br>C                 | 243.       | AAGCCAGGGGCCA<br>ACTGGAG      | 244.       | 无            | 降落     | 634        | 447/187             |
| BRCA1        | GGCGGGGAATTACAGATAA<br>ATTAAAA           | 245.       | AGGTCCCATCCTC<br>TCATACATACCA | 246.       | 无            | 降落     | 751        | 454/297             |
| DDB2         | ACCGCCCTTGGCACCAC                        | 247.       | CGGAGCTCATCTG<br>CTTCTGT      | 248.       | 无            | 降落     | 627        | 456/171             |
| EMX1         | GGAGCAGCTGGTCAGAGGG<br>G                 | 249.       | GGGAAGGGGGAC<br>ACTGGGA       | 250.       | 有            | 两步骤    | 729        | 480/249             |
| FANCF - 位点 1 | GCCCTACATCTGCTCTCCCT<br>CCA              | 251.       | GGGCCGGGAAAG<br>AGTTGCTG      | 252.       | 无            | 降落     | 634        | 361/273             |
| FANCF - 位点 2 | GCCCTACATCTGCTCTCCCT<br>CCA              | 253.       | GGGCCGGGAAAG<br>AGTTGCTG      | 254.       | 无            | 降落     | 634        | 466/168             |
| FES          | GGGAGGGAGGCTCCAGGT<br>T                  | 255.       | GGCACAATGGCTC<br>CCAAGCA      | 256.       | 无            | 降落     | 633        | 395/238             |
| GLI1         | CCTTACCCCTCCCCTCACTC<br>A                | 257.       | AGAAGGCGGGC<br>CAGACAGT       | 258.       | 无            | 降落     | 869        | 590/279             |
| MLH1         | ATATCCTTCTAGGTAGCGGG<br>CAGTAGCC         | 259.       | TCTCGGGGGAGAG<br>CGGTAAA      | 260.       | 无            | 降落     | 610        | 332/278             |
| RARA1        | CCCAGGAAAAGTGCCAGCT<br>CA                | 261.       | TGATGTCACCCC<br>AACTGGA       | 262.       | 无            | 降落     | 632        | 335/297             |
| RUNX         | AAGCGGCGCTGGCTTTT<br>GGGATGCAGGGACGGTCAA | 263.       | CCAGCACAACTTA<br>CTCGCACTTGA  | 264.       | 无            | 降落     | 626        | 389/237             |
| SS18         | GAG                                      | 265.       | GCCGCCCATCCC<br>TAGAGAAA      | 266.       | 无            | 降落     | 629        | 455/174             |
| VEGFA - 位点 1 | TCCAGATGGCACATTGTGAG                     | 267.       | AGGAGCAGGAA<br>AGTGAGGT       | 268.       | 无            | 降落     | 531        | 338/193             |
| VEGFA - 位点 2 | AGAGAAGTCGAGGAAGAGA<br>GAG               | 269.       | CAGCAGAAAAGTTC<br>ATGGTTTCG   | 270.       | 有            | 降落     | 756        | 482/274             |
| VEGFA - 位点 3 | TCCAGATGGCACATTGTGAG                     | 271.       | AGGAGCAGGAA<br>AGTGAGGT       | 272.       | 无            | 降落     | 531        | 288/243             |

[0237]

表 3

| 基因名称         | 用于深度测序的引物 1           | SEQ ID NO: | 用于深度测序的引物 2           | SEQ ID NO: |
|--------------|-----------------------|------------|-----------------------|------------|
| APC1         |                       |            |                       |            |
| BRCA1        |                       |            |                       |            |
| DDB2         | CGATGGCTCCCAAGAAACGC  | 273.       | GCAGGTAGAATGCACAGCCG  | 274.       |
| EMX1         |                       |            |                       |            |
| FANCF - 位点 1 | GCCCAGAGTCAAGGAACACG  | 275.       | AGGTAGTGCTTGAGACCGCC  | 276.       |
| FANCF - 位点 2 | CATCCATCGCGGCTTTGGTC  | 277.       | CCGGGAAAGAGTTGCTGCAC  | 278.       |
| FES          | CTCCCCGCTCTGCAGTCCATC | 279.       | CCTGCAGGACATGTGTGA    | 280.       |
| GLI1         |                       |            |                       |            |
| MLH1         |                       |            |                       |            |
| RAR1         |                       |            |                       |            |
| RUNX         | TAGGGCTAGAGGGGTGAGGC  | 281.       | CCGAGGTGAACAACAGCTGCC | 282.       |
| SS18         |                       |            |                       |            |
| VEGFA - 位点 1 | ATGAGGGCTCCAGATGGCAC  | 283.       | TTCACCCAGCTTCCCTGTGG  | 284.       |
| VEGFA - 位点 2 |                       |            |                       |            |
| VEGFA - 位点 3 |                       |            |                       |            |

[0238] 实施例2e. 单体Cas9切口酶诱导比单一gRNA/FokI-dCas9复合物高的诱变率

[0239] 如上文中所指出的, 成对Cas9切口酶法的重要弱点是单一单体切口酶可以高频率在某些靶位点上引入插入缺失突变(参见实施例1和Ran等, Cell 154, 1380-1389 (2013); Mali等, Nat Biotechnol 31, 833-838 (2013); Cho等, Genome Res (2013); 和Mali等, Science 339, 823-826 (2013)). 成对Cas9切口酶系统中该二聚化依赖性的不存在是脱靶效应的潜在来源, 因为两个单体切口酶可各自在基因组的其它地方产生不想要的插入缺失突变。有人假设, 因为RFN使用二聚化依赖性FokI核酸酶引入改变, 因此这些融合物在仅一个gRNA存在的情况下通常应当显示相较于对于单体Cas9切口酶所观察的不期望的插入缺失活性较低的不期望的插入缺失活性。

[0240] 为了测试该假设, 在单一gRNA存在的情况下在6个二聚体人基因靶位点(总共12个半位点; 表4)上比较FokI-dCas9和Cas9切口酶的活性。因为被对中的仅一个和/或另一个gRNA引导的单体Cas9切口酶可在这些靶上诱导插入缺失突变, 因此选择这些特定位点。通过使用深度测序, 在两个gRNA或仅一个或另一个gRNA存在的情况下评估FokI-dCas9或Cas9切口酶的基因组编辑活性。FokI-dCas9和Cas9切口酶在两个gRNA都存在的情况下以高效率在所有6个靶位点诱导插入缺失(表5)。如所假设的, 由12个单一gRNA引导的单体Cas9切口酶以在0.0048%至3.04%的范围内的频率诱导插入缺失(图7A和表5)。相反地, 由相同的12

个单一gRNA引起的FokI-dCas9以在0.0045%至0.473%的范围内的较低频率诱导插入缺失(图7A和表5)。通过直接比较这些数据,对于12个单一gRNA中的10个,FokI-dCas9以比Cas9切口酶低的频率诱导插入缺失(图7A和表5)。另外,当将成对的gRNA比率与单一gRNA比率相比较时,FokI-dCas9在12个半位点的11个半位点上显示比Cas9切口酶更大的插入缺失频率的倍数减少(图7B)。

[0241] 表4

| 染色体 | 位置                | 位点                | FokI-dCas9 插入缺失 | FokI-dCas9 总数 | FokI-dCas9 插入缺失频率(%) | tdTomato 对照插入缺失 | tdTomato 总数 | tdTomato 插入缺失频率(%) |
|-----|-------------------|-------------------|-----------------|---------------|----------------------|-----------------|-------------|--------------------|
| 6   | 4373<br>7290      | VEGF<br>A 位点<br>1 | 35000           | 150158        | 23.308<br>78         | 10              | 258108      | 0.00387            |
| 15  | 6563<br>7531      | OT1-3             | 1               | 169681        | 0.0005<br>8          | 1               | 139847      | 0.00071            |
| 12  | 1316<br>9018<br>2 | OT1-4             | 4               | 190111        | 0.0021<br>0          | 5               | 139762      | 0.00357            |
| 12  | 1988<br>060       | OT1-6             | 3               | 258976        | 0.0011<br>5          | 2               | 178162      | 0.00112            |
| 1   | 9934<br>7645      | OT1-1<br>1        | 4               | 235853        | 0.0016<br>9          | 4               | 186287      | 0.00214            |
| 17  | 3979<br>6322      | OT1-3<br>0        | 1               | 261605        | 0.0003<br>8          | 1               | 286850      | 0.00034            |

[0243] 表5. 利用单一和成对gRNA的FokI-dCas9、Cas9n和tdTomato对照在6个位点上的深度测序(与如图7中所示的相同的数据)

[0244]

| 核酸酶类型或对照   | 位点            | 引导RNA | 染色体 | 位置       | 插入缺失  | 总数     | 百分比     |
|------------|---------------|-------|-----|----------|-------|--------|---------|
| FokI-dCas9 | VEGFA<br>位点 1 | 两个    | 6   | 43737290 | 35000 | 150158 | 23.3088 |
| FokI-dCas9 | VEGFA<br>位点 1 | 左     | 6   | 43737290 | 5     | 95476  | 0.0052  |
| FokI-dCas9 | VEGFA<br>位点 1 | 右     | 6   | 43737290 | 9     | 91962  | 0.0098  |
| FokI-dCas9 | DDB2          | 两个    | 11  | 47236820 | 11303 | 50062  | 22.5780 |
| FokI-dCas9 | DDB2          | 左     | 11  | 47236820 | 311   | 85726  | 0.3628  |
| FokI-dCas9 | DDB2          | 右     | 11  | 47236820 | 153   | 95050  | 0.1610  |
| FokI-dCas9 | FANCF<br>位点 1 | 两个    | 11  | 22647331 | 65846 | 195311 | 33.7134 |
| FokI-dCas9 | FANCF<br>位点 1 | 左     | 11  | 22647331 | 19    | 27487  | 0.0691  |
| FokI-dCas9 | FANCF<br>位点 1 | 右     | 11  | 22647331 | 845   | 225154 | 0.3753  |
| FokI-dCas9 | FANCF<br>位点 2 | 两个    | 11  | 22647138 | 27743 | 120314 | 23.0588 |
| FokI-dCas9 | FANCF<br>位点 2 | 左     | 11  | 22647138 | 989   | 205832 | 0.4805  |
| FokI-dCas9 | FANCF<br>位点 2 | 右     | 11  | 22647138 | 142   | 165130 | 0.0860  |
| FokI-dCas9 | FES           | 两个    | 15  | 91428181 | 14260 | 125912 | 11.3254 |
| FokI-dCas9 | FES           | 左     | 15  | 91428181 | 4     | 143877 | 0.0028  |
| FokI-dCas9 | FES           | 右     | 15  | 91428181 | 7     | 145495 | 0.0048  |

[0245]

[0246] 表5. 利用单一和成对gRNA的FokI-dCas9、Cas9n和tdTomato对照在6个位点上的深度测序(与如图7中所示的相同的数据)

| 核酸酶类型或对照              | 位点         | 引导RNA | 染色体 | 位置       | 插入缺失  | 总数     | 百分比     |
|-----------------------|------------|-------|-----|----------|-------|--------|---------|
| <b>FokI-dCas9</b>     | RUNX1      | 两个    | 21  | 36421217 | 61057 | 136164 | 44.8408 |
| <b>FokI-dCas9</b>     | RUNX1      | 左     | 21  | 36421217 | 222   | 162636 | 0.1365  |
| <b>FokI-dCas9</b>     | RUNX1      | 右     | 21  | 36421217 | 109   | 169122 | 0.0645  |
| <b>Cas9n</b>          | VEGFA 位点 1 | 两个    | 6   | 43737290 | 14294 | 99036  | 14.4331 |
| <b>Cas9n</b>          | VEGFA 位点 1 | 左     | 6   | 43737290 | 573   | 82316  | 0.6961  |
| <b>Cas9n</b>          | VEGFA 位点 1 | 右     | 6   | 43737290 | 315   | 101957 | 0.3090  |
| <b>Cas9n</b>          | DDB2       | 两个    | 11  | 47236820 | 6673  | 31168  | 21.4098 |
| <b>Cas9n</b>          | DDB2       | 左     | 11  | 47236820 | 1680  | 56019  | 2.9990  |
| <b>Cas9n</b>          | DDB2       | 右     | 11  | 47236820 | 172   | 42424  | 0.4054  |
| <b>Cas9n</b>          | FANCF 位点 1 | 两个    | 11  | 22647331 | 66827 | 193111 | 34.6055 |
| <b>Cas9n</b>          | FANCF 位点 1 | 左     | 11  | 22647331 | 1565  | 109029 | 1.4354  |
| <b>Cas9n</b>          | FANCF 位点 1 | 右     | 11  | 22647331 | 2457  | 109289 | 2.2482  |
| <b>Cas9n</b>          | FANCF 位点 2 | 两个    | 11  | 22647138 | 17007 | 111468 | 15.2573 |
| <b>Cas9n</b>          | FANCF 位点 2 | 左     | 11  | 22647138 | 120   | 100591 | 0.1193  |
| <b>Cas9n</b>          | FANCF 位点 2 | 右     | 11  | 22647138 | 1063  | 93162  | 1.1410  |
| <b>Cas9n</b>          | FES        | 两个    | 15  | 91428181 | 16529 | 126597 | 13.0564 |
| <b>Cas9n</b>          | FES        | 左     | 15  | 91428181 | 6     | 125196 | 0.0048  |
| <b>Cas9n</b>          | FES        | 右     | 15  | 91428181 | 23    | 46102  | 0.0499  |
| <b>Cas9n</b>          | RUNX1      | 两个    | 21  | 36421217 | 80029 | 216800 | 36.9137 |
| <b>Cas9n</b>          | RUNX1      | 左     | 21  | 36421217 | 1106  | 108670 | 1.0178  |
| <b>Cas9n</b>          | RUNX1      | 右     | 21  | 36421217 | 2169  | 121413 | 1.7865  |
| <b>tdTomato 对照(-)</b> | VEGFA 位点 1 | 无     | 6   | 43737290 | 29    | 313517 | 0.0092  |
| <b>tdTomato 对照(-)</b> | FANCF 位点 1 | 无     | 11  | 22647331 | 18    | 578378 | 0.0031  |
| <b>tdTomato 对照(-)</b> | FANCF 位点 2 | 无     | 11  | 22647138 | 81    | 393821 | 0.0206  |
| <b>tdTomato 对照(-)</b> | FES        | 无     | 15  | 91428181 | 21    | 410620 | 0.0051  |
| <b>tdTomato 对照(-)</b> | DDB2       | 无     | 11  | 47236820 | 14    | 165314 | 0.0085  |
| <b>tdTomato 对照(-)</b> | RUNX1      | 无     | 21  | 36421217 | 13    | 511977 | 0.0025  |

[0248] 深度测序实验还揭示了某些单体Cas9切口酶的先前未描述的和预料之外的副作用：它们的靶位点内的特定位置上的点突变的引入。与针对VEGFA靶的“右”半位点的单一gRNA共表达的Cas9切口酶以10.5%的频率在识别位点的位置15上诱导碱基取代(图8A)。对于Cas9切口酶和被导向FANCF靶位点1的“右”半位点(位置16上16.3%的突变频率)(图8B)

或RUNX1靶位点的“右”半位点(位置17上2%的突变频率)(图8C)的单一gRNA观察到类似的结果。在其中在细胞中无Cas9切口酶或gRNA表达的对照样品中未观察到高于本底水平的这些位置上的点突变(图8A-8C)。有趣地,对于在其上观察到该高突变的3个位点中的2个位点,观察到的大多数取代是非靶DNA链上的C至G的颠换。在其上观察到这些点突变的位置落在靶位点的链分隔区中,已观察到所述链分隔区在体外对dCas9/gRNA/靶DNA复合物中的P1核酸酶易感。重要的是,这些点突变以低得多的频率(低至1/5至1/100)在表达FokI-dCas9蛋白和相同gRNA的细胞中发生(图8A-C)。总体上,得出通过单一gRNA引导的FokI-dCas9核酸酶通常以比匹配的单个Cas9切口酶低的频率诱导诱变的插入缺失和点突变。

[0249] 实施例2f.二聚体RFN具有高度的特异性

[0250] 预期由两个gRNA引导的二聚体RFN在人细胞中不诱导可察觉的脱靶突变。可预期由一对gRNA引导来裂解由两个半位点组成的全长序列的RFN指定靶位点中多达44bp的DNA。该长度的序列将意外地几乎总是独特的(除在其中靶可能存在于重复基因组序列中的某些情况下外)。另外,基因组中与该全长位点最密切匹配的位点在大多数情况下应当具有大量错配,预期这反过来可使RFN二聚体的裂解活性降至最低或消除裂解活性。事实上,鉴定了针对15个被RFN成功靶向的全长序列的具有0至16个错配(并允许存在长度为14至17bp的间隔区)的人基因组中的所有位点。该分析显示所有15个全长序列是独特的以及基因组中的最密切匹配的位点存在7至12个错配(表6)。含有该数目的错配的位点不应当被RFN高效诱变,并且有趣地在将来的研究中确认该假设。总之,二聚体RFN应当在人细胞中具有高度特异性但特异性的最终表征将等待可全面地在整个基因组中确定RFN特异性的无偏方法的开发。

[0251] 表6.人基因组中具有确定数目的错配的候选FokI-dCas9脱靶位点的频率

| 基因      | 0 | 7 | 8 | 9 | 10 | 11 | 12 | 13  | 14   | 15   | 16    |
|---------|---|---|---|---|----|----|----|-----|------|------|-------|
| APC     | 1 |   |   |   |    | 1  | 2  | 16  | 74   | 414  | 2254  |
| BRCA1   | 1 | 1 |   |   |    |    |    | 5   | 20   | 164  | 983   |
| DDB2    | 1 |   |   |   |    |    | 2  | 7   | 58   | 267  | 1335  |
| EMX1    | 1 |   |   | 1 |    | 2  | 8  | 40  | 175  | 828  | 3494  |
| FANCF   | 1 |   |   |   |    |    | 2  | 4   | 44   | 298  | 1639  |
| FANCF   | 1 |   |   |   |    |    | 2  | 12  | 79   | 358  | 1718  |
| FES     | 1 |   |   |   |    | 3  | 8  | 32  | 191  | 939  | 4505  |
| GLI1    | 1 |   |   |   |    | 2  | 1  | 7   | 69   | 343  | 1711  |
| MLH1    | 1 |   |   |   |    |    | 2  | 5   | 22   | 96   | 643   |
| RARA    | 1 |   |   |   | 1  | 2  | 8  | 39  | 187  | 698  | 2849  |
| RUNX1   | 1 |   |   |   |    |    |    | 3   | 25   | 145  | 800   |
| SS18    | 1 |   |   |   |    | 1  | 2  | 6   | 39   | 280  | 1207  |
| VEGFA-1 | 1 |   |   |   | 1  | 2  | 3  | 22  | 103  | 543  | 2676  |
| VEGFA-2 | 1 |   |   |   | 4  | 9  | 99 | 447 | 1675 | 5608 | 18599 |
| VEGFA-3 | 1 |   |   |   |    |    | 3  | 20  | 120  | 623  | 2783  |

[0252] 参考文献

[0254] Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res 23, 1163-1171. (2013).

- [0255] Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 31, 230–232 (2013).
- [0256] Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823 (2013).
- [0257] Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* (2013).
- [0258] Dicarlo, J.E. et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* (2013).
- [0259] Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A., and Musunuru, K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* 12, 393–394. (2013).
- [0260] Fisher, S., Barry, A., Abreu, J., Minie, B., Nolan, J., Delorey, T.M., Young, G., Fennell, T.J., Allen, A., Ambrogio, L., et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 12, R1. (2011).
- [0261] Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., and Calarco, J.A. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods* 10, 741–743. (2013).
- [0262] Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31, 822–826. (2013).
- [0263] Gabriel, R. et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 29, 816–823 (2011).
- [0264] Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442–451.
- [0265] Gratz, S.J. et al. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* (2013).
- [0266] Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* 29, 731–734 (2011).
- [0267] Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167–170 (2010).
- [0268] Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31, 827–832. (2013).
- [0269] Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31, 227–229 (2013).

- [0270] Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Kaini, P., Sander, J.D., Joung, J.K., Peterson, R.T., and Yeh, J.R. Heritable and Precise Zebrafish Genome Editing Using a CRISPR-Cas System. *PLoS One* 8, e68708. (2013a).
- [0271] Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31, 233-239 (2013).
- [0272] Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012).
- [0273] Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* 2, e00471 (2013).
- [0274] Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X., et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol* 31, 681-683. (2013a).
- [0275] Li, W., Teng, F., Li, T., and Zhou, Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol* 31, 684-686. (2013b).
- [0276] Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10, 977-979. (2013).
- [0277] Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosbumer, M., Kosuri, S., Yang, L., and Church, G.M. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 31, 833-838. (2013a).
- [0278] Mali, P., Esvelt, K.M., and Church, G.M. Cas9 as a versatile tool for engineering biology. *Nat Methods* 10, 957-963. (2013b).
- [0279] Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823-826 (2013c).
- [0280] Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* 31, 839-843. (2013).
- [0281] Pattanayak, V., Ramirez, C.L., Joung, J.K. & Liu, D.R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods* 8, 765-770 (2011).
- [0282] Perez, E.E. et al. Establishment of HIV-1 resistance in CD4+T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* 26, 808-816 (2008).
- [0283] Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10, 973-976. (2013).

- [0284] Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183. (2013).
- [0285] Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. (2013).
- [0286] Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotech* 30, 460–465 (2012).
- [0287] Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res* 38, W462–468. (2010).
- [0288] Sander, J.D., Ramirez, C.L., Linder, S.J., Pattanayak, V., Shores, N., Ku, M., Foden, J.A., Reyon, D., Bernstein, B.E., Liu, D.R., et al. In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* (2013).
- [0289] Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., and Dobbs, D. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Res* 35, W599–605. (2007).
- [0290] Shen, B. et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res* (2013).
- [0291] Sugimoto, N. et al. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* 34, 11211–11216 (1995).
- [0292] Terns, M.P. & Terns, R.M. CRISPR-based adaptive immune systems. *Curr Opin Microbiol* 14, 321–327 (2011).
- [0293] Wang, H. et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 153, 910–918 (2013).
- [0294] Wiedenheft, B., Sternberg, S.H. & Doudna, J.A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338 (2012).
- [0295] Yang, L., Guell, M., Byrne, S., Yang, J.L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A.W., Rios, X., et al. (2013). Optimization of scarless human stem cell genome editing. *Nucleic Acids Res* 41, 9049–9061.
- [0296] 其它实施方案
- [0297] 应理解, 虽然已结合其详细描述说明了本发明, 但前述说明指在举例说明而不是限制本发明的范围, 本发明的范围由所述权利要求的范围来界定。其它方面、有利方面和修改在下列权利要求的范围内。

## 序列表

|        |                                                     |    |
|--------|-----------------------------------------------------|----|
|        | <110> 通用医疗公司                                        |    |
|        | <120> 使用 RNA 引导的 FokI 核酸酶 (RFN) 提高 RNA 引导的基因组编辑的特异性 |    |
|        | <130> 00786-0895W01                                 |    |
|        | <140> PCT/US2014/028630                             |    |
|        | <141> 2014-03-14                                    |    |
|        | <150> 61/921, 007                                   |    |
|        | <151> 2013-12-26                                    |    |
|        | <150> 61/838, 148                                   |    |
|        | <151> 2013-06-21                                    |    |
|        | <150> 61/838, 178                                   |    |
|        | <151> 2013-06-21                                    |    |
| [0001] | <150> 61/799, 647                                   |    |
|        | <151> 2013-03-15                                    |    |
|        | <160> 285                                           |    |
|        | <170> PatentIn 3.5 版                                |    |
|        | <210> 1                                             |    |
|        | <211> 23                                            |    |
|        | <212> DNA                                           |    |
|        | <213> 人工序列                                          |    |
|        | <220>                                               |    |
|        | <223> 人工序列的说明：合成的<br>结合位点寡核苷酸                       |    |
|        | <400> 1                                             |    |
|        | gggcacgggc agcttgccgg tgg                           | 23 |
|        | <210> 2                                             |    |
|        | <211> 23                                            |    |
|        | <212> DNA                                           |    |

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
结合位点寡核苷酸

<400> 2

gatgccgttc ttctgcttgt cgg

23

<210> 3

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
结合位点寡核苷酸

<400> 3

ggtggtgcag atgaacttca ggg

23

[0002]

<210> 4

<211> 583

<212> PRT

<213> 海床游动微菌

<400> 4

Met Phe Leu Ser Met Val Ser Lys Ile Arg Thr Phe Gly Trp Val Gln  
1 5 10 15

Asn Pro Gly Lys Phe Glu Asn Leu Lys Arg Val Val Gln Val Phe Asp  
20 25 30

Arg Asn Ser Lys Val His Asn Glu Val Lys Asn Ile Lys Ile Pro Thr  
35 40 45

Leu Val Lys Glu Ser Lys Ile Gln Lys Glu Leu Val Ala Ile Met Asn  
50 55 60

|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|        | Gln | His | Asp | Leu | Ile | Tyr | Thr | Tyr | Lys | Glu | Leu | Val | Gly | Thr | Gly | Thr |
|        | 65  |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |     |
|        | Ser | Ile | Arg | Ser | Glu | Ala | Pro | Cys | Asp | Ala | Ile | Ile | Gln | Ala | Thr | Ile |
|        |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |     |
|        | Ala | Asp | Gln | Gly | Asn | Lys | Lys | Gly | Tyr | Ile | Asp | Asn | Trp | Ser | Ser | Asp |
|        |     |     | 100 |     |     |     |     | 105 |     |     |     |     |     | 110 |     |     |
|        | Gly | Phe | Leu | Arg | Trp | Ala | His | Ala | Leu | Gly | Phe | Ile | Glu | Tyr | Ile | Asn |
|        |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |     |
|        | Lys | Ser | Asp | Ser | Phe | Val | Ile | Thr | Asp | Val | Gly | Leu | Ala | Tyr | Ser | Lys |
|        | 130 |     |     |     |     |     | 135 |     |     |     | 140 |     |     |     |     |     |
| [0003] | Ser | Ala | Asp | Gly | Ser | Ala | Ile | Glu | Lys | Glu | Ile | Leu | Ile | Glu | Ala | Ile |
|        | 145 |     |     |     | 150 |     |     |     |     | 155 |     |     |     | 160 |     |     |
|        | Ser | Ser | Tyr | Pro | Pro | Ala | Ile | Arg | Ile | Leu | Thr | Leu | Leu | Glu | Asp | Gly |
|        |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |     |
|        | Gln | His | Leu | Thr | Lys | Phe | Asp | Leu | Gly | Lys | Asn | Leu | Gly | Phe | Ser | Gly |
|        |     |     | 180 |     |     |     |     |     | 185 |     |     |     |     | 190 |     |     |
|        | Glu | Ser | Gly | Phe | Thr | Ser | Leu | Pro | Glu | Gly | Ile | Leu | Leu | Asp | Thr | Leu |
|        |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     |
|        | Ala | Asn | Ala | Met | Pro | Lys | Asp | Lys | Gly | Glu | Ile | Arg | Asn | Asn | Trp | Glu |
|        | 210 |     |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
|        | Gly | Ser | Ser | Asp | Lys | Tyr | Ala | Arg | Met | Ile | Gly | Gly | Trp | Leu | Asp | Lys |
|        | 225 |     |     |     | 230 |     |     |     |     |     | 235 |     |     | 240 |     |     |

Leu Gly Leu Val Lys Gln Gly Lys Lys Glu Phe Ile Ile Pro Thr Leu  
245 250 255

Gly Lys Pro Asp Asn Lys Glu Phe Ile Ser His Ala Phe Lys Ile Thr  
260 265 270

Gly Glu Gly Leu Lys Val Leu Arg Arg Ala Lys Gly Ser Thr Lys Phe  
275 280 285

Thr Arg Val Pro Lys Arg Val Tyr Trp Glu Met Leu Ala Thr Asn Leu  
290 295 300

Thr Asp Lys Glu Tyr Val Arg Thr Arg Arg Ala Leu Ile Leu Glu Ile  
305 310 315 320

[0004] Leu Ile Lys Ala Gly Ser Leu Lys Ile Glu Gln Ile Gln Asp Asn Leu  
325 330 335

Lys Lys Leu Gly Phe Asp Glu Val Ile Glu Thr Ile Glu Asn Asp Ile  
340 345 350

Lys Gly Leu Ile Asn Thr Gly Ile Phe Ile Glu Ile Lys Gly Arg Phe  
355 360 365

Tyr Gln Leu Lys Asp His Ile Leu Gln Phe Val Ile Pro Asn Arg Gly  
370 375 380

Val Thr Lys Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu  
385 390 395 400

Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile  
405 410 415

|        |                                                                 |             |
|--------|-----------------------------------------------------------------|-------------|
|        | Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val |             |
|        | 420                                                             | 425 430     |
|        | Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly |             |
|        | 435                                                             | 440 445     |
|        | Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile |             |
|        | 450                                                             | 455 460     |
|        | Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn |             |
|        | 465                                                             | 470 475 480 |
|        | Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn |             |
|        |                                                                 | 485 490 495 |
| [0005] | Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr |             |
|        | 500                                                             | 505 510     |
|        | Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe |             |
|        | 515                                                             | 520 525     |
|        | Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn |             |
|        | 530                                                             | 535 540     |
|        | Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu |             |
|        | 545                                                             | 550 555 560 |
|        | Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe |             |
|        | 565                                                             | 570 575     |
|        | Asn Asn Gly Glu Ile Asn Phe                                     |             |
|        | 580                                                             |             |

&lt;210&gt; 5

&lt;211&gt; 1368

&lt;212&gt; PRT

&lt;213&gt; 化脓性链球菌

&lt;400&gt; 5

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val  
 1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe  
 20 25 30

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

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu  
 50 55 60

[0006]

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys  
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser  
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys  
 100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr  
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp  
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 145                                                             | 150 | 155 | 160 |
| Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro |     |     |     |
|                                                                 | 165 | 170 | 175 |
| Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr |     |     |     |
|                                                                 | 180 | 185 | 190 |
| Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala |     |     |     |
|                                                                 | 195 | 200 | 205 |
| Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn |     |     |     |
|                                                                 | 210 | 215 | 220 |
| Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn |     |     |     |
| 225                                                             | 230 | 235 | 240 |
| [0007]                                                          |     |     |     |
| Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe |     |     |     |
|                                                                 | 245 | 250 | 255 |
| Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp |     |     |     |
|                                                                 | 260 | 265 | 270 |
| Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp |     |     |     |
|                                                                 | 275 | 280 | 285 |
| Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp |     |     |     |
|                                                                 | 290 | 295 | 300 |
| Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys |     |     |     |

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 325                                                             | 330 | 335 |
| Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe |     |     |
| 340                                                             | 345 | 350 |
| Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser |     |     |
| 355                                                             | 360 | 365 |
| Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp |     |     |
| 370                                                             | 375 | 380 |
| Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg |     |     |
| 385                                                             | 390 | 395 |
| 400                                                             |     |     |
| Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu |     |     |
| 405                                                             | 410 | 415 |
| [0008]                                                          |     |     |
| Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe |     |     |
| 420                                                             | 425 | 430 |
| Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile |     |     |
| 435                                                             | 440 | 445 |
| Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp |     |     |
| 450                                                             | 455 | 460 |
| Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu |     |     |
| 465                                                             | 470 | 475 |
| 480                                                             |     |     |
| Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr |     |     |
| 485                                                             | 490 | 495 |
| Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser |     |     |

|                                                                 |     |         |
|-----------------------------------------------------------------|-----|---------|
| 500                                                             | 505 | 510     |
| Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys |     |         |
| 515                                                             | 520 | 525     |
| Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln |     |         |
| 530                                                             | 535 | 540     |
| Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr |     |         |
| 545                                                             | 550 | 555 560 |
| Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp |     |         |
| 565                                                             | 570 | 575     |
| Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly |     |         |
| 580                                                             | 585 | 590     |
| [0009]                                                          |     |         |
| Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp |     |         |
| 595                                                             | 600 | 605     |
| Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr |     |         |
| 610                                                             | 615 | 620     |
| Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala |     |         |
| 625                                                             | 630 | 635 640 |
| His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr |     |         |
| 645                                                             | 650 | 655     |
| Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp |     |         |
| 660                                                             | 665 | 670     |
| Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe |     |         |

| 675                                                             | 680 | 685     |
|-----------------------------------------------------------------|-----|---------|
| Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe |     |         |
| 690                                                             | 695 | 700     |
| Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu |     |         |
| 705                                                             | 710 | 715 720 |
| His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly |     |         |
| 725                                                             | 730 | 735     |
| Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly |     |         |
| 740                                                             | 745 | 750     |
| Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln |     |         |
| 755                                                             | 760 | 765     |
| [0010]                                                          |     |         |
| Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile |     |         |
| 770                                                             | 775 | 780     |
| Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro |     |         |
| 785                                                             | 790 | 795 800 |
| Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu |     |         |
| 805                                                             | 810 | 815     |
| Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg |     |         |
| 820                                                             | 825 | 830     |
| Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys |     |         |
| 835                                                             | 840 | 845     |
| Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg |     |         |

850

855

860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865

870

875

880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

885

890

895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp

900

905

910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr

915

920

925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp

930

935

940

[0011]

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser

945

950

955

960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

965

970

975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

980

985

990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe

995

1000

1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala

1010

1015

1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe

|                             |                     |             |
|-----------------------------|---------------------|-------------|
| 1025                        | 1030                | 1035        |
| Tyr Ser Asn Ile Met Asn Phe | Phe Lys Thr Glu Ile | Thr Leu Ala |
| 1040                        | 1045                | 1050        |
| Asn Gly Glu Ile Arg Lys Arg | Pro Leu Ile Glu Thr | Asn Gly Glu |
| 1055                        | 1060                | 1065        |
| Thr Gly Glu Ile Val Trp Asp | Lys Gly Arg Asp Phe | Ala Thr Val |
| 1070                        | 1075                | 1080        |
| Arg Lys Val Leu Ser Met Pro | Gln Val Asn Ile Val | Lys Lys Thr |
| 1085                        | 1090                | 1095        |
| Glu Val Gln Thr Gly Gly Phe | Ser Lys Glu Ser Ile | Leu Pro Lys |
| 1100                        | 1105                | 1110        |
| [0012]                      |                     |             |
| Arg Asn Ser Asp Lys Leu Ile | Ala Arg Lys Lys Asp | Trp Asp Pro |
| 1115                        | 1120                | 1125        |
| Lys Lys Tyr Gly Gly Phe Asp | Ser Pro Thr Val Ala | Tyr Ser Val |
| 1130                        | 1135                | 1140        |
| Leu Val Val Ala Lys Val Glu | Lys Gly Lys Ser Lys | Lys Leu Lys |
| 1145                        | 1150                | 1155        |
| Ser Val Lys Glu Leu Leu Gly | Ile Thr Ile Met Glu | Arg Ser Ser |
| 1160                        | 1165                | 1170        |
| Phe Glu Lys Asn Pro Ile Asp | Phe Leu Glu Ala Lys | Gly Tyr Lys |
| 1175                        | 1180                | 1185        |
| Glu Val Lys Lys Asp Leu Ile | Ile Lys Leu Pro Lys | Tyr Ser Leu |

|                                                             |      |      |
|-------------------------------------------------------------|------|------|
| 1190                                                        | 1195 | 1200 |
| Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly |      |      |
| 1205                                                        | 1210 | 1215 |
| Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val |      |      |
| 1220                                                        | 1225 | 1230 |
| Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser |      |      |
| 1235                                                        | 1240 | 1245 |
| Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys |      |      |
| 1250                                                        | 1255 | 1260 |
| His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys |      |      |
| 1265                                                        | 1270 | 1275 |
| [0013]                                                      |      |      |
| Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala |      |      |
| 1280                                                        | 1285 | 1290 |
| Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn |      |      |
| 1295                                                        | 1300 | 1305 |
| Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala |      |      |
| 1310                                                        | 1315 | 1320 |
| Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser |      |      |
| 1325                                                        | 1330 | 1335 |
| Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr |      |      |
| 1340                                                        | 1345 | 1350 |
| Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp |      |      |

|        | 1355                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 1360 | 1365 |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|------|
|        | <p>&lt;210&gt; 6<br/>&lt;211&gt; 262<br/>&lt;212&gt; RNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引导 RNA 多核苷酸</p> <p>&lt;220&gt;<br/>&lt;221&gt; 修饰的_碱基<br/>&lt;222&gt; (1).. (20)<br/>&lt;223&gt; a, c, u, g, 未知区域或其它区域和本区域可包括<br/>17-20 个核苷酸</p> <p>&lt;220&gt;<br/>&lt;221&gt; 修饰的_碱基<br/>&lt;222&gt; (63).. (262)<br/>&lt;223&gt; a, c, u, g, 未知区域或其它区域和本区域可包括<br/>0-200 个核苷酸</p>                                                                                                                  |      |      |
| [0014] | <p>&lt;400&gt; 6<br/>nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60<br/><br/>cgnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 120<br/><br/>nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 180<br/><br/>nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 240<br/><br/>nnnnnnnnnnn nnnnnnnnnnn nn 262</p> <p>&lt;210&gt; 7<br/>&lt;211&gt; 275<br/>&lt;212&gt; RNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的</p> |      |      |

## 引导 RNA 多核苷酸

|                                                                          |     |
|--------------------------------------------------------------------------|-----|
| <220>                                                                    |     |
| <221> 修饰的_碱基                                                             |     |
| <222> (1).. (20)                                                         |     |
| <223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸                         |     |
| <220>                                                                    |     |
| <221> 修饰的_碱基                                                             |     |
| <222> (76).. (275)                                                       |     |
| <223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>0-200 个核苷酸                         |     |
| <400> 7                                                                  |     |
| nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcugaaa agcauagcaa guuaaaauaa        | 60  |
| ggcuaguccg uuaucnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn        | 120 |
| nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn        | 180 |
| [0015] nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn | 240 |
| nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnn                                  | 275 |
| <210> 8                                                                  |     |
| <211> 287                                                                |     |
| <212> RNA                                                                |     |
| <213> 人工序列                                                               |     |
| <220>                                                                    |     |
| <223> 人工序列的说明: 合成的<br>引导 RNA 多核苷酸                                        |     |
| <220>                                                                    |     |
| <221> 修饰的_碱基                                                             |     |
| <222> (1).. (20)                                                         |     |
| <223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸                         |     |
| <220>                                                                    |     |

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| 〈221〉 修饰的_碱基                                                       |     |
| 〈222〉 (88).. (287)                                                 |     |
| 〈223〉 a, c, u, g, 未知区域或其它区域和本区域可包括<br>0-200 个核苷酸                   |     |
| 〈400〉 8                                                            |     |
| nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu uggaacaaaa acagcauagc  | 60  |
| aaguuaaaau aaggcuaguc cguuauccnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn | 120 |
| nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  | 180 |
| nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  | 240 |
| nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnn               | 287 |
| 〈210〉 9                                                            |     |
| 〈211〉 296                                                          |     |
| 〈212〉 RNA                                                          |     |
| 〈213〉 人工序列                                                         |     |
| 〈220〉                                                              |     |
| 〈223〉 人工序列的说明: 合成的<br>引导 RNA 多核苷酸                                  |     |
| 〈220〉                                                              |     |
| 〈221〉 修饰的_碱基                                                       |     |
| 〈222〉 (1).. (20)                                                   |     |
| 〈223〉 a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸                   |     |
| 〈220〉                                                              |     |
| 〈221〉 修饰的_碱基                                                       |     |
| 〈222〉 (97).. (296)                                                 |     |
| 〈223〉 a, c, u, g, 未知区域或其它区域和本区域可包括<br>0-200 个核苷酸                   |     |
| 〈400〉 9                                                            |     |
| nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc  | 60  |
| cguuaucaac uugaaaaagu ggcaccgagu cggugcnnnn nnnnnnnnnn nnnnnnnnnn  | 120 |

[0016]

|                                                                         |     |
|-------------------------------------------------------------------------|-----|
| nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn | 180 |
| nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn | 240 |
| nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnn     | 296 |
| <210> 10                                                                |     |
| <211> 96                                                                |     |
| <212> RNA                                                               |     |
| <213> 人工序列                                                              |     |
| <220>                                                                   |     |
| <223> 人工序列的说明：合成的<br>引导 RNA oligonucleotide                             |     |
| <220>                                                                   |     |
| <221> 修饰的_碱基                                                            |     |
| <222> (1).. (20)                                                        |     |
| <223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸                        |     |
| [0017]                                                                  |     |
| <400> 10                                                                |     |
| nnnnnnnnnnn nnnnnnnnnnn guuuuagagc uagaaauagc aaguuuuuuu aagguuaguc     | 60  |
| cguuuaucaac uugaaaaagu ggcaccgagu cggugc                                | 96  |
| <210> 11                                                                |     |
| <211> 106                                                               |     |
| <212> RNA                                                               |     |
| <213> 人工序列                                                              |     |
| <220>                                                                   |     |
| <223> 人工序列的说明：合成的<br>引导 RNA 多核苷酸                                        |     |
| <220>                                                                   |     |
| <221> 修饰的_碱基                                                            |     |
| <222> (1).. (20)                                                        |     |
| <223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸                        |     |

<400> 11  
nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuggaa acagcauagc aaguuuuuuu  
60  
  
aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc  
106

<210> 12  
<211> 106  
<212> RNA  
<213> 人工序列

<220>  
<223> 人工序列的说明：合成的  
引导 RNA 多核苷酸

<220>  
<221> 修饰的\_碱基  
<222> (1).. (20)  
<223> a, c, u, g, 未知区域或其它区域和本区域可包括  
17-20 个核苷酸

[0018]

<400> 12  
nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuggaa acagcauagc aaguuuuuuu  
60  
  
aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc  
106

<210> 13  
<211> 32  
<212> RNA  
<213> 人工序列

<220>  
<223> 人工序列的说明：合成的  
tracrRNA 寡核苷酸

<220>  
<221> 修饰的\_碱基  
<222> (1).. (20)  
<223> a, c, u, g, 未知区域或其它区域和本区域可包括  
17-20 个核苷酸

|        |                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------|----|
|        | <400> 13<br>nnnnnnnnnn nnnnnnnnnn guuuuagagc ua                                               | 32 |
|        | <210> 14<br><211> 42<br><212> RNA<br><213> 人工序列                                               |    |
|        | <220><br><223> 人工序列的说明：合成的<br>tracrRNA 寡核苷酸                                                   |    |
|        | <220><br><221> 修饰的_碱基<br><222> (1).. (20)<br><223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸 |    |
| [0019] | <400> 14<br>nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu ug                                    | 42 |
|        | <210> 15<br><211> 36<br><212> RNA<br><213> 人工序列                                               |    |
|        | <220><br><223> 人工序列的说明：合成的<br>tracrRNA 寡核苷酸                                                   |    |
|        | <220><br><221> 修饰的_碱基<br><222> (1).. (20)<br><223> a, c, u, g, 未知区域或其它区域和本区域可包括<br>17-20 个核苷酸 |    |
|        | <400> 15<br>nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcu                                           | 36 |

|        |                                                                                           |    |
|--------|-------------------------------------------------------------------------------------------|----|
|        | <210> 16                                                                                  |    |
|        | <211> 79                                                                                  |    |
|        | <212> RNA                                                                                 |    |
|        | <213> 人工序列                                                                                |    |
|        | <220>                                                                                     |    |
|        | <223> 人工序列的说明：合成的<br>tracrRNA 寡核苷酸                                                        |    |
|        | <400> 16                                                                                  |    |
|        | ggaacc <u>auuc</u> aaaacagcau agcaaguuaa aa <u>uaaggcua</u> guccguuau <u>c</u> aacuugaaaa | 60 |
|        | aguggcaccg agucggugc                                                                      | 79 |
|        | <210> 17                                                                                  |    |
|        | <211> 60                                                                                  |    |
|        | <212> RNA                                                                                 |    |
|        | <213> 人工序列                                                                                |    |
| [0020] | <220>                                                                                     |    |
|        | <223> 人工序列的说明：合成的<br>tracrRNA 寡核苷酸                                                        |    |
|        | <400> 17                                                                                  |    |
|        | uagcaaguua aa <u>uaaggcu</u> aguccguuau caacuugaaa aaguggcacc gagucggugc                  | 60 |
|        | <210> 18                                                                                  |    |
|        | <211> 64                                                                                  |    |
|        | <212> RNA                                                                                 |    |
|        | <213> 人工序列                                                                                |    |
|        | <220>                                                                                     |    |
|        | <223> 人工序列的说明：合成的<br>tracrRNA 寡核苷酸                                                        |    |
|        | <400> 18                                                                                  |    |
|        | agcauagcaa guu <u>aaaauaa</u> ggcuaguccg uuaucaacuu gaaaaagugg caccgagucg                 | 60 |
|        | gugc                                                                                      | 64 |

|        |                                        |    |
|--------|----------------------------------------|----|
|        | <210> 19                               |    |
|        | <211> 16                               |    |
|        | <212> DNA                              |    |
|        | <213> 人工序列                             |    |
|        | <220>                                  |    |
|        | <223> 人工序列的说明：合成的<br>引物                |    |
|        | <400> 19                               |    |
|        | gtaaacgac ggccag                       | 16 |
|        | <210> 20                               |    |
|        | <211> 28                               |    |
|        | <212> RNA                              |    |
|        | <213> 人工序列                             |    |
|        | <220>                                  |    |
|        | <223> 人工序列的说明：合成的<br>Csy4 RNA 结合位点寡核苷酸 |    |
| [0021] | <400> 20                               |    |
|        | guucacugcc guauaggcag cuaagaaa         | 28 |
|        | <210> 21                               |    |
|        | <211> 252                              |    |
|        | <212> RNA                              |    |
|        | <213> 人工序列                             |    |
|        | <220>                                  |    |
|        | <223> 人工序列的说明：合成的<br>多重引导 RNA 多核苷酸     |    |
|        | <220>                                  |    |
|        | <221> 修饰的_碱基                           |    |
|        | <222> (21)..(40)                       |    |
|        | <223> a, c, u, g, 未知的或其它               |    |
|        | <220>                                  |    |
|        | <221> 修饰的_碱基                           |    |
|        | <222> (137)..(156)                     |    |

|        |                                                                    |     |
|--------|--------------------------------------------------------------------|-----|
|        | <223> a, c, u, g, 未知的或其它                                           |     |
|        | <400> 21                                                           |     |
|        | guucacugcc guauaggcag nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc  | 60  |
|        | aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcguuc  | 120 |
|        | acugccguau aggcagnnnn nnnnnnnnnn nnnnnnguuu uagagcuaga aaauagcaagu | 180 |
|        | uaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucggu gcguucacug  | 240 |
|        | ccguauaggc ag                                                      | 252 |
|        | <210> 22                                                           |     |
|        | <211> 4                                                            |     |
|        | <212> PRT                                                          |     |
|        | <213> 人工序列                                                         |     |
|        | <220>                                                              |     |
|        | <223> 人工序列的说明：合成的<br>接头肽                                           |     |
| [0022] | <400> 22                                                           |     |
|        | Gly Gly Gly Ser                                                    |     |
|        | 1                                                                  |     |
|        | <210> 23                                                           |     |
|        | <211> 5                                                            |     |
|        | <212> PRT                                                          |     |
|        | <213> 人工序列                                                         |     |
|        | <220>                                                              |     |
|        | <223> 人工序列的说明：合成的<br>接头肽                                           |     |
|        | <400> 23                                                           |     |
|        | Gly Gly Gly Gly Ser                                                |     |
|        | 1 5                                                                |     |
|        | <210> 24                                                           |     |
|        | <211> 1602                                                         |     |

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

<223> 人工序列的说明：合成的  
FokI-dCas9 多肽

&lt;400&gt; 24

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gln | Leu | Val | Lys | Ser | Glu | Leu | Glu | Glu | Lys | Lys | Ser | Glu | Leu | Arg |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Lys | Leu | Lys | Tyr | Val | Pro | His | Glu | Tyr | Ile | Glu | Leu | Ile | Glu | Ile |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Arg | Asn | Ser | Thr | Gln | Asp | Arg | Ile | Leu | Glu | Met | Lys | Val | Met | Glu |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

[0023]

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Phe | Met | Lys | Val | Tyr | Gly | Tyr | Arg | Gly | Lys | His | Leu | Gly | Gly | Ser |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Lys | Pro | Asp | Gly | Ala | Ile | Tyr | Thr | Val | Gly | Ser | Pro | Ile | Asp | Tyr |
| 65  |     |     |     | 70  |     |     |     |     |     | 75  |     |     |     | 80  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Val | Ile | Val | Asp | Thr | Lys | Ala | Tyr | Ser | Gly | Gly | Tyr | Asn | Leu | Pro |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Gly | Gln | Ala | Asp | Glu | Met | Gln | Arg | Tyr | Val | Glu | Glu | Asn | Gln | Thr |
|     |     | 100 |     |     |     |     |     | 105 |     |     |     |     |     | 110 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Asn | Lys | His | Ile | Asn | Pro | Asn | Glu | Trp | Trp | Lys | Val | Tyr | Pro | Ser |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Val | Thr | Glu | Phe | Lys | Phe | Leu | Phe | Val | Ser | Gly | His | Phe | Lys | Gly |
|     |     | 130 |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |

|        |                                                                 |  |
|--------|-----------------------------------------------------------------|--|
|        | Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn |  |
|        | 145 150 155 160                                                 |  |
|        | Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile |  |
|        | 165 170 175                                                     |  |
|        | Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn |  |
|        | 180 185 190                                                     |  |
|        | Gly Glu Ile Asn Phe Gly Gly Gly Gly Ser Asp Lys Lys Tyr Ser Ile |  |
|        | 195 200 205                                                     |  |
|        | Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp |  |
|        | 210 215 220                                                     |  |
| [0024] | Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp |  |
|        | 225 230 235 240                                                 |  |
|        | Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser |  |
|        | 245 250 255                                                     |  |
|        | Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg |  |
|        | 260 265 270                                                     |  |
|        | Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser |  |
|        | 275 280 285                                                     |  |
|        | Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu |  |
|        | 290 295 300                                                     |  |
|        | Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe |  |
|        | 305 310 315 320                                                 |  |



|        |                                                                 |             |
|--------|-----------------------------------------------------------------|-------------|
|        | Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu |             |
|        | 500                                                             | 505 510     |
|        | Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu |             |
|        | 515                                                             | 520 525     |
|        | His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu |             |
|        | 530                                                             | 535 540     |
|        | Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr |             |
|        | 545                                                             | 550 555 560 |
|        | Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe |             |
|        | 565                                                             | 570 575     |
| [0026] | Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val |             |
|        | 580                                                             | 585 590     |
|        | Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn |             |
|        | 595                                                             | 600 605     |
|        | Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu |             |
|        | 610                                                             | 615 620     |
|        | Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys |             |
|        | 625                                                             | 630 635 640 |
|        | Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu |             |
|        | 645                                                             | 650 655     |
|        | Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu |             |
|        | 660                                                             | 665 670     |

|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|        | Thr | Ile | Thr | Pro | Trp | Asn | Phe | Glu | Glu | Val | Val | Asp | Lys | Gly | Ala | Ser |     |
|        |     |     |     |     |     |     |     | 675 |     |     |     |     |     |     |     |     | 680 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 685 |
|        | Ala | Gln | Ser | Phe | Ile | Glu | Arg | Met | Thr | Asn | Phe | Asp | Lys | Asn | Leu | Pro |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 690 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 695 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 700 |
|        | Asn | Glu | Lys | Val | Leu | Pro | Lys | His | Ser | Leu | Leu | Tyr | Glu | Tyr | Phe | Thr |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 705 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 710 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 715 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 720 |
|        | Val | Tyr | Asn | Glu | Leu | Thr | Lys | Val | Lys | Tyr | Val | Thr | Glu | Gly | Met | Arg |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 725 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 730 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 735 |
|        | Lys | Pro | Ala | Phe | Leu | Ser | Gly | Glu | Gln | Lys | Lys | Ala | Ile | Val | Asp | Leu |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 740 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 745 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 750 |
| [0027] | Leu | Phe | Lys | Thr | Asn | Arg | Lys | Val | Thr | Val | Lys | Gln | Leu | Lys | Glu | Asp |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 755 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 760 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 765 |
|        | Tyr | Phe | Lys | Lys | Ile | Glu | Cys | Phe | Asp | Ser | Val | Glu | Ile | Ser | Gly | Val |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 770 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 775 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 780 |
|        | Glu | Asp | Arg | Phe | Asn | Ala | Ser | Leu | Gly | Thr | Tyr | His | Asp | Leu | Leu | Lys |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 785 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 790 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 795 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 800 |
|        | Ile | Ile | Lys | Asp | Lys | Asp | Phe | Leu | Asp | Asn | Glu | Glu | Asn | Glu | Asp | Ile |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 805 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 810 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 815 |
|        | Leu | Glu | Asp | Ile | Val | Leu | Thr | Leu | Thr | Leu | Phe | Glu | Asp | Arg | Glu | Met |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 820 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 825 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 830 |
|        | Ile | Glu | Glu | Arg | Leu | Lys | Thr | Tyr | Ala | His | Leu | Phe | Asp | Asp | Lys | Val |     |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 835 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 840 |
|        |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 845 |

|        |      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |     |  |
|--------|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|--|
|        | Met  | Lys | Gln | Leu | Lys | Arg | Arg  | Arg | Tyr | Thr | Gly | Trp  | Gly | Arg | Leu | Ser |  |
|        | 850  |     |     |     |     |     | 855  |     |     |     |     | 860  |     |     |     |     |  |
|        | Arg  | Lys | Leu | Ile | Asn | Gly | Ile  | Arg | Asp | Lys | Gln | Ser  | Gly | Lys | Thr | Ile |  |
|        | 865  |     |     |     |     | 870 |      |     |     |     | 875 |      |     |     |     | 880 |  |
|        | Leu  | Asp | Phe | Leu | Lys | Ser | Asp  | Gly | Phe | Ala | Asn | Arg  | Asn | Phe | Met | Gln |  |
|        |      |     |     |     | 885 |     |      |     |     | 890 |     |      |     |     | 895 |     |  |
|        | Leu  | Ile | His | Asp | Asp | Ser | Leu  | Thr | Phe | Lys | Glu | Asp  | Ile | Gln | Lys | Ala |  |
|        |      |     |     | 900 |     |     |      |     | 905 |     |     |      |     | 910 |     |     |  |
|        | Gln  | Val | Ser | Gly | Gln | Gly | Asp  | Ser | Leu | His | Glu | His  | Ile | Ala | Asn | Leu |  |
|        |      | 915 |     |     |     |     | 920  |     |     |     |     |      | 925 |     |     |     |  |
| [0028] | Ala  | Gly | Ser | Pro | Ala | Ile | Lys  | Lys | Gly | Ile | Leu | Gln  | Thr | Val | Lys | Val |  |
|        | 930  |     |     |     |     |     | 935  |     |     |     |     | 940  |     |     |     |     |  |
|        | Val  | Asp | Glu | Leu | Val | Lys | Val  | Met | Gly | Arg | His | Lys  | Pro | Glu | Asn | Ile |  |
|        | 945  |     |     |     |     | 950 |      |     |     | 955 |     |      |     |     | 960 |     |  |
|        | Val  | Ile | Glu | Met | Ala | Arg | Glu  | Asn | Gln | Thr | Thr | Gln  | Lys | Gly | Gln | Lys |  |
|        |      |     |     | 965 |     |     |      |     | 970 |     |     |      |     | 975 |     |     |  |
|        | Asn  | Ser | Arg | Glu | Arg | Met | Lys  | Arg | Ile | Glu | Glu | Gly  | Ile | Lys | Glu | Leu |  |
|        |      |     |     | 980 |     |     |      |     | 985 |     |     |      |     | 990 |     |     |  |
|        | Gly  | Ser | Gln | Ile | Leu | Lys | Glu  | His | Pro | Val | Glu | Asn  | Thr | Gln | Leu | Gln |  |
|        |      | 995 |     |     |     |     | 1000 |     |     |     |     | 1005 |     |     |     |     |  |
|        | Asn  | Glu | Lys | Leu | Tyr | Leu | Tyr  | Tyr | Leu | Gln | Asn | Gly  | Arg | Asp | Met |     |  |
|        | 1010 |     |     |     |     |     | 1015 |     |     |     |     | 1020 |     |     |     |     |  |

|        |      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |
|--------|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|
|        | Tyr  | Val | Asp | Gln | Glu | Leu | Asp  | Ile | Asn | Arg | Leu | Ser  | Asp | Tyr | Asp |
|        | 1025 |     |     |     |     |     | 1030 |     |     |     |     | 1035 |     |     |     |
|        | Val  | Asp | Ala | Ile | Val | Pro | Gln  | Ser | Phe | Leu | Lys | Asp  | Asp | Ser | Ile |
|        | 1040 |     |     |     |     |     | 1045 |     |     |     |     | 1050 |     |     |     |
|        | Asp  | Asn | Lys | Val | Leu | Thr | Arg  | Ser | Asp | Lys | Asn | Arg  | Gly | Lys | Ser |
|        | 1055 |     |     |     |     |     | 1060 |     |     |     |     | 1065 |     |     |     |
|        | Asp  | Asn | Val | Pro | Ser | Glu | Glu  | Val | Val | Lys | Lys | Met  | Lys | Asn | Tyr |
|        | 1070 |     |     |     |     |     | 1075 |     |     |     |     | 1080 |     |     |     |
|        | Trp  | Arg | Gln | Leu | Leu | Asn | Ala  | Lys | Leu | Ile | Thr | Gln  | Arg | Lys | Phe |
|        | 1085 |     |     |     |     |     | 1090 |     |     |     |     | 1095 |     |     |     |
| [0029] | Asp  | Asn | Leu | Thr | Lys | Ala | Glu  | Arg | Gly | Gly | Leu | Ser  | Glu | Leu | Asp |
|        | 1100 |     |     |     |     |     | 1105 |     |     |     |     | 1110 |     |     |     |
|        | Lys  | Ala | Gly | Phe | Ile | Lys | Arg  | Gln | Leu | Val | Glu | Thr  | Arg | Gln | Ile |
|        | 1115 |     |     |     |     |     | 1120 |     |     |     |     | 1125 |     |     |     |
|        | Thr  | Lys | His | Val | Ala | Gln | Ile  | Leu | Asp | Ser | Arg | Met  | Asn | Thr | Lys |
|        | 1130 |     |     |     |     |     | 1135 |     |     |     |     | 1140 |     |     |     |
|        | Tyr  | Asp | Glu | Asn | Asp | Lys | Leu  | Ile | Arg | Glu | Val | Lys  | Val | Ile | Thr |
|        | 1145 |     |     |     |     |     | 1150 |     |     |     |     | 1155 |     |     |     |
|        | Leu  | Lys | Ser | Lys | Leu | Val | Ser  | Asp | Phe | Arg | Lys | Asp  | Phe | Gln | Phe |
|        | 1160 |     |     |     |     |     | 1165 |     |     |     |     | 1170 |     |     |     |
|        | Tyr  | Lys | Val | Arg | Glu | Ile | Asn  | Asn | Tyr | His | His | Ala  | His | Asp | Ala |
|        | 1175 |     |     |     |     |     | 1180 |     |     |     |     | 1185 |     |     |     |

|        |      |     |     |     |     |     |      |     |     |     |     |     |      |     |     |
|--------|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|------|-----|-----|
|        | Tyr  | Leu | Asn | Ala | Val | Val | Gly  | Thr | Ala | Leu | Ile | Lys | Lys  | Tyr | Pro |
|        | 1190 |     |     |     |     |     | 1195 |     |     |     |     |     | 1200 |     |     |
|        | Lys  | Leu | Glu | Ser | Glu | Phe | Val  | Tyr | Gly | Asp | Tyr | Lys | Val  | Tyr | Asp |
|        | 1205 |     |     |     |     |     | 1210 |     |     |     |     |     | 1215 |     |     |
|        | Val  | Arg | Lys | Met | Ile | Ala | Lys  | Ser | Glu | Gln | Glu | Ile | Gly  | Lys | Ala |
|        | 1220 |     |     |     |     |     | 1225 |     |     |     |     |     | 1230 |     |     |
|        | Thr  | Ala | Lys | Tyr | Phe | Phe | Tyr  | Ser | Asn | Ile | Met | Asn | Phe  | Phe | Lys |
|        | 1235 |     |     |     |     |     | 1240 |     |     |     |     |     | 1245 |     |     |
|        | Thr  | Glu | Ile | Thr | Leu | Ala | Asn  | Gly | Glu | Ile | Arg | Lys | Arg  | Pro | Leu |
|        | 1250 |     |     |     |     |     | 1255 |     |     |     |     |     | 1260 |     |     |
| [0030] | Ile  | Glu | Thr | Asn | Gly | Glu | Thr  | Gly | Glu | Ile | Val | Trp | Asp  | Lys | Gly |
|        | 1265 |     |     |     |     |     | 1270 |     |     |     |     |     | 1275 |     |     |
|        | Arg  | Asp | Phe | Ala | Thr | Val | Arg  | Lys | Val | Leu | Ser | Met | Pro  | Gln | Val |
|        | 1280 |     |     |     |     |     | 1285 |     |     |     |     |     | 1290 |     |     |
|        | Asn  | Ile | Val | Lys | Lys | Thr | Glu  | Val | Gln | Thr | Gly | Gly | Phe  | Ser | Lys |
|        | 1295 |     |     |     |     |     | 1300 |     |     |     |     |     | 1305 |     |     |
|        | Glu  | Ser | Ile | Leu | Pro | Lys | Arg  | Asn | Ser | Asp | Lys | Leu | Ile  | Ala | Arg |
|        | 1310 |     |     |     |     |     | 1315 |     |     |     |     |     | 1320 |     |     |
|        | Lys  | Lys | Asp | Trp | Asp | Pro | Lys  | Lys | Tyr | Gly | Gly | Phe | Asp  | Ser | Pro |
|        | 1325 |     |     |     |     |     | 1330 |     |     |     |     |     | 1335 |     |     |
|        | Thr  | Val | Ala | Tyr | Ser | Val | Leu  | Val | Val | Ala | Lys | Val | Glu  | Lys | Gly |
|        | 1340 |     |     |     |     |     | 1345 |     |     |     |     |     | 1350 |     |     |

|        |         |                     |                     |             |
|--------|---------|---------------------|---------------------|-------------|
|        | Lys Ser | Lys Lys Leu Lys Ser | Val Lys Glu Leu Leu | Gly Ile Thr |
|        | 1355    | 1360                | 1365                |             |
|        | Ile Met | Glu Arg Ser Ser Phe | Glu Lys Asn Pro Ile | Asp Phe Leu |
|        | 1370    | 1375                | 1380                |             |
|        | Glu Ala | Lys Gly Tyr Lys Glu | Val Lys Lys Asp Leu | Ile Ile Lys |
|        | 1385    | 1390                | 1395                |             |
|        | Leu Pro | Lys Tyr Ser Leu Phe | Glu Leu Glu Asn Gly | Arg Lys Arg |
|        | 1400    | 1405                | 1410                |             |
|        | Met Leu | Ala Ser Ala Gly Glu | Leu Gln Lys Gly Asn | Glu Leu Ala |
|        | 1415    | 1420                | 1425                |             |
| [0031] | Leu Pro | Ser Lys Tyr Val Asn | Phe Leu Tyr Leu Ala | Ser His Tyr |
|        | 1430    | 1435                | 1440                |             |
|        | Glu Lys | Leu Lys Gly Ser Pro | Glu Asp Asn Glu Gln | Lys Gln Leu |
|        | 1445    | 1450                | 1455                |             |
|        | Phe Val | Glu Gln His Lys His | Tyr Leu Asp Glu Ile | Ile Glu Gln |
|        | 1460    | 1465                | 1470                |             |
|        | Ile Ser | Glu Phe Ser Lys Arg | Val Ile Leu Ala Asp | Ala Asn Leu |
|        | 1475    | 1480                | 1485                |             |
|        | Asp Lys | Val Leu Ser Ala Tyr | Asn Lys His Arg Asp | Lys Pro Ile |
|        | 1490    | 1495                | 1500                |             |
|        | Arg Glu | Gln Ala Glu Asn Ile | Ile His Leu Phe Thr | Leu Thr Asn |
|        | 1505    | 1510                | 1515                |             |

Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp  
1520 1525 1530

Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu  
1535 1540 1545

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu  
1550 1555 1560

Ser Gln Leu Gly Gly Asp Gly Ser Pro Lys Lys Lys Arg Lys Val  
1565 1570 1575

Ser Ser Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp  
1580 1585 1590

[0032] Ile Asp Tyr Lys Asp Asp Asp Asp Lys  
1595 1600

<210> 25

<211> 4809

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
FokI-dCas9 多核苷酸

<400> 25

atgcaactag tcaaaagtga actggaggag aagaaatctg aacttcgtca taaattgaaa 60

tatgtgcctc atgaatatat tgaattaatt gaaattgcc aaaattccac tcaggataga 120

attcttgaaa tgaaggtaat ggaatttttt atgaaagttt atggatatag aggtaaacat 180

ttgggtggat caaggaaacc ggacggagca atttatactg tcggatctcc tattgattac 240

ggtgtgatcg tggatactaa agcttatagc ggagggtata atctgccaat tggccaagca 300

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
|        | gatgaaatgc aacgatatgt cgaagaaaat caaacacgaa acaaacatat caacccta    | 360  |
|        | gaatgggtgga aagtctatcc atcttctgta acggaattta agtttttatt tgtgagtgg  | 420  |
|        | cactttaaag gaaactacaa agctcagctt acacgattaa atcatatcac taattgta    | 480  |
|        | ggagctgttc ttagtgtaga agagctttta attgggtggag aaatgattaa agccggcaca | 540  |
|        | ttaaccttag aggaagtcag acggaaattt aataacggcg agataaactt tgggtggcgt  | 600  |
|        | ggatccgata aaaagtattc tattggttta gccatcgga ctaattccgt tggatgggct   | 660  |
|        | gtcataaccg atgaatacaa agtaccttca aagaaattta aggtgttggg gaacacagac  | 720  |
|        | cgtcattcga ttaaaaagaa tcttatcggt gccctctat tcgatagtgg cgaaacggca   | 780  |
|        | gaggcgactc gcctgaaacg aaccgctcgg agaaggtata cacgtcgaa gaaccgaata   | 840  |
|        | tgttacttac aagaaatttt tagcaatgag atggccaaag ttgacgattc tttctttcac  | 900  |
|        | cgtttggaag agtccttcct tgtcgaagag gacaagaaac atgaacggca ccccatctt   | 960  |
| [0033] | ggaaacatag tagatgaggt ggcatatcat gaaaagtacc caacgattta tcacctcaga  | 1020 |
|        | aaaaagctag ttgactcaac tgataaagcg gacctgaggt taatctactt ggctcttgcc  | 1080 |
|        | catatgataa agttccgtgg gcactttctc attgagggtg atctaaatcc ggacaactcg  | 1140 |
|        | gatgtcgaca aactgttcat ccagttagta caaacctata atcagttgtt tgaagagaac  | 1200 |
|        | cctataaatg caagtggcgt ggatgcgaag gctattctta gcgcccgcct ctctaaatcc  | 1260 |
|        | cgacggctag aaaacctgat cgcacaatta cccggagaga agaaaaatgg gttgttcggt  | 1320 |
|        | aaccttatag cgtctcact aggcctgaca ccaaatttta agtcgaactt cgacttagct   | 1380 |
|        | gaagatgcca aattgcagct tagtaaggac acgtacgatg acgatctcga caatctactg  | 1440 |
|        | gcacaaattg gagatcagta tgcggactta tttttggctg ccaaaaacct tagcgatgca  | 1500 |
|        | atctctctat ctgacatact gagagttaat actgagatta ccaaggcgcc gttatccgct  | 1560 |
|        | tcaatgatca aaaggtacga tgaacatcac caagacttga cacttctcaa ggccctagtc  | 1620 |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
|        | cgtcagcaac tgcctgagaa atataaggaa atattctttg atcagtcgaa aaacgggtac | 1680 |
|        | gcaggttata ttgacggcgg agcgagtcaa gaggaattct acaagtttat caaacccata | 1740 |
|        | ttagagaaga tggatgggac ggaagagttg cttgtaaaac tcaatcgca agatctactg  | 1800 |
|        | cgaaagcagc ggactttcga caacggtagc attccacatc aaatccactt aggccaattg | 1860 |
|        | catgctatac ttagaaggca ggaggatttt tatccgttcc tcaaagacaa tcgtgaaaag | 1920 |
|        | attgagaaaa tcctaacctt tcgcatacct tactatgtgg gacccctggc ccgagggaac | 1980 |
|        | tctcggttcg catggatgac aagaaagtcc gaagaaacga ttactccatg gaattttgag | 2040 |
|        | gaagttgtcg ataaaggtgc gtcagctcaa tcgttcacg agaggatgac caactttgac  | 2100 |
|        | aagaatttac cgaacgaaaa agtattgcct aagcacagtt tactttacga gtatttcaca | 2160 |
|        | gtgtacaatg aactcacgaa agttaagtat gtcactgagg gcatgcgtaa acccgccttt | 2220 |
|        | ctaagcggag aacagaagaa agcaatagta gatctgttat tcaagaccaa ccgcaaagtg | 2280 |
| [0034] | acagttaagc aattgaaaga ggactacttt aagaaaattg aatgcttcga ttctgtcgag | 2340 |
|        | atctccgggg tagaagatcg atttaatgcg tcacttggtg cgtatcatga cctcctaaag | 2400 |
|        | ataattaaag ataaggactt cctggataac gaagagaatg aagatatctt agaagatata | 2460 |
|        | gtgttgactc ttaccctctt tgaagatcgg gaaatgattg aggaaagact aaaaacatac | 2520 |
|        | gctcacctgt tcgacgataa ggttatgaaa cagttaaaga ggcgtcgcta tacgggctgg | 2580 |
|        | ggacgattgt cgcggaaact tatcaacggg ataagagaca agcaaagtgg taaaactatt | 2640 |
|        | ctcgattttc taaagagcga cggcttcgcc aataggaact ttatgcagct gatccatgat | 2700 |
|        | gactctttaa cttcaaaaga ggatatacaa aaggcacagg tttccggaca aggggactca | 2760 |
|        | ttgcacgaac atattgcgaa tcttgctggt tcgccagcca tcaaaaaggg catactccag | 2820 |
|        | acagtcaaag tagtgatga gctagttaag gtcatgggac gtcacaaacc ggaaaacatt  | 2880 |
|        | gtaatcgaga tggcacgcga aaatcaaacg actcagaagg ggcaaaaaaa cagtcgagag | 2940 |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
|        | cggatgaaga gaatagaaga gggattataa gaactgggca gccagatctt aaaggagcat | 3000 |
|        | cctgtggaaa atacccaatt gcagaacgag aaactttacc tctattacct acaaaatgga | 3060 |
|        | agggacatgt atgttgatca ggaactggac ataaaccgtt tatctgatta cgacgtcgat | 3120 |
|        | gccattgtac cccaatcctt ttggaaggac gattcaatcg acaataaagt gcttacacgc | 3180 |
|        | tcggataaga accgagggaa aagtgacaat gttccaagcg aggaagtcgt aaagaaaatg | 3240 |
|        | aagaactatt ggcggcagct cctaaatgcg aaactgataa cgcaaagaaa gttcgataac | 3300 |
|        | ttaactaaag ctgagagggg tggtttgtct gaacttgaca aggccggatt tattaacgt  | 3360 |
|        | cagctcgtgg aaaccgcca aatcacaaag catgttgac agatactaga ttcccgaatg   | 3420 |
|        | aatacgaaat acgacgagaa cgataagctg attcggaag tcaaagtaat cactttaaag  | 3480 |
|        | tcaaaattgg tgtcgactt cagaaaggat ttcaattct ataaagttag ggagataaat   | 3540 |
|        | aactaccacc atgcgcacga cgcttatctt aatgccgtcg tagggaccgc actcattaag | 3600 |
| [0035] | aaataccga agctagaaag tgagtttgtg tatggtgatt acaaagtta tgacgtccgt   | 3660 |
|        | aagatgatcg cgaaaagcga acaggagata ggcaaggcta cagccaaata cttctttat  | 3720 |
|        | tctaacatta tgaatttctt taagacggaa atcactctgg caaacggaga gatacgcaa  | 3780 |
|        | cgaccttta ttgaaaccaa tggggagaca ggtgaaatcg tatgggataa gggccgggac  | 3840 |
|        | ttcgcgacgg tgagaaaagt ttgtccatg cccaagtca acatagtaaa gaaaactgag   | 3900 |
|        | gtgcagaccg gagggttttc aaaggaatcg attcttccaa aaaggaatag tgataagctc | 3960 |
|        | atcgtctgta aaaaggactg ggacccgaaa aagtacgtg gcttcgatag ccctacagtt  | 4020 |
|        | gcctattctg tcctagtagt ggcaaaagtt gagaaggga aatccaagaa actgaagtca  | 4080 |
|        | gtcaaagaat tattggggat aacgattatg gagegtcgt cttttgaaaa gaacccatc   | 4140 |
|        | gacttccttg aggcgaaagg ttacaaggaa gtaaaaaagg atctcataat taaactacca | 4200 |
|        | aagtatagtc tgtttgagtt agaaaatggc cgaaaacgga tgttggctag cgccggagag | 4260 |

```

cttcaaaagg ggaacgaact cgactaccg tctaaatacg tgaatttcct gtatttagcg      4320

tcccattacg agaagttgaa aggttcacct gaagataacg aacagaagca actttttgtt      4380

gagcagcaca aacattatct cgacgaaatc atagagcaaa tttcggaatt cagtaagaga      4440

gtcatcctag ctgatgccaa tctggacaaa gtattaagcg catacaacaa gcacagggat      4500

aaaccatac  gtgagcaggc ggaaaatatt atccatttgt ttactcttac caacctcggc      4560

gctccagccg cattcaagta ttttgacaca acgatagatc gcaaacgata cacttctacc      4620

aaggagggtgc tagacgcgac actgattcac caatccatca cgggattata tgaaactcgg      4680

atagatttgt cacagcttgg gggtgacgga tcccccaaga agaagaggaa agtctcgagc      4740

gactacaaag accatgacgg tgattataaa gatcatgaca tcgattacaa ggatgacgat      4800

gacaagtga                                         4809

```

[0036]

<210> 26  
 <211> 1401  
 <212> PRT  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列的说明：合成的  
       Nls-FokI-dCas9-nls 多肽

```

<400> 26
Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1             5             10             15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
                20             25             30

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

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

```

50

55

60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys  
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser  
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys  
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr  
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp  
130 135 140

[0037]

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His  
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro  
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr  
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala  
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn  
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 225                                                             | 230 | 235 | 240 |
| Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe |     |     |     |
| 245                                                             | 250 | 255 |     |
| Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp |     |     |     |
| 260                                                             | 265 | 270 |     |
| Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp |     |     |     |
| 275                                                             | 280 | 285 |     |
| Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp |     |     |     |
| 290                                                             | 295 | 300 |     |
| Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| [0038]                                                          |     |     |     |
| Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys |     |     |     |
| 325                                                             | 330 | 335 |     |
| Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe |     |     |     |
| 340                                                             | 345 | 350 |     |
| Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser |     |     |     |
| 355                                                             | 360 | 365 |     |
| Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp |     |     |     |
| 370                                                             | 375 | 380 |     |
| Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg |     |     |     |
| 385                                                             | 390 | 395 | 400 |
| Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu |     |     |     |

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
|                                                                 | 405 | 410 | 415 |
| Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe |     |     |     |
|                                                                 | 420 | 425 | 430 |
| Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile |     |     |     |
|                                                                 | 435 | 440 | 445 |
| Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp |     |     |     |
|                                                                 | 450 | 455 | 460 |
| Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu |     |     |     |
|                                                                 | 465 | 470 | 475 |
| Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr |     |     |     |
|                                                                 | 485 | 490 | 495 |
| [0039]                                                          |     |     |     |
| Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser |     |     |     |
|                                                                 | 500 | 505 | 510 |
| Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys |     |     |     |
|                                                                 | 515 | 520 | 525 |
| Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln |     |     |     |
|                                                                 | 530 | 535 | 540 |
| Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr |     |     |     |
|                                                                 | 545 | 550 | 555 |
| Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp |     |     |     |
|                                                                 | 565 | 570 | 575 |
| Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly |     |     |     |

| 580                                                             | 585 | 590     |
|-----------------------------------------------------------------|-----|---------|
| Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp |     |         |
| 595                                                             | 600 | 605     |
| Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr |     |         |
| 610                                                             | 615 | 620     |
| Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala |     |         |
| 625                                                             | 630 | 635 640 |
| His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr |     |         |
| 645                                                             | 650 | 655     |
| Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp |     |         |
| 660                                                             | 665 | 670     |
| [0040]                                                          |     |         |
| Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe |     |         |
| 675                                                             | 680 | 685     |
| Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe |     |         |
| 690                                                             | 695 | 700     |
| Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu |     |         |
| 705                                                             | 710 | 715 720 |
| His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly |     |         |
| 725                                                             | 730 | 735     |
| Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly |     |         |
| 740                                                             | 745 | 750     |
| Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln |     |         |

| 755                                                                    | 760 | 765     |
|------------------------------------------------------------------------|-----|---------|
| Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile<br>770 | 775 | 780     |
| Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro<br>785 | 790 | 795 800 |
| Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu<br>805 | 810 | 815     |
| Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg<br>820 | 825 | 830     |
| Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys<br>835 | 840 | 845     |
| [0041]                                                                 |     |         |
| Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg<br>850 | 855 | 860     |
| Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys<br>865 | 870 | 875 880 |
| Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys<br>885 | 890 | 895     |
| Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp<br>900 | 905 | 910     |
| Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr<br>915 | 920 | 925     |
| Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp        |     |         |

|                                                                 |      |         |
|-----------------------------------------------------------------|------|---------|
| 930                                                             | 935  | 940     |
| Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser |      |         |
| 945                                                             | 950  | 955 960 |
| Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg |      |         |
| 965                                                             | 970  | 975     |
| Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val |      |         |
| 980                                                             | 985  | 990     |
| Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe |      |         |
| 995                                                             | 1000 | 1005    |
| Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala     |      |         |
| 1010                                                            | 1015 | 1020    |
| [0042]                                                          |      |         |
| Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe     |      |         |
| 1025                                                            | 1030 | 1035    |
| Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala     |      |         |
| 1040                                                            | 1045 | 1050    |
| Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu     |      |         |
| 1055                                                            | 1060 | 1065    |
| Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val     |      |         |
| 1070                                                            | 1075 | 1080    |
| Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr     |      |         |
| 1085                                                            | 1090 | 1095    |
| Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys     |      |         |

|                             |                     |             |
|-----------------------------|---------------------|-------------|
| 1100                        | 1105                | 1110        |
| Arg Asn Ser Asp Lys Leu Ile | Ala Arg Lys Lys Asp | Trp Asp Pro |
| 1115                        | 1120                | 1125        |
| Lys Lys Tyr Gly Gly Phe Asp | Ser Pro Thr Val Ala | Tyr Ser Val |
| 1130                        | 1135                | 1140        |
| Leu Val Val Ala Lys Val Glu | Lys Gly Lys Ser Lys | Lys Leu Lys |
| 1145                        | 1150                | 1155        |
| Ser Val Lys Glu Leu Leu Gly | Ile Thr Ile Met Glu | Arg Ser Ser |
| 1160                        | 1165                | 1170        |
| Phe Glu Lys Asn Pro Ile Asp | Phe Leu Glu Ala Lys | Gly Tyr Lys |
| 1175                        | 1180                | 1185        |
| [0043]                      |                     |             |
| Glu Val Lys Lys Asp Leu Ile | Ile Lys Leu Pro Lys | Tyr Ser Leu |
| 1190                        | 1195                | 1200        |
| Phe Glu Leu Glu Asn Gly Arg | Lys Arg Met Leu Ala | Ser Ala Gly |
| 1205                        | 1210                | 1215        |
| Glu Leu Gln Lys Gly Asn Glu | Leu Ala Leu Pro Ser | Lys Tyr Val |
| 1220                        | 1225                | 1230        |
| Asn Phe Leu Tyr Leu Ala Ser | His Tyr Glu Lys Leu | Lys Gly Ser |
| 1235                        | 1240                | 1245        |
| Pro Glu Asp Asn Glu Gln Lys | Gln Leu Phe Val Glu | Gln His Lys |
| 1250                        | 1255                | 1260        |
| His Tyr Leu Asp Glu Ile Ile | Glu Gln Ile Ser Glu | Phe Ser Lys |

|                                                             |      |      |
|-------------------------------------------------------------|------|------|
| 1265                                                        | 1270 | 1275 |
| Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala |      |      |
| 1280                                                        | 1285 | 1290 |
| Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn |      |      |
| 1295                                                        | 1300 | 1305 |
| Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala |      |      |
| 1310                                                        | 1315 | 1320 |
| Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser |      |      |
| 1325                                                        | 1330 | 1335 |
| Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr |      |      |
| 1340                                                        | 1345 | 1350 |
| [0044]                                                      |      |      |
| Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp |      |      |
| 1355                                                        | 1360 | 1365 |
| Gly Ser Pro Lys Lys Lys Arg Lys Val Ser Ser Asp Tyr Lys Asp |      |      |
| 1370                                                        | 1375 | 1380 |
| His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp |      |      |
| 1385                                                        | 1390 | 1395 |
| Asp Asp Lys                                                 |      |      |
| 1400                                                        |      |      |
| <210> 27                                                    |      |      |
| <211> 4836                                                  |      |      |
| <212> DNA                                                   |      |      |
| <213> 人工序列                                                  |      |      |

&lt;220&gt;

&lt;223&gt; 人工序列的说明：合成的

Nls-FokI-dCas9-nls 多核苷酸

&lt;400&gt; 27

|                                                                          |      |
|--------------------------------------------------------------------------|------|
| atgcctaaga agaagcgga ggtgagcagc caacttgtga agtctgaact cgaggagaaa         | 60   |
| aatcagagt tgagacacaa gttgaagtac gtgccacacg aatacatcga gcttatcgag         | 120  |
| atgccagaa acagtaccca ggataggatc cttgagatga aagtcatgga gttctttatg         | 180  |
| aaggtctacg gttatagagg aaagcacctt ggcggtagca gaaagcccga tggcgccatc        | 240  |
| tatactgtcg gatctcctat cgattatggg gtgatcgtgg ataccaaagc ttactcaggc        | 300  |
| gggtacaact tgcccatagg acaagccgac gagatgcagc ggtatgtcga agagaaccag        | 360  |
| acgcgcaaca agcacatcaa cccaatgaa tgggtgaaag tgtacccaag tagtgtgact         | 420  |
| gagttcaagt tcctgtttgt ctccggccac ttaaggga attataaagc tcagctcact          | 480  |
| [0045] agactcaatc acatcacaaa ctgcaacgga gctgtgttgt cagtggagga gctcctgatt | 540  |
| ggaggcgaga tgatcaaagc cggcaccctt acactggagg aggtgcggcg gaagttcaac        | 600  |
| aatggagaga tcaacttcgg tggcggtgga tccgataaaa agtattctat tggtttagcc        | 660  |
| atcggcacta attccgttg atgggctgtc ataaccgatg aatacaaagt accttcaaag         | 720  |
| aaatttaagg tgttggggaa cacagaccgt cattcgatta aaaagaatct tatcggtgcc        | 780  |
| ctcctattcg atagtggcga aacggcagag gcgactcgcc tgaaacgaac cgctcggaga        | 840  |
| aggtatacac gtcgcaagaa ccgaatatgt tacttacaag aaatttttag caatgagatg        | 900  |
| gccaaagttg acgattcttt ctttcaccgt ttggaagagt ccttccttgt cgaagaggac        | 960  |
| aagaaacatg aacggcaccc catctttgga aacatagtag atgaggtggc atatcatgaa        | 1020 |
| aagtacccaa cgatttatca cctcagaaaa aagctagttg actcaactga taaagcggac        | 1080 |
| ctgagggttaa tctacttggc tcttgcccat atgataaagt tccgtgggca ctttctcatt       | 1140 |
| gagggtgatc taaatccgga caactcgat gtcgacaaac tgttcatcca gttagtacaa         | 1200 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
|        | acctataatc agttgtttga agagaaccct ataaatgcaa gtggcgtgga tgcgaaggct  | 1260 |
|        | attcttagcg cccgcctctc taaatcccga cggctagaaa acctgatcgc acaattaccc  | 1320 |
|        | ggagagaaga aaaatgggtt gttcggtaac cttatagcgc tctcactagg cctgacacca  | 1380 |
|        | aattttaagt cgaacttcga cttagctgaa gatgccaaat tgcagcttag taaggacacg  | 1440 |
|        | tacgatgacg atctcgacaa tctactggca caaattggag atcagtatgc ggacttattt  | 1500 |
|        | ttggctgcca aaaaccttag cgatgcaatc ctctatctg acatactgag agttaatact   | 1560 |
|        | gagattacca aggcgccgtt atccgcttca atgatcaaaa ggtacgatga acatcaccaa  | 1620 |
|        | gacttgacac ttctcaagc cctagtccgt cagcaactgc ctgagaaata taaggaaata   | 1680 |
|        | ttctttgatc agtcgaaaaa cgggtacgca gggttatattg acggcggagc gagtcaagag | 1740 |
|        | gaattctaca agtttatcaa acccatatta gagaagatgg atgggacgga agagttgctt  | 1800 |
|        | gtaaaactca atcggaaga tctactgca aagcagcgga ctttcgacaa cggtagcatt    | 1860 |
| [0046] | ccacatcaaa tccacttagg cgaattgcat gctatactta gaaggcagga ggatttttat  | 1920 |
|        | ccgttcctca aagacaatcg tgaaaagatt gagaaaatcc taacctttcg cataccttac  | 1980 |
|        | tatgtgggac ccctggcccg agggaaactct cggttcgcat ggatgacaag aaagtccgaa | 2040 |
|        | gaaacgatta ctccatggaa ttttgaggaa gttgtcgata aaggtgcgtc agctcaatcg  | 2100 |
|        | ttcatcgaga ggatgaccaa ctttgacaag aatttaccga acgaaaaagt attgcctaag  | 2160 |
|        | cacagtttac tttacgagta tttcacagtg tacaatgaac tcacgaaagt taagtatgtc  | 2220 |
|        | actgagggca tgcgtaaacc cgcctttcta agcggagaac agaagaaagc aatagtagat  | 2280 |
|        | ctgttattca agaccaaccg caaagtgaca gttaagcaat tgaaagagga ctactttaag  | 2340 |
|        | aaaattgaat gcttcgattc tgtcgagatc tccggggtag aagatcgatt taatgcgtca  | 2400 |
|        | cttggtacgt atcatgacct cctaaagata attaaagata aggacttcct ggataacgaa  | 2460 |
|        | gagaatgaag atatcttaga agatatagtg ttgactctta ccctctttga agatcgggaa  | 2520 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
|        | atgattgagg aaagactaaa aacatacgtc cacctgttcg acgataaggt tatgaaacag  | 2580 |
|        | ttaaagaggc gtcgtatac gggctgggga cgattgtcgc ggaaacttat caacgggata   | 2640 |
|        | agagacaagc aaagtggtaa aactattctc gattttctaa agagcgacgg cttcgccaat  | 2700 |
|        | aggaacttta tgcagctgat ccatgatgac tctttaacct tcaaagagga tatacaaaaag | 2760 |
|        | gcacaggttt ccggacaagg ggactcattg cacgaacata ttgcgaatct tgctggttcg  | 2820 |
|        | ccagccatca aaaaggcat actccagaca gtcaaagtag tggatgagct agttaaggtc   | 2880 |
|        | atgggacgtc acaaaccgga aaacattgta atcgagatgg cacgcgaaaa tcaaacgact  | 2940 |
|        | cagaaggggc aaaaaaacag tcgagagcgg atgaagagaa tagaagaggg tattaagaa   | 3000 |
|        | ctgggcagcc agatcttaaa ggagcatcct gtggaaaata cccaattgca gaacgagaaa  | 3060 |
|        | ctttacctct attacctaca aaatggaagg gacatgtatg ttgatcagga actggacata  | 3120 |
|        | aaccgtttat ctgattacga cgtcgatgcc attgtacccc aatccttttt gaaggacgat  | 3180 |
| [0047] | tcaatcgaca ataaagtgct tacacgctcg gataagaacc gagggaaaag tgacaatgtt  | 3240 |
|        | ccaagcgagg aagtcgtaaa gaaaatgaag aactattggc ggcagctcct aaatgcgaaa  | 3300 |
|        | ctgataacgc aaagaaagt cgaataacta actaaagctg agaggggttg cttgtctgaa   | 3360 |
|        | cttgacaagg ccgatttat taaacgtcag ctcgtggaaa cccgccaaat cacaagcat    | 3420 |
|        | gttgcacaga tactagattc ccgaatgaat acgaaatacg acgagaacga taagctgatt  | 3480 |
|        | cggggaagtca aagtaatcac tttaaagtca aaattggtgt cggacttcag aaaggathtt | 3540 |
|        | caattctata aagttaggga gataaataac taccacatg cgcacgacgc ttatcttaat   | 3600 |
|        | gccgtcgtag ggaccgcact cattaagaaa taccgaagc tagaaagtga gtttgtgtat   | 3660 |
|        | ggtgattaca aagtttatga cgtccgtaag atgatcgca aaagcgaaca ggagatagcg   | 3720 |
|        | aaggctacag ccaataactt cttttattct aacattatga atttctttaa gacggaaatc  | 3780 |
|        | actctgcaa acggagagat acgcaaacga ctttaattg aaaccaatgg ggagacaggt    | 3840 |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
|        | gaaatcgtat gggataaggg ccgggacttc gcgacggtga gaaaagtttt gtccatgccc | 3900 |
|        | caagtcaaca tagtaaagaa aactgaggtg cagaccggag ggttttcaaa ggaatcgatt | 3960 |
|        | cttccaaaaa ggaatagtga taagtcctc gctcgtaaaa aggactggga cccgaaaaag  | 4020 |
|        | tacggtggct tcgatagccc tacagttgcc tattctgtcc tagtagtggc aaaagttgag | 4080 |
|        | aagggaat ccaagaaact gaagtcagtc aaagaattat tggggataac gattatggag   | 4140 |
|        | cgctcgtctt ttgaaaagaa ccccatcgac ttccttgagg cgaaaggta caaggaagta  | 4200 |
|        | aaaaaggatc tcataattaa actaccaaag tatagtctgt ttgagttaga aaatggccga | 4260 |
|        | aaacggatgt tggctagcgc cggagagctt caaaaggga acgaactgc actaccgtct   | 4320 |
|        | aaatacgtga atttctgta tttagcgtcc cattacgaga agttgaaagg ttcacctgaa  | 4380 |
|        | gataacgaac agaagcaact tttgttgag cagcaciaaac attatctcga cgaaatcata | 4440 |
|        | gagcaaat ttcggaattcag taagagagtc atcctagctg atgccaatct ggacaaagta | 4500 |
| [0048] | ttaagcgc atacaagca cagggataaa cccatagctg agcaggcgga aaatattatc    | 4560 |
|        | catttgctta ctcttaccac cctcggcgct ccagccgcat tcaagtattt tgacacaacg | 4620 |
|        | atagatcgca aacgatacac ttctaccaag gaggtgctag acgcgacact gattcaccaa | 4680 |
|        | tccatcacgg gatttatga aactcggata gatttgtcac agcttggggg tgacggatcc  | 4740 |
|        | cccaagaaga agaggaaagt ctcgagcgac tacaagacc atgacggtga ttataaagat  | 4800 |
|        | catgacatcg attacaagga tgacgatgac aagtga                           | 4836 |

<210> 28

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|        |                                                                                   |    |
|--------|-----------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 28<br/>gagctggacg gcgacgtaaa cgg</p>                               | 23 |
|        | <p>&lt;210&gt; 29<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                             |    |
|        | <p>&lt;400&gt; 29<br/>cgccggacac gctgaacttg tgg</p>                               | 23 |
|        | <p>&lt;210&gt; 30<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0049] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                             |    |
|        | <p>&lt;400&gt; 30<br/>ccggcaagct gcccggtgcc tgg</p>                               | 23 |
|        | <p>&lt;210&gt; 31<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                             |    |
|        | <p>&lt;400&gt; 31<br/>ggtcaggggtg gtcacgaggg tgg</p>                              | 23 |

|        |                           |    |
|--------|---------------------------|----|
|        | <210> 32                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 32                  |    |
|        | cgaggagctg ttcaccgggg tgg | 23 |
|        | <210> 33                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0050] | <400> 33                  |    |
|        | ccgtccagct cgaccaggat ggg | 23 |
|        | <210> 34                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 34                  |    |
|        | cgaggagctg ttcaccgggg tgg | 23 |
|        | <210> 35                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |

|        |                           |    |  |
|--------|---------------------------|----|--|
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 35                  |    |  |
|        | gccgtccagc tcgaccagga tgg | 23 |  |
|        | <210> 36                  |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 36                  |    |  |
|        | ccggcaagct gcccgtgccc tgg | 23 |  |
| [0051] | <210> 37                  |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 37                  |    |  |
|        | gtaggtcagg gtggtcacga ggg | 23 |  |
|        | <210> 38                  |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 38                  |    |  |

|        |                           |    |
|--------|---------------------------|----|
|        | gggcgaggag ctgttcaccg ggg | 23 |
|        | <210> 39                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 39                  |    |
|        | ccgtccagct cgaccaggat ggg | 23 |
|        | <210> 40                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
| [0052] | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 40                  |    |
|        | agggcgagga gctgttcacc ggg | 23 |
|        | <210> 41                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 41                  |    |
|        | ccgtccagct cgaccaggat ggg | 23 |
|        | <210> 42                  |    |
|        | <211> 23                  |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 42                  |    |
|        | aagggcgagg agctgttcac cgg | 23 |
|        | <210> 43                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 43                  |    |
| [0053] | ccgtccagct cgaccaggat ggg | 23 |
|        | <210> 44                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 44                  |    |
|        | aagggcgagg agctgttcac cgg | 23 |
|        | <210> 45                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的         |    |

|        |                           |    |
|--------|---------------------------|----|
|        | 引物                        |    |
|        | <400> 45                  |    |
|        | gccgtccagc tcgaccagga tgg | 23 |
|        | <210> 46                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 46                  |    |
|        | cagcgtgtcc ggcgagggcg agg | 23 |
|        | <210> 47                  |    |
|        | <211> 23                  |    |
| [0054] | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 47                  |    |
|        | cttcagggtc agcttgccgt agg | 23 |
|        | <210> 48                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 48                  |    |
|        | gggcgaggag ctgttcaccg ggg | 23 |

|        |                           |    |
|--------|---------------------------|----|
|        | <210> 49                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 49                  |    |
|        | cgtcgccgtc cagctcgacc agg | 23 |
|        | <210> 50                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0055] | <400> 50                  |    |
|        | agggcgagga gctgttcacc ggg | 23 |
|        | <210> 51                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 51                  |    |
|        | cgtcgccgtc cagctcgacc agg | 23 |
|        | <210> 52                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 52                  |    |
|        | aagggcgagg agctgttcac cgg | 23 |
|        | <210> 53                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 53                  |    |
|        | cgtcgccgtc cagctcgacc agg | 23 |
| [0056] | <210> 54                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 54                  |    |
|        | ctgaagtta tctgcaccac cgg  | 23 |
|        | <210> 55                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |

|        |                                                                                   |    |
|--------|-----------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 55<br/>gtggtcacga gggtagggcca ggg</p>                              | 23 |
|        | <p>&lt;210&gt; 56<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                             |    |
|        | <p>&lt;400&gt; 56<br/>aagttcagcg tgtccggcga ggg</p>                               | 23 |
|        | <p>&lt;210&gt; 57<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0057] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                             |    |
|        | <p>&lt;400&gt; 57<br/>cttcagggtc agcttgccgt agg</p>                               | 23 |
|        | <p>&lt;210&gt; 58<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                             |    |
|        | <p>&lt;400&gt; 58<br/>caagttcagc gtgtccggcg agg</p>                               | 23 |
|        | <p>&lt;210&gt; 59</p>                                                             |    |

|        |                            |    |
|--------|----------------------------|----|
|        | <211> 23                   |    |
|        | <212> DNA                  |    |
|        | <213> 人工序列                 |    |
|        | <220>                      |    |
|        | <223> 人工序列的说明：合成的<br>引物    |    |
|        | <400> 59                   |    |
|        | cttcagggtc agcttgccgt agg  | 23 |
|        | <210> 60                   |    |
|        | <211> 23                   |    |
|        | <212> DNA                  |    |
|        | <213> 人工序列                 |    |
|        | <220>                      |    |
|        | <223> 人工序列的说明：合成的<br>引物    |    |
| [0058] | <400> 60                   |    |
|        | ggtgcccattc ctggtcgagc tgg | 23 |
|        | <210> 61                   |    |
|        | <211> 23                   |    |
|        | <212> DNA                  |    |
|        | <213> 人工序列                 |    |
|        | <220>                      |    |
|        | <223> 人工序列的说明：合成的<br>引物    |    |
|        | <400> 61                   |    |
|        | cgccggacac gctgaacttg tgg  | 23 |
|        | <210> 62                   |    |
|        | <211> 23                   |    |
|        | <212> DNA                  |    |
|        | <213> 人工序列                 |    |
|        | <220>                      |    |

|        |                                                                                                                                                                                             |    |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 62<br/>gagctggacg gcgacgtaaa cgg</p>                                                                                                   | 23 |
|        | <p>&lt;210&gt; 63<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 63<br/>ggcatcgccc tcgccctcgc cgg</p> | 23 |
| [0059] | <p>&lt;210&gt; 64<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 64<br/>ggagcgcacc atcttcttca agg</p> | 23 |
|        | <p>&lt;210&gt; 65<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 65<br/>ctcgaacttc acctcggcgc ggg</p> | 23 |

|        |                                                                                                                                                                                                       |    |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 66</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 66</p> <p>ctgaagttca tctgcaccac cgg</p> | 23 |
|        | <p>&lt;210&gt; 67</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 67</p> <p>gtcagggtgg tcacgagggt ggg</p> | 23 |
| [0060] | <p>&lt;210&gt; 68</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 68</p> <p>ggccacaagt tcagcgtgtc cgg</p> | 23 |
|        | <p>&lt;210&gt; 69</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p>                                                                                                                                    |    |

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 69

cttcagggtc agcttgccgt agg

23

<210> 70

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 70

ctcgtgacca ccctgacctg cgg

23

[0061]

<210> 71

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 71

cgtgctgctt catgtggtcg ggg

23

<210> 72

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|  |                                                                                          |    |
|--|------------------------------------------------------------------------------------------|----|
|  | <p>&lt;400&gt; 72<br/>ccggcaagct gcccggtgcc tgg</p>                                      | 23 |
|  | <p>&lt;210&gt; 73<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p>        |    |
|  | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                                    |    |
|  | <p>&lt;400&gt; 73<br/>gctgaagcac tgcacgccgt agg</p>                                      | 23 |
|  | <p>&lt;210&gt; 74<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>[0062] &lt;213&gt; 人工序列</p> |    |
|  | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                                    |    |
|  | <p>&lt;400&gt; 74<br/>agcgtgtccg gcgagggcga ggg</p>                                      | 23 |
|  | <p>&lt;210&gt; 75<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p>        |    |
|  | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                                    |    |
|  | <p>&lt;400&gt; 75<br/>ggtggtgcag atgaacttca ggg</p>                                      | 23 |

|        |                                                                                                                                                                                                       |    |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 76</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 76</p> <p>cagcgtgtcc ggcgagggcg agg</p> | 23 |
|        | <p>&lt;210&gt; 77</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p>                                                        |    |
| [0063] | <p>&lt;400&gt; 77</p> <p>ggtggtgcag atgaacttca ggg</p>                                                                                                                                                | 23 |
|        | <p>&lt;210&gt; 78</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 78</p> <p>caccggggtg gtgcccatcc tgg</p> | 23 |
|        | <p>&lt;210&gt; 79</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p>                                                                                                            |    |

|        |                           |    |  |
|--------|---------------------------|----|--|
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 79                  |    |  |
|        | cgccggacac gctgaacttg tgg | 23 |  |
|        | <210> 80                  |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 80                  |    |  |
|        | ggcgagggcg atgccaccta cgg | 23 |  |
| [0064] | <210> 81                  |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 81                  |    |  |
|        | gggcacgggc agcttgccgg tgg | 23 |  |
|        | <210> 82                  |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 82                  |    |  |

|        |                           |    |
|--------|---------------------------|----|
|        | ctcgtgacca ccctgacctg cgg | 23 |
|        | <210> 83                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 83                  |    |
|        | agaagtcgtg ctgcttcacg tgg | 23 |
|        | <210> 84                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
| [0065] | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 84                  |    |
|        | caactacaag acccgcgccg agg | 23 |
|        | <210> 85                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 85                  |    |
|        | cgatgccctt cagctcgatg cgg | 23 |
|        | <210> 86                  |    |
|        | <211> 23                  |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 86                  |    |
|        | cccatcctgg tcgagctgga cgg | 23 |
|        | <210> 87                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 87                  |    |
| [0066] | ggcatcgccc tcgccctcgc cgg | 23 |
|        | <210> 88                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 88                  |    |
|        | caagttcagc gtgtccggcg agg | 23 |
|        | <210> 89                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的         |    |

## 引物

|  |                                                 |    |
|--|-------------------------------------------------|----|
|  | <400> 89<br>ggtggtgcag atgaacttca ggg           | 23 |
|  | <210> 90<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|  | <220><br><223> 人工序列的说明：合成的<br>引物                |    |
|  | <400> 90<br>gagctggacg gcgacgtaaa cgg           | 23 |
|  | <210> 91<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|  | <220><br><223> 人工序列的说明：合成的<br>引物                |    |
|  | <400> 91<br>gaccaggatg ggcaccaccc cgg           | 23 |
|  | <210> 92<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|  | <220><br><223> 人工序列的说明：合成的<br>引物                |    |
|  | <400> 92<br>accatcttct tcaaggacga cgg           | 23 |

[0067]

|        |                           |    |
|--------|---------------------------|----|
|        | <210> 93                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 93                  |    |
|        | cgctcctgga cgtagccttc ggg | 23 |
|        | <210> 94                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0068] | <400> 94                  |    |
|        | ggcgagggcg atgccaccta cgg | 23 |
|        | <210> 95                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 95                  |    |
|        | cgccggacac gctgaacttg tgg | 23 |
|        | <210> 96                  |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |

<220>

<223> 人工序列的说明：合成的  
引物

<400> 96

ttcaagtccg ccatgcccga agg

23

<210> 97

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 97

gtcgtgctgc ttcattgtgt cgg

23

[0069]

<210> 98

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 98

ttcaagtccg ccatgcccga agg

23

<210> 99

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|        |                                                                                    |    |
|--------|------------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 99<br/>tcgtgctgct tcatgtggc ggg</p>                                 | 23 |
|        | <p>&lt;210&gt; 100<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 100<br/>ggccacaagt tcagcgtgic cgg</p>                               | 23 |
|        | <p>&lt;210&gt; 101<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0070] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 101<br/>cgtcgccgic cagctcagc agg</p>                                | 23 |
|        | <p>&lt;210&gt; 102<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 102<br/>ctcgtgacca ccctgaccta cgg</p>                               | 23 |
|        | <p>&lt;210&gt; 103</p>                                                             |    |

|        |                                                                                                                                                                                                         |    |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 103</p> <p>ccagggcacg ggcagcttgc cgg</p>                        | 23 |
|        | <p>&lt;210&gt; 104</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 104</p> <p>cagccacaac gtctatatca tgg</p> | 23 |
| [0071] | <p>&lt;210&gt; 105</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 105</p> <p>tgtactccag cttgtgcccc agg</p> | 23 |
|        | <p>&lt;210&gt; 106</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p>                                                                                          |    |

|        |                                                                                                                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 106<br/>ggccacaagt tcagcgtgtc cgg</p>                                                                                                    | 23 |
|        | <p>&lt;210&gt; 107<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 107<br/>gccgtccagc tcgaccagga tgg</p> | 23 |
| [0072] | <p>&lt;210&gt; 108<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 108<br/>ggccacaagt tcagcgtgtc cgg</p> | 23 |
|        | <p>&lt;210&gt; 109<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 109<br/>ccgtccagct cgaccaggat ggg</p> | 23 |

|        |                                                                                                                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 110<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 110<br/>aagttcagcg tgtccggcga ggg</p> | 23 |
| [0073] | <p>&lt;210&gt; 111<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 111<br/>cgtcgccgtc cagctcgacc agg</p> | 23 |
|        | <p>&lt;210&gt; 112<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 112<br/>cgcgccgagg tgaagttcga ggg</p> | 23 |
|        | <p>&lt;210&gt; 113<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA</p>                                                                                                                                 |    |

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 113

gccgtcgtcc ttgaagaaga tgg

23

<210> 114

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 114

ccggcaagct gcccggtccc tgg

23

[0074]

<210> 115

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 115

cttcagggtc agcttgccgt agg

23

<210> 116

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|        |                                                                                    |    |
|--------|------------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 116<br/>caagttcagc gtgtccggcg agg</p>                               | 23 |
|        | <p>&lt;210&gt; 117<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 117<br/>gccgtccagc tcgaccagga tgg</p>                               | 23 |
|        | <p>&lt;210&gt; 118<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0075] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 118<br/>caagttcagc gtgtccggcg agg</p>                               | 23 |
|        | <p>&lt;210&gt; 119<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 119<br/>ccgtccagct cgaccaggat ggg</p>                               | 23 |

|        |                           |    |
|--------|---------------------------|----|
|        | <210> 120                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 120                 |    |
|        | aagttcagcg tgtccggcga ggg | 23 |
|        | <210> 121                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0076] | <400> 121                 |    |
|        | ccgtccagct cgaccaggat ggg | 23 |
|        | <210> 122                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 122                 |    |
|        | agcgtgtccg gcgagggcga ggg | 23 |
|        | <210> 123                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |

|        |                           |    |  |
|--------|---------------------------|----|--|
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 123                 |    |  |
|        | cgtcgccgtc cagctcgacc agg | 23 |  |
|        | <210> 124                 |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 124                 |    |  |
|        | ctgaagttca tctgcaccac cgg | 23 |  |
| [0077] | <210> 125                 |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 125                 |    |  |
|        | ggcatcgccc tcgccctcgc cgg | 23 |  |
|        | <210> 126                 |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 126                 |    |  |

|        |                           |    |
|--------|---------------------------|----|
|        | cagcgtgtcc ggcgagggcg agg | 23 |
|        | <210> 127                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 127                 |    |
|        | gccgtccagc tcgaccagga tgg | 23 |
|        | <210> 128                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
| [0078] | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 128                 |    |
|        | ggccacaagt tcagcgtgtc cgg | 23 |
|        | <210> 129                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 129                 |    |
|        | gaccaggatg ggcaccaccc cgg | 23 |
|        | <210> 130                 |    |
|        | <211> 23                  |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 130                 |    |
|        | agcgtgtccg gcgagggcga ggg | 23 |
|        | <210> 131                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 131                 |    |
| [0079] | ccgtccagct cgaccaggat ggg | 23 |
|        | <210> 132                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 132                 |    |
|        | caactacaag acccgcgccg agg | 23 |
|        | <210> 133                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的         |    |

## 引物

|        |                                                                                    |    |
|--------|------------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 133<br/>gcgctcctgg acgtagcctt cgg</p>                               | 23 |
|        | <p>&lt;210&gt; 134<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 134<br/>caactacaag acccgcgccg agg</p>                               | 23 |
| [0080] | <p>&lt;210&gt; 135<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 135<br/>cgctcctgga cgtagccttc ggg</p>                               | 23 |
|        | <p>&lt;210&gt; 136<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 136<br/>gctgaagggc atcgacttca agg</p>                               | 23 |

|        |                                                                                                                                                                                                         |    |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 137</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 137</p> <p>cctcgaactt cacctcggcg cgg</p> | 23 |
|        | <p>&lt;210&gt; 138</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p>                                                         |    |
| [0081] | <p>&lt;400&gt; 138</p> <p>caagttcagc gtgtccggcg agg</p>                                                                                                                                                 | 23 |
|        | <p>&lt;210&gt; 139</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 139</p> <p>gaccaggatg ggcaccaccc cgg</p> | 23 |
|        | <p>&lt;210&gt; 140</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p>                                                                                                             |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 140                 |    |
|        | aagttcagcg tgtccggcga ggg | 23 |
|        | <210> 141                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 141                 |    |
|        | gaccaggatg ggcaccaccc cgg | 23 |
| [0082] | <210> 142                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 142                 |    |
|        | gaaggcatc gacttcaagg agg  | 23 |
|        | <210> 143                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |

|        |                                                                                    |    |
|--------|------------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 143<br/>cctcgaactt cacctcggcg cgg</p>                               | 23 |
|        | <p>&lt;210&gt; 144<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 144<br/>caactacaag acccgcgccg agg</p>                               | 23 |
|        | <p>&lt;210&gt; 145<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0083] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 145<br/>ctggacgtag ccttcgggca tgg</p>                               | 23 |
|        | <p>&lt;210&gt; 146<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 146<br/>ggagcgcacc atcttcttca agg</p>                               | 23 |
|        | <p>&lt;210&gt; 147</p>                                                             |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 147                 |    |
|        | agaagtcgtg ctgcttcatg tgg | 23 |
|        | <210> 148                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0084] | <400> 148                 |    |
|        | caagttcagc gtgtccggcg agg | 23 |
|        | <210> 149                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 149                 |    |
|        | cgtcgccgtc cagctcgacc agg | 23 |
|        | <210> 150                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |

|        |                                                                                                                                                                                           |    |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 150<br/>catgcccgaa ggctacgtcc agg</p>                                                                                                     | 23 |
|        | <p>&lt;210&gt; 151<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 151<br/>agaagtcgtg ctgcttcattg tgg</p> | 23 |
| [0085] | <p>&lt;210&gt; 152<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 152<br/>caactacaag acccgcgccg agg</p>  | 23 |
|        | <p>&lt;210&gt; 153<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 153<br/>tgaagaagat ggtgcgctcc tgg</p>  | 23 |

|        |                                                                                                                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 154<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 154<br/>ggtgaaccgc atcgagctga agg</p> | 23 |
|        | <p>&lt;210&gt; 155<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                                                      |    |
| [0086] | <p>&lt;400&gt; 155<br/>cctcgaactt cacctcggcg cgg</p>                                                                                                                                          | 23 |
|        | <p>&lt;210&gt; 156<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 156<br/>ttcaagtccg ccatgcccga agg</p> | 23 |
|        | <p>&lt;210&gt; 157<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA</p>                                                                                                                                 |    |

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 157

gcttcatgtg gtcggggtag cgg

23

<210> 158

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 158

ccgcgccgag gtgaagttcg agg

23

[0087]

<210> 159

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 159

gccgtcgtcc ttgaagaaga tgg

23

<210> 160

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|        |                                                                                    |    |
|--------|------------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 160<br/>ggtgaaccgc atcgagctga agg</p>                               | 23 |
|        | <p>&lt;210&gt; 161<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 161<br/>ctcgaacttc acctcggcgc ggg</p>                               | 23 |
|        | <p>&lt;210&gt; 162<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0088] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 162<br/>gtgaaccgca tcgagctgaa ggg</p>                               | 23 |
|        | <p>&lt;210&gt; 163<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 163<br/>cctcgaactt cacctcggcg cgg</p>                               | 23 |

|        |                                                                                                                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 164<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 164<br/>caagatccgc cacaacatcg agg</p> | 23 |
|        | <p>&lt;210&gt; 165<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                                                      |    |
| [0089] | <p>&lt;400&gt; 165<br/>gatgccgttc ttctgcttgt cgg</p>                                                                                                                                          | 23 |
|        | <p>&lt;210&gt; 166<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 166<br/>gtgaaccgca tcgagctgaa ggg</p> | 23 |
|        | <p>&lt;210&gt; 167<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p>                                                                                                            |    |

|        |                           |    |  |
|--------|---------------------------|----|--|
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 167                 |    |  |
|        | ctcgaacttc acctcggcgc ggg | 23 |  |
|        | <210> 168                 |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 168                 |    |  |
|        | caaggaggac ggcaacatcc tgg | 23 |  |
| [0090] | <210> 169                 |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 169                 |    |  |
|        | tcagctcgat gcggttcacc agg | 23 |  |
|        | <210> 170                 |    |  |
|        | <211> 23                  |    |  |
|        | <212> DNA                 |    |  |
|        | <213> 人工序列                |    |  |
|        | <220>                     |    |  |
|        | <223> 人工序列的说明：合成的引物       |    |  |
|        | <400> 170                 |    |  |

|        |                           |    |
|--------|---------------------------|----|
|        | catgcccgaaggctacgtccagg   | 23 |
|        | <210> 171                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 171                 |    |
|        | gtcgtgctgc ttcattgtgt cgg | 23 |
|        | <210> 172                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
| [0091] | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 172                 |    |
|        | caaggaggac ggcaacatcc tgg | 23 |
|        | <210> 173                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 173                 |    |
|        | cagctcgatg cggttcacca ggg | 23 |
|        | <210> 174                 |    |
|        | <211> 23                  |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 174                 |    |
|        | aaggaggacg gcaacatcct ggg | 23 |
|        | <210> 175                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 175                 |    |
| [0092] | tcagctcgat gcggttcacc agg | 23 |
|        | <210> 176                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 176                 |    |
|        | aagttcagcg tgtccggcga ggg | 23 |
|        | <210> 177                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的         |    |

## 引物

|        |                                                  |    |
|--------|--------------------------------------------------|----|
|        | <400> 177<br>gccgtccagc tcgaccagga tgg           | 23 |
|        | <210> 178<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                 |    |
|        | <400> 178<br>catgcccgaaggctacgtcc agg            | 23 |
|        | <210> 179<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                 |    |
| [0093] | <400> 179<br>tcgtgctgct tcatgtggtc ggg           | 23 |
|        | <210> 180<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                 |    |
|        | <400> 180<br>aaggaggacg gcaacatcct ggg           | 23 |

|        |                           |    |
|--------|---------------------------|----|
|        | <210> 181                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 181                 |    |
|        | cagctcgatg cggttcacca ggg | 23 |
|        | <210> 182                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0094] | <400> 182                 |    |
|        | aggaggacgg caacatcctg ggg | 23 |
|        | <210> 183                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 183                 |    |
|        | tcagctcgat gcggttcacc agg | 23 |
|        | <210> 184                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |

|        |                           |    |
|--------|---------------------------|----|
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 184                 |    |
|        | atccgccaca acatcgagga cgg | 23 |
|        | <210> 185                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 185                 |    |
|        | gatgccgttc ttctgcttgt cgg | 23 |
| [0095] | <210> 186                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |
|        | <400> 186                 |    |
|        | cagcgtgtcc ggcgagggcg agg | 23 |
|        | <210> 187                 |    |
|        | <211> 23                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的引物       |    |

|        |                                                  |    |
|--------|--------------------------------------------------|----|
|        | <400> 187<br>cgtcgccgtc cagctcgacc agg           | 23 |
|        | <210> 188<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                 |    |
|        | <400> 188<br>ccggcaagct gcccgtgccc tgg           | 23 |
|        | <210> 189<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
| [0096] | <220><br><223> 人工序列的说明：合成的<br>引物                 |    |
|        | <400> 189<br>cagggtcagc ttgccgtagg tgg           | 23 |
|        | <210> 190<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                 |    |
|        | <400> 190<br>catgcccgaaggctacgtcc agg            | 23 |
|        | <210> 191                                        |    |

|        |                                                                                                                                                                                                         |    |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 191</p> <p>cgtgctgctt catgtggtcg ggg</p>                        | 23 |
|        | <p>&lt;210&gt; 192</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 192</p> <p>aggaggacgg caacatcctg ggg</p> | 23 |
| [0097] | <p>&lt;210&gt; 193</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 193</p> <p>cagctcgatg cggttcacca ggg</p> | 23 |
|        | <p>&lt;210&gt; 194</p> <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p>                                                                                          |    |

|        |                                                                                                                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 194<br/>caacatcctg gggcacaagc tgg</p>                                                                                                    | 23 |
|        | <p>&lt;210&gt; 195<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 195<br/>cgatgccctt cagctcgatg cgg</p> | 23 |
| [0098] | <p>&lt;210&gt; 196<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 196<br/>gctgaagggc atcgacttca agg</p> | 23 |
|        | <p>&lt;210&gt; 197<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 197<br/>gtcgccctcg aacttcacct cgg</p> | 23 |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
|        | <210> 198                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
|        | <400> 198                                                         |    |
|        | ccagaagtac gagcgccgcc cgg                                         | 23 |
|        | <210> 199                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
| [0099] | <400> 199                                                         |    |
|        | tggcaggtga gtgaggctgc agg                                         | 23 |
|        | <210> 200                                                         |    |
|        | <211> 62                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                  |    |
|        | <400> 200                                                         |    |
|        | ccgggcggcg ctcgtacttc tggccactgg gcgagcgtct ggcaggtgag tgaggctgca | 60 |
|        | gg                                                                | 62 |
|        | <210> 201                                                         |    |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
|        | <400> 201                                                         |    |
|        | gaatacccat ctgtcagctt cgg                                         | 23 |
|        | <210> 202                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
| [0100] | <400> 202                                                         |    |
|        | ggcggaacct gagaggcgta agg                                         | 23 |
|        | <210> 203                                                         |    |
|        | <211> 62                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                  |    |
|        | <400> 203                                                         |    |
|        | ccgaagctga cagatgggta ttctttgacg gggggtaggg gcggaacctg agaggcgtaa | 60 |
|        | gg                                                                | 62 |
|        | <210> 204                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |

|        |                                                                   |    |  |
|--------|-------------------------------------------------------------------|----|--|
|        | <220>                                                             |    |  |
|        | <223> 人工序列的说明：合成的引物                                               |    |  |
|        | <400> 204                                                         |    |  |
|        | aatattcaag cagcaggcac agg                                         | 23 |  |
|        | <210> 205                                                         |    |  |
|        | <211> 23                                                          |    |  |
|        | <212> DNA                                                         |    |  |
|        | <213> 人工序列                                                        |    |  |
|        | <220>                                                             |    |  |
|        | <223> 人工序列的说明：合成的引物                                               |    |  |
|        | <400> 205                                                         |    |  |
|        | ctcgcgcagg aggctgcagc ggg                                         | 23 |  |
| [0101] | <210> 206                                                         |    |  |
|        | <211> 62                                                          |    |  |
|        | <212> DNA                                                         |    |  |
|        | <213> 人工序列                                                        |    |  |
|        | <220>                                                             |    |  |
|        | <223> 人工序列的说明：合成的RFN靶位点寡核苷酸                                       |    |  |
|        | <400> 206                                                         |    |  |
|        | cctgtgcctg ctgcttgaat atttccgcct tttagggtgc tcgcgcagga ggctgcagcg | 60 |  |
|        | gg                                                                | 62 |  |
|        | <210> 207                                                         |    |  |
|        | <211> 23                                                          |    |  |
|        | <212> DNA                                                         |    |  |
|        | <213> 人工序列                                                        |    |  |
|        | <220>                                                             |    |  |
|        | <223> 人工序列的说明：合成的                                                 |    |  |

|        |                                                                               |    |
|--------|-------------------------------------------------------------------------------|----|
|        | 引物                                                                            |    |
|        | <400> 207<br>cccaaagcct ggccagggag tgg                                        | 23 |
|        | <210> 208<br><211> 23<br><212> DNA<br><213> 人工序列                              |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                                              |    |
|        | <400> 208<br>gccccacagg gcttgaagcc cgg                                        | 23 |
|        | <210> 209<br><211> 62<br><212> DNA<br><213> 人工序列                              |    |
| [0102] | <220><br><223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                     |    |
|        | <400> 209<br>ccactccctg gccaggcttt ggggaggcct ggagtcattg cccacaggg cttgaagccc | 60 |
|        | gg                                                                            | 62 |
|        | <210> 210<br><211> 23<br><212> DNA<br><213> 人工序列                              |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                                              |    |
|        | <400> 210                                                                     |    |

|                                                                   |    |
|-------------------------------------------------------------------|----|
| ccctacttcc gctttcacct tgg                                         | 23 |
| <210> 211                                                         |    |
| <211> 23                                                          |    |
| <212> DNA                                                         |    |
| <213> 人工序列                                                        |    |
| <220>                                                             |    |
| <223> 人工序列的说明：合成的<br>引物                                           |    |
| <400> 211                                                         |    |
| ggaatccctt ctgcagcacc tgg                                         | 23 |
| <210> 212                                                         |    |
| <211> 62                                                          |    |
| <212> DNA                                                         |    |
| <213> 人工序列                                                        |    |
| [0103]                                                            |    |
| <220>                                                             |    |
| <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                  |    |
| <400> 212                                                         |    |
| ccaaggtgaa agcggaagta gggccttcgc gcacctcatg gaatcccttc tgcagcacct | 60 |
| gg                                                                | 62 |
| <210> 213                                                         |    |
| <211> 23                                                          |    |
| <212> DNA                                                         |    |
| <213> 人工序列                                                        |    |
| <220>                                                             |    |
| <223> 人工序列的说明：合成的<br>引物                                           |    |
| <400> 213                                                         |    |
| cgctccagag ccgtgcgaat ggg                                         | 23 |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
|        | <210> 214                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
|        | <400> 214                                                         |    |
|        | tggaggcaag agggcggctt tgg                                         | 23 |
|        | <210> 215                                                         |    |
|        | <211> 62                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                  |    |
| [0104] | <400> 215                                                         |    |
|        | cccattcgca cggctctgga gcggcggctg cacaaccagt ggaggcaaga gggcggcttt | 60 |
|        | gg                                                                | 62 |
|        | <210> 216                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
|        | <400> 216                                                         |    |
|        | cgaggagact ggggactgta ggg                                         | 23 |
|        | <210> 217                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |

|        |                                                                    |    |
|--------|--------------------------------------------------------------------|----|
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |
|        | <223> 人工序列的说明：合成的<br>引物                                            |    |
|        | <400> 217                                                          |    |
|        | ccagctgctg ccttgccctcc agg                                         | 23 |
|        | <210> 218                                                          |    |
|        | <211> 62                                                           |    |
|        | <212> DNA                                                          |    |
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |
|        | <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                   |    |
|        | <400> 218                                                          |    |
|        | ccctacagtc cccagccctcc tcgtcccatg cctccgtctc cagctgctgc cttgcctcca | 60 |
| [0105] | gg                                                                 | 62 |
|        | <210> 219                                                          |    |
|        | <211> 23                                                           |    |
|        | <212> DNA                                                          |    |
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |
|        | <223> 人工序列的说明：合成的<br>引物                                            |    |
|        | <400> 219                                                          |    |
|        | catagctact gattggtggt ggg                                          | 23 |
|        | <210> 220                                                          |    |
|        | <211> 23                                                           |    |
|        | <212> DNA                                                          |    |
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |

<223> 人工序列的说明：合成的  
引物

<400> 220

ccggccccctc cccagtcagg ggg

23

<210> 221

<211> 62

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
RFN 靶位点寡核苷酸

<400> 221

cccaccacca atcagtagct atggcgagcc ctgctgtctc cggccccctcc ccagtcaggg

60

gg

62

[0106]

<210> 222

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 222

ggaaacgtct agatgctcaa cgg

23

<210> 223

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|        |                                                                               |    |
|--------|-------------------------------------------------------------------------------|----|
|        | <400> 223<br>caaaatgtcg ttcgtggcag tgg                                        | 23 |
|        | <210> 224<br><211> 62<br><212> DNA<br><213> 人工序列                              |    |
|        | <220><br><223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                     |    |
|        | <400> 224<br>ccgttgagca tctagacgtt tccttggtc ttctggcgcc aaaatgtcgt tcgtggcagg | 60 |
|        | gg                                                                            | 62 |
| [0107] | <210> 225<br><211> 23<br><212> DNA<br><213> 人工序列                              |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                                              |    |
|        | <400> 225<br>ctgttgctgg ccatgccaag cgg                                        | 23 |
|        | <210> 226<br><211> 23<br><212> DNA<br><213> 人工序列                              |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                                              |    |
|        | <400> 226<br>cctggggcg ggcacctcaa tgg                                         | 23 |

|        |                                                                    |    |
|--------|--------------------------------------------------------------------|----|
|        | <210> 227                                                          |    |
|        | <211> 62                                                           |    |
|        | <212> DNA                                                          |    |
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |
|        | <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                   |    |
|        | <400> 227                                                          |    |
|        | ccgcttggca tggccagcaa cagcagctcc tgcccagacac ctgggggcgg gcacctcaat | 60 |
|        | gg                                                                 | 62 |
|        | <210> 228                                                          |    |
|        | <211> 23                                                           |    |
|        | <212> DNA                                                          |    |
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |
| [0108] | <223> 人工序列的说明：合成的<br>引物                                            |    |
|        | <400> 228                                                          |    |
|        | ttcggagcga aaaccaagac agg                                          | 23 |
|        | <210> 229                                                          |    |
|        | <211> 23                                                           |    |
|        | <212> DNA                                                          |    |
|        | <213> 人工序列                                                         |    |
|        | <220>                                                              |    |
|        | <223> 人工序列的说明：合成的<br>引物                                            |    |
|        | <400> 229                                                          |    |
|        | gagtcccccg ccttcagaag agg                                          | 23 |
|        | <210> 230                                                          |    |
|        | <211> 62                                                           |    |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                  |    |
|        | <400> 230                                                         |    |
|        | cctgtcttgg ttttcgctcc gaaggtaaaa gaaatcattg agtccccgcg cttcagaaga | 60 |
|        | gg                                                                | 62 |
|        | <210> 231                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
| [0109] | <400> 231                                                         |    |
|        | ggcccggtcg actccgggcc cgg                                         | 23 |
|        | <210> 232                                                         |    |
|        | <211> 23                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |
|        | <220>                                                             |    |
|        | <223> 人工序列的说明：合成的<br>引物                                           |    |
|        | <400> 232                                                         |    |
|        | tgctgggaat cagcagtgtt tgg                                         | 23 |
|        | <210> 233                                                         |    |
|        | <211> 62                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人工序列                                                        |    |

|                                                                    |    |
|--------------------------------------------------------------------|----|
| <220>                                                              |    |
| <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                   |    |
| <400> 233                                                          |    |
| ccgggccccgg agtcgaccgg gccgaggcgg aggcgggcct gctgggaatc agcagtgttt | 60 |
| gg                                                                 | 62 |
| <210> 234                                                          |    |
| <211> 23                                                           |    |
| <212> DNA                                                          |    |
| <213> 人工序列                                                         |    |
| <220>                                                              |    |
| <223> 人工序列的说明：合成的<br>引物                                            |    |
| <400> 234                                                          |    |
| gggtgggggg agtttgctcc tgg                                          | 23 |
| [0110]                                                             |    |
| <210> 235                                                          |    |
| <211> 23                                                           |    |
| <212> DNA                                                          |    |
| <213> 人工序列                                                         |    |
| <220>                                                              |    |
| <223> 人工序列的说明：合成的<br>引物                                            |    |
| <400> 235                                                          |    |
| tccctcttta gccagagccg ggg                                          | 23 |
| <210> 236                                                          |    |
| <211> 62                                                           |    |
| <212> DNA                                                          |    |
| <213> 人工序列                                                         |    |
| <220>                                                              |    |
| <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                   |    |

|        |                                                                              |    |
|--------|------------------------------------------------------------------------------|----|
|        | <400> 236<br>ccaggagcaa actccccca cccctttcc aaagccatt ccctcttag ccagagccgg   | 60 |
|        | gg                                                                           | 62 |
|        | <210> 237<br><211> 23<br><212> DNA<br><213> 人工序列                             |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                                             |    |
|        | <400> 237<br>gccgccggcc ggggaggagg tgg                                       | 23 |
| [0111] | <210> 238<br><211> 23<br><212> DNA<br><213> 人工序列                             |    |
|        | <220><br><223> 人工序列的说明：合成的<br>引物                                             |    |
|        | <400> 238<br>ggcgagccgc gggcaggggc cgg                                       | 23 |
|        | <210> 239<br><211> 62<br><212> DNA<br><213> 人工序列                             |    |
|        | <220><br><223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                    |    |
|        | <400> 239<br>ccacctctc cccggccggc ggcggacagt ggacgcggcg gcgagccgc ggcaggggcc | 60 |
|        |                                                                              |    |
|        |                                                                              |    |
|        |                                                                              |    |
|        |                                                                              |    |
|        |                                                                              |    |

|                                                                  |    |
|------------------------------------------------------------------|----|
| gg                                                               | 62 |
| <210> 240                                                        |    |
| <211> 23                                                         |    |
| <212> DNA                                                        |    |
| <213> 人工序列                                                       |    |
| <220>                                                            |    |
| <223> 人工序列的说明：合成的<br>引物                                          |    |
| <400> 240                                                        |    |
| ccgtctgcac accccggctc tgg                                        | 23 |
| <210> 241                                                        |    |
| <211> 23                                                         |    |
| <212> DNA                                                        |    |
| <213> 人工序列                                                       |    |
| [0112]                                                           |    |
| <220>                                                            |    |
| <223> 人工序列的说明：合成的<br>引物                                          |    |
| <400> 241                                                        |    |
| ctcggccacc acaggaagc tgg                                         | 23 |
| <210> 242                                                        |    |
| <211> 62                                                         |    |
| <212> DNA                                                        |    |
| <213> 人工序列                                                       |    |
| <220>                                                            |    |
| <223> 人工序列的说明：合成的<br>RFN 靶位点寡核苷酸                                 |    |
| <400> 242                                                        |    |
| ccagagccgg ggtgtgcaga cggcagtcac tagggggcgc tcggccacca caggaagct | 60 |
| gg                                                               | 62 |

|        |                              |    |
|--------|------------------------------|----|
|        | <210> 243                    |    |
|        | <211> 20                     |    |
|        | <212> DNA                    |    |
|        | <213> 人工序列                   |    |
|        | <220>                        |    |
|        | <223> 人工序列的说明：合成的<br>引物      |    |
|        | <400> 243                    |    |
|        | ggctgtggga agccagcaac        | 20 |
|        | <210> 244                    |    |
|        | <211> 20                     |    |
|        | <212> DNA                    |    |
|        | <213> 人工序列                   |    |
|        | <220>                        |    |
|        | <223> 人工序列的说明：合成的<br>引物      |    |
| [0113] | <400> 244                    |    |
|        | aagccagggg ccaactggag        | 20 |
|        | <210> 245                    |    |
|        | <211> 26                     |    |
|        | <212> DNA                    |    |
|        | <213> 人工序列                   |    |
|        | <220>                        |    |
|        | <223> 人工序列的说明：合成的<br>引物      |    |
|        | <400> 245                    |    |
|        | gcgcgggaat tacagataaa ttaaaa | 26 |
|        | <210> 246                    |    |
|        | <211> 25                     |    |
|        | <212> DNA                    |    |
|        | <213> 人工序列                   |    |

|        |                             |    |
|--------|-----------------------------|----|
|        | <220>                       |    |
|        | <223> 人工序列的说明：合成的引物         |    |
|        | <400> 246                   |    |
|        | aggtcccatc ctctcataca tacca | 25 |
|        | <210> 247                   |    |
|        | <211> 18                    |    |
|        | <212> DNA                   |    |
|        | <213> 人工序列                  |    |
|        | <220>                       |    |
|        | <223> 人工序列的说明：合成的引物         |    |
|        | <400> 247                   |    |
|        | accgcccctt ggcaccac         | 18 |
| [0114] | <210> 248                   |    |
|        | <211> 21                    |    |
|        | <212> DNA                   |    |
|        | <213> 人工序列                  |    |
|        | <220>                       |    |
|        | <223> 人工序列的说明：合成的引物         |    |
|        | <400> 248                   |    |
|        | cggagetcat ctgcttctg t      | 21 |
|        | <210> 249                   |    |
|        | <211> 20                    |    |
|        | <212> DNA                   |    |
|        | <213> 人工序列                  |    |
|        | <220>                       |    |
|        | <223> 人工序列的说明：合成的引物         |    |

|        |                                                                                    |    |
|--------|------------------------------------------------------------------------------------|----|
|        | <p>&lt;400&gt; 249<br/>ggagcagctg gtcagagggg</p>                                   | 20 |
|        | <p>&lt;210&gt; 250<br/>&lt;211&gt; 20<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 250<br/>gggaaggggg acactgggga</p>                                   | 20 |
|        | <p>&lt;210&gt; 251<br/>&lt;211&gt; 23<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| [0115] | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 251<br/>gccctacatc tgctctccct cca</p>                               | 23 |
|        | <p>&lt;210&gt; 252<br/>&lt;211&gt; 20<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
|        | <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p>                              |    |
|        | <p>&lt;400&gt; 252<br/>gggccgggaa agagttgctg</p>                                   | 20 |
|        | <p>&lt;210&gt; 253</p>                                                             |    |

|        |                                                                                                                                                                                                     |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;211&gt; 23</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 253</p> <p>gccctacatc tgetctccct cca</p>                    | 23 |
|        | <p>&lt;210&gt; 254</p> <p>&lt;211&gt; 20</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 254</p> <p>gggccgggaa agagttgctg</p> | 20 |
| [0116] | <p>&lt;210&gt; 255</p> <p>&lt;211&gt; 20</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 255</p> <p>ggggaggag gctccaggtt</p>  | 20 |
|        | <p>&lt;210&gt; 256</p> <p>&lt;211&gt; 20</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p>                                                                                      |    |

|        |                                                                                                                                                                                                    |    |
|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 256<br/>ggcacaatgg ctccaagca</p>                                                                                                              | 20 |
|        | <p>&lt;210&gt; 257<br/>&lt;211&gt; 21<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 257<br/>ccttaccctt cccctcactc a</p>        | 21 |
| [0117] | <p>&lt;210&gt; 258<br/>&lt;211&gt; 20<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 258<br/>agaagggcgg gccagacagt</p>          | 20 |
|        | <p>&lt;210&gt; 259<br/>&lt;211&gt; 28<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 259<br/>atatccttct aggtagcggg cagtagcc</p> | 28 |

|        |                                                                                                                                                                                             |    |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;210&gt; 260<br/>&lt;211&gt; 20<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 260<br/>tctcggggga gagcggtaaa</p>   | 20 |
| [0118] | <p>&lt;210&gt; 261<br/>&lt;211&gt; 21<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 261<br/>cccaggaaaa gtgccagctc a</p> | 21 |
|        | <p>&lt;210&gt; 262<br/>&lt;211&gt; 20<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的说明：合成的<br/>引物</p> <p>&lt;400&gt; 262<br/>tgatggtcac cccaactgga</p>   | 20 |
|        | <p>&lt;210&gt; 263<br/>&lt;211&gt; 19<br/>&lt;212&gt; DNA</p>                                                                                                                               |    |

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 263

aaggcgcgcg tggttttt

19

<210> 264

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 264

ccagcacaac ttactcgcac ttga

24

[0119]

<210> 265

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

<400> 265

gggatgcagg gacggtcaag

20

<210> 266

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明：合成的  
引物

|  |                                                         |    |
|--|---------------------------------------------------------|----|
|  | <400> 266<br>gccgccccat ccctagagaa a                    | 21 |
|  | <210> 267<br><211> 20<br><212> DNA<br><213> 人工序列        |    |
|  | <220><br><223> 人工序列的说明：合成的<br>引物                        |    |
|  | <400> 267<br>tccagatggc acattgtcag                      | 20 |
|  | <210> 268<br><211> 20<br><212> DNA<br>[0120] <213> 人工序列 |    |
|  | <220><br><223> 人工序列的说明：合成的<br>引物                        |    |
|  | <400> 268<br>agggagcagg aaagtgaggt                      | 20 |
|  | <210> 269<br><211> 22<br><212> DNA<br><213> 人工序列        |    |
|  | <220><br><223> 人工序列的说明：合成的<br>引物                        |    |
|  | <400> 269<br>agagaagtcg aggaagagag ag                   | 22 |

|        |                           |    |
|--------|---------------------------|----|
|        | <210> 270                 |    |
|        | <211> 22                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 270                 |    |
|        | cagcagaaag ttcattggttt cg | 22 |
|        | <210> 271                 |    |
|        | <211> 20                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
| [0121] | <400> 271                 |    |
|        | tccagatggc acattgtcag     | 20 |
|        | <210> 272                 |    |
|        | <211> 20                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |
|        | <220>                     |    |
|        | <223> 人工序列的说明：合成的<br>引物   |    |
|        | <400> 272                 |    |
|        | agggagcagg aaagtgaggt     | 20 |
|        | <210> 273                 |    |
|        | <211> 20                  |    |
|        | <212> DNA                 |    |
|        | <213> 人工序列                |    |

|        |                       |    |  |
|--------|-----------------------|----|--|
|        | <220>                 |    |  |
|        | <223> 人工序列的说明：合成的引物   |    |  |
|        | <400> 273             |    |  |
|        | cgatggctcc caagaaacgc | 20 |  |
|        | <210> 274             |    |  |
|        | <211> 20              |    |  |
|        | <212> DNA             |    |  |
|        | <213> 人工序列            |    |  |
|        | <220>                 |    |  |
|        | <223> 人工序列的说明：合成的引物   |    |  |
|        | <400> 274             |    |  |
|        | gcaggtagaa tgcacagccg | 20 |  |
| [0122] | <210> 275             |    |  |
|        | <211> 20              |    |  |
|        | <212> DNA             |    |  |
|        | <213> 人工序列            |    |  |
|        | <220>                 |    |  |
|        | <223> 人工序列的说明：合成的引物   |    |  |
|        | <400> 275             |    |  |
|        | gcccagagtc aaggaacacg | 20 |  |
|        | <210> 276             |    |  |
|        | <211> 20              |    |  |
|        | <212> DNA             |    |  |
|        | <213> 人工序列            |    |  |
|        | <220>                 |    |  |
|        | <223> 人工序列的说明：合成的引物   |    |  |
|        | <400> 276             |    |  |

|        |                         |    |
|--------|-------------------------|----|
|        | aggtagtgtct tgagaccgcc  | 20 |
|        | <210> 277               |    |
|        | <211> 20                |    |
|        | <212> DNA               |    |
|        | <213> 人工序列              |    |
|        | <220>                   |    |
|        | <223> 人工序列的说明：合成的<br>引物 |    |
|        | <400> 277               |    |
|        | catccatcgg cgctttggtc   | 20 |
|        | <210> 278               |    |
|        | <211> 20                |    |
|        | <212> DNA               |    |
|        | <213> 人工序列              |    |
| [0123] | <220>                   |    |
|        | <223> 人工序列的说明：合成的<br>引物 |    |
|        | <400> 278               |    |
|        | ccgggaaaga gttgctgcac   | 20 |
|        | <210> 279               |    |
|        | <211> 20                |    |
|        | <212> DNA               |    |
|        | <213> 人工序列              |    |
|        | <220>                   |    |
|        | <223> 人工序列的说明：合成的<br>引物 |    |
|        | <400> 279               |    |
|        | ctccccgtct gcagtccatc   | 20 |
|        | <210> 280               |    |
|        | <211> 20                |    |

|        |                                                                                                                                                                                                     |    |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
|        | <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 280</p> <p>cctgcagga catgtggtga</p>                                               | 20 |
|        | <p>&lt;210&gt; 281</p> <p>&lt;211&gt; 20</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 281</p> <p>tagggctaga ggggtgaggc</p> | 20 |
| [0124] | <p>&lt;210&gt; 282</p> <p>&lt;211&gt; 20</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的引物</p> <p>&lt;400&gt; 282</p> <p>ccgaggtgaa acaagctgcc</p> | 20 |
|        | <p>&lt;210&gt; 283</p> <p>&lt;211&gt; 20</p> <p>&lt;212&gt; DNA</p> <p>&lt;213&gt; 人工序列</p> <p>&lt;220&gt;</p> <p>&lt;223&gt; 人工序列的说明：合成的</p>                                                       |    |

|        |                                                                   |     |
|--------|-------------------------------------------------------------------|-----|
|        | 引物                                                                |     |
|        | <400> 283                                                         |     |
|        | atgagggtc cagatggcac                                              | 20  |
|        | <210> 284                                                         |     |
|        | <211> 20                                                          |     |
|        | <212> DNA                                                         |     |
|        | <213> 人工序列                                                        |     |
|        | <220>                                                             |     |
|        | <223> 人工序列的说明: 合成的<br>引物                                          |     |
|        | <400> 284                                                         |     |
|        | ttcaccagc ttccctgtgg                                              | 20  |
|        | <210> 285                                                         |     |
|        | <211> 1752                                                        |     |
|        | <212> DNA                                                         |     |
| [0125] | <213> 海床游动细菌                                                      |     |
|        | <400> 285                                                         |     |
|        | atgtttttga gtatggtttc taaaataaga actttcggtt gggttcaaaa tccagtaaa  | 60  |
|        | tttgagaatt taaaacgagt agttcaagta ttgatagaa attctaaagt acataatgaa  | 120 |
|        | gtgaaaaata taaagatacc aaccctagtc aaagaaagta agatccaaaa agaactagtt | 180 |
|        | gctattatga atcaacatga ttgtatttat acatataaag agttagtagg aacaggaact | 240 |
|        | tcaatacggt cagaagcacc atgcgatgca attattcaag caacaatagc agatcaagga | 300 |
|        | aataaaaaag gctatatcga taattggica tctgacggtt ttttgcgttg ggcacatgct | 360 |
|        | ttaggattta ttgaatatat aaataaaagt gattcttttg taataactga tgttggactt | 420 |
|        | gcttactcta aatcagctga cggcagcgcc attgaaaaag agattttgat tgaagcgata | 480 |
|        | tcattctatc ctccagcgat tcgtatttta actttgctag aagatggaca acatttgaca | 540 |
|        | aagtttgatc ttggcaagaa tttaggtttt agtggagaaa gtggatttac ttctctaccg | 600 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
|        | gaaggaattc ttttagatac tctagctaata gctatgccta aagataaagg cgaaattcgt | 660  |
|        | aataattggg aaggatcttc agataagtac gcaagaatga taggtggttg gctggataaa  | 720  |
|        | ctaggattag taaagcaagg aaaaaaagaa tttatcattc ctactttggg taagccggac  | 780  |
|        | aataaagagt ttatatccca cgctttttaa attactggag aaggtttgaa agtactgcgt  | 840  |
|        | cgagcaaaag gctctacaaa atttacacgt gtacctaaaa gagtatatg ggaaatgctt   | 900  |
|        | gctacaaacc taaccgataa agagtatgta agaacaagaa gagctttgat ttagaaata   | 960  |
|        | ttaatcaaag ctggatcatt aaaaatagaa caaatacaag acaacttgaa gaaattagga  | 1020 |
|        | tttgatgaag ttatagaaac tattgaaaat gatatcaaag gcttaattaa cacaggtata  | 1080 |
|        | tttatagaaa tcaaagggcg attttatcaa ttgaaagacc atattcttca atttgtaata  | 1140 |
|        | cctaactcgtg gtgtgactaa gcaactagtc aaaagtgaac tggaggagaa gaaatctgaa | 1200 |
| [0126] | cttcgtcata aattgaaata tgtgcctcat gaatatattg aattaattga aattgccaga  | 1260 |
|        | aattccactc aggatagaat tcttgaaatg aaggtaatgg aattttttat gaaagtttat  | 1320 |
|        | ggatatagag gtaaacattt ggggtgatca aggaaaccgg acggagcaat ttatactgtc  | 1380 |
|        | ggatctccta ttgattacgg tgtgatcgtg gatactaaag cttatagcgg aggttataat  | 1440 |
|        | ctgccaatg gccaaagcaga tgaaatgcaa cgatatgtcg aagaaaatca aacacgaaac  | 1500 |
|        | aaacatatca accctaata atggtggaaa gtctatccat cttctgtaac ggaatttaag   | 1560 |
|        | tttttatattg tgagtggta ctttaaagga aactacaaag ctcagcttac acgattaaat  | 1620 |
|        | catatcacta attgtaatgg agctgttctt agtgtagaag agcttttaat tgggtggagaa | 1680 |
|        | atgattaaag ccggcacatt aaccttagag gaagtgaagac ggaaatttaa taacggcgag | 1740 |
|        | ataaactttt aa                                                      | 1752 |

## CRISPR/Cas9 - RGN

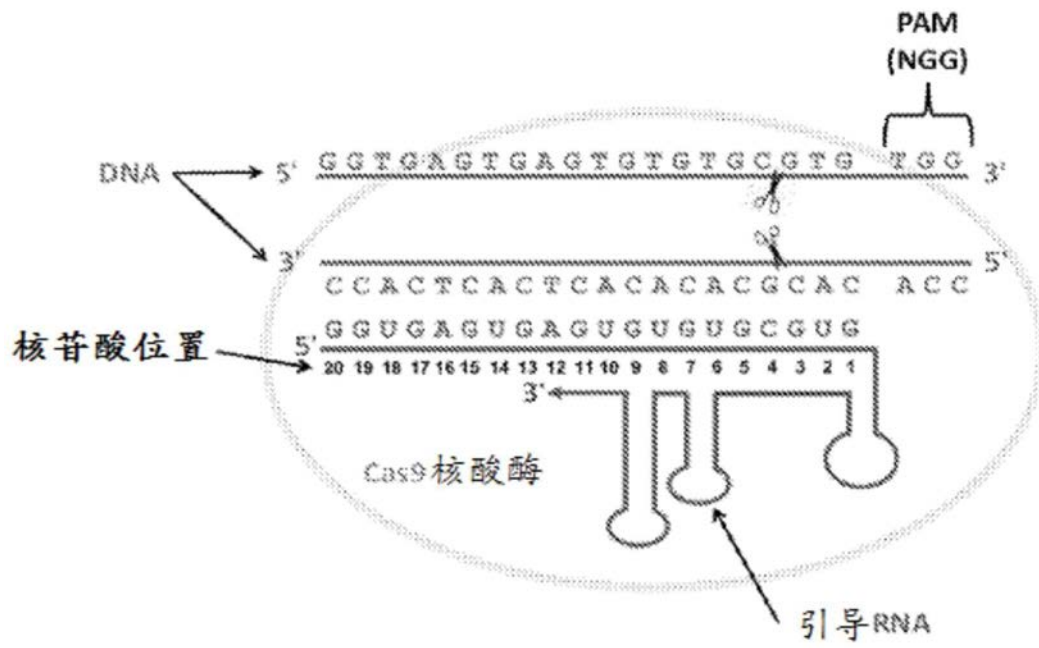


图1

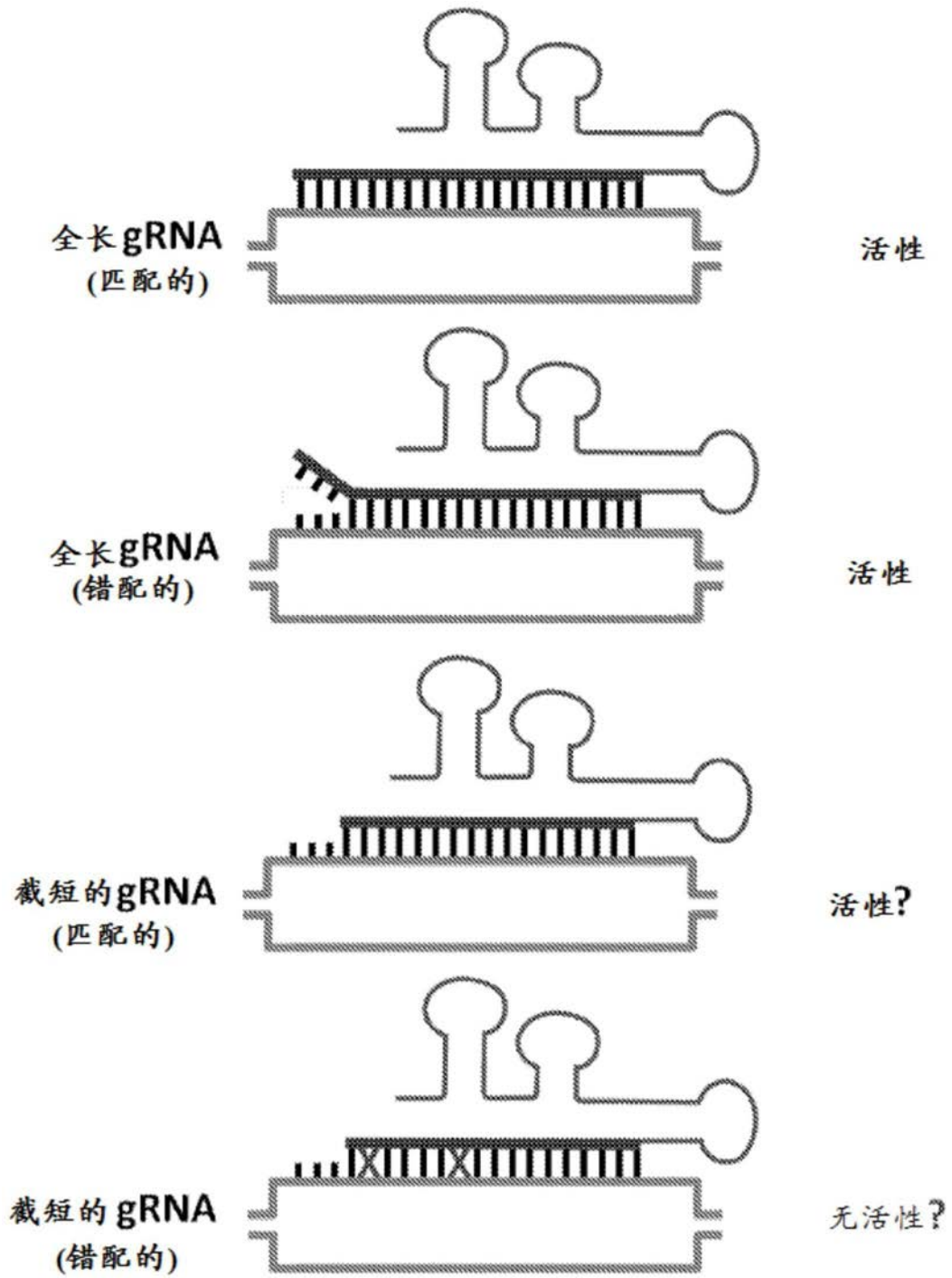


图2A

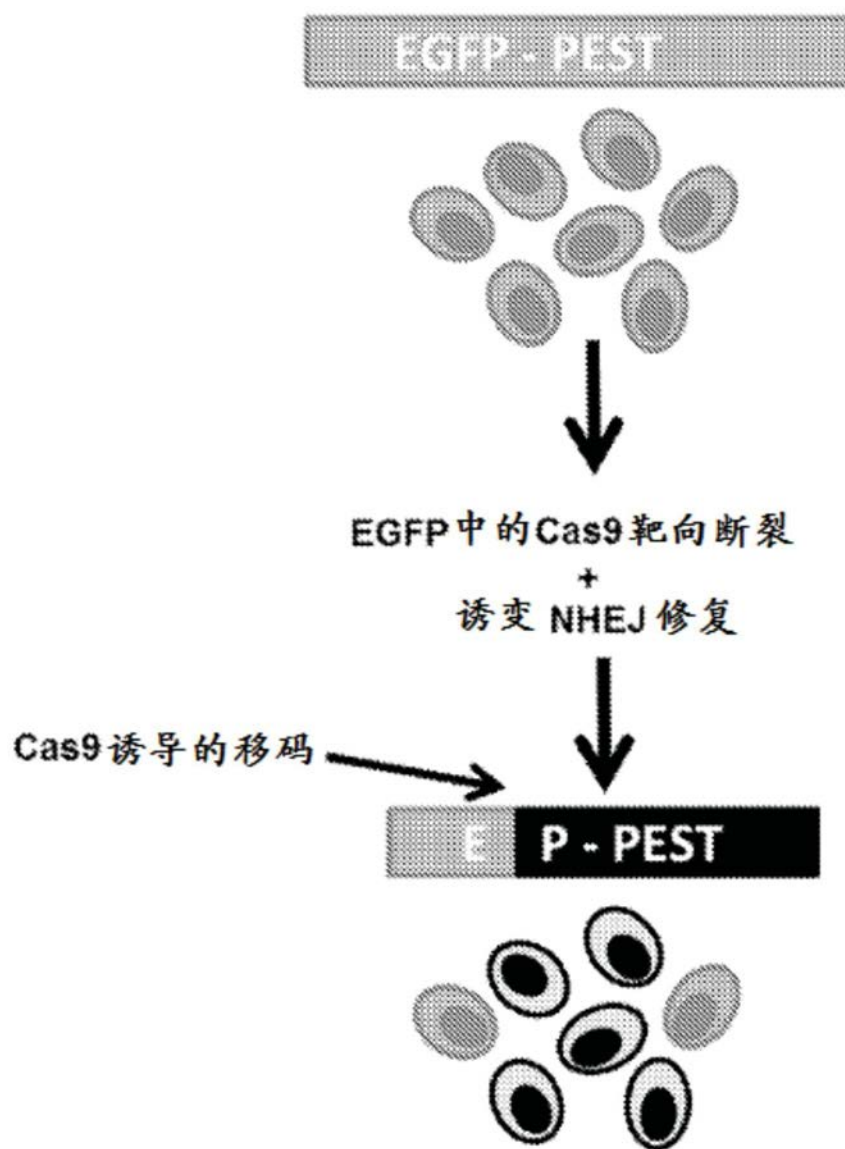


图2B

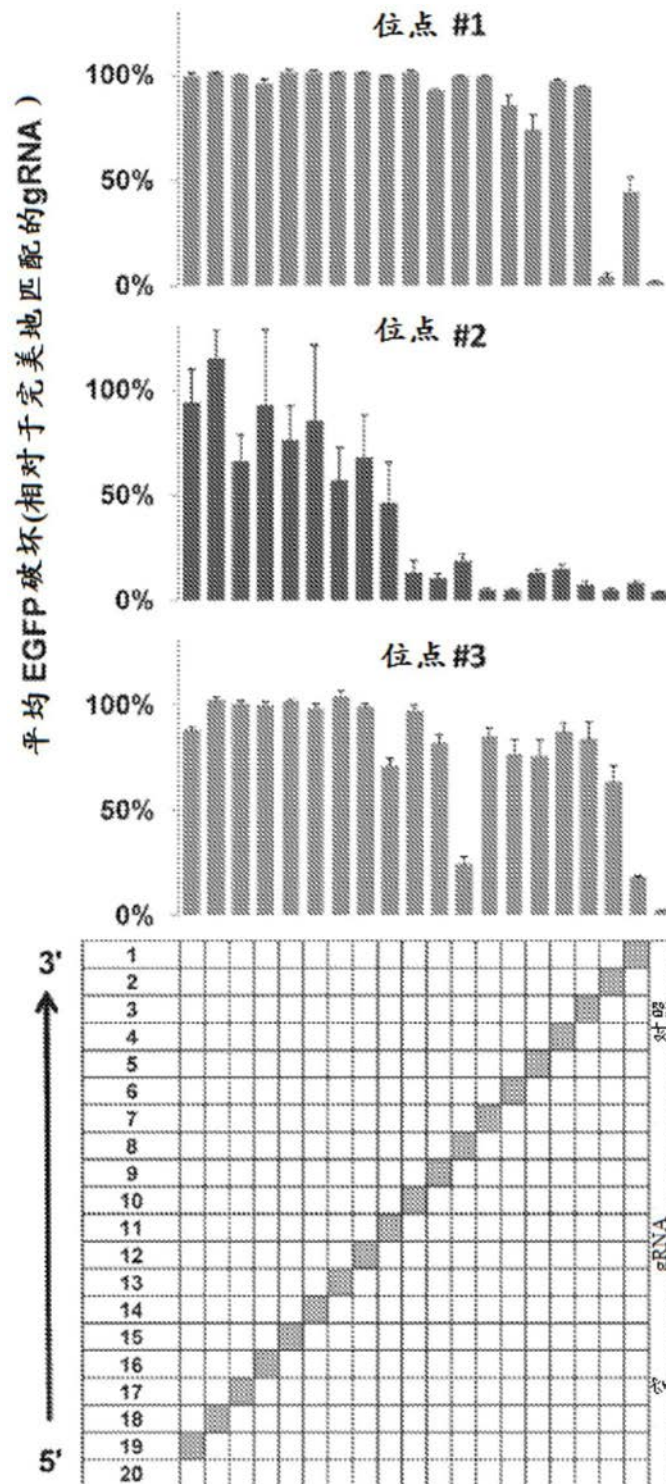


图2C

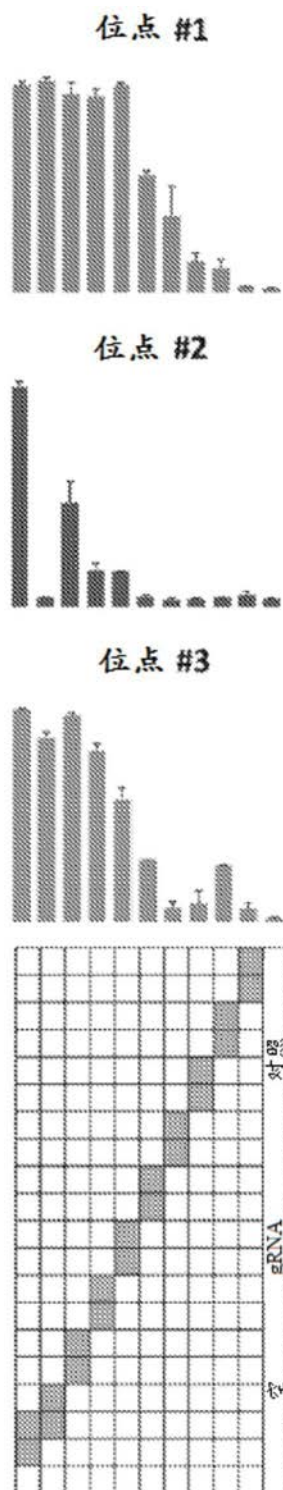


图2D

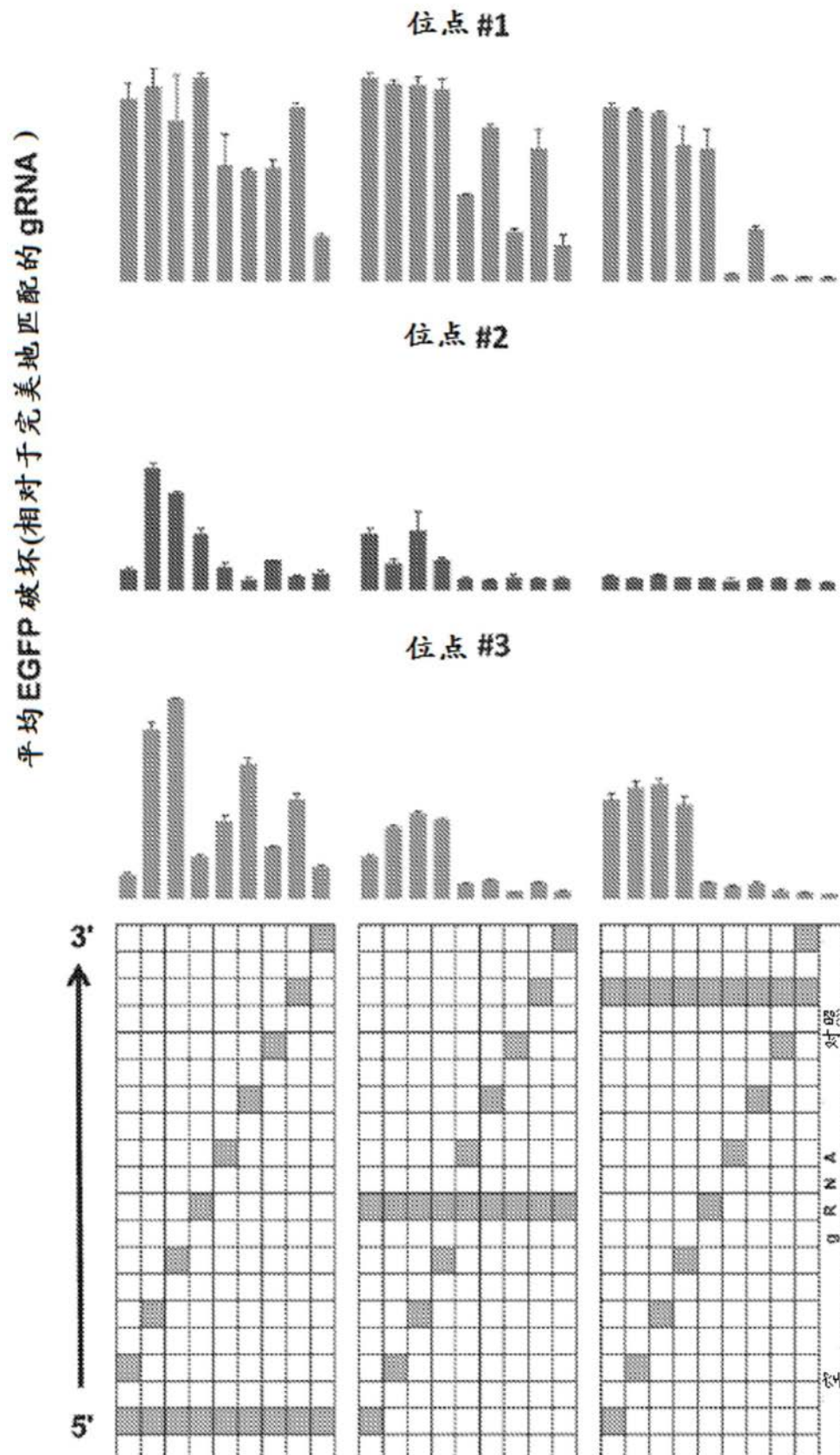


图2E

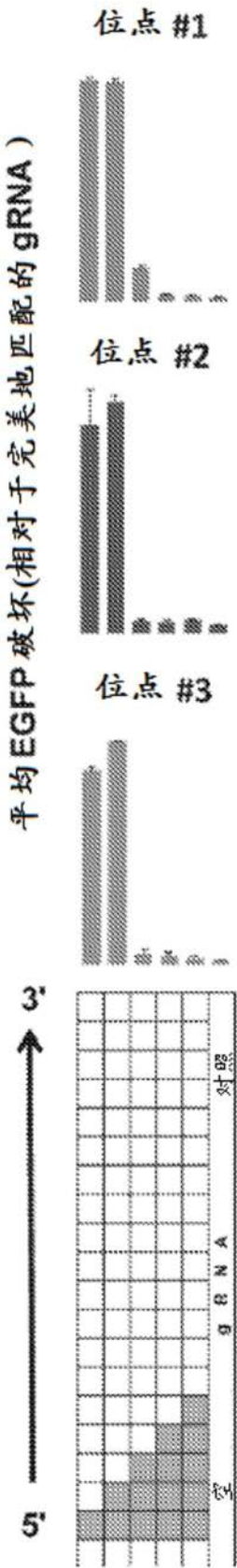


图2F

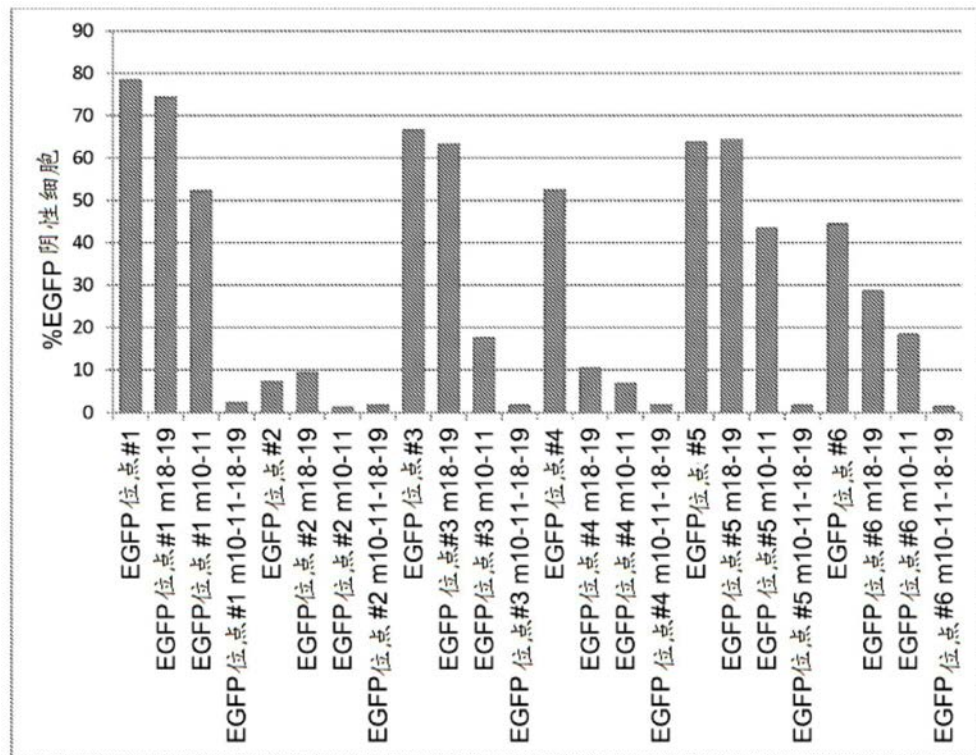


图2G

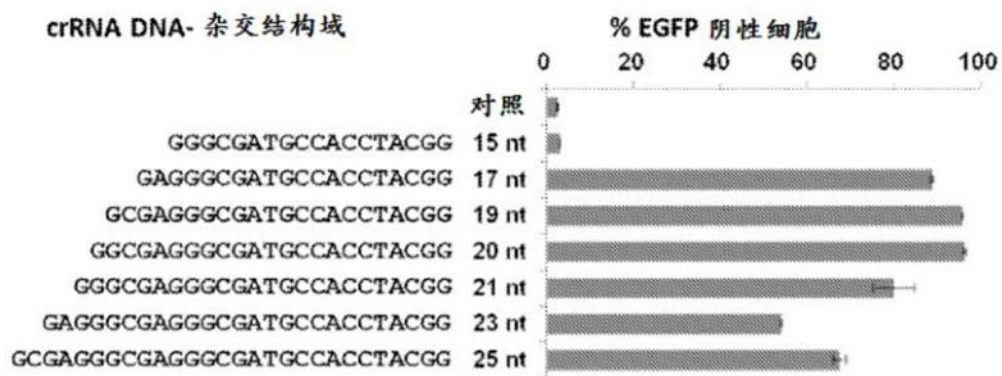


图2H

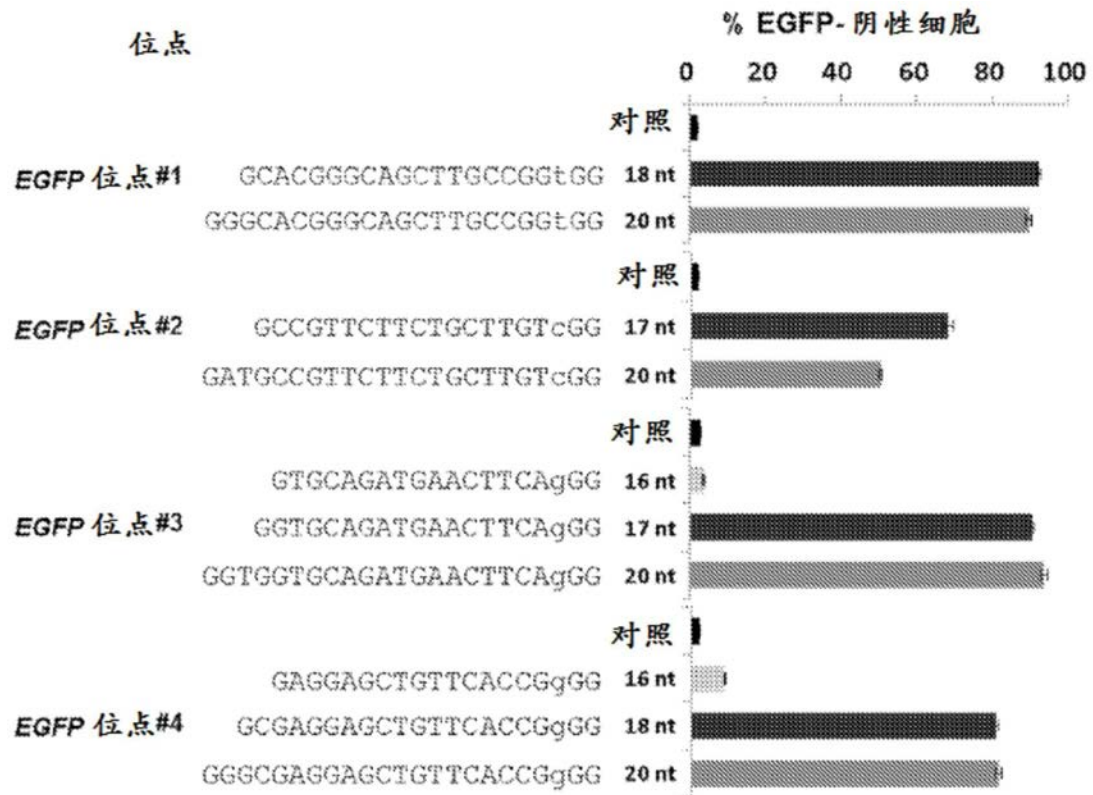


图3A

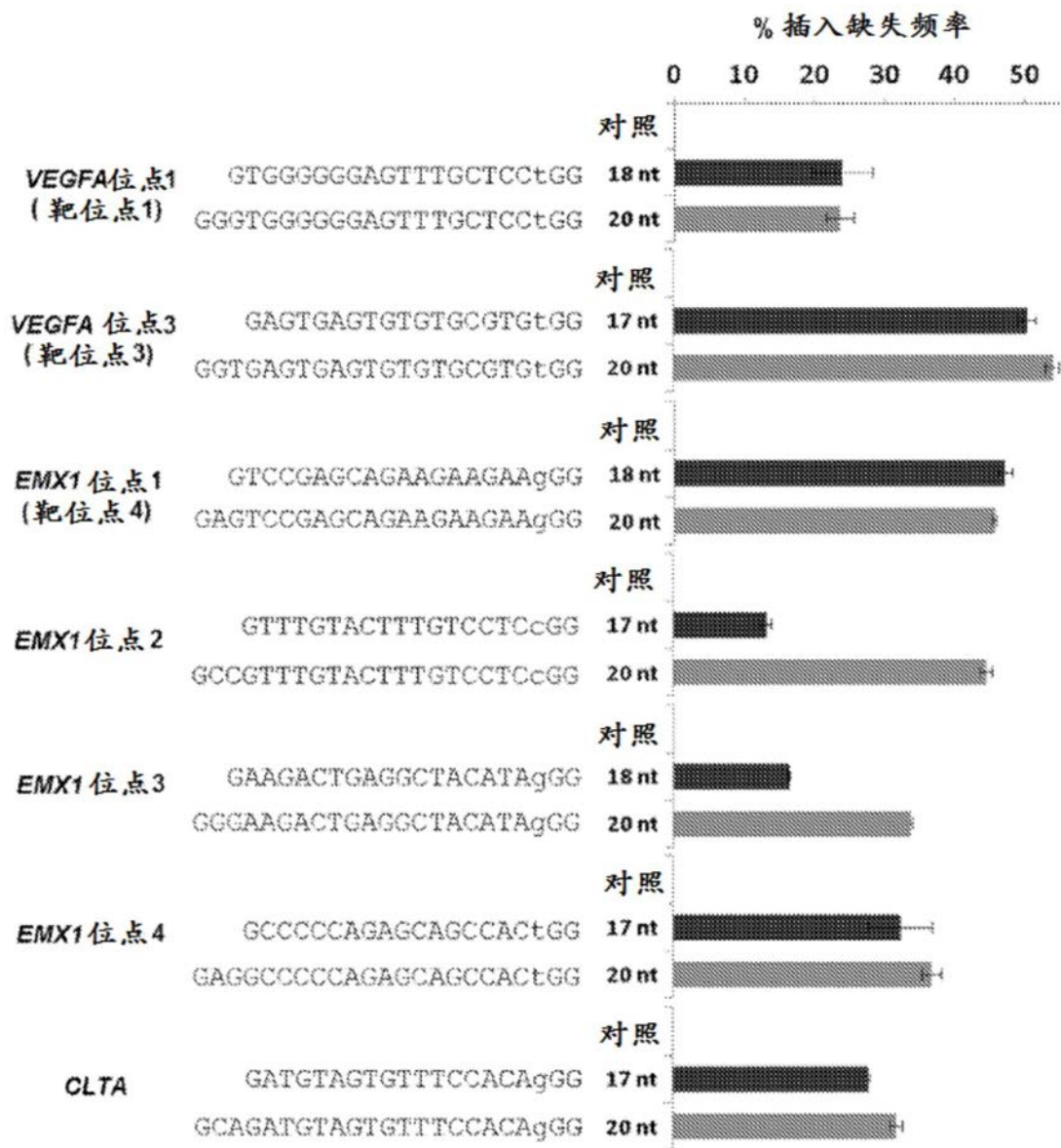


图3B

**EMX1 截短的 gRNA**

```

GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAgGGCTCCCATCACATCAACCGGTGG 野生型 x24
GAAGCTGGAGGAGGA-----> Δ365
-----TCAACCGGTGG Δ181
GAAGCTGGAGGAGGAAGG-----> Δ138
-----> Δ126
-----GGGCTCCCATCACATCAACCGGTGG Δ114
GAAGCTGGAGGAGGAAGGGCCCTGA-----> Δ101
GAAGCTGGAGGA-----GG Δ53
GAAGCTGGAGGAGGAAGGG-----> CCCATCACATCAACCGGTGG Δ28
GAAGCTGGAGGAGGAAGGGC-----> TCGCACATCAACCGGTGG Δ27
GAAGCTGGAGGAGGAAGGGC-----> CTTCCATCACATCAACCGGTGG Δ25
GAAGCTGGAGGAGGAAGGGCCCTGAG-----> TCCCATCACATCAACCGGTGG Δ21 x2
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAG-----> TCCCATCACATCAACCGGTGG Δ13
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAG-----> TCCCATCACATCAACCGGTGG Δ9
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGCGCTTTGTAG-----> CCATCACATCAACCGGTGG Δ8
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAGA-----> GCTCCCATCACATCAACCGGTGG Δ6
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAGAA-----> CTCCTCCATCACATCAACCGGTGG Δ6
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGC-----AGAAGAAGGGCTCCCATCACATCAACCGGTGG Δ3 x3
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAGA-----AAGGGCTCCCATCACATCAACCGGTGG Δ2
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAGAAGCAGAAGGGCTCCCATCACATCAACCGGT +2

```

**EMX1 全长 gRNA**

```

GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAGAAGAAgGGCTCCCATCACATCAACCGGTGG 野生型 x35
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAG-----> Δ202
-----> Δ115
GAA-----> Δ94
-----> Δ79
GAAGCTGGAGG-----> Δ72
GAAGCTGGA-----GG Δ56
GAAGCTGGAGGAGGAAGGGCCCTGA-----> GTGG Δ39
GAAGCTGGAGGAG-----GAAGGGCTCCCATCACATCAACCGGTGG Δ26 x2
GAAGCTGGAGGAGGAAGGGCCCTGAGT-----> CCATCACATCAACCGGTGG Δ22
GAAGCTGGAGGAGGAAGGGCCCTGAG-----> TCCCATCACATCAACCGGTGG Δ21 x3
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAG-----> CATCACATCAACCGGTGG Δ13
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGA-----> GCTCCCATCACATCAACCGGTGG Δ14
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGC-----AGAAGGGCTCCCATCACATCAACCGGTGG Δ6 x3
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGC-----AGAAGAAGGGCTCCCATCACATCAACCGGTGG Δ3 x3
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGA-----AAGAAGGGCTCCCATCACATCAACCGGTGG Δ2 x2
GAAGCTGGAGGAGGAAGGGCCCTGAGTCCGAGCAGAAGAAGCAGAAGGGCTCCCATCACATCAACCGGT +2

```

图3C

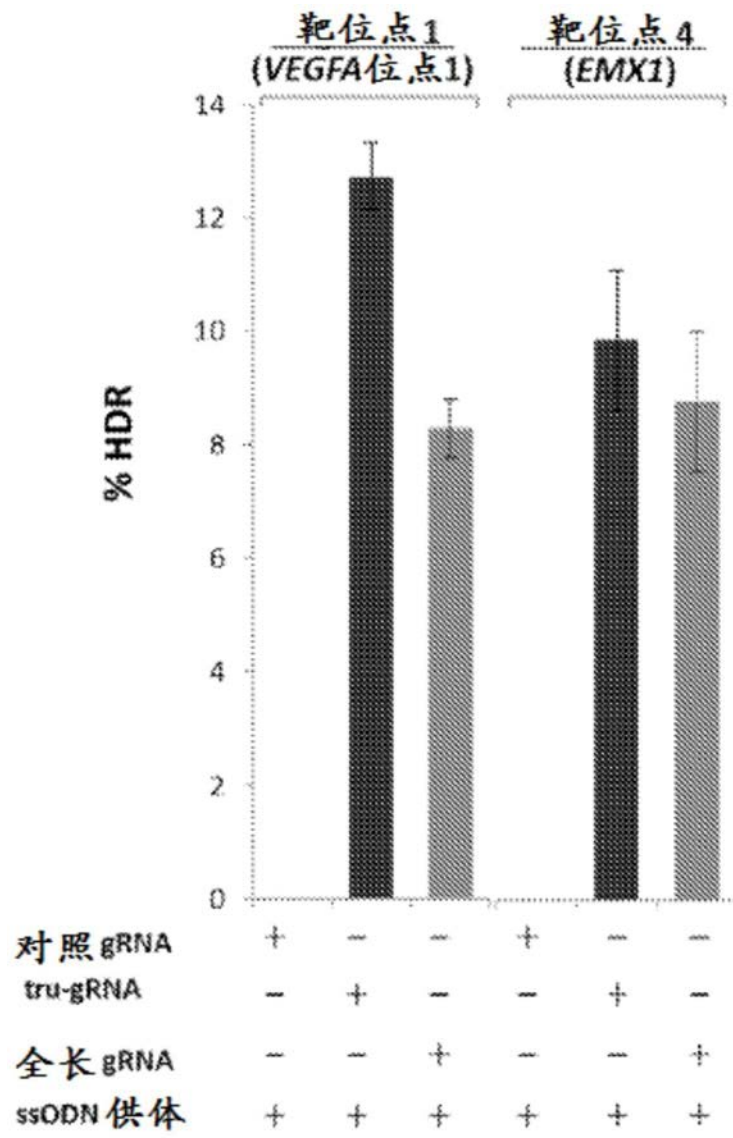


图3D

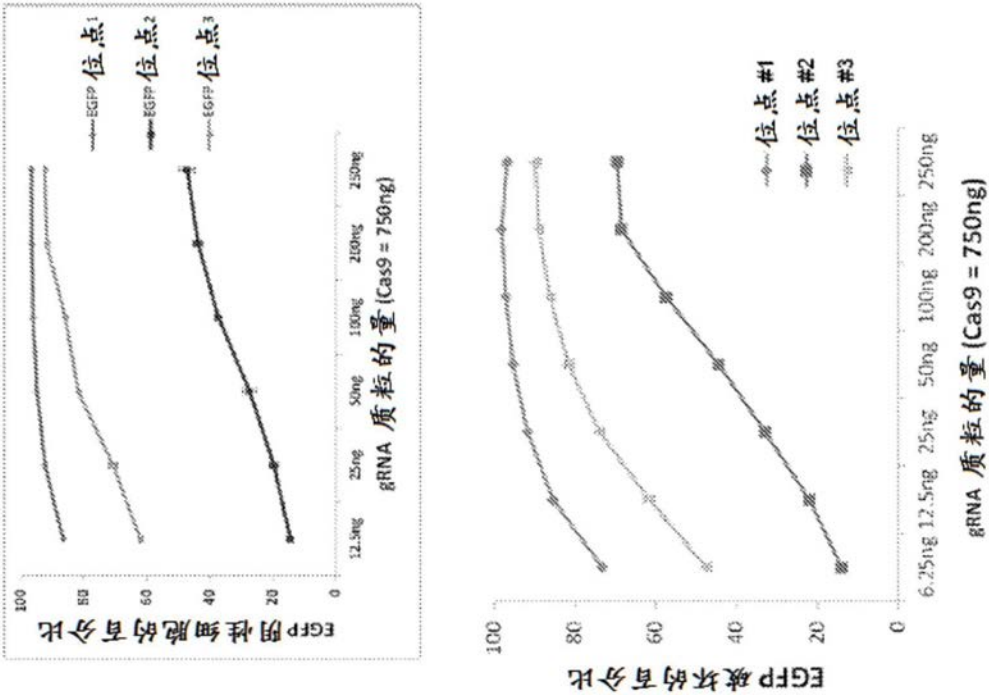


图3E

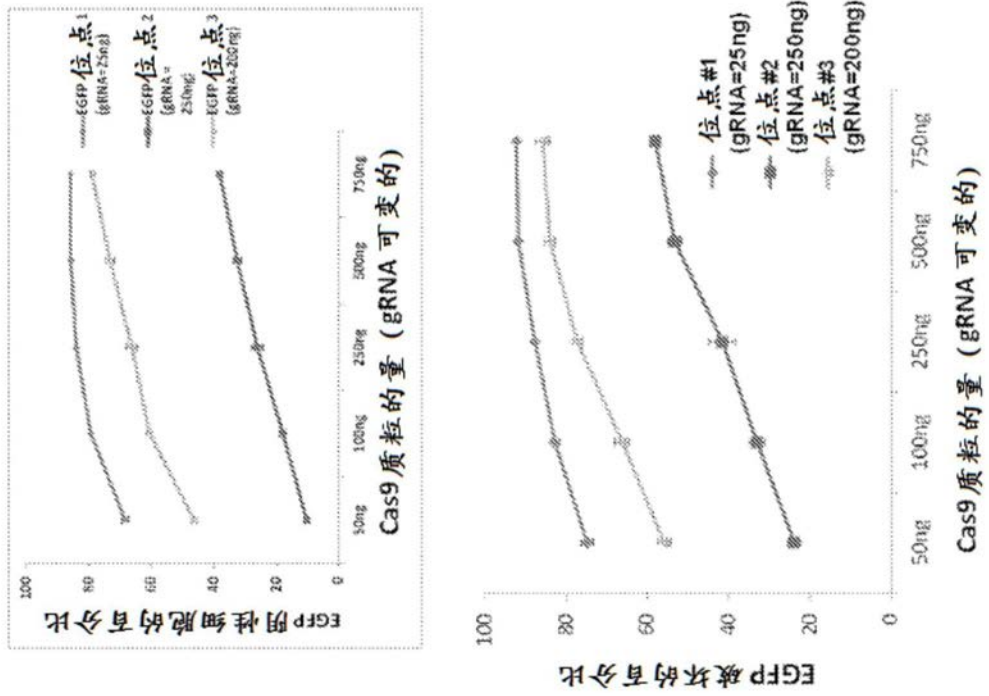


图3F

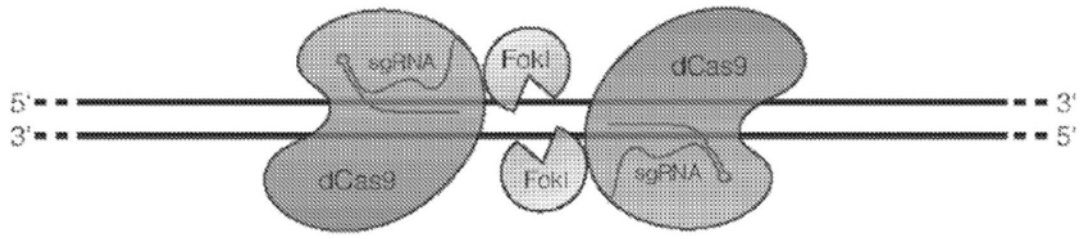


图4A

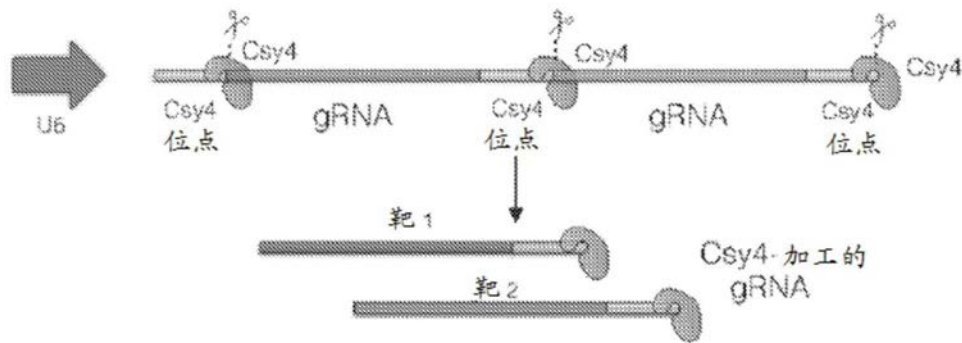


图4B

EGFP

图4C

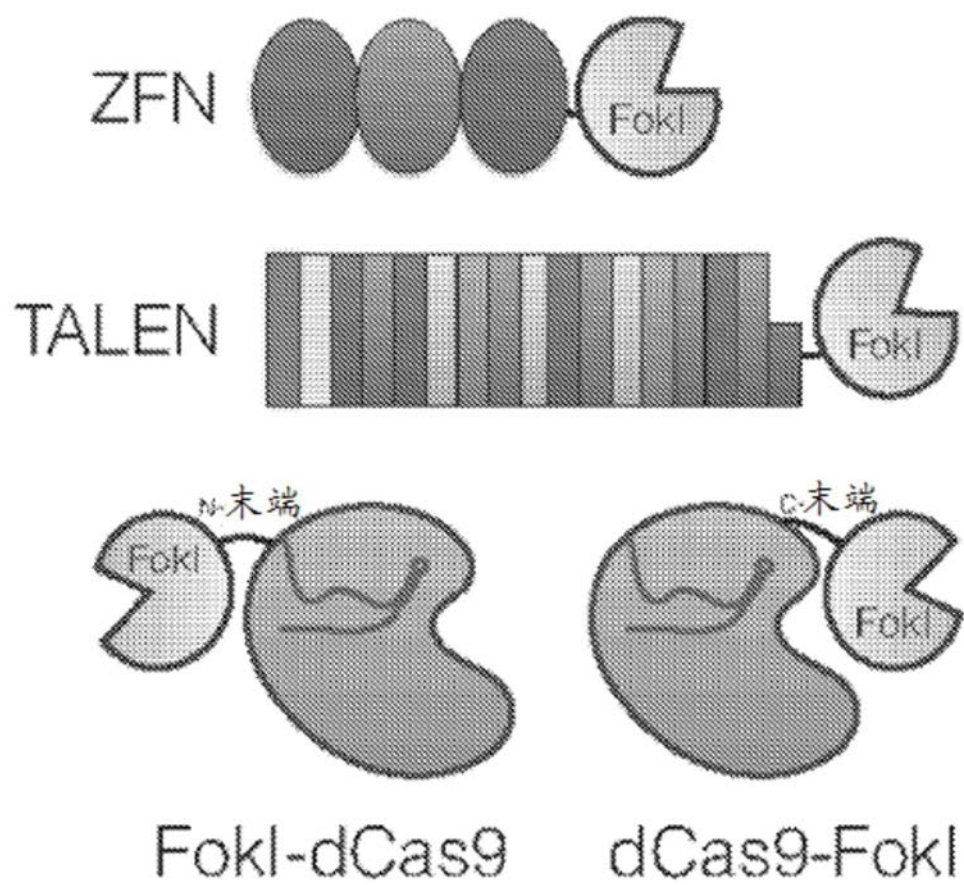


图5A

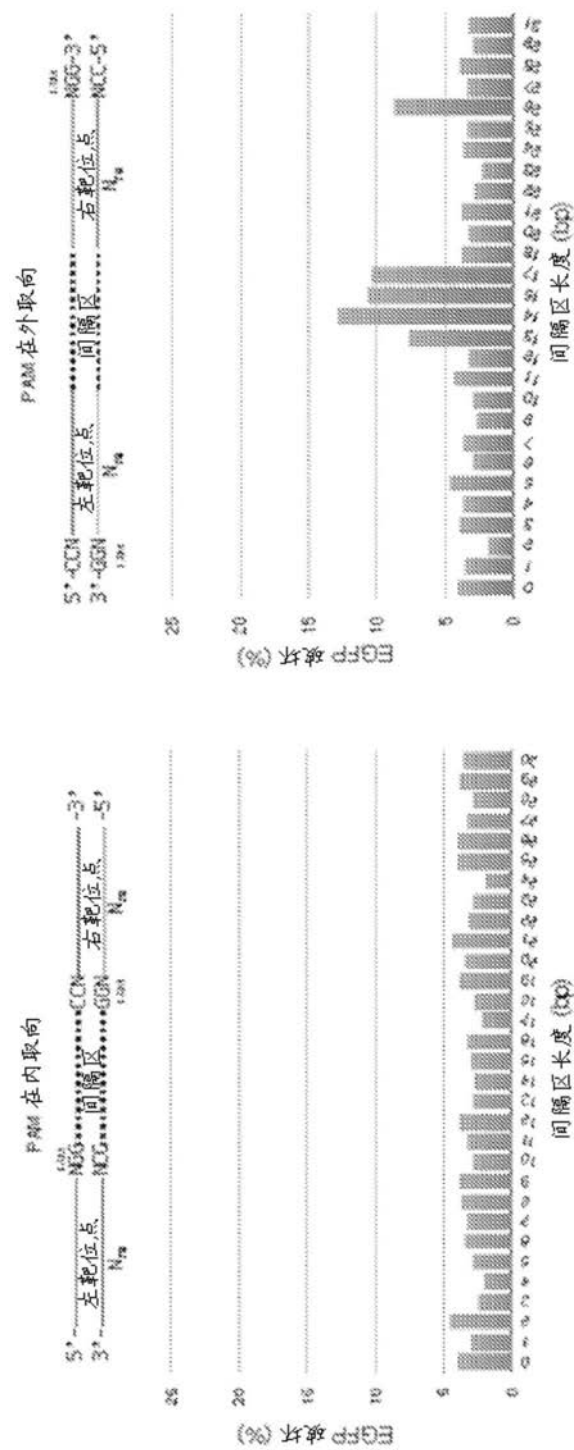


图5B

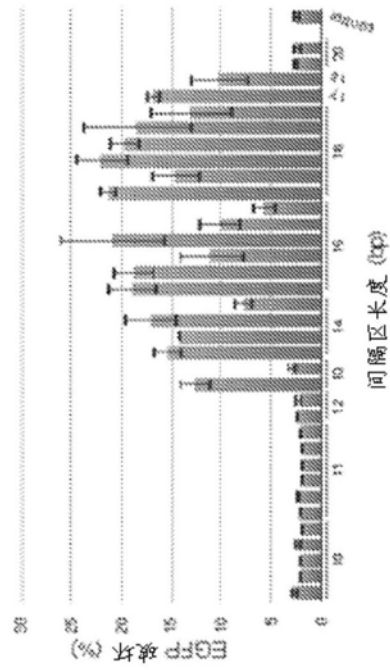


图5C

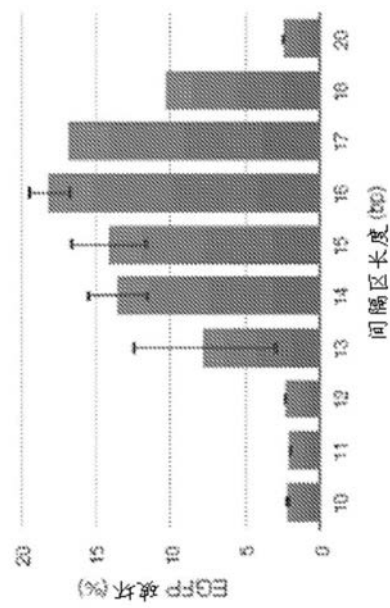


图5D

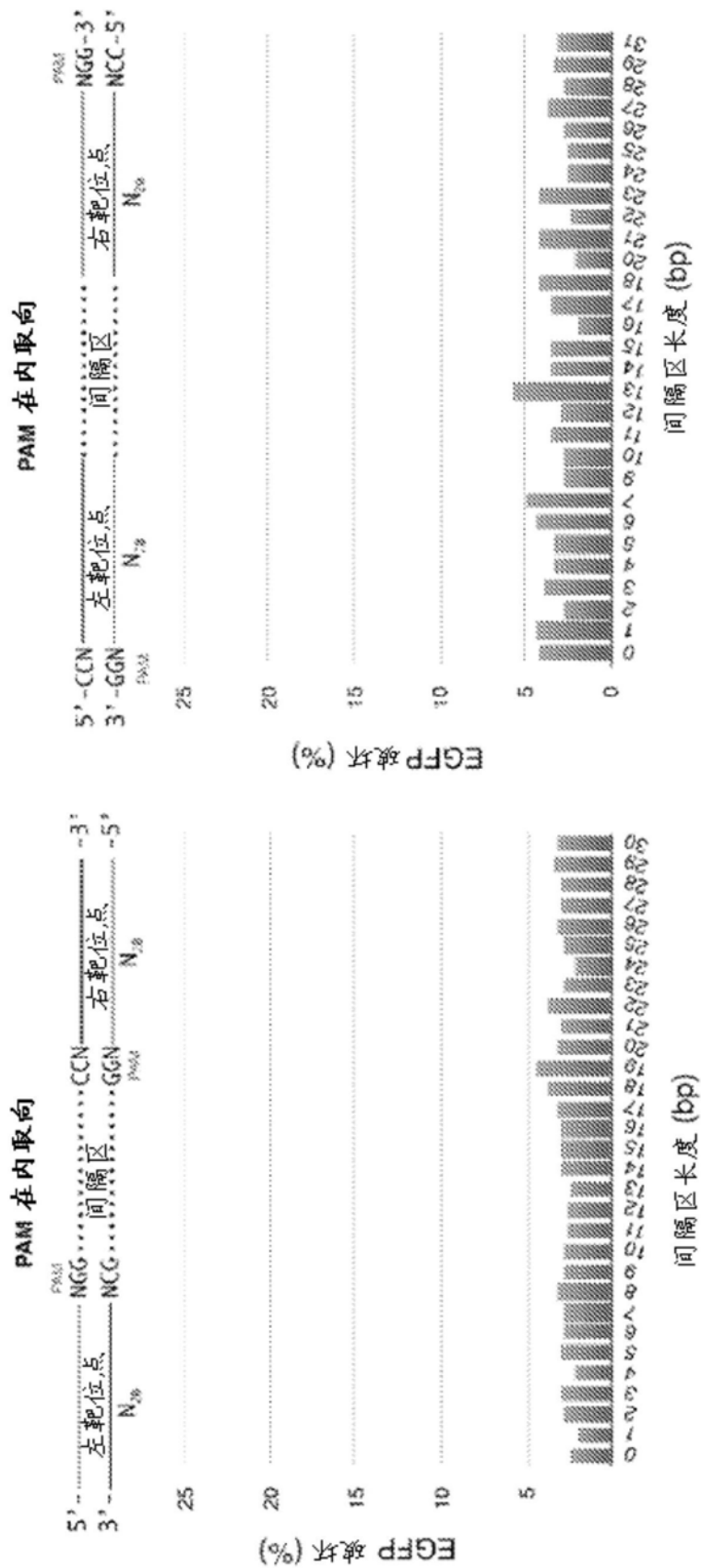


图5E

APC  
23个序列中的9个中的突变 ≈ 39.1%

CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTTGGCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC WT  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ1  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCC-----TCTGGCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ12  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ12  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ16  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ17  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ19  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ23  
CAGGACGAAGAAGAGCCGGGGGGGGCTGGTGGCCACTGGCGAGGCTGGCGAAGCGTCTGGCCAGGTGAGTGAAGCTGCAAGCATTGACGTCCTCTC Δ34  
ACCGCGGCGACAGG----- / -----TGAGTGAGGCTGCAGGCAATTGACGTCCTCTC Δ73

BRCA1  
86个序列中的7个中的突变 ≈ 8.1%

TGGGAGAGTGGAATTCGGAAGCTGACAGATGGGTATCTTTGAGGGGGGTAGGGCGGAACTTGAGAGCGCTAAGGCGTTGTGAACCCCTGG WT  
TGGGAGAGTGGAATTCGGAAGCTGACAGATGGGTATCTTTGA-----GTAGGGGCGGAACCTGAGAGGCGTAAGCGTTGTGAACCCCTGG Δ6  
TGGGAGAGTGGAATTCGGAAGCTGACAGATGGGTATCTTTGAGGCGTAAAGCGCTTAAGCGTTGTGAACCCCTGG Δ8  
TGGGAGAGTGGAATTCGGAAGCTGACAGATGGGTATCTTTGAGGCGGCGAACCCTGAGAGCGCTTAAGCGTTGTGAACCCCTGG Δ8  
TGGGAGAGTGGAATTCGGAAGCTGACAGATGGGTATCTTTGAGGCGG-----GAACTGAGAGGCGTAAGCGCTTGTGAACCCCTGG Δ10  
TGGGAGAGTGGAATTCGGAAGCTGACAGATGGGTATCTTTG-----GGCGAACCCTGAGAGGCGTAAGCGCTTGTGAACCCCTGG Δ12  
TGGGAGAGTGGAATTCGGAAGCTGACAGATGGG-----GGCGAACCCTGAGAGGCGTAAGCGCTTGTGAACCCCTGG Δ22

EMX1  
91个序列中的2个中的突变 ≈ 2.2%

CAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTT-----ATGGCCCCACAGGCTTGAAGCCCGGGGGCCGCCATTGAC WT  
CAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTT-----ATGGCCCCACAGGCTTGAAGCCCGGGGGCCGCCATTGAC Δ16

FANCF 位点1  
名称/名称 29个序列中的9个中的突变 ≈ 31%

AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAGGGCCCTTCGCGACCTCATGGAATCCCTTCTGCAGACACCTTGGATCGCTTTTCCGAGC WT  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAG-----CGACACTCATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ8  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAGGGCC-----CTCATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ9  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAGGGCCCT-----CATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ10  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAGGGCCCT-----CATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ10  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAG-----CACCTCATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ10  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAG-----CGACACTCATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ11  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAGTAGGGCCCTTCGCGA-----TTCATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ14  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAG-----CCTCATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ15  
AGAGAGTCGCGCTCCAAAGGTGAAAGCGGAAG-----CATGGAATCCCTTCTGCAGACACCTGGATCGCTTTTCCGAGC Δ31

图5F-1

**FANCE**位点2  
71个序列中的13个中的突变  $\approx 18.3\%$

```

TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGTGCACACCACTGGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC WT
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTCAACCACTGGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ3
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTAACCACTGGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ4
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTCAGTGGAGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ8
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ9
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ11
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ12 (Δ14 +2)
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ13
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ14
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ15
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ18
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ23
TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC Δ97

TTTGGTCGGCATGGCCCATTCGACGGCTCTGGAGCGGCGGCTGAGGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC +168(Δ9 +177)
Tgaaggggttcctcatgatgtcccatataattttggcagagagaaagatctcagtggtattgtgagcagggcatttggccacacacagccac
caccttttgatagcagcctgcaactgaggagtgcgAGTGGAGCAAGAGGGGGCTTTGGGCGGGGTCCAGTTCC

```

**GLI1**  
66个序列中的4个中的突变  $\approx 6.1\%$

```

GTTCAACTCGATGACCCCAACCAATCAGTAGCTATGGCGAGCCCTGCTGCTCCGGCCCTCCCCAGTCAGGGGCCCCCAGTGTGGGA WT
GTTCAACTCGATGACCCCAACCAATCAGTAGCTATGCTGCTGCTCCGGCCCTCCCCAGTCAGGGGCCCCCAGTGTGGGA Δ7
GTTCAACTCGATGACCCCAACCAATCAGTAGCTATGCTCCGGCCCTCCCCAGTCAGGGGCCCCCAGTGTGGGA Δ13
GTTCAACTCGATGACCCCAACCAATCAGTAGCTATGCTCCGGCCCTCCCCAGTCAGGGGCCCCCAGTGTGGGA Δ17
GTTCAACTCGATGACCCCAACCAATCAGTAGCTATGCTCCGGCCCTCCCCAGTCAGGGGCCCCCAGTGTGGGA Δ18

```

**MLH1**  
20个序列中的4个中的突变  $\approx 20\%$

```

TGGCTGAAGGCACATTCGGTTGAGCATCTAGACGTTTCGTTGGCTCTTCTGGGCGCAAAATGTCGTTGGTGGCAGGGTTATTTCGGCGGCTGG WT
TGGCTGAAGGCACATTCGGTTGAGCATCTAGACGTTTCCTTTGGCGCCAAATGCTCGTTCGTCGAGGGGTTATTTCGGCGGCTGG Δ9
TGGCTGAAGGCACATTCGGTTGAGCATCTA---TTCTGGCGCCAAATGCTCGTTCGTCGAGGGGTTATTTCGGCGGCTGG Δ16
TGGCTGAAGGCACATTCGGTTGAGCATCTAGACGTTTCCTTGG---ATGTCGTTTCGTCGAGGGGTTATTTCGGCGGCTGG Δ16
TGGCTGAAGGCACATTCGGTTGAGCATCTAGACGTTT---GTCGTTTCGTCGAGGGGTTATTTCGGCGGCTGG Δ24

```

图5F-2

**RARA**  
86个序列中的4个中的突变  $\approx 4.7\%$

CCCTTCTGACTGTGGCGCTTGGCATGGGCGAGCAAGACAGCTCTGCCCCACAGCTGGGGGGGGGACCTCAATGGGTATACCCGGTGCCCTCC WT  
CCCTTCTGACTGTGGCGCTTGGCATGGGCGAGCAACA-----CCTGCGGACACCTGGGGGGGGGACCTCAATGGGTATACCCGGTGCCCTCC  $\Delta 6$   
CCCTTCTGACTGTGGCGCTTGGCATGGGCGAGCAACA-----CTGCGGACACCTGGGGGGGGGACCTCAATGGGTATACCCGGTGCCCTCC  $\Delta 7$   
CCCTTCTGACTGTGGCGCTTGGCATGGGCGAGCA-----GCCGACACCTGGGGGGGGGACCTCAATGGGTATACCCGGTGCCCTCC  $\Delta 12$   
CCCTTCTGACTGTGGCGCTTGGCATGGGCGAG-----GCCGACACCTGGGGGGGGGACCTCAATGGGTATACCCGGTGCCCTCC  $\Delta 14$

**VEGFA 位点 1**  
74个序列中的19个中的突变  $\approx 25.7\%$

TCAGAAATAGGGGGTCCAGAGCAAACTGCCCAACGCCCTTTCCAAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT WT  
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTTCCAAAGGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCA  $\Delta 2$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTTCC-----GCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 3$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----AGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 5$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----AAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 5$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----AAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 6$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTTCCA-----TTCCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 7$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTTCTC-----ATTCCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 8$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTTAC-----AAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 9$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----ATTCCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 14$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----CAAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 17$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----CAAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 18$   
TCAGAAATAGGGGGTCCAGAGCAAACTCCCGCCACCCCTTT-----AGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 20$   
TCAGAAATAGGGGGTCCAGG-----CAAAGCCCATTCCTCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 24$   
TCAGAAATAGGGGGTCCAGGAGC-----TCCTTTAGCCAGAGCGCGGGTGTGCAGACGGCAGT  $\Delta 35$

**VEGFA 位点 2**  
80个序列中的26个中的突变  $\approx 32.5\%$

CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGGCGGAGACAGTGGACGCGCGCGCGAGCGCGGGGCGGCGCGGAGCCCCCGCCCGGAG WT  
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGGCGG-----TGGACGCGCGCGCGAGCGCGGGCGGCGCGGAGCCCCCGCCCGGAG  $\Delta 8$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGG-----ACGCGCGCGCGAGCGCGGGCGGCGCGGAGCCCCCGCCCGGAG  $\Delta 13$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGC-----TGGACGCGCGCGCGAGCGCGGGCGGCGCGGAGCCCCCGCCCGGAG  $\Delta 13$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGGCGGACAG-----CCGCGGGCAGGCGCGGGCGGAGCCCCCGCCCGGAG  $\Delta 16$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGG-----ACGCGCGCGCGAGCGCGGGCGGAGCCCCCGCCCGGAG  $\Delta 17$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGG-----AGCCCGCGCGGAG-----AGCCCGCGCGGAG  $\Delta 43$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGGCGGAGCTGGA-----/-----GGGGTCTGGGGCTCG  $\Delta 54$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGG-----/-----/-----GAG  $\Delta 56$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGGCGG-----/-----GGTGGAGGGGTCGG  $\Delta 56$   
CCCCAGCCCCAGCTACCTCTCTCCCGGGCGGCGGACa-----/-----CaCGCGCTCGCAC  $\Delta 69$   
CCCCAGCCCCAGCTACCTCTCTCCCGGGCGGCGGCGG-----/-----GGAG  $\Delta 87$   
CCCCAGCCCCAGCTACCACTCTCTCCCGGGCGGCGGCGG-----/-----CTGAACTTTTCGTC  $\Delta 89$   
CCGAGCGCGGGCTG-----/-----GAGGGGTCTGGGGCT  $\Delta 165$

图5F-3

VEGFA位点3  
77个序列中的31个中的突变 $\approx 40.3\%$

```

CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCACTAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC WT
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCACT---TAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ3
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCA---CTAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ4
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---GGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ5
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCA---TAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ5
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGAC---CTAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ8
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGAC---GGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ9
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGAC---TAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ9
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGAC---CTAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ10
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGAC---AGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ11
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGAC---CTAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ11
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGT---TCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ13
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGT---CTAGGGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ14
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGT---GGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ15
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGT---GGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ16
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---TCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ16
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---CTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ17
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---GGGGCGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ19
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---GCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ21
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---AGGGAAGCTGGGTGAATGGAGCGAGC Δ21
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---CACCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ21
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---GGGAAGCTGGGTGAATGGAGCGAGC Δ23
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---GCCACACAGGGAAGCTGGGTGAATGGAGCGAGC Δ24
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---CACCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ24
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---CACCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ27
CCATTCCCTCTTTAGCCAGAGCCGGGGTGTGCAGACGGCAGTCA---AGTGACTGCTCGGCCACCCACAGGGAAGCTGGGTGAATGGAGCGAGC Δ34

```

图5F-4

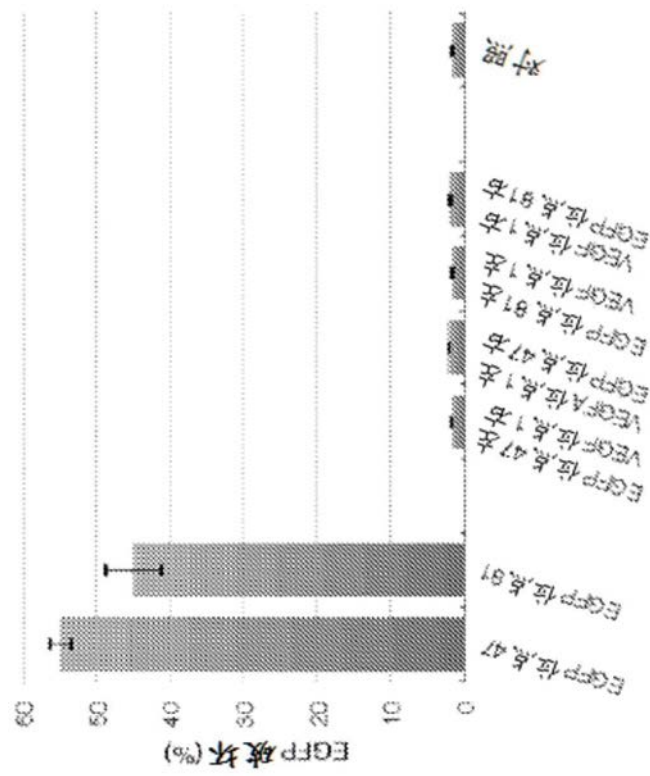


图6A

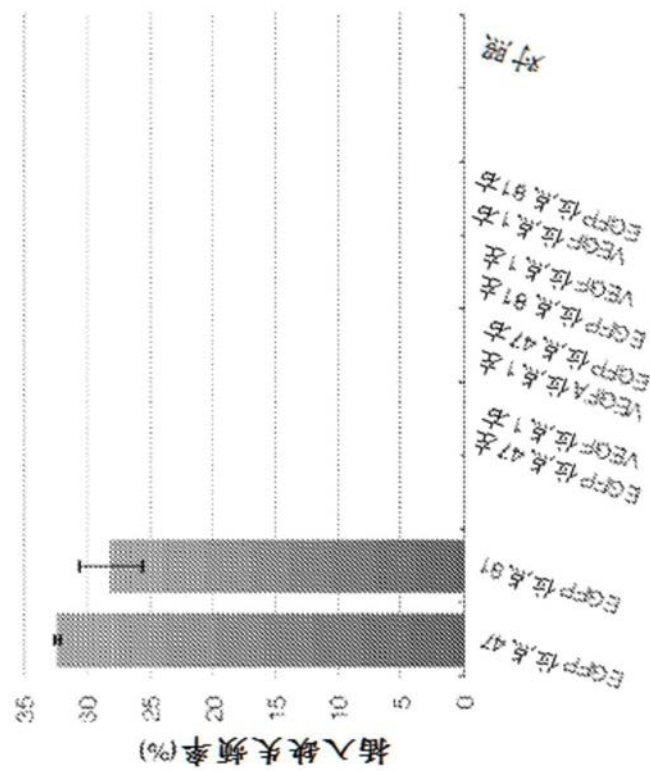


图6B

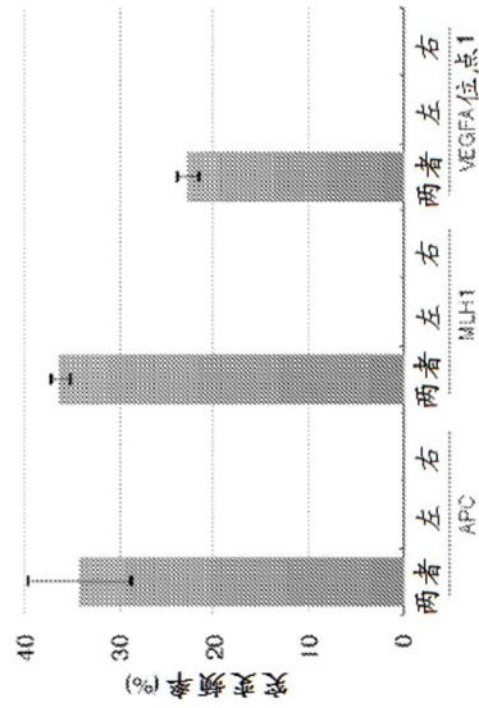


图6C

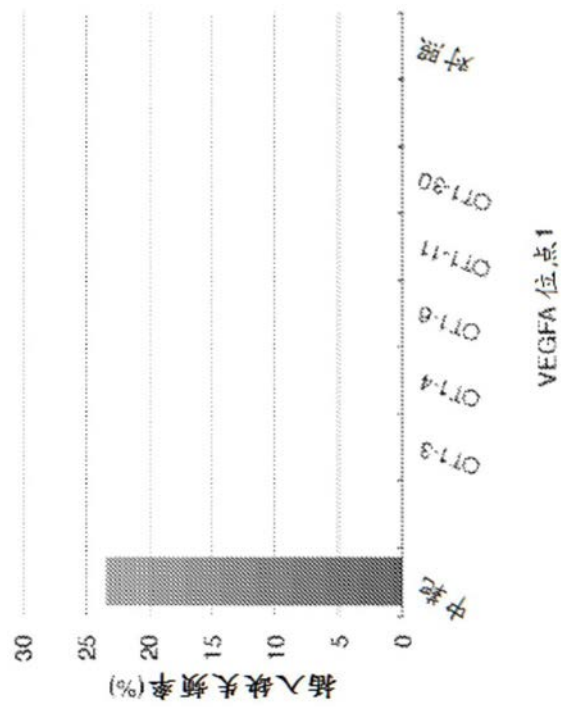


图6D

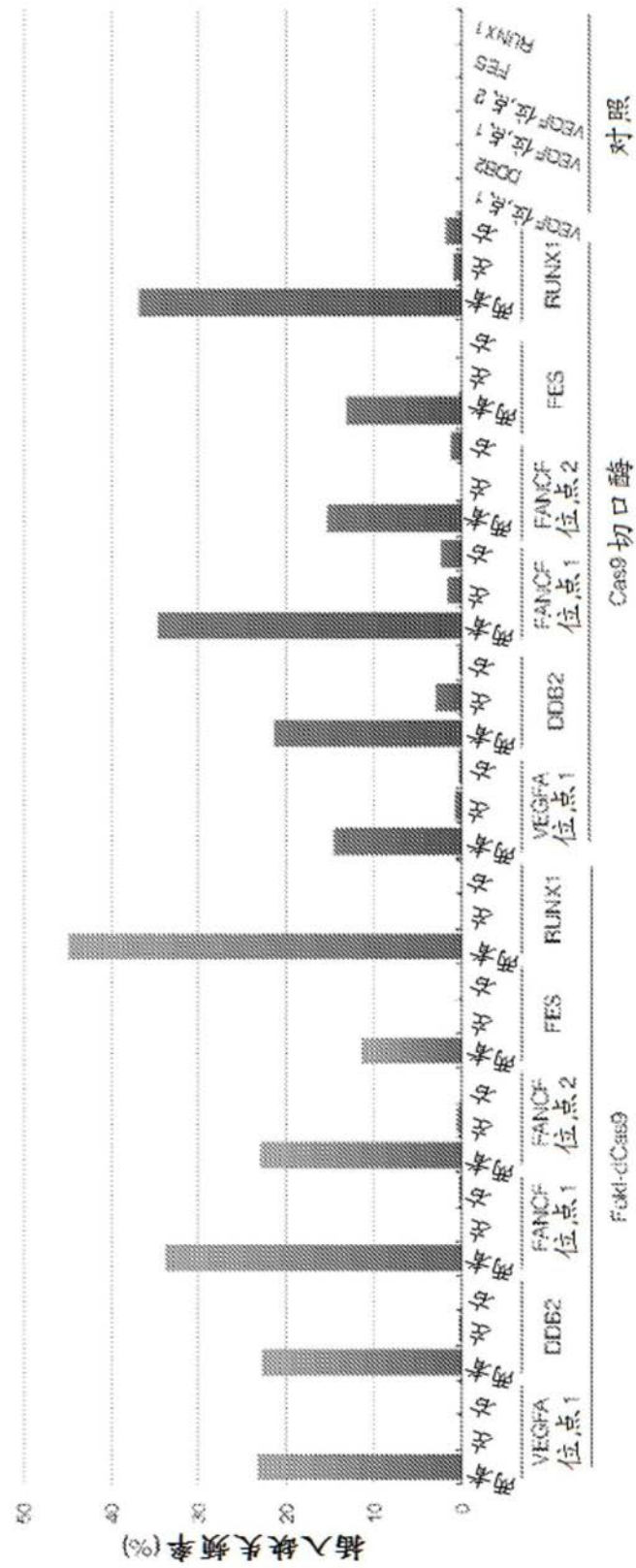


图7A

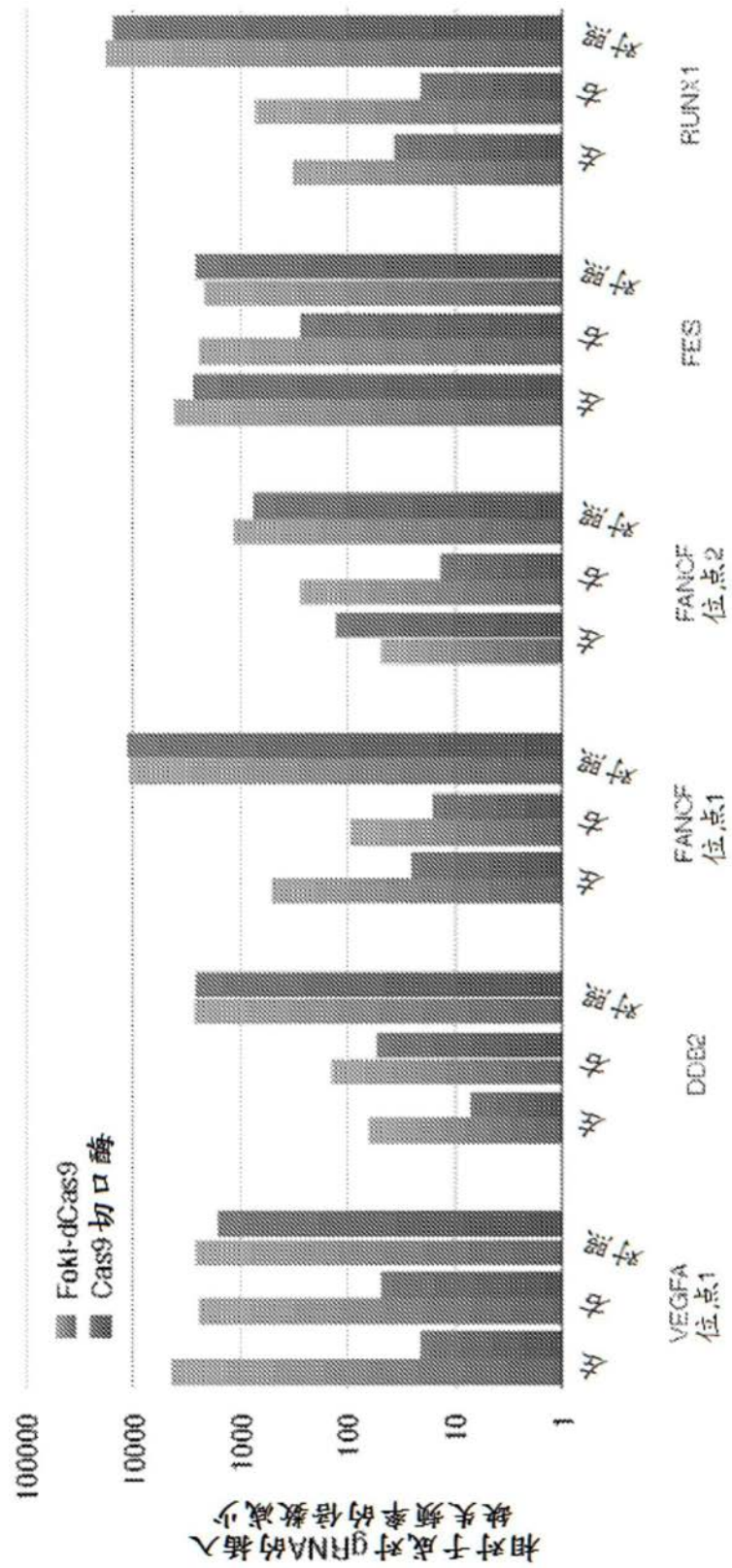


图7B

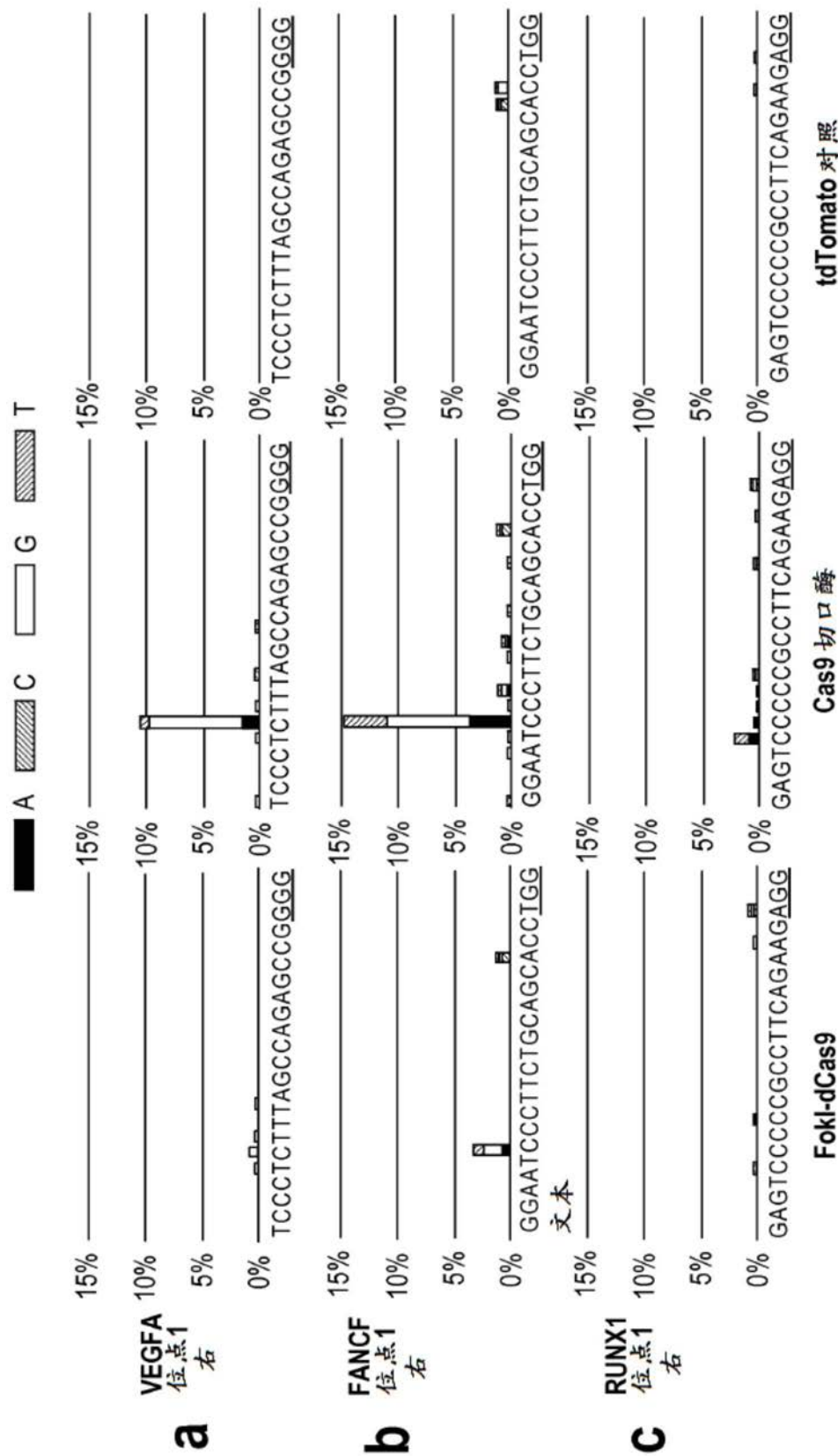


图8A-C