(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110214180 A
(43)申请公布日 2019.09.06

(21)申请号 201780076474.6

(22)申请日 2017.10.13

(30)优先权数据
62/408,575 2016.10.14 US
62/475,780 2017.03.23 US

(85)PCT国际申请进入国家阶段日
2019.06.11

(86)PCT国际申请的申请数据
PCT/US2017/056671 2017.10.13

(87)PCT国际申请的公布数据
WO2018/071868 EN 2018.04.19

(71)申请人 哈佛大学的校长及成员们
地址 美国马萨诸塞州

(72)发明人 D.R.刘 J.M.莱维 W.H.叶

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 张晓飞

(51)Int.Cl.
C12N 7/00(2006.01)
C12N 15/864(2006.01)
C12N 9/22(2006.01)
C12N 5/10(2006.01)
A61K 48/00(2006.01)
C12N 15/10(2006.01)
C12N 15/79(2006.01)

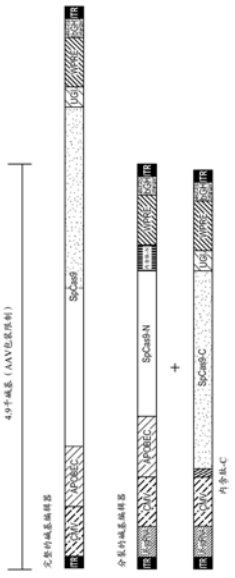
权利要求书4页 说明书153页
序列表（电子公布） 附图13页

(54)发明名称

核碱基编辑器的AAV递送

(57)摘要

本文提供了将“分裂的”Cas9蛋白或核碱基编辑器递送到细胞中的方法，例如经由重组腺相关病毒(rAAV)，以形成完整且功能性的Cas9蛋白或核碱基编辑器。将Cas9蛋白或核碱基编辑器分裂成两个部分，每个部分与内含肽系统的一个部分融合(例如，分别由dnaEn和dnaEc编码的内含肽-N和内含肽-C)。共表达后，经由内含肽介导的蛋白质剪接将Cas9蛋白或核碱基编辑器的两个部分连接在一起。还提供了用于递送分裂的Cas9蛋白或核碱基编辑器的重组AAV载体和颗粒，以及使用此类AAV载体和颗粒的方法。



CN 110214180 A

1.组合物,其包含:

(i)第一核苷酸序列,其编码在C-末端与内含肽-N融合的Cas9蛋白的N-末端部分;和

(ii)第二核苷酸序列,其编码与所述Cas9蛋白的C-末端部分的N-末端融合的内含肽-C,

其中所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号的核苷酸序列可操作地连接。

2.权利要求1的组合物,其中所述Cas9蛋白的N-末端部分包含对应于SEQ ID NO:1的氨基酸1-573或1-637的SEQ ID NO:1-275和394-397的任一个的部分。

3.权利要求1或权利要求2的组合物,其中所述Cas9蛋白的C-末端部分包含对应于SEQ ID NO:1的氨基酸574-1368或638-1368的SEQ ID NO:1-275和394-397的任一个的部分。

4.权利要求1-3中任一项的组合物,其中所述内含肽-N包含如SEQ ID NO:350-351和354-355的任一个中所示的氨基酸序列。

5.权利要求1-4中任一项的组合物,其中所述内含肽-C包含如SEQ ID NO:352-353和356-357的任一个中所示的氨基酸序列。

6.权利要求1-5中任一项的组合物,其中所述第一核苷酸序列或所述第二核苷酸序列进一步包含与启动子可操作地连接的编码指导RNA(gRNA)的核苷酸。

7.权利要求1-6中任一项的组合物,其中所述第一核苷酸序列或所述第二核苷酸序列进一步包含转录终止子。

8.权利要求7的组合物,其中所述转录终止子是来自bGH基因、hGH基因或SV40基因的转录终止子。

9.权利要求8的组合物,其中所述转录终止子是来自bGH基因的转录终止子。

10.权利要求1-9中任一项的组合物,其中所述第一核苷酸序列或所述第二核苷酸序列进一步包含插入所述转录终止子的5′的土拨鼠肝炎转录后调控元件(WPRE)。

11.权利要求1-10中任一项的组合物,其中所述二分核定位信号包含选自下组的氨基酸序列:KRPAATKKAGQAKKKK(SEQ ID NO:344)、KKTELQTTNAENKTKKL(SEQ ID NO:345)、KRGINDRNFWRGENGRKTR(SEQ ID NO:346)和RKSGKIAAIVVKRPRK(SEQ ID NO:347)。

12.权利要求11的组合物,其中所述二分核定位信号包含如SEQ ID NO:344中所示的氨基酸序列。

13.权利要求1-12中任一项的组合物,其中所述Cas9蛋白是催化无活性的Cas9(dCas9)或Cas9切口酶(nCas9),并且其中(i)的所述第一核苷酸序列进一步包含编码核碱基修饰酶的核苷酸序列,所述核碱基修饰酶与所述Cas9蛋白的N-末端部分的N-末端融合。

14.权利要求1-12中任一项的组合物,其中所述Cas9蛋白是催化无活性的Cas9(dCas9)或Cas9切口酶(nCas9),并且其中(ii)的所述第二核苷酸序列进一步包含编码核碱基修饰酶的核苷酸序列,所述核碱基修饰酶与所述Cas9蛋白的C-末端部分的C-末端融合。

15.权利要求13或14的组合物,其中所述核碱基修饰酶是脱氨酶。

16.权利要求15的组合物,其中所述脱氨酶是胞嘧啶脱氨酶。

17.权利要求15的组合物,其中所述脱氨酶是腺苷脱氨酶。

18.权利要求14-17中任一项的组合物,其中(ii)的所述第二核苷酸序列进一步包含在所述第二核苷酸序列的3′末端处的编码尿嘧啶糖基化酶抑制剂(UGI)的核苷酸序列。

19.权利要求13-17中任一项的组合物,其中(i)的所述第一核苷酸序列进一步包含在所述第一核苷酸序列的5′末端处的编码尿嘧啶糖基化酶抑制剂(UGI)的核苷酸序列。

20.权利要求18或19的组合物,其中所述UGI包含SEQ ID NO:299-302的氨基酸序列。

21.权利要求1-20中任一项的组合物,其中所述第一核苷酸序列和所述第二核苷酸序列在不同的载体上。

22.权利要求21的组合物,其中每个所述不同的载体是重组腺相关病毒(rAAV)的基因组。

23.权利要求22的组合物,其中每个载体包装在rAAV颗粒中。

24.组合物,其包含:

(i)第一重组腺相关病毒(rAAV)颗粒,其包含编码在C-末端处与内含肽-N融合的Cas9蛋白的N-末端部分的第一核苷酸序列;和

(ii)第二重组腺相关病毒(rAAV)颗粒,其包含编码与所述Cas9蛋白的C-末端部分的N-末端融合的内含肽-C的第二核苷酸序列,

其中所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号的核苷酸序列可操作地连接。

25.细胞,其包含权利要求1-24中任一项的组合物。

26.权利要求25的细胞,其中所述Cas9蛋白的N-末端部分和所述Cas9蛋白的C-末端部分接合在一起形成所述Cas9蛋白。

27.权利要求25或26的细胞,其中所述细胞是原核细胞。

28.权利要求27的细胞,其中所述细胞是细菌细胞。

29.权利要求25或26的细胞,其中所述细胞是真核细胞。

30.权利要求29的细胞,其中所述细胞是酵母细胞、植物细胞或哺乳动物细胞。

31.权利要求29的细胞,其中所述细胞是人细胞。

32.试剂盒,其包含权利要求1-24中任一项的组合物。

33.组合物,其包含:

(i)第一核苷酸序列,其编码在C-末端处与内含肽-N融合的核碱基编辑器的N-末端部分;和

(ii)第二核苷酸序列,其编码与所述核碱基编辑器的C-末端部分的N-末端融合的内含肽-C。

34.权利要求33的组合物,其中所述内含肽-N包含如SEQ ID NO:350-351和354-355的任一个中所示的氨基酸序列。

35.权利要求33或34的组合物,其中所述内含肽-C包含如SEQ ID NO:352-353和356-357的任一个中所示的氨基酸序列。

36.权利要求33-35中任一项的组合物,其中所述第一核苷酸序列或所述第二核苷酸序列进一步包含与启动子可操作地连接的编码指导RNA(gRNA)的核苷酸。

37.权利要求33-36中任一项的组合物,其中所述第一核苷酸序列或所述第二核苷酸序列进一步包含转录终止子。

38.权利要求37的组合物,其中所述转录终止子是来自bGH基因、hGH基因或SV40基因的转录终止子。

39.权利要求38的组合物,其中所述转录终止子来自bGH基因。

40.权利要求33-39中任一项的组合物,其中所述第一核苷酸序列或所述第二核苷酸序列进一步包含插入所述转录终止子的5′的土拨鼠肝炎转录后调控元件(WPRE)。

41.权利要求33-40中任一项的组合物,其中所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号的核苷酸序列可操作地连接。

42.权利要求33-41中任一项的组合物,其中所述二分核定位信号包含选自下组的氨基酸序列:KRPAATKKAGQAKKKK(SEQ ID NO:344)、KKTELQTTNAENKTKKL(SEQ ID NO:345)、KRGINDRNFWRGENGRKTR(SEQ ID NO:346)和RKSGKIAAIVVKRPRK(SEQ ID NO:347)。

43.权利要求42的组合物,其中所述二分核定位信号包含如SEQ ID NO:344中所示的氨基酸序列。

44.权利要求33-43中任一项的组合物,其中所述核碱基编辑器包含与催化无活性的Cas9或Cas9切口酶的N-末端融合的胞嘧啶脱氨酶。

45.权利要求44的组合物,其中所述胞嘧啶脱氨酶选自下组:APOBEC1、APOBEC3、AID和pmCDA1。

46.权利要求44或45的组合物,其中所述核碱基编辑器进一步包含尿嘧啶糖基化酶抑制剂(UGI)。

47.权利要求46的组合物,其中所述UGI包含SEQ ID NO:299-302的氨基酸序列。

48.权利要求33-47中任一项的组合物,其中所述第一核苷酸序列和所述第二核苷酸序列在不同的载体上。

49.权利要求48的组合物,其中每个所述不同的载体是重组腺相关病毒(rAAV)的基因组。

50.权利要求49的组合物,其中所述载体包装在rAAV颗粒中。

51.组合物,其包含:

(i)第一重组腺相关病毒(rAAV)颗粒,其包含编码在C-末端处与内含肽-N融合的核碱基编辑器的N-末端部分的第一核苷酸序列;和

(ii)第二重组腺相关病毒(rAAV)颗粒,其包含编码与所述核碱基编辑器的C-末端部分的N-末端融合的内含肽-C的第二核酸。

52.细胞,其包含权利要求33-51中任一项的组合物。

53.权利要求52的细胞,其中所述核碱基编辑器的N-末端部分和所述核碱基编辑器的C-末端部分接合在一起形成所述核碱基编辑器。

54.权利要求52或权利要求53的细胞,其中所述细胞是原核细胞。

55.权利要求54的细胞,其中所述细胞是细菌细胞。

56.权利要求52或权利要求53的细胞,其中所述细胞是真核细胞。

57.权利要求56的细胞,其中所述细胞是酵母细胞、植物细胞或哺乳动物细胞。

58.权利要求56的细胞,其中所述细胞是人细胞。

59.试剂盒,其包含权利要求33-51中任一项的组合物。

60.方法,其包括:

使细胞与权利要求1-24中任一项的组合物接触,其中所述接触导致所述第一核苷酸序列和所述第二核苷酸序列递送到所述细胞中,并且其中所述Cas9蛋白的N-末端部分和所述

Cas9蛋白的C-末端部分接合以形成Cas9蛋白。

61.方法,其包括:

使细胞与权利要求33-51中任一项的组合物接触,其中所述接触导致所述第一核苷酸序列和所述第二核苷酸序列递送到所述细胞中,并且其中所述核碱基编辑器的N-末端部分和所述核碱基编辑器的C-末端部分接合以形成核碱基编辑器。

62.方法,其包括:

向有此需要的受试者施用治疗有效量的权利要求1-24和33-51中任一项的组合物。

63.权利要求62的方法,其中所述受试者具有疾病或病症。

64.权利要求63的方法,其中所述疾病或病症选自下组:囊性纤维化、苯丙酮尿症、表皮松解性角化过度(EHK)、慢性阻塞性肺病(COPD)、Charcot-Marie-Toot疾病4J.型、神经母细胞瘤(NB)、血管性血友病(vWD)、先天性肌强直、遗传性肾淀粉样变性、扩张型心肌病、遗传性淋巴水肿、家族性阿尔茨海默病、朊病毒病、慢性婴儿神经性皮肤关节综合征(CINCA)和结蛋白相关性肌病(DRM)。

核碱基编辑器的AAV递送

[0001] 相关申请

[0002] 根据35U.S.C.§119(e),本申请要求2016年10月14日提交的美国临时申请U.S.S.N.62/408,575和2017年3月23日提交的美国临时申请U.S.S.N.62/475,780的优先权,其每一个通过引用并入本文。

[0003] 政府支持

[0004] 本发明是在国立卫生研究院(National Institutes of Health)授予的拨款号R35 GM118062和R01 EB022376下得到政府支持完成的。政府拥有本发明的某些权利。

[0005] 背景

[0006] 最近已经在包括基因疗法在内的广泛应用中探索了使用CRISPR/Cas9系统的精确基因组靶向技术。Cas9和基于Cas9的基因组编辑剂在基因疗法中的应用的主要限制是Cas9的大小(>4kb),阻碍其经由重组腺相关病毒(rAAV)的有效递送。

[0007] 发明概述

[0008] 本文描述了用于将Cas9蛋白或核碱基编辑器递送至细胞(例如经由重组腺相关病毒载体)的系统、组合物、试剂盒和方法。典型地,Cas9蛋白或核碱基编辑器"分裂"成N-末端部分和C-末端部分。Cas9蛋白或核碱基编辑器的N-末端部分或C-末端部分可以分别与内含肽系统的一个成员融合。当在分开的载体(例如分开的rAAV载体)上递送到一个细胞中并共表达时,得到的融合蛋白可以接合以形成完整的和功能性的Cas9蛋白或核碱基编辑器(例如,经由内含肽介导的蛋白质剪接)。本文进一步提供了在用于分裂的Cas9蛋白或核碱基编辑器的高表达水平的递送载体中调节元件的经验测试。

[0009] 本公开的一些方面提供组合物,其包含:(i)第一核苷酸序列,其编码在其C-末端处与内含肽-N融合的Cas9蛋白的N-末端部分;和(ii)第二核苷酸序列,其编码与所述Cas9蛋白的C-末端部分的N-末端融合的内含肽-C,其中所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号的核苷酸序列可操作地连接。

[0010] 在一些实施方案中,所述Cas9蛋白的N-末端部分包含对应于SEQ ID NO:1的氨基酸1-573或1-637的SEQ ID NO:1-275和394-397的任一个的部分。在一些实施方案中,所述Cas9蛋白的C-末端部分包含对应于SEQ ID NO:1的氨基酸574-1368或638-1368的SEQ ID NO:1-275和394-397的任一个的部分。在一些实施方案中,所述内含肽-N包含如SEQ ID NO:350-351和354-355中所示的氨基酸序列。在一些实施方案中,所述内含肽-C包含如SEQ ID NO:352-353和356-357中所示的氨基酸序列。

[0011] 在一些实施方案中,所述第一核苷酸序列或所述第二核苷酸序列进一步包含与启动子可操作地连接的编码指导RNA(gRNA)的核苷酸。

[0012] 在一些实施方案中,所述第一核苷酸序列或所述第二核苷酸序列进一步包含转录终止子。在一些实施方案中,所述转录终止子是来自bGH基因、hGH基因或SV40基因的转录终止子。在一些实施方案中,所述转录终止子是来自bGH基因的转录终止子。

[0013] 在一些实施方案中,所述第一核苷酸序列或所述第二核苷酸序列进一步包含插入所述转录终止子的5′的土拨鼠肝炎转录后调控元件(WPRE)。

[0014] 在一些实施方案中，所述二分核定位信号包含选自下组的氨基酸序列：KRPAATKKAGQAKKKK(SEQ ID NO:344)、KKTELQTTNAENKTKKL(SEQ ID NO:345)、KRGINDRNFWRGENGRKTR(SEQ ID NO:346)和RKSGKIAAIVVKRPRK(SEQ ID NO:347)。在一些实施方案中，所述二分核定位信号包含如SEQ ID NO:344中所示的氨基酸序列。

[0015] 在一些实施方案中，所述Cas9蛋白是催化无活性的Cas9(dCas9)或Cas9切口酶(nCas9)，并且其中(i)的所述第一核苷酸序列进一步包含编码核碱基修饰酶的核苷酸序列，所述核碱基修饰酶与所述Cas9蛋白的N-末端部分的N-末端融合。

[0016] 在一些实施方案中，所述Cas9蛋白是催化无活性的Cas9(dCas9)或Cas9切口酶(nCas9)，并且其中(ii)的所述第二核苷酸序列进一步包含编码核碱基修饰酶的核苷酸序列，所述核碱基修饰酶与所述Cas9蛋白的C-末端部分的C-末端融合。

[0017] 在一些实施方案中，所述核碱基修饰酶是脱氨酶。在一些实施方案中，所述脱氨酶是胞嘧啶脱氨酶。在一些实施方案中，所述脱氨酶是腺苷脱氨酶。在一些实施方案中，(ii)的所述第二核苷酸序列进一步包含在所述第二核苷酸序列的3′末端处融合的编码尿嘧啶糖基化酶抑制剂(UGI)的核苷酸序列。在一些实施方案中，(i)的所述第一核苷酸序列进一步包含在所述第一核苷酸序列的5′末端处的编码尿嘧啶糖基化酶抑制剂(UGI)的核苷酸序列。在一些实施方案中，所述UGI包含SEQ ID NO:299-302的氨基酸序列。

[0018] 在一些实施方案中，所述第一核苷酸序列和所述第二核苷酸序列在不同的载体上。在一些实施方案中，每个所述不同的载体是重组腺相关病毒(rAAV)的基因组。在一些实施方案中，每个载体包装在rAAV颗粒中。

[0019] 本公开的其他方面提供组合物，其包含：(i)第一重组腺相关病毒(rAAV)颗粒，其包含编码在其C-末端处与内含肽-N融合的Cas9蛋白的N-末端部分的第一核苷酸序列；和(ii)第二重组腺相关病毒(rAAV)颗粒，其包含编码与所述Cas9蛋白的C-末端部分的N-末端融合的内含肽-C的第二核苷酸序列，其中所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号的核苷酸序列可操作地连接。

[0020] 提供细胞，其包含本文所述的组合物。在一些实施方案中，所述Cas9蛋白的N-末端部分和所述Cas9蛋白的C-末端部分接合在一起形成所述Cas9蛋白。在一些实施方案中，所述细胞是原核细胞。在一些实施方案中，所述细胞是细菌细胞。在一些实施方案中，所述细胞是真核细胞。在一些实施方案中，所述细胞是酵母细胞、植物细胞或哺乳动物细胞。在一些实施方案中，所述细胞是人细胞。

[0021] 本文进一步提供的是试剂盒，其包含本文所述的任何组合物。

[0022] 本公开的一些方面提供组合物，其包含：(i)第一核苷酸序列，其编码在其C-末端处与内含肽-N融合的核碱基编辑器的N-末端部分；和(ii)第二核苷酸序列，其编码与所述核碱基编辑器的C-末端部分的N-末端融合的内含肽-C。

[0023] 在一些实施方案中，所述内含肽-N包含如SEQ ID NO:350-351和354-355中所示的氨基酸序列。在一些实施方案中，所述内含肽-C包含如SEQ ID NO:352-353和356-357中所示的氨基酸序列。在一些实施方案中，所述第一核苷酸序列或所述第二核苷酸序列进一步包含与启动子可操作地连接的编码指导RNA(gRNA)的核苷酸。

[0024] 在一些实施方案中，所述第一核苷酸序列或所述第二核苷酸序列进一步包含转录终止子。在一些实施方案中，所述转录终止子是来自bGH基因、hGH基因或SV40基因的转录终

止子。在一些实施方案中，所述转录终止子来自bGH基因。

[0025] 在一些实施方案中，所述第一核苷酸序列或所述第二核苷酸序列进一步包含插入所述转录终止子的5′的土拨鼠肝炎转录后调控元件(WPRE)。

[0026] 在一些实施方案中，所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号的核苷酸序列可操作地连接。在一些实施方案中，所述二分核定位信号包含选自下组的氨基酸序列:KRPAATKKAGQAKKKK(SEQ ID NO:344)、KKTELQTTNAENKTKKL(SEQ ID NO:345)、KRGINDRNFWRGENGRKTR(SEQ ID NO:346)和RKSGKIAAIVVKRPRK(SEQ ID NO:347)。在一些实施方案中，所述二分核定位信号包含如SEQ ID NO:344中所示的氨基酸序列。

[0027] 在一些实施方案中，所述核碱基编辑器包含与催化无活性的Cas9或Cas9切口酶的N-末端融合的胞嘧啶脱氨酶。在一些实施方案中，所述胞嘧啶脱氨酶选自下组:APOBEC1、APOBEC3、AID和pmCDA1。在一些实施方案中，所述核碱基编辑器进一步包含尿嘧啶糖基化酶抑制剂(UGI)。在一些实施方案中，所述UGI包含SEQ ID NO:299-302的氨基酸序列。

[0028] 在一些实施方案中，所述第一核苷酸序列和所述第二核苷酸序列在不同的载体上。在一些实施方案中，每个所述不同的载体是重组腺相关病毒(rAAV)的基因组。在一些实施方案中，所述载体包装在rAAV颗粒中。

[0029] 本公开的其他方面提供组合物，其包含:(i)第一重组腺相关病毒(rAAV)颗粒，其包含编码在其C-末端处与内含肽-N融合的核碱基编辑器的N-末端部分的第一核苷酸序列;和(ii)第二重组腺相关病毒(rAAV)颗粒，其包含编码内含肽-C的第二核酸，所述内含肽-C与所述核碱基编辑器的C-末端部分的N-末端融合。

[0030] 提供细胞，其包含本文所述的任何组合物。在一些实施方案中，所述核碱基编辑器的N-末端部分和所述核碱基编辑器的C-末端部分接合在一起形成所述核碱基编辑器。在一些实施方案中，所述细胞是原核细胞。在一些实施方案中，所述细胞是细菌细胞。在一些实施方案中，所述细胞是真核细胞。在一些实施方案中，所述细胞是酵母细胞、植物细胞或哺乳动物细胞。在一些实施方案中，所述细胞是人细胞。

[0031] 本文进一步提供的是试剂盒，其包含本文所述的任何组合物。

[0032] 本公开的其他方面提供方法，其包括:使细胞与本文所述的任何组合物接触，其中所述接触导致所述第一核苷酸序列和所述第二核苷酸序列递送到所述细胞中，并且其中所述核碱基编辑器的N-末端部分和所述核碱基编辑器的C-末端部分接合以形成核碱基编辑器。

[0033] 本公开的其他方面提供方法，其包括:向有此需要的受试者施用治疗有效量的本文所述的任何组合物。在一些实施方案中，所述受试者具有疾病或病症。

[0034] 在一些实施方案中，所述疾病或病症选自下组:囊性纤维化、苯丙酮尿症、表皮松解性角化过度(EHK)、慢性阻塞性肺病(COPD)、Charcot-Marie-Toot疾病4J型、神经母细胞瘤(NB)、血管性血友病(vWD)、先天性肌强直、遗传性肾淀粉样变性、扩张型心肌病、遗传性淋巴水肿、家族性阿尔茨海默病、朊病毒病、慢性婴儿神经性皮肤关节综合征(CINCA)和结蛋白相关性肌病(DRM)。

[0035] 如下所述，在某些实施方案的详述中阐述了本发明的某些实施方案的细节。根据定义、实施例、附图和权利要求，本发明的其他特征、目的和优点将显而易见。

[0036] 附图简述

[0037] 构成本申请的一部分的附图示出了本发明的若干实施方案，并与说明书一起用于解释本发明的原理。

[0038] 图1A-1C是显示用于使用重组腺相关病毒(rAAV)载体递送到细胞中的“分裂的核碱基编辑器”的图。图1A是核碱基编辑器如何分裂成两部分的示意图。图1B显示AAV递送的分裂的核碱基编辑器可以在细胞中表达两半时进行蛋白质剪接，以形成完整的核碱基编辑器，其具有与作为整体表达的核碱基编辑器相当的活性。图1C显示了经由通过DnaE内含肽介导的蛋白质剪接从两半形成完整的核碱基编辑器。

[0039] 图2显示U1118细胞被含有编码mCherry的核酸的AAV2有效转染。测试了不同的病毒滴度(2.5-10μl，4.5x 10^{11}vg/ml*)，并且所有都导致U118细胞的有效转染。*vg/ml意指每微升含有病毒基因组的颗粒。

[0040] 图3A-3B是显示U118和HEK细胞中通过rAAV递送的分裂的核碱基编辑器的核碱基编辑的高通量序列(HTS)结果的图。脂质转染的核碱基编辑器用作对照。使用靶向PRNP基因中的R37的sgRNA，并对PRNP基因座进行测序。图3A显示了HTS读出，且图3B总结了碱基编辑结果。

[0041] 图4是显示编码分裂的核碱基编辑器的AAV构建体中使用的转录终止子的优化的图。测试了不同大小和来源的转录终止子。bGH转录终止子相对较短并且与较长的终止子序列相比有效地终止转录。因此选择其用于下游实验。

[0042] 图5A-5B是显示在表达人ApoE4 cDNA的小鼠星形胶质细胞中编码分裂的核碱基编辑器的AAV长期(长达15天)转导的核碱基编辑的结果的图。靶碱基在ApoE4中的精氨酸112和精氨酸158的密码子中，其在碱基编辑时转化为半胱氨酸。图5A显示当小鼠星形胶质细胞在10^{10}vg转导时，精氨酸158的编辑随时间增加，而精氨酸112的编辑保持最小。精氨酸158序列的密码子3′的核苷酸序列特征在于侧翼NGG PAM，其允许SpCas9的高活性(具有指导序列GAAGCGCCTGGCAGTGTACC，SEQ ID NO:348)，而精氨酸112的密码子3′的核苷酸序列含有侧翼NAG PAM，其不允许高活性(具有指导序列GACGTGCGCGGCCGCCTGGTG，SEQ ID NO:349)。图5B显示用编码mCherry的rAAV在10^{10}vg转导的细胞(对照)。

[0043] 图6是编码分裂的核碱基编辑器的AAV构建体中核定位信号的优化的示意图。核定位信号控制核输入，其对于作为编辑的先决条件的重组核碱基编辑器与基因组DNA缔合必须发生，并且是该过程中潜在的限速步骤。该示意图显示NLS(和NLS优化)对于将核碱基编辑器输入到细胞核中至关重要。

[0044] 图7是显示使用含有不同核定位信号(NLS)的不同rAAV分裂的核碱基编辑器构建体的碱基编辑的结果的图。

[0045] 图8A-8B是显示使用AAV编码的分裂的核碱基编辑器在解离的小鼠皮质神经元中编辑DNMT1基因的图。

[0046] 图9A-9B是显示使用AAV编码的分裂的核碱基编辑器或脂质转染的DNA编码的核碱基编辑器在小鼠Neuro-2a细胞系中编辑DNMT1基因的图。

[0047] 定义

[0048] 如本文所用，除非上下文另外明确指出，否则单数形式“一种”、“一个”和“该/所述”包括单数和复数。因此，例如，提及“一种试剂”包括单一试剂和多个此类试剂。

[0049] 如本文所用，术语“Cas9”、“Cas9蛋白”或“Cas9核酸酶”是指包含Cas9蛋白(例如来

自多种细菌物种的Cas9核酸酶)、其片段、变体(例如催化无活性的Cas9或Cas9切口酶)或融合蛋白(例如与另一个蛋白结构域融合的Cas9)的RNA引导的核酸酶。Cas9核酸酶有时也称为casn1核酸酶或CRISPR(聚簇规则间隔短回文重复)相关核酸酶。CRISPR是适应性免疫系统,其提供针对移动遗传元件(病毒、可转座元件和接合质粒)的保护。CRISPR簇含有间隔区,与先前的移动元件互补的序列,并靶向侵入核酸。CRISPR簇得以转录并加工成CRISPR RNA(crRNA)。在II型CRISPR系统中,对pre-crRNA的正确加工需要反式编码的小RNA(tracrRNA)、内源性核糖核酸酶3(rne)和Cas9蛋白。tracrRNA充当用于pre-crRNA的核糖核酸酶3辅助加工的指导。随后,Cas9/crRNA/tracrRNA以内切核水解方式切割与间隔区互补的线性或环状dsDNA靶标。在自然界中,DNA结合和切割通常需要蛋白质和这两种RNA。然而,单一指导RNA("sgRNA"或简称"gNRA")可以经工程化以将crRNA和tracrRNA两者的方面并入单一RNA种类中。实施例1中提供了Cas9蛋白及其各自的氨基酸序列的非限制性实例。

[0050] 核酸酶无活性的Cas9蛋白可以可互换地称为"dCas9"蛋白(代表核酸酶-"死亡的"Cas9)。用于生成具有无活性的DNA切割结构域的Cas9蛋白(或其片段)的方法是已知的(参见例如Jinek et al.,Science.337:816-821(2012);Qi et al.,(2013)Cell.28;152(5):1173-83,通过引用并入本文)。例如,已知Cas9的DNA切割结构域包括两个亚结构域,即HNH核酸酶亚结构域和RuvC1亚结构域。HNH亚结构域切割与gRNA互补的链,而RuvC1亚结构域切割非互补链。这些亚结构域内的突变可以沉默Cas9的核酸酶活性。例如,突变D10A和H840A使酿脓链球菌Cas9的核酸酶活性完全失活(Jinek et al.,Science.337:816-821(2012);Qi et al.,Cell.28;152(5):1173-83(2013)。基于本公开和本领域的知识,另外的合适的核酸酶无活性的dCas9结构域对于本领域技术人员而言将是显而易见的,并且在本公开的范围内。此类另外的示例性合适的核酸酶无活性的Cas9结构域包括但不限于D10A/H840A、D10A/D839A/H840A和D10A/D839A/H840A/N863A突变体结构域(参见例如Prashant et al.,Nature Biotechnology.2013;31(9):833-838,通过引用并入本文)。

[0051] 在一些实施方案中,Cas9切口酶用作核碱基编辑器的一部分。Cas9切口酶能够切割双链DNA的一条链。可以通过将失活的突变引入到HNH结构域或RuvC1结构域中来生成Cas9切口酶。例如,可以在酿脓链球菌Cas9的RuvC1结构域中引入失活的突变(D10A),而HNH结构域保持有活性,即位置840处的残基保持为组氨酸。此类Cas9变体能够基于gRNA确定的靶序列在特定位置处生成单链DNA断裂(切口)。本领域技术人员能够鉴定任何已知Cas9蛋白的RuvC1和HNH结构域中的催化残基并引入失活的突变以生成相应的dCas9或nCas9。

[0052] "分裂的Cas9蛋白"或"分裂的Cas9"是指作为由两个单独的核苷酸序列编码的N-末端部分(也称为N-末端半)和C-末端部分(也称为C-末端半)提供的Cas9蛋白。对应于Cas9蛋白的N-末端部分和C-末端部分的多肽可以组合(接合)以形成完整的Cas9蛋白。已知Cas9蛋白由通过无序接头连接的双叶结构组成(例如,如Nishimasu et al.,Cell,Volume 156,Issue 5,pp.935-949,2014中所述,其通过引用并入本文)。在一些实施方案中,"分裂"发生在两个叶之间,生成Cas9蛋白质的两个部分,每个部分含有一个叶。

[0053] "内含肽"是蛋白质的区段,其能够切除自身并在已知为蛋白质剪接的过程中用肽键接合剩余部分(外显肽)。内含肽也称为"蛋白质内含子"。内含肽切除自身并接合蛋白质的剩余部分的过程在本文中称为"蛋白质剪接"或"内含肽介导的蛋白质剪接"。在一些实施方案中,前体蛋白(在内含肽介导的蛋白质剪接之前含有内含肽的蛋白质)的内含肽来自两

个基因。此类内含肽在本文中称为分裂的内含肽。例如，在蓝细菌中，DnaE，DNA聚合酶III的催化亚基α，由两个单独的基因dnaE-n和dnaE-c编码。由dnaE-n基因编码的内含肽在本文中称为“内含肽-N”。由dnaE-c基因编码的内含肽在本文中称为“内含肽-C”。

[0054] 也可以使用其他内含肽系统。例如，已经描述了基于dnaE内含肽、Cfa-N和Cfa-C内含肽对的合成内含肽(例如在Stevens et al.,J Am Chem Soc.2016Feb 24;138(7):2162-5中，其通过引用并入本文)。可以根据本公开使用的内含肽对的非限制性实例包括：Cfa DnaE内含肽、Ssp GyrB内含肽、Ssp DnaX内含肽、Ter DnaE3内含肽、Ter ThyX内含肽、Rma DnaB内含肽和Cne Prp8内含肽(例如，如美国专利8,394,604中所述，其通过引用并入本文。

[0055] 提供了内含肽的示例性核苷酸和氨基酸序列。

[0056] DnaE内含肽-N DNA：

[0057]

TGCCTGTCATACGAAACCGAGATACTGACAGTAGAATATGGCCTTCTG
CCAATCGGGAAGATTGTGGAGAAACGGATAGAATGCACAGTTTACTCT
GTCGATAACAATGGTAACATTTATACTCAGCCAGTTGCCCAGTGGCAC
GACCGGGGAGAGCAGGAAGTATTCGAATACTGTCTGGAGGATGGAAG
TCTCATTAGGGCCACTAAGGACCACAAATTTATGACAGTCGATGGCCA
GATGCTGCCTATAGACGAAATCTTTGAGCGAGAGTTGGACCTCATGCG
AGTTGACAACCTTCCTAAT (SEQ ID NO: 350)

[0058] DnaE内含肽-N蛋白质：

[0059]

CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRG
EQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLP
N (SEQ ID NO: 351)

[0060] DnaE内含肽-C DNA：

[0061]

ATGATCAAGATAGCTACAAGGAAGTATCTTGGCAAACAAAACGTTTAT
GATATTGGAGTCGAAAGAGATCACAACTTTGCTCTGAAGAACGGATTC
ATAGCTTCTAAT (SEQ ID NO: 352)

[0062] 内含肽-C：

[0063]

MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN (SEQ ID NO: 353)

[0064] Cfa-N DNA：

[0065]

TGCCTGTCTTATGATACCGAGATACTTACCGTTGAATATGGCTTCTTGC
CTATTGGAAAGATTGTCGAAGAGAGAATTGAATGCACAGTATATACTG
TAGACAAGAATGGTTTCGTTTACACACAGCCCATTGCTCAATGGCACA
ATCGCGGCGAACAAGAAGTATTTGAGTACTGTCTCGAGGATGGAAGCA
TCATACGAGCAACTAAAGATCATAAATTCATGACCACTGACGGGCAGA
TGTTGCCAATAGATGAGATATTCGAGCGGGGCTTGGATCTCAAACAAG
TGGATGGATTGCCA (SEQ ID NO: 354)

[0066] Cfa-N蛋白质：

[0067]

CLSYDTEILTVEYGFLPIGKIVEERIECTVYTVDKNGFVYTQPIAQWHNRG EQEVFEYCLEDGSIIRATKDHKFMTTDGQMLPIDEIFERGLDLKQVDGLP (SEQ ID NO: 355)

[0068] Cfa-C DNA：

[0069]

ATGAAGAGGACTGCCGATGGATCAGAGTTTGAATCTCCCAAGAAGAA GAGGAAAGTAAAGATAATATCTCGAAAAAGTCTTGGTACCCAAAATGT CTATGATATTGGAGTGGAGAAAGATCACAACTTCCTTCTCAAGAACGG TCTCGTAGCCAGCAAC (SEQ ID NO: 356)

[0070] Cfa-C蛋白质：

[0071]

MKRTADGSEFESPKKKRKVKIISRKSLGTQNVYDIGVEKDHNFLLKNGLV ASN (SEQ ID NO: 357)

[0072] 内含肽-N和内含肽-C可以分别与分裂的Cas9的N-末端部分和分裂的Cas9的C-末端部分融合，用于接合分裂的Cas9的N-末端部分和分裂的Cas9的C-末端部分。例如，在一些实施方案中，内含肽-N与分裂的Cas9的N-末端部分的C-末端融合，即形成N-[分裂的Cas9的N-末端部分]-[内含肽-N]-C的结构。在一些实施方案中，内含肽-C与分裂的Cas9的C-末端部分的N-末端融合，即形成N-[内含肽-C]-[分裂的Cas9的C-末端部分]-C的结构。用于接合内含肽融合的蛋白质(例如分裂的Cas9)的内含肽介导的蛋白质剪接的机制是本领域已知的，例如，如Shah et al.,Chem Sci.2014;5(1):446-461中所述，其通过引用并入本文。

[0073] 本文中，“核碱基编辑器”是指编辑核苷酸碱基的蛋白质。“编辑”是指将一个核碱基转化为另一个核碱基(例如，A至G、A至C、A至T、C至T、C至G、C至A、G至A、G至C、G至T、T至A、T至C、T至G)。在一些实施方案中，核碱基编辑器是大分子或大分子复合物，其主要导致(例如超出80％、超出85％、超出90％、超出95％、超出99％、超出99.9％或100％)多核酸序列中的核碱基转化为另一个核碱基(即，转换或颠换)，使用以下的组合：1)核苷酸-、核苷-或核碱基-修饰酶和2)可以编程以与特定核酸序列结合的核酸结合蛋白。

[0074] 在一些实施方案中，核碱基编辑器包含将其引导至靶序列的DNA结合结构域(例如，可编程DNA结合结构域，如dCas9或nCas9)。在一些实施方案中，核碱基编辑器包含与可编程DNA结合结构域(例如dCas9或nCas9)融合的核碱基修饰酶。“核碱基修饰酶”是可以修饰核碱基并将一个核碱基转化为另一个核碱基的酶(例如，脱氨酶如胞嘧啶脱氨酶或腺苷脱氨酶)。在一些实施方案中，核碱基编辑器可以靶向核酸序列中的胞嘧啶(C)碱基并将C转化为胸腺嘧啶(T)碱基。在一些实施方案中，C至T编辑通过脱氨酶，例如胞嘧啶脱氨酶进行。还可以考虑可以进行其他类型的碱基转化(例如，腺苷(A)至鸟嘌呤(G)，C至G)的碱基编辑器。

[0075] 在一些实施方案中，将C转化为T的核碱基编辑器包含胞嘧啶脱氨酶。“胞嘧啶脱氨酶”是指催化化学反应“胞嘧啶+H_2O→尿嘧啶+NH_3”或“5-甲基-胞嘧啶+H_2O→胸腺嘧啶+NH_3”的酶。如从反应式中可以明显看出的，此类化学反应导致C至U/T核碱基变化。在基因的背景

中，此类核苷酸变化或突变可以继而导致蛋白质中的氨基酸变化，其可能影响蛋白质的功能，例如功能的丧失或功能的获得。在一些实施方案中，C至T核碱基编辑器包含与胞嘧啶脱氨酶融合的dCas9或nCas9。在一些实施方案中，胞嘧啶脱氨酶结构域与dCas9或nCas9的N-末端融合。在一些实施方案中，核碱基编辑器进一步包含抑制尿嘧啶糖基化酶和/或核定位信号的结构域。此类核碱基编辑器已在本领域中描述，例如在2015年6月18日公开的美国专利9,068,179、美国专利申请公开US 2015/0166980；2015年6月18日公开的US 2015/0166981；2015年6月18日公开的US 2015/0166982；2015年6月18日公开的US 2015/0166984；和2015年6月18日公开的US2015/0165054；以及2015年10月23日提交美国临时申请，U.S.S.N.62/245,828；2016年1月15日提交的U.S.S.N.62/279,346；2016年3月22日提交的U.S.S.N.62/311,763；2016年4月13日提交的U.S.S.N.62/322,178；2016年6月30日提交的U.S.S.N.62/357,352；2016年8月3日提交的U.S.S.N.62,370,700；2016年9月22日提交的U.S.S.N.62/398,490；和2016年10月14日提交的U.S.S.N.62/408,686；2016年10月22日提交的PCT申请PCT/US2016/058344，2016年10月22日提交的美国专利申请，U.S.S.N.15/311,852中；以及在Komor et al.，Nature，Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage，533，420-424(2016)中，其每一个的全部内容通过引用并入本文。

[0076] 在一些实施方案中，核碱基编辑器将A转化为G。在一些实施方案中，核碱基编辑器包含腺苷脱氨酶。“腺苷脱氨酶”是参与嘌呤代谢的酶。需要它用于从食物中分解腺苷以及用于组织中核酸的转换。它在人中的主要功能是免疫系统的发育和维持。腺苷脱氨酶在DNA的背景中催化腺苷的水解脱氨基作用(形成肌苷，其如鸟嘌呤(G)一样碱基配对)。没有已知的作用于DNA的腺苷脱氨酶。相反，已知的腺苷脱氨酶仅作用于RNA(tRNA或mRNA)。已经描述了接受DNA底物并将dA脱氨基成脱氧肌苷，并且在此用于腺苷核碱基编辑器中的演化的脱氧腺苷脱氨酶，例如在2016年8月3日提交的美国临时申请，U.S.S.N.62/370,684；2017年2月3日提交的美国临时申请，U.S.S.N.62/370,684，2017年3月20日提交的美国临时申请，U.S.S.N.62/473,714，和2017年8月3日提交的PCT申请PCT/US2017/045381；其每一个通过引用并入本文。实施例1中提供了接受DNA作为底物的演化的腺苷脱氨酶的非限制性实例。

[0077] 在一些实施方案中，腺苷脱氨酶是大肠杆菌TadA(SEQ ID NO:314)。表2中提供了ecTadA中可能的突变和表达包含经修饰的ecTadA的核碱基编辑器的构建体。实施例1中提供了示例性EcTadA突变体和包含此类突变体的核碱基编辑器的序列。

[0078] 表2:用于A至G核碱基编辑器的EcTadA突变体

[0079]

名称	构建体构造	TadA中的突变
pNMG-142	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	野生型
pNMG-143	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	D108N
pNMG-144	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	A106V_D108N
pNMG-145	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	D108G
pNMG-146	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	R107C_D108N
pNMG-147	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	D108V
pNMG-155	pCMV_ecTadA_XTEN_死亡的Cas9_SGGS_UGI_NLS	D108N
pNMG-156	pCMV_ecTadA_XTEN_nCas9_SGGS_UGI_SGGS_NLS	D108N
pNMG-157	pCMV_ecTadA_XTEN_死亡的Cas9_SGGS_UGI_SGGS_NLS	D108G
pNMG-158	pCMV_ecTadA_XTEN_nCas9_SGGS_UGI_SGGS_NLS	D108G
pNMG-160	pCMV_ecTadA_XTEN_nCas9_SGGS_AAG*(E125Q)_SGGS_NLS	D108N
pNMG-161	pCMV_ecTadA_XTEN_Cas9n_SGGS_EndoV*(D35A)_NLS	D108N
pNMG-162	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	H8Y_D108N_S127S_D147Y_Q154H
pNMG-163	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	H8Y_R24W_D108N_N127S_D147Y_E155V

[0080]

pNMG-164	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	D108N_D147Y_E155V
pNMG-165	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	H8Y_D108N_S127S
pNMG-171	pCMV_Cas9n_XTEN_ecTadA_SGGS_NLS	野生型
pNMG-172	pCMV_Cas9n_XTEN_ecTadA_SGGS_NLS	D108N
pNMG-173	pCMV_Cas9n_XTEN_ecTadA_SGGS_NLS	H8Y_D108N_N127S_D147Y_Q154H
pNMG-174	pCMV_Cas9n_XTEN_ecTadA_SGGS_NLS	H8Y_R24W_D108N_N127S_D147Y_E155V
pNMG-175	pCMV_Cas9n_XTEN_ecTadA_SGGS_NLS	D108N_D147Y_E155V
pNMG-176	pCMV_Cas9n_XTEN_ecTadA_SGGS_NLS	H8Y_D108N_S127S
pNMG-177	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-178	pCMV_ecTadA_XTEN_Cas9n_SGGS_UGI_SGGS_NLS	D108N_D147Y_E155V
pNMG-179	pCMV_ecTadA_XTEN_Cas9n_SGGS_AAG*(E125Q)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-180	pCMV_ecTadA_XTEN_Cas9n_SGGS_UGI_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-181	pCMV_ecTadA_XTEN_Cas9n_SGGS_AAG*(E125Q)_SGGS_NLS	D108N_D147Y_E155V
pNMG-182	pCMV_ecTadA_SGGS_nCas9_SGGS_NLS	D108N_D147Y_E155V
pNMG-183	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	D108N_D147Y_E155V
pNMG-235	pCMV_ecTadA_XTEN_Cas9n_XTEN_AAG*(E125A)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-236	pCMV_ecTadA_XTEN_Cas9n_XTEN_AAG*(E125Q)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-237	pCMV_ecTadA_XTEN_Cas9n_XTEN_AAG*(wt)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-238	pCMV_AAG*(E125A)_XTEN_ecTadA_XTEN_Cas9n_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-239	pCMV_AAG*(wt)_XTEN_ecTadA_XTEN_Cas9n_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-240	pCMV_ecTadA_XTEN_Cas9n_XTEN_EndoV*(D35A)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-241	pCMV_ecTadA_XTEN_Cas9n_XTEN_EndoV*(wt)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-242	pCMV_EndoV*(D35A)_XTEN_ecTadA_XTEN_Cas9n_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-243	pCMV_EndoV*(wt)_XTEN_ecTadA_XTEN_Cas9n_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-247	pCMV_ecTadA_XTEN_Cas9（野生型）_SGGS_NLS	野生型
pNMG-248	pCMV_ecTadA_XTEN_Cas9（野生型）_SGGS_NLS	D108N_D147Y_E155V
pNMG-249	pCMV_ecTadA_XTEN_Cas9（野生型）_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-250	pCMV_ecTadA_XTEN_Cas9（野生型）_SGGS_UGI_SGGS_NLS	D108N_D147Y_E155V

[0081]

pNMG-251	pCMV_ecTadA_XTEN_Cas9 (野生型)_SGGS_AAG*(E125Q)_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-274	pCMV_ecTadA_SGGS_NLS (无Cas9融合)	野生型
pNMG-275	pCMV_ecTadA_SGGS_NLS (无Cas9融合)	A106V_D108N_D147Y_E155V
pNMG-276	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(野生型)+(野生型)
pNMG-277	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-278	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	D108Q_D147Y_E155V
pNMG-279	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	D108M_D147Y_E155V
pNMG-280	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	D108L_D147Y_E155V
pNMG-281	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	D108K_D147Y_E155V
pNMG-282	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	D108I_D147Y_E155V
pNMG-283	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	D108F_D147Y_E155V
pNMG-284	pCMV_ecTadA_较长接头(92 a.a.)_ecTadA_XTEN_nCas9_SGGS_NLS	(野生型) + (A106V_D108N_D147Y_E155V)
pNMG-285	pCMV_ecTadA_较长接头(92 a.a.)_ecTadA_XTEN_nCas9_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y)
pNMG-285b	pCMV_ecTadA_较长接头(92 a.a.)_ecTadA_XTEN_nCas9_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-286	pCMV_ecTadA_XTEN_nCas9_SGGS_NLS	A106V_D108M_D147Y_E155V
pNMG-287	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN-nCas9 (金黄色葡萄球菌)_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-289	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_UGI_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-290	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-293	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	E59A_A106V_D108N_D147Y_E155V
pNMG-294	pCMV_ecTadA_XTEN_Cas9n_SGGS_NLS	E59A
pNMG-295	pCMV_ecTadA_SGGS_NLS (无Cas9融合)	E59A
pNMG-296	pCMV_ecTadA_SGGS_NLS (无Cas9融合)	E59A cat 死亡的 _A106V_D108N_D147Y_E155V
pNMG-297	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (野生型)
pNMG-298	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(D108M_D147Y_E155V) + (D108M_D147Y_E155V)
pNMG-320	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(野生型) + (A106V_D108N_D147Y_E155V)

[0082]

pNMG-321	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(E59A_A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-322	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_XTEN_nCas9_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (E59A_A106V_D108N_D147Y_E155V)
pNMG-335	pCMV_TadA3p-XTEN-TadA2p-XTEN-nCas9-NLS	野生型
pNMG-336	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y
pNMG-337	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_SGGS_NLS	A106V_D108N_D147Y_E155V
pNMG-338	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_D147Y_E155V_I156F
pNMG-339	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y)
pNMG-340	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-341	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_UGI_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-345	pCMV_金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-nCas9_SGGS_NLS	野生型
pNMG-346	pCMV_金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-nCas9_SGGS_NLS	(D108N) + (D108N)
pNMG-347	pCMV_金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-nCas9_SGGS_NLS	(D107A_D018N) + (D107A_D108N)
pNMG-348	pCMV_金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-nCas9_SGGS_NLS	(G26P_D107A_D108N) + (G26P_D107A_D108N)
pNMG-349	pCMV_金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-nCas9_sGGS_NLS	(G26P_D107A_D108N_S142A) +(G26P_D107A_D108N_S142A)

[0083]

pNMG-350	pCMV_金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-金黄色葡萄球菌TadA-(SGGS)2-XTEN-(SGGS)2-nCas9_SGGS_NLS	(D104A_D108N_S142A) + (D107A_D108N_S142A)
pNMG-351	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F)
pNMG-352	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F)
pNMG-353	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D147Y_E155V_I156F)
pNMG-354	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F)
pNMG-355	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F)
pNMG-356	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F)
pNMG-357	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_A142N_A143L_D147Y_E155V_I156F)
pNMG-358	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26G_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F)
pNMG-359	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25A_R26G_L84F_A106V_R107N_D108N_H123Y_A142N_A143E_D147Y_E155V_I156F)
pNMG-360	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F) +(R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F)
pNMG-361	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F) X 2
pNMG-362	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D147Y_E155V_I156F) X 2
pNMG-363	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F) X 2
pNMG-364	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F) X 2

[0084]

pNMG-365	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F) X 2
pNMG-366	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_A142N_A143L_D147Y_E155V_I156F) X 2
pNMG-367	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R26G_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F) X 2
pNMG-368	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25A_R26G_L84F_A106V_R107N_D108N_H123Y_A142N_A143E_D147Y_E155V_I156F) X 2
pNMG-369	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y)
pNMG-370	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(A106V_D108N_D147Y_E155V) + (A106V_D108N_D147Y_E155V)
pNMG-371	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-372	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	A106V_D108N_A142N_D147Y_E155V
pNMG-373	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	R26G_A106V_D108N_A142N_D147Y_E155V
pNMG-374	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	E25D_R26G_A106V_R107K_D108N_A142N_A143G_D147Y_E155V
pNMG-375	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	R26G_A106V_D108N_R107H_A142N_A143D_D147Y_E155V
pNMG-376	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	E25D_R26G_A106V_D108N_A142N_D147Y_E155V
pNMG-377	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	A106V_R107K_D108N_A142N_D147Y_E155V
pNMG-378	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	A106V_D108N_A142N_A143G_D147Y_E155V
pNMG-379	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	A106V_D108N_A142N_A143L_D147Y_E155V
pNMG-382	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	A106V_D108N_A142N_D147Y_E155V X 2
pNMG-383	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	R26G_A106V_D108N_A142N_D147Y_E155V X 2
pNMG-384	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	E25D_R26G_A106V_R107K_D108N_A142N_A143G_D147Y_E155V X 2

[0085]

pNMG-385	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	R26G_A106V_D108N_R107H_A142N_A143D_D147Y_E155V X 2
pNMG-386	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	E25D_R26G_A106V_D108N_A142N_D147Y_E155V X 2
pNMG-387	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	A106V_R107K_D108N_A142N_D147Y_E155V X 2
pNMG-388	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	A106V_D108N_A142N_A143G_D147Y_E155V X 2
pNMG-389	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	A106V_D108N_A142N_A143L_D147Y_E155V X 2
pNMG-391	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N
pNMG-392	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	N37T_P48T_M70L_L84F_A106V_D108N_H123Y_D147Y_I49V_E155V_I156F
pNMG-393	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T
pNMG-394	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_L84F_A106V_D108N_H123Y_D147Y_Q154H_E155V_I156F
pNMG-395	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	N72S_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F
pNMG-396	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_P48L_L84F_A106V_D108N_H123Y_E134G_D147Y_E155V_I156F
pNMG-397	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N
pNMG-398	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
pNMG-399	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T
pNMG-400	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	N37S_R51H_D77G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F
pNMG-401	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	R51L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N

[0086]

pNMG-402	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N) x 2
pNMG-403	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(N37T_P48T_M70L_L84F_A106V_D108N_H123Y_D147Y_I49V_E155V_I156F) x 2
pNMG-404	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T) x 2
pNMG-405	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_L84F_A106V_D108N_H123Y_D147Y_Q154H_E155V_I156F) x 2
pNMG-406	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(N72S_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F) x 2
pNMG-407	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_P48L_L84F_A106V_D108N_H123Y_E134G_D147Y_E155V_I156F) x 2
pNMG-408	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N) x 2
pNMG-409	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F) x 2
pNMG-410	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T) x 2
pNMG-411	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(N37S_R51H_D77G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) x 2
pNMG-412	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R51L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N) x 2
pNMG-440	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E
pNMG-441	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_G67V_L84F_A106V_D108N_H123Y_S146T_D147Y_E155V_I156F
pNMG-442	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	Q71L_L84F_A106V_D108N_H123Y_L137M_A143E_D147Y_E155V_I156F
pNMG-443	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	E25G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L
pNMG-444	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A91T_F104I_A106V_D108N_H123Y_D147Y_E155V_I156F

[0087]

pNMG-445	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	N72D_L84F_A106V_D108N_H123Y_G125A_D147Y_E155V_I156F
pNMG-446	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	P48S_L84F_S97C_A106V_D108N_H123Y_D147Y_E155V_I156F
pNMG-447	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	W23G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F
pNMG-448	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	D24G_P48L_Q71R_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L
pNMG-449	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E) x 2
pNMG-450	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_G67V_L84F_A106V_D108N_H123Y_S146T_D147Y_E155V_I156F) x 2
pNMG-451	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(Q71L_L84F_A106V_D108N_H123Y_L137M_A143E_D147Y_E155V_I156F) x 2
pNMG-452	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(E25G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L) x 2
pNMG-453	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A91T_F104I_A106V_D108N_H123Y_D147Y_E155V_I156F) x 2
pNMG-454	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(N72D_L84F_A106V_D108N_H123Y_G125A_D147Y_E155V_I156F) x 2
pNMG-455	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48S_L84F_S97C_A106V_D108N_H123Y_D147Y_E155V_I156F) x 2
pNMG-456	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(W23G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) x 2
pNMG-457	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(D24G_P48L_Q71R_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L) x 2
pNMG-473	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F
pNMG-474	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F) x 2
pNMG-475	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(野生型) + (A106V_D108N_D147Y_E155V)

[0088]

pNMG-476	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(野生型) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-477	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(野生型) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-478	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(野生型) + (N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T)
pNMG-479	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(野生型) + (L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T)
pNMG-480	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	野生型
pNMG-481	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_NLS	A106V_D108N
pNMG-482	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	野生型 + 野生型
pNMG-483	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(A106V_D108N) x2
pNMG-484	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(野生型) + (A106V_D108N)
pNMG-485	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N
pNMG-486	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K161T
pNMG-487	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_D147Y_E155V_I156F
pNMG-488	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K161T
pNMG-489	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T
pNMG-490	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K161T
pNMG-491	pCMV_ecTadA_(SGGS)2-XTEN-(SGGS)2_Cas9n_SGGS_UGI_SGGS_NLS	L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E
pNMG-492	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F)
pNMG-493	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E)

[0089]

pNMG-494	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)
pNMG-495	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K161T)
pNMG-496	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) +(L84F_A106V_D108N_D147Y_E155V_I156F)
pNMG-497	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K161T)
pNMG-498	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T)
pNMG-499	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K161T)
pNMG-500	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E)
pNMG-513	pCMV_ecTadA-92 a.a.-ecTadA-32 a.a._nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-514	pCMV_ecTadA-92 a.a.-ecTadA-32 a.a._nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-515	pCMV_ecTadA-92 a.a.-ecTadA- 64 a.a._nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-516	pCMV_ecTadA-92 a.a.-ecTadA- 64 a.a._nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-517	pCMV_ecTadA- 32 a.a.-ecTadA- 64 a.a._nCas9_SGGS_NLS	(wt) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-518	pCMV_ecTadA- 32 a.a.-ecTadA- 64 a.a._nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-519	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	R74Q
pNMG-520	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	R74Q L84F_A106V_D108N_H123Y_D147Y_E155V_I156F
pNMG-521	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F
pNMG-522	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	R98Q
pNMG-523	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	R129Q

[0090]

pNMG-524	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt + R74Q) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-525	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt + R74Q) + (R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-526	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-527	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt + R98Q) + (L84F_R98Q_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-528	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt + R129Q) + (L84F_A106V_D108N_H123Y_R129Q_D147Y_E155V_I156F)
pNMG-529	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-530	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-543	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F)
pNMG-544	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N)
pNMG-545	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	P48S_A142N
pNMG-546	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	P48T_I49V_A142N
pNMG-547	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F)
pNMG-548	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F) + (P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F))
pNMG-549	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48S_A142N) + (P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F))
pNMG-550	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48S_A142N) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-551	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L1

[0091]

		57N)
pNMG-552	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N) + (P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N)
pNMG-553	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48T_I49V_A142N) + (P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N)
pNMG-554	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(P48T_I49V_A142N) + (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)
pNMG-555	pCMV_ecTadA-24 a.a. 接头-ecTadA-24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-556	pCMV_ecTadA-24 a.a. 接头-ecTadA- 32 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-557	pCMV_ecTadA-24 a.a. 接头-ecTadA- 40 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-558	pCMV_ecTadA- 32 a.a. 接头-ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-559	pCMV_ecTadA- 32 a.a. 接头-ecTadA- 40 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-560	pCMV_ecTadA- 40 a.a. 接头-ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-561	pCMV_ecTadA- 40 a.a. 接头-ecTadA- 32 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-562	pCMV_ecTadA- 40 a.a. 接头-ecTadA- 40 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-563	pCMV_ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	野生型
pNMG-564	pCMV_ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-565	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_XTEN_MBD4_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F

[0092]

		_K157N)
pNMG-566	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_XTEN_TDG_SGGS_NLS	(wt) + (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-572	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-573	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)
pNMG-574	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-575	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)
pNMG-576	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-577	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)
pNMG-578	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-579	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)
pNMG-580	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N) + (H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-581	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-583	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)
pNMG-586	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)

[0093]

pNMG-588	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)
pNMG-603	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-604	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-605	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T)
pNMG-606	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152H_E155V_I156F_K157N)
pNMG-607	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)
pNMG-608	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)
pNMG-609	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_E155V_I156F_K157N)
pNMG-610	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152P_E155V_I156F_K157N)
pNMG-611	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-612	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)
pNMG-613	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) +(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T)
pNMG-614	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152H_E155V_I156F_K157N)
pNMG-615	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)
pNMG-616	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)
pNMG-617	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D1

[0094]

		47Y_E155V_I156F _K157N)
pNMG-618	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152P_E155V_I156F _K157N)
pNMG-619	pCMV_ecTadA- 32 a.a.-_nCas9_SGGS_NLS	(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N)
pNMG-620	pCMV_ecTadA-(SGGS)2-XTEN-(SGGS)2-ecTadA-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS	(wt) + (W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N)
pNMG-621	pCMV_ecTadA- 32 a.a. 接头-ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N)
pNMG-622	pCMV_ecTadA- 32 a.a. 接头-ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P_E155V_I156F _K157N)
pNMG-623	pCMV_ecTadA- 32 a.a. 接头-ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N)
pNMG-624	pCMV_ecTadA- 32 a.a. 接头-ecTadA- 24 a.a. 接头_nCas9_SGGS_NLS	(wt) + (W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N)

[0095] 在一些实施方案中,A至G核碱基编辑器包含与腺苷脱氨酶融合的dCas9或nCas9。此类核碱基编辑器描述于2016年8月3日提交的美国临时申请,U.S.S.N.62/370,684;2017年2月3日提交的美国临时申请,U.S.S.N.62/370,684,2017年3月20日提交的美国临时申请,U.S.S.N.62/473,714,和2017年8月3日提交的PCT申请PCT/US2017/045381中;其每一个通过引用并入本文。

[0096] 在一些实施方案中,A至G核碱基编辑器包含NH2-[第二腺苷脱氨酶]-[第一腺苷脱氨酶]-[dCas9]-COOH的结构。在一些实施方案中,第二腺苷脱氨酶是野生型ecTadA(SEQ ID NO:314)。在一些实施方案中,在每个结构域之间使用接头。在一些实施方案中,接头为32个氨基酸长并且包含SGGSSGGSSGSETPGTSESATPESSGGSSGGS(SEQ ID NO:384)的氨基酸序列。

[0097] 在一些实施方案中,腺苷脱氨酶包含以下各项的一个或多个:SEQ ID NO:314中的W23X、H36X、N37X、P48X、I49X、R51X、N72X、L84X、S97X、A106X、D108X、H123X、G125X、A142X、S146X、D147X、R152X、E155X、I156X、K157X和/或K161X突变,或在另一个腺苷脱氨酶中的一个或多个相应的突变,其中X的存在表示除了野生型腺苷脱氨酶中相应的氨基酸之外的任何氨基酸。在一些实施方案中,腺苷脱氨酶包含以下各项的一个或多个:SEQ ID NO:314中的W23L、W23R、H36L、P48S、P48A、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、R152P、E155V、I156F和/或K157N突变,或在另一个腺苷脱氨酶中的一个或多个相应的突变。

[0098] 在一些实施方案中,腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个或十二个突变或者由之组成:SEQ ID NO:314中的

H36X、P48X、R51X、L84X、A106X、D108X、H123X、S146X、D147X、E155X、I156X和K157X，或在另一个腺苷脱氨酶中的相应的一个或多个突变，其中X表示存在除野生型腺苷脱氨酶中相应的氨基酸之外的任何氨基酸。在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个或十二个突变或者由之组成：SEQ ID NO:314中的H36L、P48S、R51L、L84F、A106V、D108N、H123Y、S146C、D147Y、E155V、I156F和K157N，或在另一个腺苷脱氨酶中的相应的一个或多个突变。在一些实施方案中，腺苷脱氨酶包含以下各项或者由之组成：SEQ ID NO:314中的H36L、P48S、R51L、L84F、A106V、D108N、H123Y、S146C、D147Y、E155V、I156F和K157N突变，或在另一个腺苷脱氨酶中的相应的突变。

[0099] 在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个或十三个突变或者由之组成：SEQ ID NO:314中的H36X、P48X、R51X、L84X、A106X、D108X、H123X、A142X、S146X、D147X、E155X、I156X和K157X，或在另一个腺苷脱氨酶中的相应的一个或多个突变，其中X表示存在除野生型腺苷脱氨酶中相应的氨基酸之外的任何氨基酸。在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个或十三个突变或者由之组成：SEQ ID NO:314中的H36L、P48S、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、E155V、I156F和K157N，或在另一个腺苷脱氨酶中的相应的一个或多个突变。在一些实施方案中，腺苷脱氨酶包含以下各项或者由之组成：SEQ ID NO:314中的H36L、P48S、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、E155V、I156F和K157N突变，或在另一个腺苷脱氨酶中的相应的突变。

[0100] 在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个或十四个突变或者由之组成：SEQ ID NO:314中的W23X、H36X、P48X、R51X、L84X、A106X、D108X、H123X、A142X、S146X、D147X、E155X、I156X和K157X，或在另一个腺苷脱氨酶中的相应的一个或多个突变，其中X表示存在除野生型腺苷脱氨酶中相应的氨基酸之外的任何氨基酸。在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个或十四个突变或者由之组成：SEQ ID NO:314中的W23L、H36L、P48A、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、E155V、I156F和K157N，或在另一个腺苷脱氨酶中的相应的一个或多个突变。在一些实施方案中，腺苷脱氨酶包含以下各项或者由之组成：SEQ ID NO:314中的W23L、H36L、P48A、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、E155V、I156F和K157N突变，或在另一个腺苷脱氨酶中的相应的突变。

[0101] 在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个、十四个或十五个突变或者由之组成：SEQ ID NO:314中的W23X、H36X、P48X、R51X、L84X、A106X、D108X、H123X、A142X、S146X、D147X、R152X、E155X、I156X和K157X，或在另一个腺苷脱氨酶中的相应的一个或多个突变，其中X表示存在除野生型腺苷脱氨酶中相应的氨基酸之外的任何氨基酸。在一些实施方案中，腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个、十四个或十五个突变或者由之组成：SEQ ID NO:314中的W23L、H36L、P48A、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、R152P、E155V、I156F和K157N，或在另一个腺苷脱氨酶中的相应的一个或多个突变。在一些实施方案中，腺

苷脱氨酶包含以下各项或者由之组成:SEQ ID NO:314中的W23L、H36L、P48A、R51L、L84F、A106V、D108N、H123Y、A142N、S146C、D147Y、R152P、E155V、I156F和K157N突变,或在另一个腺苷脱氨酶中的相应的突变。

[0102] 在一些实施方案中,腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个或十四个突变或者由之组成:SEQ ID NO:314中的W23X、H36X、P48X、R51X、L84X、A106X、D108X、H123X、S146X、D147X、R152X、E155X、I156X和K157X,或在另一个腺苷脱氨酶中的相应的一个或多个突变,其中X表示存在除野生型腺苷脱氨酶中相应的氨基酸之外的任何氨基酸。在一些实施方案中,腺苷脱氨酶包含选自以下各项的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个或十四个突变或者由之组成:SEQ ID NO:314中的W23R、H36L、P48A、R51L、L84F、A106V、D108N、H123Y、S146C、D147Y、R152P、E155V、I156F和K157N,或在另一个腺苷脱氨酶中的相应的一个或多个突变。在一些实施方案中,腺苷脱氨酶包含以下各项或者由之组成:SEQ ID NO:314中的W23R、H36L、P48A、R51L、L84F、A106V、D108N、H123Y、S146C、D147Y、R152P、E155V、I156F和K157N突变,或在另一个腺苷脱氨酶中的相应的突变。

[0103] 在一些实施方案中,核碱基编辑器将C转化为G。此类核碱基编辑器描述于2017年3月10日提交的美国临时申请U.S.S.N.62/470,175,2017年3月10日提交的美国临时申请U.S.S.N.62/470,175,通过引用并入本文。

[0104] 实施例1中提供了非限制性、示例性类型的核碱基编辑器(包括C至T、A至G和C至G核碱基编辑器)及其各自的序列。在一些实施方案中,核碱基编辑器是本文所述的核碱基编辑器的变体。例如,在一些实施方案中,核碱基编辑器与本文所述的核碱基编辑器(实施例1中提供的示例性序列)至少50%、至少55%、至少60%、至少65%、至少70%、至少75%至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同。在一些实施方案中,核碱基编辑器包含长度上比本文提供的任何核碱基编辑器更短或更长的氨基酸序列(例如,不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中,核碱基编辑器包含长度上比本文提供的任何核碱基编辑器更短或更长的氨基酸序列(例如,不超过500个氨基酸、不超过450个氨基酸、不超过400个氨基酸、不超过350个氨基酸、不超过300个氨基酸、不超过250个氨基酸、不超过200个氨基酸、不超过200个氨基酸、不超过150个氨基酸、不超过100个氨基酸、不超过50个氨基酸、不超过10个氨基酸、不超过5个氨基酸更长或更短)。

[0105] “脱氨酶”是指催化从分子中除去胺基或催化脱氨基作用(例如通过水解)的酶。在一些实施方案中,脱氨酶是胞苷脱氨酶,分别催化胞苷(C)至尿苷(U)、脱氧胞苷(dC)至脱氧尿苷(dU)或5-甲基-胞苷至胸苷(T,5-甲基-U)的脱氨基作用。随后的DNA修复机制确保dU被T替换,如Komor等人(Nature,Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage,533,420-424(2016),其通过引用并入本文)中所述。在一些实施方案中,脱氨酶是胞嘧啶脱氨酶,催化并促进胞嘧啶至尿嘧啶(例如在RNA中)或胸腺嘧啶(例如在DNA中)的转化。在一些实施方案中,脱氨酶是将A转化为G的腺苷脱氨酶。在一些实施方案中,脱氨酶是来自生物体(例如人、黑猩猩、大猩猩、猴、牛、狗、大鼠或小鼠)的天然存在的脱氨酶。在一些实施方案中,脱氨酶是来自生物体的天然存在的脱氨酶的变体,并且该变体不在自然界中存在。例如,在一些实施方案中,脱氨酶或脱氨酶结构域

与来自生物体的天然存在的脱氨酶至少50%、至少55%、至少60%、至少65%、至少70%、至少75%至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同。在一些实施方案中，脱氨酶包含长度上比本文提供的任何脱氨酶更短或更长的氨基酸序列(例如，不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中，脱氨酶包含长度上比本文提供的任何脱氨酶更短或更长的氨基酸序列(例如，不超过100个氨基酸、不超过90个氨基酸、不超过80个氨基酸、不超过70个氨基酸、不超过60个氨基酸、不超过50个氨基酸、不超过40个氨基酸、不超过30个氨基酸、不超过20个氨基酸、不超过10个氨基酸、不超过5个氨基酸、不超过2个氨基酸更长或更短)。

[0106] “分裂的核碱基编辑器”是指作为由两个单独的核酸编码的N-末端部分(也称为N-末端半)和C-末端部分(也称为C-末端半)提供的核碱基编辑器。对应于核碱基编辑器的N-末端部分和C-末端部分的多肽可以组合以形成完整的核碱基编辑器。在一些实施方案中，对于包含dCas9或nCas9的核碱基编辑器，“分裂”位于dCas9或nCas9结构域中，在分裂的Cas9中如本文所述的位置处。因此，在一些实施方案中，核碱基编辑器的N-末端部分含有分裂的Cas9的N-末端部分，并且核碱基编辑器的C-末端部分含有分裂的Cas9的C-末端部分。类似地，内含肽-N或内含肽-C可以分别与核碱基编辑器的N-末端部分或C-末端部分融合，用于接合核碱基编辑器的N-和C-末端部分以形成完整的核碱基编辑器。

[0107] 当形成连接两个蛋白质或两个蛋白质结构域的肽键时，认为两个蛋白质或蛋白质结构域是“融合的”。在一些实施方案中，接头(例如肽接头)存在于两个蛋白质或两个蛋白质结构域之间。如本文所用，术语“接头”是指连接两个分子或部分，例如融合蛋白的两个结构域，诸如例如核酸酶无活性的Cas9结构域和核酸编辑结构域(例如脱氨酶结构域)的化学基团或分子。通常，接头位于两个基团、分子或其他部分之间或侧翼有两个基团、分子或其他部分，并且经由共价键与每一个连接，从而连接两者。在一些实施方案中，接头是一个氨基酸或多个氨基酸(例如肽或蛋白质)。在一些实施方案中，接头是有机分子、基团、聚合物或化学部分。在一些实施方案中，接头的长度为5-100个氨基酸，例如长度为1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、30-35、35-40、40-45、45-50、50-60、60-70、70-80、80-90、90-100、100-150或150-200个氨基酸。也考虑了更长或更短的接头。

[0108] “尿嘧啶糖基化酶抑制剂(UGI)”是指抑制尿嘧啶-DNA糖基化酶的活性的蛋白质。用于根据本公开的用途的合适的UGI蛋白包括例如Wang et al.,J.Biol.Chem.264:1163-1171(1989);Lundquist et al.,J.Biol.Chem.272:21408-21419(1997);Ravishankar et al.,Nucleic Acids Res.26:4880-4887(1998);和Putnam et al.,J.Mol.Biol.287:331-346(1999)中发表的那些，其每一个通过引用并入本文。实施例1中提供了可以用作本公开的UGI的非限制性、示例性蛋白质及其各自的序列。在一些实施方案中，UGI是来自生物体的天然存在的脱氨酶的变体，并且变体不在自然界中存在。例如，在一些实施方案中，UGI与来自生物体的天然存在的UGI或本文(例如实施例1中)提供的任何UGI至少50%、至少55%、至少60%、至少65%、至少70%、至少75%至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同。在一些实施方案中，UGI包含长度上比本文提供的任何UGI更短或更长的氨基酸序列(例如，不超过30%、不超过25%、不超过

20％、不超过15％、不超过10％、不超过5％、不超过1％更长或更短)。在一些实施方案中，UGI包含长度上比本文提供的任何UGI更短或更长的氨基酸序列(例如，不超过20个氨基酸、不超过15个氨基酸、不超过10个氨基酸、不超过5个氨基酸、不超过2个氨基酸更长或更短)。

[0109] gRNA是CRISPR/Cas系统的组分。本文中的“gRNA”(指导核糖核酸)是指CRISPR靶向RNA(crRNA)和反式激活crRNA(tracrRNA)的融合，为Cas9核酸酶提供靶向特异性和支架/结合能力两者。“crRNA”是赋予靶特异性并且需要tracrRNA与Cas9结合的细菌RNA。“tracrRNA”是将crRNA与Cas9核酸酶连接的细菌RNA，并且通常可以结合任何crRNA。Cas DNA结合蛋白的序列特异性由gRNA确定，gRNA具有与靶DNA序列的核苷酸碱基配对互补性。天然gRNA包含20个核苷酸(nt)的特异性确定序列(SDS)，其指定待靶向的DNA序列，并且紧接着是80nt的支架序列，其将gRNA与Cas9缔合。在一些实施方案中，本公开的SDS具有15至100个核苷酸或更多个的长度。例如，SDS可以具有15至90、15至85、15至80、15至75、15至70、15至65、15至60、15至55、15至50、15至45、15至40、15至35、15至30或15至20个核苷酸的长度。在一些实施方案中，SDS为20个核苷酸长。例如，SDS可以为15、16、17、18、19、20、21、22、23、24或25个核苷酸长。靶DNA序列的至少一部分与gRNA的SDS互补。为了使Cas9成功结合DNA靶序列，靶序列的一个区域与gRNA序列的SDS互补，并且紧接着是正确的前间隔区相邻基序(PAM)序列(例如，用于Cas9的NGG和用于Cpf1的TTN、TTTN或YTN)。在一些实施方案中，SDS与其靶序列100％互补。在一些实施方案中，SDS序列与其靶序列小于100％互补性，因此认为与其靶序列部分互补。例如，靶向序列可以与其靶序列99％、98％、97％、96％、95％、94％、93％、92％、91％或90％互补。在一些实施方案中，模板DNA或靶DNA的SDS可以与gRNA的互补区域相差1、2、3、4或5个核苷酸。

[0110] 除了SDS之外，gRNA包含与Cas9缔合所需的支架序列(对应于天然CRISPR/Cas系统中的tracrRNA)(本文称为“gRNA柄(gRNA handle)”)。在一些实施方案中，gRNA包含结构5′-[SDS]-[gRNA柄]-3′。在一些实施方案中，支架序列包含5′-guuuuagagcuagaaauagcaaguuaaaauaaaggcuaguccguuaucaacuugaaaaaguggcaccgagucggugcuuuuu-3′(SEQ ID NO:358)的核苷酸序列。可以根据本公开使用的其他非限制性、合适的gRNA柄序列列于表1中。

[0111] 表1:指导RNA柄序列

[0112]

生物体	gRNA柄序列	SEQ ID NO
酿脓链球菌	GUUUAAGAGCUAUGCUGGAAAGCCACGGUGAAAAAGUUCAACUAUUGCCUGAUCGGAAUAAAUUUGAACGAUACGACAGUCGGUGCUUUUUUU	359
酿脓链球菌	GUUUAAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUU	360
嗜热链球菌CRISPR1	GUUUUUGUACUCUCAAGAUUCAAUAAUCUUGCAGAAGCUACAAAGAUAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUAUGGCAGGGUGUUUU	361
嗜热链球菌CRISPR3	GUUUUAGAGCUGUGUUGUUUGUUAAAACAACACAGCGAGUUAAAAUAAGGCUUAGUCCGUACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU	362
空肠弯曲杆菌(*C. jejuni*)	AAGAAAUUUAAAAAGGGACUAAAAUAAAGAGUUUGCGGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU	363
新凶手弗朗西斯菌(*F. novicida*)	AUCUAAAAUUAUAAAUGUACCAAAUAAUUAAUGCUCUGUAAUCAUUUAAAAGUAUUUUGAACGGACCUCUGUUUGACACGUCUGAAUAACUAAAA	364

[0113]

嗜热链球菌 *2*	UGUAAGGGACGCCUUACACAGUUACUUA AAUCUUGCAGAAGCUACAAAGAUAAGGC UUCAUGCCGAAAUCAACACCCUGUCAUU UUAUGGCAGGGUGUUUUCGUUAUUU	365
M. mobile	UGUAUUUCGAAAUACAGAUGUACAGUUA AGAAUACAUAAGAAUGAUACAUCACUAA AAAAAGGCUUUAUGCCGUAACUACUACU UAUUUUCAAAAUAAGUAGUUUUUUUU	366
无害李斯特氏菌 (*L. innocua*)	AUUGUUAGUAUUCAAAAUAACAUAGCAA GUUAAAAUAAGGCUUUGUCCGUUAUCAA CUUUUAAUUAAGUAGCGCUGUUUCGGCG CUUUUUUU	367
酿脓链球菌	GUUGGAACCAUUCAAAACAGCAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCUU UUUUU	368
变异链球菌 (*S. mutans*)	GUUGGAAUCAUUCGAAACAACACAGCAA GUUAAAAUAAGGCAGUGAUUUUUAAUCC AGUCCGUACACAACUUGAAAAAGUGCGC ACCGAUUCGGUGCUUUUUUAUUU	369
嗜热链球菌	UUGUGGUUUGAAACCAUUCGAAACAACA CAGCGAGUUAAAAUAAGGCUUAGUCCGU ACUCAACUUGAAAAGGUGGCACCGAUUC GGUGUUUUUUUU	370
脑膜炎奈瑟氏球菌 (*N. meningitides*)	ACAUAUUGUCGCACUGCGAAAUGAGAAC CGUUGCUACAAUAAGGCCGUCUGAAAAG AUGUGCCGCAACGCUCUGCCCCUUAAAG CUUCUGCUUUAAGGGGCA	371
多杀巴斯德氏菌 (*P.*	GCAUAUUGUUGCACUGCGAAAUGAGAGA CGUUGCUACAAUAAGGCUUCUGAAAAGA	372

[0114]

multocida)	AUGACCGUAACGCUCUGCCCCUUGUGAU UCUUAAUUGCAAGGGGCAUCGUUUUU	

[0115] 在一些实施方案中，指导RNA约15-120个核苷酸长并且包含与靶序列互补的至少10个连续核苷酸的序列。在一些实施方案中，指导RNA为15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、97、98、99、100、101、102、103、104、105、106、107、108、109、110、111、112、113、114、115、116、117、118、119或120个核苷酸长。在一些实施方案中，指导RNA包含与靶序列互补的10、11、12、13、14、15、16、17、18、19、20或更多个连续核苷酸的序列。序列互补性是指腺嘌呤和胸腺嘧啶(DNA)或尿嘧啶(RNA)之间以及鸟嘌呤和胞嘧啶之间的独特的相互作用。

[0116] “前间隔区相邻基序”(PAM)通常是位于相邻的核苷酸的序列(例如，在靶序列的10、9、8、7、6、5、4、3、3或1个核苷酸内)。如果PAM序列与靶序列邻接(即，如果没有核苷酸位于PAM序列和靶序列之间)，则PAM序列与靶序列“紧邻”。在一些实施方案中，PAM序列是野生型PAM序列。PAM序列的实例包括但不限于，NGG，NGR，NNGRR(T/N)，NNNNGATT，NNAGAAW，NGGAG，NAAAAC，AWG和CC。在一些实施方案中，PAM序列获得自酿脓链球菌(Streptococcus pyogenes)(例如NGG或NGR)。在一些实施方案中，PAM序列获得自金黄色葡萄球菌(Staphylococcus aureus)(例如NNGRR(T/N))。在一些实施方案中，PAM序列获得自脑膜炎奈瑟氏球菌(Neisseria meningitides)(例如NNNNGATT)。在一些实施方案中，PAM序列获得自嗜热链球菌(Streptococcus thermophiles)(例如NNAGAAW或NGGAG)。在一些实施方案中，PAM序列获得自齿垢密螺旋体(Treponema denticola)(例如NAAAAC)。在一些实施方案中，PAM序列获得自大肠杆菌(Escherichia coli)(例如AWG)。在一些实施方案中，PAM序列获得自铜绿假单胞菌(Pseudomonas auruginosa)(例如CC)。考虑了其他PAM序列。PAM序列通常位于靶序列的下游(即3′)，但在一些实施方案中，PAM序列可以位于靶序列的上游(即5′)。

[0117] “核定位信号”或“NLS”是指“标记”蛋白质用于通过核转运的输入到细胞核中的氨基酸序列。通常，该信号由暴露在蛋白质表面上的一个或多个带正电荷的赖氨酸或精氨酸的短序列组成。可以将一个或多个NLS添加到蛋白质的N-或C-末端，或内部(例如，在两个蛋白质结构域之间)。例如，可以将一个或多个NLS添加到核碱基编辑器的N-或C-末端，或者在核碱基编辑器中的Cas9和脱氨酶之间。在一些实施方案中，可以添加1、2、3、4、5或更多个NLS。核定位序列是本领域已知的，并且对于熟练技术人员将是显而易见的。例如，NLS序列描述于2000年11月23日提交的Plank等人，PCT/EP2000/011690，其内容通过引用并入本文，用于其对示例性的核定位序列的公开内容。在一些实施方案中，NLS包含氨基酸序列PKKKRKV(SEQ ID NO:373)或MDSLLMNRRKFLYQFKNVRWAKGRRETYLC(SEQ ID NO:374)。在一些实施方案中，接头插入在Cas9和脱氨酶之间。

[0118] NLS可以分类为单分或二分。单分NLS的非限制性实例是SV40大T抗原中的序列PKKKRKV(SEQ ID NO:373)。二分NLS通常含有两个碱性氨基酸的簇，由约10个氨基酸的间隔

区分开。二分NLS的一个非限制性实例是核质蛋白的NLS,KRPAATKKAGQAKKKK(间隔区加下划线)(SEQ ID NO:344)。在一些实施方案中,根据本公开使用的NLS是包含KRPAATKKAGQAKKKK(SEQ ID NO:344)的氨基酸序列的核质蛋白的NLS。可以根据本公开使用的其他二分NLS包括但不限于:SV40二分NLS(KRTADGSEFESPKKKRKV(SEQ ID NO:375),例如如Hodel et al.,J Biol Chem.2001 Jan 12;276(2):1317-25中所述,通过引用并入本文);Kanadaptin二分NLS(KKTELQTTNAENKTKKL(SEQ ID NO:345),例如如Hubner et al.,Biochem J.2002Jan 15;361(Pt 2):287-96中所述,通过引用并入本文);甲型流感病毒核蛋白二分NLS(KRGINDRNFWRGENGRKTR(SEQ ID NO:346),例如如Ketha et al.,BMC Cell Biology.2008;9:22中所述,通过引用并入本文);和ZO-2二分NLS(RKSGKIAAIVVKRPRK(SEQ ID NO:347),例如如Quiros et al.,Nusrat A,ed.Molecular Biology of the Cell.2013;24(16):2528-2543中所述,通过引用并入本文)。

[0119] 当两个编码序列“彼此符合读框(in-frame with each other)”并且翻译为融合两个序列的单个多肽时,编码NLS的核苷酸序列与编码与NLS融合的蛋白质(例如Cas9或核碱基编辑器)的核苷酸序列“可操作地连接”。

[0120] 本公开的核酸可以包括一个或多个遗传元件。“遗传元件”是指在核酸表达中起作用的特定核苷酸序列(例如,启动子、增强子、终止子)或编码工程化的核酸的离散的产物(例如,编码指导RNA、蛋白质和/或RNA干扰分子的核苷酸序列)。

[0121] “启动子”是指核酸序列的控制区域,在该控制区域处核酸序列的其余部分的起始和转录速率受到控制。启动子还可以含有亚区域,在该亚区域处调节蛋白和分子可以结合例如RNA聚合酶和其他转录因子。启动子可以是组成型的、诱导型的、可激活的、可抑制的、组织特异性的或其任何组合。启动子驱动表达或驱动其调节的核酸序列的转录。本文中,当启动子处于与其调节的核酸序列缔合的正确的功能位置和方向以控制(“驱动”)该序列的转录起始和/或表达时,认为启动子是“可操作地连接的”。

[0122] 启动子可以是与基因或序列天然缔合的启动子,如可以通过分离位于给定基因或序列的编码区段的上游的5′非编码序列而获得的。此类启动子称为“内源启动子”。在一些实施方案中,编码核酸序列可以位于重组或异源启动子的控制下,所述启动子是指在其天然环境中通常不与编码序列缔合的启动子。此类启动子可以包括其他基因的启动子;从任何其他细胞分离的启动子;和非“天然存在的”合成启动子或增强子,诸如例如含有不同转录调节区域的不同元件和/或通过本领域已知的基因工程化的方法改变表达的突变的合成启动子或增强子。除了以合成方式产生启动子和增强子的核酸序列之外,可以使用重组克隆和/或核酸扩增技术(包括聚合酶链式反应(PCR))产生序列。

[0123] 在一些实施方案中,根据本公开使用的启动子是“诱导型启动子”,其是特征在于当存在诱导物信号、受诱导物信号影响或接触诱导物信号时调节(例如,起始或激活)转录活性的启动子。诱导物信号可以是内源的或通常地外源的条件(例如光)、化合物(例如,化学或非化学化合物)或接触诱导型启动子的蛋白质,以此方式以便有效调节来自诱导型启动子的转录活性。因此,核酸的“调节转录的信号”是指作用于诱导型启动子的诱导物信号。根据所使用的调节系统,调节转录的信号可以激活或灭活转录。转录的激活可以涉及直接作用于启动子以驱动转录或通过使阻止启动子驱动转录的阻遏物失活而间接作用于启动子。相反,转录的失活可以涉及直接作用于启动子以阻止转录或通过激活随后作用于启动

子的阻遏物间接作用于启动子。

[0124] “转录终止子”是导致转录停止的核酸序列。转录终止子可以是单向的或双向的。它包含参与通过RNA聚合酶特异性终止RNA转录物的DNA序列。转录终止子序列通过上游启动子阻止下游核酸序列的转录激活。转录终止子在体内可以是必需的，以达到所期望的表达水平或避免某些序列的转录。当转录终止子能够终止与其连接的序列的转录时，认为转录终止子与核苷酸序列“可操作地连接”。

[0125] 终止子的最常用的类型是正向终止子。当置于通常转录的核酸序列的下游时，正向转录终止子将导致转录中止。在一些实施方案中，提供双向转录终止子，其通常导致正向和反向链两者上的转录终止。在一些实施方案中，提供反向转录终止子，其通常仅终止反向链上的转录。

[0126] 在原核系统中，终止子通常分为两类(1)rho非依赖性终止子和(2)rho依赖性终止子。Rho非依赖性终止子通常由回文序列组成，所述回文序列形成富含G-C碱基对的茎环，然后是几个T碱基。不希望受理论束缚，转录终止的常规模型是茎环导致RNA聚合酶暂停，并且poly-A尾的转录导致RNA:DNA双链体解开并从RNA聚合酶解离。

[0127] 在真核系统中，终止子区域可以包含特定的DNA序列，其允许新转录物的位点特异性切割，从而暴露多腺苷酸化位点。这表示专门的内源聚合酶将一段约200个A残基(polyA)添加至转录物的3′末端。用该polyA尾修饰的RNA分子似乎更稳定并且更有效地翻译。因此，在涉及真核生物的一些实施方案中，终止子可以包含用于切割RNA的信号。在一些实施方案中，终止子信号促进消息(message)的多腺苷酸化。终止子和/或多腺苷酸化位点元件可以用于增强输出核酸水平和/或最小化核酸之间的读数。

[0128] 根据本公开使用的终止子包括本文所述的或本领域普通技术人员已知的任何转录的终止子。终止子的实例包括但不限于基因的终止序列，诸如例如牛生长激素终止子，和病毒终止序列，诸如例如SV40终止子、spy、yejM、secG-leuU、thrLABC、rrnB T1、hisLGDCBHAFI、metZWV、rrnC、xapR、aspA和arcA终止子。在一些实施方案中，终止信号可以是不能转录或翻译的序列，例如由序列截短产生的序列。

[0129] “土拨鼠肝炎病毒(WHP)转录后调节元件(WPRE)”是其在转录时产生增强表达的三级结构的DNA序列。通常用于分子生物学中以增加病毒载体递送的基因的表达。WPRE是具有gamma、alpha和beta组分的三分调节元件。完整的WPRE序列为609bp长：

[0130]

GCTTATCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACT
GGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTT
AATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCT
CCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGT
TGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCC
CACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTC
GCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTG
CCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG
TGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTATGTTGC
CACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTC
AATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTC
TTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGC
CGCCTCCCCGCATCGATACCG (SEQ ID NO:376)。

[0131] “腺相关病毒”或“AAV”是感染人和一些其他灵长类动物的病毒。野生型AAV基因组是单链脱氧核糖核酸(ssDNA),正义或负义。基因组包含两个反向末端重复(ITR)(DNA链的每个末端处一个)和两个开放阅读框(ORF):在ITR之间的rep和cap。rep ORF包含编码AAV生命周期所需的Rep蛋白的四个重叠基因。cap ORF包含编码衣壳蛋白(VP1、VP2和VP3,它们一起相互作用形成病毒衣壳)的重叠基因。VP1、VP2和VP3翻译自一个mRNA转录物,其可以用两种不同的方式进行剪接:可以切除更长或更短的内含子,从而形成mRNA的两个同种型:~2.3kb和~2.6kb长的mRNA同种型。衣壳将大约60个单独的衣壳蛋白亚基的超分子组装形成为能够保护AAV基因组的无包膜、T-1二十面体晶格。成熟衣壳由VP1、VP2和VP3(分别为约87、73和62kDa的分子质量)以1:1:10的比例组成。

[0132] rAAV颗粒可以包含核酸载体(例如,重组基因组),其可以最低限度包含:(a)一个或多个异源核酸区域,其包含编码感兴趣的蛋白质或多肽(例如,分裂的Cas9或分裂的核碱基)的序列或感兴趣的RNA(例如,gRNA),或一个或多个核酸区域,其包含编码Rep蛋白的序列;和(b)一个或多个区域,其包含侧翼于一个或多个核酸区域(例如,异源核酸区域)的反向末端重复(ITR)序列(例如,野生型ITR序列或工程化的ITR序列)。在一些实施方案中,核酸载体的大小为4kb和5kb之间(例如,大小为4.2至4.7kb)。在一些实施方案中,核酸载体进一步包含编码Rep蛋白的区域。在一些实施方案中,核酸载体是环状的。在一些实施方案中,核酸载体是单链的。在一些实施方案中,核酸载体是双链的。在一些实施方案中,双链核酸载体可以是,例如,自身互补载体,其含有与核酸载体的另一个区域互补的核酸载体的区域,起始形成核酸载体的双链。

[0133] 如本文所用,术语“核酸”和“多核苷酸”是指包含核碱基和酸性部分(例如核苷酸或核苷酸的聚合物)的化合物。通常,聚合核酸,例如包含三个或更多个核苷酸的核酸分子是线性分子,其中相邻的核苷酸经由磷酸二酯连接彼此连接。在一些实施方案中,“核酸”是指个别的核酸残基(例如核苷酸和/或核苷)。在一些实施方案中,“核酸”是指包含三个或更多个个别核苷酸残基的寡核苷酸链。如本文所用,术语“寡核苷酸”和“多核苷酸”可以可互换地使用以指核苷酸的聚合物(例如,至少三个核苷酸的串)。在一些实施方案中,“核酸”涵盖RNA以及单链和/或双链DNA。核酸可以是天然存在的,例如在基因组、转录物、mRNA、tRNA、

rRNA、siRNA、snRNA、质粒、粘粒、染色体、染色单体或其他天然存在的核酸分子的背景中。另一方面，核酸分子可以是非天然存在的分子，例如重组DNA或RNA、人工染色体、工程化的基因组(例如工程化的病毒载体)、工程化的载体或其片段，或合成DNA、RNA或DNA/RNA杂合体，任选地包括非天然存在的核苷酸或核苷。此外，术语“核酸”、“DNA”、“RNA”和/或类似术语包括核酸类似物，例如具有除磷酸二酯主链之外的类似物。核酸可以从天然来源纯化，使用重组表达系统产生并任选地纯化、化学合成等。在适当的情况下，例如在化学合成分子的情况下，核酸可以包含核苷类似物，例如具有化学修饰的碱基或糖，和主链修饰的类似物。除非另有说明，核酸序列以5′至3′方向呈现。在一些实施方案中，核酸是或包含天然核苷(例如腺苷、胸苷、鸟苷、胞苷、尿苷、脱氧腺苷、脱氧胸苷、脱氧鸟苷和脱氧胞苷)；核苷类似物(例如2-氨基腺苷、2-硫代胸苷、肌苷、吡咯并嘧啶、3-甲基腺苷、5-甲基胞苷、2-氨基腺苷、C5-溴尿苷、C5-氟尿苷、C5-碘尿苷、C5-丙炔基-尿苷、C5-丙炔基-胞苷、C5-甲基胞苷、2-氨基腺苷、7-脱氮腺苷、7-脱氮鸟苷、8-氧代腺苷、8-氧代鸟苷、O(6)-甲基鸟嘌呤和2-硫代胞苷)；化学修饰的碱基；生物修饰的碱基(例如甲基化碱基)；插入的碱基；修饰的糖(例如2′-氟核糖、核糖、2′-脱氧核糖、阿拉伯糖和己糖)；和/或修饰的磷酸基团(例如硫代磷酸酯和5′-N-亚磷酰胺连接)。

[0134] 术语“蛋白质”、“肽”和“多肽”在本文中可互换使用，并且是指通过肽(酰胺)键连接在一起的氨基酸残基的聚合物。该术语是指任何大小、结构或功能的蛋白质、肽或多肽。通常，蛋白质、肽或多肽将是至少三个氨基酸长。蛋白质、肽或多肽可以指个别的蛋白质或蛋白质的集合。蛋白质、肽或多肽中的一个或多个氨基酸可以被修饰，例如通过添加化学实体如碳水化合物基团、羟基、磷酸基团、法呢基、异法呢基、脂肪酸基团，用于缀合、官能化的接头或其他修饰等。蛋白质、肽或多肽也可以是单个分子或者可以是多分子复合物。蛋白质、肽或多肽可以仅仅是天然存在的蛋白质或肽的片段。蛋白质、肽或多肽可以是天然存在的、重组的或合成的，或其任何组合。如本文所用的术语“融合蛋白”是指包含来自至少两种不同蛋白质的蛋白质结构域的杂合多肽。一种蛋白质可以位于融合蛋白的氨基-末端(N-末端)部分或位于羧基-末端(C-末端)蛋白质，从而分别形成“氨基-末端融合蛋白”或“羧基-末端融合蛋白”。蛋白质可以包含不同的结构域，例如核酸结合结构域(例如引导蛋白质与靶位点结合的Cas9的gRNA结合结构域)和核酸编辑蛋白的核酸切割结构域或催化结构域。在一些实施方案中，蛋白质与核酸例如RNA或DNA复合或缔合。本文提供的任何蛋白质可以通过本领域已知的任何方法产生。例如，本文提供的蛋白质可以经由重组蛋白质表达和纯化产生，其特别适用于包含肽接头的融合蛋白。用于重组蛋白质表达和纯化的方法是熟知的，并且包括Green and Sambrook,Molecular Cloning:A Laboratory Manual(第4版，Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.(2012))描述的那些，其通过引用并入本文。

[0135] 如本文所用，术语“受试者”是指个体生物体，例如个体哺乳动物。在一些实施方案中，受试者是人。在一些实施方案中，受试者是非人哺乳动物。在一些实施方案中，受试者是非人灵长类动物。在一些实施方案中，受试者是啮齿动物(例如小鼠、大鼠)。在一些实施方案中，受试者是驯养的动物。在一些实施方案中，受试者是绵羊、山羊、牛、猫或狗。在一些实施方案中，受试者是研究动物。在一些实施方案中，受试者是经基因工程化的，例如基因工程化的非人受试者。受试者可以是任何一个性别和处于任何发展阶段的。

[0136] 如本文中在蛋白质或核酸的背景中使用，术语“重组”是指自然界中不存在，但是作为人工程化的产物的蛋白质或核酸。例如，在一些实施方案中，重组蛋白质或核酸分子包含氨基酸或核苷酸序列，其相比于任何天然存在的序列包含至少一个、至少两个、至少三个、至少四个、至少五个、至少六个或至少七个突变。本文所述的融合蛋白(例如碱基编辑器)是以重组方式制备的。重组技术是本领域技术人员熟悉的。

[0137] 术语“药学上可接受的载体”意指药学上可接受的材料、组合物或媒介物，例如液体或固体填充剂、稀释剂、赋形剂、制造助剂(例如，润滑剂、滑石粉、硬脂酸镁、钙或锌或硬脂酸)或溶剂包封材料，涉及将化合物从身体的一个部位(例如，递送部位)运送或运输到另一个部位(例如，器官、组织或身体的一部分)。药学上可接受的载体是“可接受的”，意思是与制剂的其他成分相容并且对受试者的组织无害(例如，生理学相容的、无菌的、生理学的pH等)。

[0138] 如本文所用，“治疗有效量”是指本公开中描述的需要赋予受试者治疗效果(单独或与一种或多种其他治疗剂组合)的每种治疗剂(例如核碱基编辑器、rAAV)的量。如本领域技术人员所认识到的，有效量变化取决于所治疗的具体状况，状况的严重程度，个体受试者参数包括年龄、身体状况、大小、性别和体重，治疗的持续时间，同步疗法的性质(如果有的话)，特定的施用途径以及健康从业者的知识和专业技能内的相似因素。这些因素是本领域普通技术人员所熟知的，并且可以仅通过常规实验来解决。通常优选使用最大剂量的各个组分或其组合，即根据合理的医学判断的最高安全剂量。然而，本领域普通技术人员将理解，受试者可以出于医学原因、心理原因或实际上任何其他原因而坚持较低剂量或可耐受剂量。经验考虑，例如半衰期，通常将有助于确定剂量。例如，与人免疫系统相容的治疗剂(例如包含来自人源化抗体或完全人抗体的区域的多肽)可以用于延长多肽的半衰期并防止多肽被宿主的免疫系统攻击。

[0139] “有此需要的受试者”是指具有疾病、疾病的征兆和/或症状，或对疾病的倾向的个体，其目的是治疗、治愈、减轻、缓解、改变、补救、改善、改进或影响疾病、疾病的症状或对疾病的倾向。在一些实施方案中，受试者是哺乳动物。在一些实施方案中，受试者是非人灵长类动物。在一些实施方案中，受试者是人。在一些实施方案中，哺乳动物是啮齿动物。在一些实施方案中，啮齿动物是小鼠。在一些实施方案中，啮齿动物是大鼠。在一些实施方案中，哺乳动物是伴侣动物。“伴侣动物”是指宠物和其他家畜。伴侣动物的非限制性实例包括狗和猫；牲畜，如马、牛、猪、绵羊、山羊和鸡；和其他动物，如小鼠、大鼠、豚鼠和仓鼠。

[0140] 术语“治疗/处理”是指如本文所述旨在逆转、减轻疾病或病症或其一种或多种症状、延迟疾病或病症或其一种或多种症状的发作或抑制疾病或病症或其一种或多种症状进展的临床干预。如本文所用，术语“治疗/处理”是指如本文所述旨在逆转、减轻疾病或病症或其一种或多种症状、延迟疾病或病症或其一种或多种症状的发作或抑制疾病或病症或其一种或多种症状进展的临床干预。在一些实施方案中，可以在一种或多种症状已经得以形成之后和/或疾病已经得到诊断之后施用治疗。在其他实施方案中，可以在没有症状的情况下施用治疗，例如用于预防或延迟症状的发作或抑制疾病的发作或进展。例如，可以在症状发作之前(例如，鉴于症状的历史和/或鉴于遗传或其他易感性因素)施用治疗于易感个体。治疗也可以在症状消退后继续进行，例如以预防或延迟其复发。

[0141] 某些实施方案的详述

[0142] 本文提供的是包含编码分裂的Cas9蛋白或核碱基编辑器的核酸的组合物(例如载体、重组病毒)和试剂盒,以及使用此类核酸将核碱基编辑器或Cas9蛋白递送到细胞中的方法。核碱基编辑器或Cas9蛋白的N-末端部分和C-末端部分在分开的核酸上编码并递送到细胞中,例如经由重组腺相关病毒(rAAV颗粒)递送。对应于核碱基编辑器或Cas9蛋白的N-末端部分和C-末端部分的多肽可以接合以形成完整的核碱基编辑器或Cas9蛋白,例如经由内含肽介导的蛋白质剪接。

[0143] 因此,本公开的一些方面涉及组合物,其包含(i)第一核苷酸序列,其编码在其C-末端处与内含肽-N融合的Cas9蛋白的N-末端部分,所述Cas9蛋白的N-末端部分;和(ii)第二核苷酸序列,其编码与所述Cas9蛋白的C-末端部分的N-末端融合的内含肽-C,其中所述第一核苷酸序列或第二核苷酸序列与编码至少一个二分核定位信号(NLS)的核苷酸序列可操作地连接。

[0144] 由第一和第二核苷酸序列编码的Cas9蛋白在本文中称为“分裂的Cas9”。已知Cas9蛋白具有通过无序接头连接的N-末端叶和C-末端叶(例如,如Nishimasu et al.,Cell,Volume 156,Issue 5,pp.935-949,2014中所述,通过引用并入本文)。在一些实施方案中,分裂的Cas9蛋白的N-末端部分包含Cas9蛋白的N-末端叶。在一些实施方案中,分裂的Cas9的C-末端部分包含Cas9蛋白的C-末端叶。在一些实施方案中,分裂的Cas9的N-末端部分包含SEQ ID NO:1-275和394-397的任一个的一部分,其对应于SEQ ID NO:1中的氨基酸1-(550-650)。“1-(550-650)”意指从氨基酸1开始并在氨基酸550-650(包括端点)之间的任何地方结束。例如,分裂的Cas9的N-末端部分可以包含SEQ ID NO:1-275和394-397的任一个的一部分,其对应于SEQ ID NO:1中的氨基酸1-550、1-551、1-552、1-553、1-554、1-555、1-556、1-557、1-558、1-559、1-560、1-561、1-562、1-563、1-564、1-565、1-566、1-567、1-568、1-569、1-570、1-571、1-572、1-573、1-574、1-575、1-576、1-577、1-578、1-579、1-580、1-581、1-582、1-583、1-584、1-585、1-586、1-587、1-588、1-589、1-590、1-591、1-592、1-593、1-594、1-595、1-596、1-597、1-598、1-599、1-600、1-601、1-602、1-603、1-604、1-605、1-606、1-607、1-608、1-609、1-610、1-611、1-612、1-613、1-614、1-615、1-616、1-617、1-618、1-619、1-620、1-621、1-622、1-623、1-624、1-625、1-626、1-627、1-628、1-629、1-630、1-631、1-632、1-633、1-634、1-635、1-636、1-637、1-638、1-639、1-640、1-641、1-642、1-643、1-644、1-645、1-646、1-647、1-648、1-649或1-650。在一些实施方案中、分裂的Cas9蛋白的N-末端部分包含SEQ ID NO:1-275和394-397的任一个的一部分、其对应于SEQ ID NO:1的氨基酸1-573或1-637。

[0145] 分裂的Cas9的C-末端部分可以与分裂的Cas9的N-末端部分接合以形成完整的Cas9蛋白。在一些实施方案中、Cas9蛋白的C-末端部分从Cas9蛋白的N-末端部分结束的位置开始。因此、在一些实施方案中、分裂的Cas9的C-末端部分包含SEQ ID NO:1-275和394-397的任一个的一部分、其对应于SEQ ID NO:1的氨基酸(551-651)-1368。“(551-651)-1368”意指在氨基酸551-651(包括端点)之间的氨基酸开始、并且在氨基酸1368结束。例如、分裂的Cas9的C-末端部分可以包含SEQ ID NO:1-275和394-397的任一个的一部分、其对应于SEQ ID NO:1中的氨基酸551-1368、552-1368、553-1368、554-1368、555-1368、556-1368、557-1368、558-1368、559-1368、560-1368、561-1368、562-1368、563-1368、564-1368、565-1368、566-1368、567-1368、568-1368、569-1368、570-1368、571-1368、572-1368、573-1368、

574-1368、575-1368、576-1368、577-1368、578-1368、579-1368、580-1368、581-1368、582-1368、583-1368、584-1368、585-1368、586-1368、587-1368、588-1368、589-1368、590-1368、591-1368、592-1368、593-1368、594-1368、595-1368、596-1368、597-1368、598-1368、599-1368、600-1368、601-1368、602-1368、603-1368、604-1368、605-1368、606-1368、607-1368、608-1368、609-1368、610-1368、611-1368、612-1368、613-1368、614-1368、615-1368、616-1368、617-1368、618-1368、619-1368、620-1368、621-1368、622-1368、623-1368、624-1368、625-1368、626-1368、627-1368、628-1368、629-1368、630-1368、631-1368、632-1368、633-1368、634-1368、635-1368、636-1368、637-1368、638-1368、639-1368、640-1368、641-1368、642-1368、643-1368、644-1368、645-1368、646-1368、647-1368、648-1368、649-1368、650-1368或651-1368。在一些实施方案中、分裂的Cas9蛋白的C-末端部分包含SEQ ID NO:1-275和394-397的任一个的一部分、其对应于SEQ ID NO:1的氨基酸574-1368或638-1368。

[0146] 还可以使用本文描述的方法将Cas9变体递送至细胞。例如,Cas9变体也可以如本文所述“分裂”。Cas9变体可以包含与本文提供的Cas9序列的任一个至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。在一些实施方案中,Cas9变体包含长度上比本文(例如实施例1中)提供的任何Cas9蛋白更短或更长的氨基酸序列(例如,不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中,UGI包含长度上比本文提供的任何Cas9蛋白更短或更长的氨基酸序列(例如,不超过200个氨基酸、不超过150个氨基酸、不超过100个氨基酸、不超过50个氨基酸、不超过10个氨基酸、不超过5个氨基酸或不超过2个氨基酸更长或更短)。

[0147] 在一些实施方案中,分裂的Cas9的N-末端部分包含与本文(例如实施例1中)提供的Cas9序列的任一个的相应的部分至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。在一些实施方案中,分裂的Cas9的N-末端部分包含长度上比本文提供的任何Cas9蛋白的相应的部分更短或更长的氨基酸序列(例如,不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中,分裂的Cas9的N-末端部分包含长度上比本文提供的任何Cas9蛋白的相应的部分更短或更长的氨基酸序列(例如,不超过200个氨基酸、不超过150个氨基酸、不超过100个氨基酸、不超过50个氨基酸、不超过10个氨基酸、不超过5个氨基酸或不超过2个氨基酸更长或更短)。

[0148] 在一些实施方案中,分裂的Cas9的C-末端部分包含与本文(例如实施例1中)提供的Cas9序列的任一个的相应的部分至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。在一些实施方案中,分裂的Cas9的C-末端部分包含长度上比本文提供的任何Cas9蛋白的相应的部分更短或更长的氨基酸序列(例如,不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中,分裂的Cas9的C-末端部分包含长度上比本文提供的任何Cas9蛋白的相应的部分更短或更长的氨基酸序列(例如,不超过200个氨基酸、不超过150个氨基酸、不超过100个氨基酸、不超过50个氨基酸、不超过10个氨基酸、不超过5个氨基酸或不超过2个氨基酸更长或更短)。

[0149] 在一些实施方案中,Cas9变体是dCas9或nCas9。在一些实施方案中,分裂的Cas9的

N-末端部分包含对应于SEQ ID NO:1中的D10A突变的突变。在一些实施方案中，分裂的Cas9的N-末端部分包含对应于SEQ ID NO:1中的D10A突变的突变并且分裂的Cas9的C-末端部分包含对应于SEQ ID NO:1中的H840A突变的突变。在一些实施方案中，分裂的Cas9的N-末端部分包含对应于SEQ ID NO:1中的D10A突变的突变，并且分裂的Cas9的C-末端部分包含在对应于SEQ ID NO:1中的位置840的位置处的组氨酸。

[0150] 在一些实施方案中，为了接合Cas9蛋白的N-末端部分和Cas9蛋白的C-末端部分，可以使用内含肽系统。在一些实施方案中，Cas9的N-末端部分与内含肽-N融合。在一些实施方案中，内含肽-N与Cas9的N-末端部分的C-末端融合，以形成NH_2-[Cas9的N-末端部分]-[内含肽-N]-COOH的结构。在一些实施方案中，内含肽-N由dnaE-n基因编码。在一些实施方案中，内含肽-N包含SEQ ID NO:350-351和354-355的任一个的氨基酸序列。在一些实施方案中，Cas9的C-末端部分与内含肽-C融合，并且内含肽-C与Cas9的C-末端部分的N-末端融合以形成NH_2-[内含肽-C]-[Cas9的C-末端部分]-COOH的结构。在一些实施方案中，内含肽-C由dnaE-c基因编码。在一些实施方案中，内含肽-C包含SEQ ID NO:352-353和356-357的任一个的氨基酸序列。其他分裂的内含肽系统也可以用于本公开中并且是本领域中已知的。

[0151] 分裂的核碱基编辑器可以用于本公开中。本公开的一些方面涉及组合物，其包含(i)第一核苷酸序列，其在其C-末端处与内含肽-N融合的编码核碱基编辑器的N-末端部分；和(ii)第二核苷酸序列，其编码与所述核碱基编辑器的C-末端部分的N-末端融合的内含肽-C。

[0152] 考虑了核碱基编辑器变体。例如，核碱基编辑器变体也可以如本文所述“分裂”。核碱基编辑器变体可以包含与本文提供的核碱基编辑器序列(SEQ ID NO:X-X)的任一个至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。

[0153] 在一些实施方案中，分裂的核碱基编辑器的N-末端部分包含与本文(例如实施例1中)提供的核碱基编辑器的任一个的相应的部分至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。在一些实施方案中，分裂的核碱基编辑器的N-末端部分包含长度上比本文提供的任何核碱基编辑器的相应的部分更短或更长的氨基酸序列(例如，不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中，分裂的核碱基编辑器的N-末端部分包含长度上比本文提供的任何核碱基编辑器的相应的部分更短或更长的氨基酸序列(例如，不超过200个氨基酸、不超过150个氨基酸、不超过100个氨基酸、不超过50个氨基酸、不超过10个氨基酸、不超过5个氨基酸或不超过2个氨基酸更长或更短)。

[0154] 在一些实施方案中，分裂的核碱基编辑器的C-末端部分包含与本文(例如实施例1中)提供的核碱基编辑器的任一个的相应的部分至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。在一些实施方案中，分裂的核碱基编辑器的C-末端部分包含长度上比本文提供的任何核碱基编辑器的相应的部分更短或更长的氨基酸序列(例如，不超过30%、不超过25%、不超过20%、不超过15%、不超过10%、不超过5%、不超过1%更长或更短)。在一些实施方案中，分裂的核碱基编辑器的C-末端部分包含长度上比本文提供的任

何核碱基编辑器的相应的部分更短或更长的氨基酸序列(例如,不超过200个氨基酸、不超过150个氨基酸、不超过100个氨基酸、不超过50个氨基酸、不超过10个氨基酸、不超过5个氨基酸或不超过2个氨基酸更长或更短)。

[0155] 如本文所述,核碱基编辑器的N-末端部分包含核酸酶无活性的Cas9蛋白(dCas9)或Cas9切口酶(nCas9)的N-末端部分。在一些实施方案中,核碱基编辑器的N-末端部分进一步包含核碱基修饰酶(例如,核酸酶、切口酶、重组酶、脱氨酶、DNA修复酶、DNA损伤酶、歧化酶、烷基化酶、脱嘌呤酶、氧化酶、嘧啶二聚体形成酶、整合酶、转座酶、聚合酶、连接酶、解旋酶、光裂合酶、糖基化酶、表观遗传修饰物如甲基化酶、乙酰化酶、甲基转移酶、去甲基化酶等)。在一些实施方案中,核碱基修饰酶是脱氨酶(例如,胞嘧啶脱氨酶或腺苷脱氨酶,或其功能变体)。在一些实施方案中,核碱基修饰酶与分裂的dCas9或分裂的nCas9的N-末端部分的N-末端融合。在一些实施方案中,核碱基编辑器的N-末端部分具有以下结构:NH_2-[核碱基修饰酶]-[dCas9或nCas9的N-末端部分]-COOH。在一些实施方案中,核碱基编辑器的N-末端部分与内含肽N融合。在一些实施方案中,内含肽-N与核碱基编辑器的N-末端部分的C-末端融合。

[0156] 在一些实施方案中,第一核苷酸序列编码包含结构NH_2-[核碱基修饰酶]-[dCas9或nCas9的N-末端部分]-[内含肽-N]-COOH的多肽。

[0157] 在一些实施方案中,核碱基编辑器的C-末端部分包含核酸酶无活性的Cas9蛋白(dCas9)或Cas9切口酶(nCas9)的C-末端部分。在一些实施方案中,核碱基修饰酶与分裂的dCas9或分裂的nCas9的C-末端部分的C-末端融合.在一些实施方案中,核碱基编辑器的C-末端部分是以下结构:NH_2-[dCas9或nCas9的C-末端部分]-[核碱基修饰酶]-COOH。在一些实施方案中,核碱基编辑器的C-末端部分包含与Cas9蛋白的C-末端部分融合的内含肽-C。在一些实施方案中,内含肽-C与核碱基编辑器的C-末端部分的N-末端融合。在一些实施方案中,第二核苷酸序列编码以下结构的多肽:NH_2-[内含肽-C]-[Cas9蛋白的C-末端部分]-COOH。

[0158] 在一些实施方案中,分裂的核碱基编辑器的N-末端部分进一步包含尿嘧啶糖基化酶的抑制剂(UGI)。在一些实施方案中,第一核苷酸序列编码以下结构的多肽:NH_2-[UGI]-[核碱基修饰酶]-[dCas9或nCas9的N-末端部分]-[内含肽-N]。在一些实施方案中,第一核苷酸序列编码以下结构的多肽:NH_2-[核碱基修饰酶]-[UGI]-[dCas9或nCas9的N-末端部分]-[内含肽-N]。

[0159] 在一些实施方案中,分裂的核碱基编辑器的C-末端部分进一步包含抑制尿嘧啶糖基化酶的活性的酶(UGI)。在一些实施方案中,第二核苷酸序列编码以下结构的多肽:NH_2-[内含肽-C]-[dCas9或nCas9的C-末端部分]-[UGI]-COOH。在一些实施方案中,第二核苷酸序列编码以下结构的多肽:NH_2-[内含肽-C]-[dCas9或nCas9的C-末端部分]-[核碱基修饰酶]-[UGH-COOH。在一些实施方案中,第二核苷酸序列编码以下结构的多肽:NH_2-[内含肽-C]-[dCas9或nCas9的C-末端部分]-[UGI]-[核碱基修饰酶]-COOH。

[0160] 在一些实施方案中,当核碱基的N-末端部分和C-末端部分接合,以形成完整的分裂的核碱基编辑器。在一些实施方案中,分裂的核碱基编辑器可以包含以下结构的任一个:

[0161] NH_2-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0162] NH_2-[UGI]-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0163] NH_2-[核碱基修饰酶]-[UGI]-[dCas9或nCas9]-COOH

[0164] NH_2-[核碱基修饰酶]-[dCas9或nCas9]-[UGH-COOH

[0165] NH_2-[dCas9或nCas9]-[核碱基修饰酶]-COOH

[0166] NH_2-[UGI]-[dCas9或nCas9]-[核碱基修饰酶]-COOH

[0167] NH_2-[dCas9或nCas9]-[UGI]-[核碱基修饰酶]-COOH或

[0168] NH_2-[dCas9或nCas9]-[核碱基修饰酶]-[UGI]-COOH。

[0169] 在一些实施方案中，第一核苷酸序列或第二核苷酸序列(编码分裂的Cas9蛋白或分裂的核碱基编辑器)与编码至少一个二分核定位信号(NLS)的核苷酸序列可操作地连接。例如，第一核苷酸序列可以与编码一个或多个(例如2、3、4、5或更多个)二分NLS的核苷酸序列可操作地连接。在一些实施方案中，第二核苷酸序列可以与编码一个或多个(例如2、3、4、5或更多个)二分NLS的核苷酸序列可操作地连接。因此，通过接合N-末端部分和C-末端部分形成的分裂的Cas9或分裂的核碱基编辑器可以包含一个或多个二分NLS。例如，分裂的Cas9或分裂的核碱基编辑器可以包含以下结构的任一个(bNLS意指一个或多个二分核定位信号)：

[0170] NH_2-bNLS-[Cas9]-COOH

[0171] NH_2-[Cas9]-bNLS-COOH

[0172] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0173] NH_2-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-COOH

[0174] NH_2-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-COOH

[0175] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-COOH

[0176] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0177] NH_2-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0178] NH_2-bNLS-[UGI]-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0179] NH_2-[UGI]-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0180] NH_2-[UGI]-[核碱基修饰酶]-bNLS[dCas9或nCas9]-COOH

[0181] NH_2-[UGI]-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-COOH

[0182] NH_2-bNLS-[UGI]-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-COOH

[0183] NH_2-bNLS-[UGI]-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-COOH

[0184] NH_2-bNLS-[UGI]-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-COOH

[0185] NH_2-[UGI]-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-COOH

[0186] NH_2-[UGI]-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-COOH

[0187] NH_2-[UGI]-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0188] NH_2-bNLS-[UGI]-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-COOH

[0189] NH_2-bNLS-[UGI]-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0190] NH_2-[UGI]-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0191] NH_2-bNLS-[UGI]-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0192] NH_2-bNLS-[核碱基修饰酶]-[UGI]-[dCas9或nCas9]-COOH

[0193] NH_2-[核碱基修饰酶]-bNLS-[UGI]-[dCas9或nCas9]-COOH

[0194] NH_2-[核碱基修饰酶]-[UGI]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0195] NH_2-bNLS-[核碱基修饰酶]-bNLS-[UGI]-[dCas9或nCas9]-COOH

[0196] NH_2-bNLS-[核碱基修饰酶]-[UGI]-bNLS-[dCas9或nCas9]-COOH

[0197] NH_2-bNLS-[核碱基修饰酶]-[UGI]-[dCas9或nCas9]-bNLS-COOH

[0198] NH_2-[核碱基修饰酶]-bNLS-[UGI]-bNLS-[dCas9或nCas9]-COOH

[0199] NH_2-[核碱基修饰酶]-bNLS-[UGI]-[dCas9或nCas9]-bNLS-COOH

[0200] NH_2-[核碱基修饰酶]-[UGI]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0201] NH_2-bNLS-[核碱基修饰酶]-bNLS-[UGI]-bNLS-[dCas9或nCas9]-COOH

[0202] NH_2-bNLS-[核碱基修饰酶]-bNLS-[UGI]-[dCas9或nCas9]-bNLS-COOH

[0203] NH_2-[核碱基修饰酶]-bNLS-[UGI]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0204] NH_2-bNLS-[核碱基修饰酶]-bNLS-[UGI]-bNLS-[dCas9或nCas9]-bNLS-COOH

[0205] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-[UGI]-COOH

[0206] NH_2-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-[UGI]-COOH

[0207] NH_2-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-[UGI]-COOH

[0208] NH_2-[核碱基修饰酶]-[dCas9或nCas9]-[UGI]-bNLS-COOH

[0209] NH_2-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-[UGI]-COOH

[0210] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-[UGI]-COOH

[0211] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-[UGI]-bNLS-COOH

[0212] NH_2-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-[UGI]-COOH

[0213] NH_2-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-[UGI]-bNLS-COOH

[0214] NH_2-bNLS-[核碱基修饰酶]-[dCas9或nCas9]-bNLS-[UGI]-bNLS-COOH

[0215] NH_2-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-[UGI]-bNLS-COOH

[0216] NH_2-bNLS-[核碱基修饰酶]-bNLS-[dCas9或nCas9]-bNLS-[UGI]-bNLS-COOH

[0217] NH_2-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-COOH

[0218] NH_2-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-COOH

[0219] NH_2-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-COOH

[0220] NH_2-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-COOH

[0221] NH_2-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-COOH

[0222] NH_2-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0223] NH_2-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0224] NH_2-bNLS-[UGI]-[dCas9或nCas9]-[核碱基修饰酶]-COOH

[0225] NH_2-[UGI]-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-COOH

[0226] NH_2-[UGI]-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-COOH

[0227] NH_2-[UGI]-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-COOH

[0228] NH_2-bNLS-[UGI]-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-COOH

[0229] NH_2-bNLS-[UGI]-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-COOH

[0230] NH_2-bNLS-[UGI]-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-COOH

[0231] NH_2-[UGI]-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-COOH

[0232] NH_2-[UGI]-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-COOH

[0233] NH_2-[UGI]-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0234] NH_2-bNLS-[UGI]-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-COOH

[0235] NH_2-bNLS-[UGI]-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0236] NH_2-bNLS-[UGI]-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-COOH

[0237] NH_2-[UGI]-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0238] NH_2-bNLS-[UGI]-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0239] NH_2-bNLS-[dCas9或nCas9]-[UGI]-[核碱基修饰酶]-COOH

[0240] NH_2-[dCas9或nCas9]-bNLS-[UGI]-[核碱基修饰酶]-COOH

[0241] NH_2-[dCas9或nCas9]-[UGI]-bNLS-[核碱基修饰酶]-COOH

[0242] NH_2-[dCas9或nCas9]-[UGI]-[核碱基修饰酶]-bNLS-COOH

[0243] NH_2-bNLS-[dCas9或nCas9]-bNLS[UGI]-[核碱基修饰酶]-COOH

[0244] NH_2-bNLS-[dCas9或nCas9]-[UGI]-bNLS-[核碱基修饰酶]-COOH

[0245] NH_2-bNLS-[dCas9或nCas9]-[UGI]-[核碱基修饰酶]-bNLS-COOH

[0246] NH_2-[dCas9或nCas9]-bNLS-[UGI]-bNLS-[核碱基修饰酶]-COOH

[0247] NH_2-[dCas9或nCas9]-bNLS-[UGI]-[核碱基修饰酶]-bNLS-COOH

[0248] NH_2-[dCas9或nCas9]-[UGI]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0249] NH_2-bNLS-[dCas9或nCas9]-bNLS-[UGI]-bNLS-[核碱基修饰酶]-COOH

[0250] NH_2-bNLS-[dCas9或nCas9]-bNLS-[UGI]-[核碱基修饰酶]-bNLS-COOH

[0251] NH_2-bNLS-[dCas9或nCas9]-[UGI]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0252] NH_2-[dCas9或nCas9]-bNLS-[UGI]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0253] NH_2-bNLS-[dCas9或nCas9]-bNLS-[UGI]-bNLS-[核碱基修饰酶]-bNLS-COOH

[0254] NH_2-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-[UGI]-COOH

[0255] NH_2-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-[UGI]-COOH

[0256] NH_2-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-[UGI]-COOH

[0257] NH_2-[dCas9或nCas9]-[核碱基修饰酶]-[UGI]-bNLS-COOH

[0258] NH_2-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-[UGI]-COOH

[0259] NH_2-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-[UGI]-COOH

[0260] NH_2-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-[UGI]-bNLS-COOH

[0261] NH_2-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-[UGI]-COOH

[0262] NH_2-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-[UGI]-bNLS-COOH

[0263] NH_2-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-[UGI]-bNLS-COOH

[0264] NH_2-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-[UGI]-COOH

[0265] NH_2-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-[UGI]-bNLS-COOH

[0266] NH_2-bNLS-[dCas9或nCas9]-[核碱基修饰酶]-bNLS-[UGI]-bNLS-COOH

[0267] NH_2-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-[UGI]-bNLS-COOH或者

[0268] NH_2-bNLS-[dCas9或nCas9]-bNLS-[核碱基修饰酶]-bNLS-[UGI]-bNLS-COOH

[0269] 本文中，“NH_2-”表示蛋白质或多肽的N-末端，而“-COOH”表示蛋白质或多肽的C-末端。“]-[”表示肽键或接头。在一些实施方案中，接头可以用于连接本文所述的任何蛋白质或蛋白质结构域。接头可以像共价键一样简单，或者其可以是长度为多个原子的聚合物接头。在一些实施方案中，接头是多肽或基于氨基酸。在一些实施方案中，接头不是肽样的。在

一些实施方案中，接头是共价键(例如，碳-碳键、二硫键、碳-杂原子键等)。在一些实施方案中，接头是酰胺连接的碳-氮键。在一些实施方案中，接头是环状或非环状的、取代或未取代的、支化或未支化的脂族或杂脂族接头。在一些实施方案中，接头是聚合物(例如，聚乙烯、聚乙二醇、聚酰胺、聚酯等)。在一些实施方案中，接头包含氨基链烷酸的单体、二聚体或聚合物。在一些实施方案中，接头包含氨基链烷酸(例如甘氨酸、乙酸、丙氨酸、beta-丙氨酸、3-氨基丙酸、4-氨基丁酸、5-戊酸等)。在一些实施方案中，接头包含氨基己酸(Ahx)的单体、二聚体或聚合物。在一些实施方案中，接头基于碳环部分(例如，环戊烷、环己烷)。在一些实施方案中，接头包含聚乙二醇部分(PEG)。在一些实施方案中，接头包含氨基酸。在一些实施方案中，接头包含肽。在一些实施方案中，接头包含芳基或杂芳基部分。在一些实施方案中，接头基于苯环。接头可以包括官能化部分以促进亲核试剂(例如，硫醇、氨基)从肽附接到接头。任何亲电试剂可以用作接头的一部分。示例性亲电试剂包括但不限于活化酯、活化酰胺、迈克尔受体、烷基卤化物、芳基卤化物、酰基卤化物和异硫氰酸酯。

[0270] 在一些实施方案中，接头是一个氨基酸或多个氨基酸(例如肽或蛋白质)。在一些实施方案中，接头是键(例如，共价键)、有机分子、基团、聚合物或化学部分。在一些实施方案中，接头的长度为5-100个氨基酸，例如长度为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、30-35、35-40、40-45、45-50、50-60、60-70、70-80、80-90、90-100、100-110、110-120、120-130、130-140、140-150或150-200个氨基酸。也考虑了更长或更短的接头。在一些实施方案中，接头包含氨基酸序列SGSETPGTSESATPES(SEQ ID NO:377)，其也可以称为XTEN接头。在一些实施方案中，接头包含氨基酸序列:SGGS(SEQ ID NO:378)。在一些实施方案中，接头包含氨基酸序列:$(SGGS)_n$(SEQ ID NO:379)、$(GGGS)_n$(SEQ ID NO:380)、$(GGGGS)_n$(SEQ ID NO:381)、$(G)_n$(SEQ ID NO:390)、$(EAAAK)_n$(SEQ ID NO:382)、$(GGS)_n$、SGSETPGTSESATPES(SEQ ID NO:377)或$(XP)_n$基序或任何这些的组合，其中n独立地是1和30之间的整数，包括端点，并且其中X是任何氨基酸。在一些实施方案中，n是1、2、3、4、5、6、7、8、9、10、11、12、13、14或15。在一些实施方案中，接头包含氨基酸序列:SGSETPGTSESATPES(SEQ ID NO:377)和SGGS(SEQ ID NO:378)。在一些实施方案中，接头包含氨基酸序列:SGGSSGSETPGTSESATPESSGGS(SEQ ID NO:383)。在一些实施方案中，接头包含氨基酸序列:SGGSSGGSSGSETPGTSESATPESSGGSSGGS(SEQ ID NO:384)。在一些实施方案中，接头包含氨基酸序列:GGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS(SEQ ID NO:385)。

[0271] 在一些实施方案中，接头长度为24个氨基酸。在一些实施方案中，接头包含氨基酸序列SGGSSGGSSGSETPGTSESATPES(SEQ ID NO:343)。在一些实施方案中，接头长度为40个氨基酸。在一些实施方案中，接头包含氨基酸序列SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS(SEQ ID NO:391)。在一些实施方案中，接头长度为64个氨基酸。在一些实施方案中，接头包含氨基酸序列SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGS(SEQ ID NO:392)。在一些实施方案中，接头长度为92个氨基酸。在一些实施方案中，接头包含氨基酸序列PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS(SEQ ID NO:393)。

[0272] 在一些实施方案中，第一和第二核苷酸序列在相同的核酸载体上。在一些实施方案中，第一和第二核苷酸序列在不同的核酸载体上。在一些实施方案中，载体是质粒。在一

些实施方案中，核酸载体是腺相关病毒(rAAV)的重组基因组。在一些实施方案中，核酸载体是包装在rAAV颗粒中的腺相关病毒的基因组。在一些实施方案中，第一和/或第二核苷酸序列与启动子可操作地连接。在一些实施方案中，核酸载体进一步包含编码与启动子可操作地连接的一个或多个(例如，1、2、3、4、5、6、7、8、9、10个或更多个)gRNA的核苷酸序列。在一些实施方案中，启动子是组成型启动子。在一些实施方案中，启动子是诱导型启动子。

[0273] 本公开的诱导型启动子可以由一种或多种生理条件的变化诱导(或抑制)，所述生理条件例如光、pH、温度、辐射、渗透压、盐水梯度、细胞表面结合，和一种或多种外在或内在诱导剂的浓度。外在诱导物信号或诱导剂可以包含但不限于氨基酸和氨基酸类似物、糖类和多糖、核酸、蛋白质转录激活物和阻遏物、细胞因子、毒素，石油基化合物、含金属化合物、盐、离子、酶底物类似物、激素或其组合。

[0274] 本公开的诱导型启动子包括本文所述的或本领域普通技术人员已知的任何诱导型启动子。诱导型启动子的实例包括但不限于化学/生物化学调节和物理调节的启动子，例如醇调节的启动子、四环素调节的启动子(例如，无水四环素(aTc)-响应性启动子和其他四环素响应性启动子系统，其包括四环素阻遏蛋白(tetR)、四环素操纵子序列(tetO)和四环素反式激活因子融合蛋白(tTA))、类固醇调节的启动子(例如，基于大鼠糖皮质激素受体、人雌激素受体、蛾蜕皮激素受体的启动子和来自类固醇/类视黄醇/甲状腺受体超家族的启动子)、金属调节的启动子(例如，衍生自来自酵母、小鼠和人的金属硫蛋白(结合并多价螯合(sequester)金属离子的蛋白质)基因的启动子)、发病机制调节的启动子(例如，由水杨酸、乙烯或苯并噻二唑(BTH)诱导)、温度/热诱导型启动子(例如热激启动子)和光调节的启动子(例如，来自植物细胞的光响应性启动子)。其他诱导型启动子系统是本领域已知的，并且可以根据本公开使用。

[0275] 在一些实施方案中，本公开的诱导型启动子在原核细胞(例如细菌细胞)中起作用。用于原核细胞的诱导型启动子的实例包括但不限于，噬菌体启动子(例如Pls1con、T3、T7、SP6、PL)和细菌启动子(例如Pbad、PmgrB、Ptrc2、Plac/ara、Ptac、Pm)或其杂合体(例如PLlacO、PLtetO)。根据本公开使用的细菌启动子的实例包括但不限于，正调节的大肠杆菌启动子，例如正调节的σ70启动子(例如诱导型pBad/araC启动子、Lux盒右启动子、经修饰的lamdba Prm启动子、plac Or2-62(正)，具有额外的REN位点的pBad/AraC、pBad、P(Las)TetO、P(Las)CIO、P(Rh1)、Pu、FecA、pRE、cadC、hns、pLas、pLux)、σS启动子(例如Pdps)、σ532启动子(例如热激)和σ54启动子(例如glnAp2)；负调节的大肠杆菌启动子，例如负调节的σ70启动子(例如启动子(PRM+)、经修饰的lamdba Prm启动子、TetR-TetR-4C P(Las)TetO、P(Las)CIO、P(Lac)IQ、RecA_DlexO_DLacO1、dapAp、FecA、Pspac-hy、pcI、plux-cI、plux-lac、CinR、CinL、葡萄糖控制的、经修饰的Pr、经修饰的Prm+、FecA、Pcya、rec A(SOS)、Rec A(SOS)、EmrR_调节的、BetI_调节的、pLac_lux、pTet_Lac、pLac/Mnt、pTet/Mnt、LsrA/cI、pLux/cI、LacI、LacIQ、pLacIQ1、pLas/cI、pLas/Lux、pLux/Las、具有LexA结合位点的pRecA、反向的BBa_R0011、pLacI/ara-1、pLacIq、rmB P1、cadC、hns、PfhuA、pBad/araC、nhaA、OmpF、RcnR)、σS启动子(例如具有替代的sigma因子σ38的Lutz-Bujard LacO)、σ32启动子(例如具有替代的sigma因子σ32的Lutz-Bujard LacO)和σ54启动子(例如glnAp2)；负调节的枯草芽孢杆菌启动子，例如可抑制的枯草芽孢杆菌σA启动子(例如革兰氏阳性IPTG诱导型、Xy1、hyper-spank)和σB启动子。其他诱导型微生物启动子可以根据本公开使用。

[0276] 在一些实施方案中，本公开的诱导型启动子在真核细胞(例如哺乳动物细胞)中起作用。用于真核细胞的诱导型启动子的实例包括但不限于，化学调节的启动子(例如，醇调节的启动子、四环素调节的启动子、类固醇调节的启动子、金属调节的启动子和发病机制相关的(PR)启动子)和物理调节的启动子(例如，温度调节的启动子和光调节的启动子)。

[0277] 重组腺相关病毒(rAAV)

[0278] 本公开的一些方面涉及使用重组腺相关病毒载体用于将分裂的Cas9蛋白或分裂核碱基编辑器递送到细胞中。由于全长Cas9蛋白或核碱基编辑器超过rAAV的包装限制(~4.9kb)，Cas9蛋白或核碱基编辑器的N-末端部分和Cas9蛋白或核碱基编辑器的C-末端部分通过分开的rAAV载体或颗粒递送到相同的细胞中。

[0279] 因此，在一些实施方案中，提供了用于将分裂的Cas9蛋白或分裂的核碱基编辑器递送到细胞(例如，哺乳动物细胞、人细胞)中的组合物。在一些实施方案中，本公开的组合物包含：(i)第一重组腺相关病毒(rAAV)颗粒，其包含编码在其C-末端处与内含肽-N融合的Cas9蛋白或核碱基编辑器的N-末端部分的第一核苷酸序列；和(ii)第二重组腺相关病毒(rAAV)颗粒，其包含编码与所述Cas9蛋白或核碱基编辑器的C-末端部分的N-末端融合的内含肽-C的第二核苷酸序列。本公开的rAAV颗粒包含包在病毒衣壳蛋白中的rAAV载体(即rAAV的重组基因组)。

[0280] 在一些实施方案中，rAAV载体包含：(1)异源核酸区域，其包含编码以本文所述任何形式的分裂的Cas9蛋白或分裂的核碱基编辑器的N-末端部分或C-末端部分的第一或第二核苷酸序列，(2)一个或多个核苷酸序列，其包含促进异源核酸区域的表达的序列(例如，启动子)，和(3)一个或多个核酸区域，其包含促进将异源核酸区域(任选地具有一个或多个核酸区域，其包含促进表达的序列)整合到细胞的基因组中的序列。在一些实施方案中，促进整合的病毒序列包含反向末端重复(ITR)序列。在一些实施方案中，编码分裂的Cas9蛋白或分裂的核碱基编辑器的N-末端部分或C-末端部分的第一或第二核苷酸序列在每一侧侧翼为ITR序列。在一些实施方案中，核酸载体进一步包含编码如本文所述的AAV Rep蛋白的区域，其包含在侧翼为ITR的区域内或区域外。ITR序列可以衍生自任何AAV血清型(例如，1、2、3、4、5、6、7、8、9或10)，或者可以衍生自多于一种血清型。在一些实施方案中，ITR序列衍生自AAV2或AAV6。

[0281] ITR序列和含有ITR序列的质粒是本领域已知的并且可商购获得(参见例如可得自以下的产品和服务：Vector Biolabs，Philadelphia，PA；Cellbiolabs，San Diego，CA；Agilent Technologies，Santa Clara，Ca；和Addgene，Cambridge，MA；以及Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein.Kessler PD，Podsakoff GM，Chen X，McQuiston SA，Colosi PC，Matelis LA，Kurtzman GJ，Byrne BJ.Proc Natl Acad Sci USA.1996Nov 26；93(24)：14082-7；和Curtis A.Machida.Methods in Molecular Medicine™.Viral Vectors for Gene Therapy Methods and Protocols.10.1385/1-59259-304-6:201 ©Humana Press Inc.2003.Chapter 10.Targeted Integration by Adeno-Associated Virus.Matthew D.Weitzman，Samuel M.Young Jr.，Toni Cathomen and Richard Jude Samulski；美国专利号5,139,941和5,962,313，其全部通过引用并入本文)。以下提供示例性ITR序列。

[0282] AAV2：

[0283]

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT (SEQ ID NO: 386)

[0284] AAV3：

[0285]

TTGGCCACTCCCTCTATGCGCACTCGCTCGCTCGGTGGGGCCTGGCGACCAAAGGTCGCCAGACGGACGTGCTTTGCACGTCCGGCCCCACCGAGCGAGCGAGTGCGCATAGAGGGAGTGGCCAACTCCATCACTAGAGGTATGGC (SEQ ID NO: 387)

[0286] AAV5：

[0287]

CTCTCCCCCCTGTCGCGTTCGCTCGCTCGCTGGCTCGTTTGGGGGGGTGGCAGCTCAAAGAGCTGCCAGACGACGGCCCTCTGGCCGTCGCCCCCCCAAACGAGCCAGCGAGCGAGCGAACGCGACAGGGGGGAGAGTGCCACACTCTCAAGCAAGGGGGTTTTGTA (SEQ ID NO: 388)

[0288] AAV6：

[0289]

TTGCCCACTCCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCCTGCGGACCAAAGGTCCGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGAGCGCGCATAGAGGGAGTGGGCAACTCCATCACTAGGGGTA (SEQ ID NO: 389)

[0290] 在一些实施方案中，本公开的rAAV载体包含一个或多个调节元件以控制异源核酸区域的表达(例如，启动子、转录终止子和/或其他调节元件)。在一些实施方案中，第一和/或第二核苷酸序列与一个或多个(例如，1、2、3、4、5或更多个)转录终止子可操作地连接。可以根据本公开使用的转录终止子的非限制性实例包括牛生长激素基因(bGH)、人生长激素基因(hGH)、SV40、CW3、Φ或其组合的转录终止子。已经测试了几种转录终止子的效率以确定它们在分裂的Cas9蛋白或分裂的核碱基编辑器的表达水平中的各自的作用(例如，参见图4)。在一些实施方案中，本公开中使用的转录终止子是bGH转录终止子。在一些实施方案中，rAAV载体进一步包含土拨鼠肝炎病毒转录后调节元件(WPRE)。在一些实施方案中，WPRE插入转录终止子的5′。

[0291] 在一些实施方案中，包含rAAV颗粒(以本文考虑的任何形式)的组合物进一步包含药学上可接受的载体。在一些实施方案中，将组合物配制在适当的药学媒介物中，用于施用于人或动物受试者。

[0292] 可以充当药学上可接受的载体的材料的一些实例包括：(1)糖，例如乳糖、葡萄糖和蔗糖；(2)淀粉，如玉米淀粉和马铃薯淀粉；(3)纤维素及其衍生物，如羧甲基纤维素钠、甲基纤维素、乙基纤维素、微晶纤维素和醋酸纤维素；(4)粉末黄蓍胶；(5)麦芽；(6)明胶；(7)润滑剂，如硬脂酸镁、十二烷基硫酸钠和滑石粉；(8)赋形剂，如可可脂和栓剂蜡；(9)油，如花生油、棉籽油、红花油、芝麻油、橄榄油、玉米油和豆油；(10)二醇，如丙二醇；(11)多元醇，

如甘油、山梨糖醇、甘露醇和聚乙二醇(PEG)；(12)酯类，如油酸乙酯和月桂酸乙酯；(13)琼脂；(14)缓冲剂，如氢氧化镁和氢氧化铝；(15)海藻酸；(16)无热原水；(17)等渗盐水；(18)林格氏液；(19)乙醇；(20)pH缓冲溶液；(21)聚酯，聚碳酸酯和/或聚酸酐；(22)增量剂(bulking agent)，如多肽和氨基酸(23)血清成分，如血清白蛋白、HDL和LDL；(22)C2-C12醇，如乙醇；和(23)药物制剂中采用的其他无毒相容物质。润湿剂、着色剂、脱模剂、包衣剂、甜味剂、调味剂、芳香剂、防腐剂和抗氧化剂也可以存在于制剂中。诸如“赋形剂”、“载体”、“药学上可接受的载体”等术语在本文中可互换使用。

[0293] 使用方法

[0294] 本公开的其他方面提供了将分裂的Cas9蛋白或分裂的核碱基编辑器递送到细胞中以形成完整且功能性的Cas9蛋白或核碱基编辑器的方法。例如，在一些实施方案中，使细胞与本文所述的组合物(例如，包含编码分裂的Cas9或分裂的核碱基编辑器的核苷酸序列或含有包含此类核苷酸序列的核酸载体的AAV颗粒的组合物)接触。在一些实施方案中，接触导致将此类核苷酸序列递送到细胞中，其中Cas9蛋白或核碱基编辑器的N-末端部分和Cas9蛋白或核碱基编辑器的C-末端部分在细胞中表达并接合以形成完整的Cas9蛋白或完整的核碱基编辑器。

[0295] 与原始Cas9蛋白或核碱基编辑器(即，递送至细胞或在细胞中整体表达的未分裂的蛋白质)相比，使用本文所述方法递送的分裂的Cas9蛋白或分裂的核碱基编辑器优选具有相当的活性。例如，分裂的Cas9蛋白或分裂的核碱基编辑器保留至少50％(例如，至少50％、至少60％、至少70％、至少80％、至少90％、至少95％、至少98％、至少99％或100％)的原始Cas9蛋白或核碱基编辑器的活性。在一些实施方案中，分裂的Cas9蛋白或分裂的核碱基编辑器比原始Cas9蛋白或核碱基编辑器更有活性(例如，2倍、5倍、10倍、100倍、1000倍或更多)。

[0296] 可以以治疗有效量向有此需要的受试者施用本文所述的组合物，以治疗和/或预防受试者患有的疾病或病症。可以使用基于CRISPR/Cas9的基因组编辑技术治疗和/或预防的任何疾病或病症可以通过本文所述的分裂的Cas9蛋白或分裂的核碱基编辑器进行治疗。应理解，如果编码分裂的Cas9蛋白或核碱基编辑器的核苷酸序列不进一步编码gRNA，则编码gRNA的分开的核酸载体可以与本文所述的组合物一起施用。

[0297] 示例性合适的疾病和病症包括但不限于囊性纤维化(参见例如，Schwank et al.，Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients.Cell stem cell.2013;13:653-658；和Wu et.al.，Correction of a genetic disease in mouse via use of CRISPR-Cas9.Cell stem cell.2013;13:659-662，两者都不使用脱氨酶融合蛋白来纠正遗传缺陷)；苯丙酮尿症-例如，苯丙氨酸羟化酶基因中位置835(小鼠)或240(人)或同源残基处的苯丙氨酸至丝氨酸突变(T＞C突变)-参见例如，McDonald et al.，Genomics.1997;39:402-405；巨血小板综合征(BSS)-例如，血小板膜糖蛋白IX中位置55或同源残基处的苯丙氨酸至丝氨酸突变，或者残基24或同源残基处的半胱氨酸至精氨酸(T＞C突变)-参见例如，Noris et al.，British Journal of Haematology.1997;97:312-320，和Ali et al.，Hematol.2014;93:381-384；表皮松解性角化过度(EHK)-例如，角蛋白1中位置160或161(如果算上起始子甲硫氨酸)或同源残基处的亮氨酸至脯氨酸突变(T＞C突变)-参见例如，Chipev et al.，Cell.1992;70:

821-828,也参见位于www[dot]uniprot[dot]org的UNIPROT数据库中的登录号P04264;慢性阻塞性肺病(COPD)-例如,加工形式的α_1-抗胰蛋白酶中位置54或55(如果算上起始子甲硫氨酸)或同源残基或者未加工形式中的残基78或同源残基处的亮氨酸至脯氨酸突变(T>C突变)-参见例如,Poller et al.,Genomics.1993;17:740-743,也参见UNIPROT数据库中的登录号P01011;Charcot-Marie-Toot疾病4J型-例如,FIG4中位置41或同源残基处的异亮氨酸至苏氨酸突变(T>C突变)-参见例如,Lenk et al.,PLoS Genetics.2011;7:e1002104;神经母细胞瘤(NB)-例如,胱天蛋白酶-9中位置197或同源残基处的亮氨酸至脯氨酸突变(T>C突变)-参见例如,Kundu et al.,3 Biotech.2013,3:225-234;血管性血友病(vWD)-例如,加工形式的血管性血友病因子中位置509或同源残基处,或者未加工形式的血管性血友病因子中位置1272或同源残基处的半胱氨酸至精氨酸突变(T>C突变)-参见例如,Lavergne et al.,Br.J.Haematol.1992,也参见UNIPROT数据库中的登录号P04275;82:66-72;先天性肌强直-例如,肌肉氯通道基因CLCN1中位置277或同源残基处的半胱氨酸至精氨酸突变(T>C突变)-参见例如,Weinberger et al.,The J.of Physiology.2012;590:3449-3464;遗传性肾淀粉样变性-例如,加工形式的载脂蛋白AII中位置78或同源残基处,或者未加工形式中的位置101或同源残基处的终止密码子至精氨酸突变(T>C突变)-参见例如,Yazaki et al.,Kidney Int.2003;64:11-16;扩张型心肌病(DCM)-例如,FOXD4基因中位置148或同源残基处的色氨酸至精氨酸突变(T>C突变),参见例如,Minoretti et.al.,Int.J.of Mol.Med.2007;19:369-372;遗传性淋巴水肿-例如,VEGFR3酪氨酸激酶中位置1035或同源残基处组氨酸至精氨酸突变(A>G突变),参见例如,Irrthum et al.,Am.J.Hum.Genet.2000;67:295-301;家族性阿尔茨海默病-例如,衰老蛋白1中位置143或同源残基处的异亮氨酸至缬氨酸突变(A>G突变),参见例如,Gallo et.al.,J.Alzheimer's disease.2011;25:425-431;朊病毒病-例如,朊病毒蛋白中位置129或同源残基处甲硫氨酸至缬氨酸突变(A>G突变)-参见例如,Lewis et.al.,J.of General Virology.2006;87:2443-2449;慢性婴儿神经性皮肤关节综合征(CINCA)-例如,cryopyrin中位置570或同源残基处的酪氨酸至半胱氨酸突变(A>G突变)-参见例如,Fujisawa et.al.Blood.2007;109:2903-2911;和结蛋白相关性肌病(DRM)-例如,αβ晶体蛋白中位置120或同源残基处的精氨酸至甘氨酸突变(A>G突变)-参见例如,Kumar et al.,J.Biol.Chem.1999;274:24137-24141。所有参考文献和数据库条目的全部内容通过引用并入本文。

[0298] 施用组合物用于疼痛抑制的合适途径包括但不限于:局部、皮下、经皮、皮内、病灶内、关节内、腹膜内、膀胱内、经粘膜、牙龈、牙内、耳蜗内、经鼓室、器官内、硬膜外、鞘内、肌肉内、静脉内、血管内、骨内(intraosseus)、眼周、肿瘤内、脑内和脑室内施用。

[0299] 例如,本公开组合物可以作为单位剂量施用或包装为单位剂量。当用于提及本公开的药物组合物时,术语"单位剂量"是指适合作为受试者的单一剂量的物理上离散的单位,每个单位含有预定量的活性物质,其经计算与所需稀释剂(即载体或媒介物)联合产生所期望的治疗效果。

[0300] 疾病或病症的治疗包括延迟疾病的发展或进展,或降低疾病严重性。治疗疾病并不一定需要有治愈结果。

[0301] 如其中所用,"延迟"疾病的发展意指推迟、阻碍、减缓、延缓、稳定和/或延期疾病的进展。该延迟可以是不同的时间长度的,取决于疾病的历史和/或被治疗的个体。"延迟"

或减轻疾病的发展或延迟疾病的发作的方法是当与不使用该方法相比，降低在给定时间范围中发展一种或多种疾病的症状的可能性和/或降低在给定时间范围中症状的程度。此类比较通常基于临床研究，使用足以给出统计学显著结果的多个受试者。

[0302] 疾病的“发展”或“进展”意指疾病的初始表现和/或随后的进展。可以使用本领域熟知的标准临床技术检测和评估疾病的发展。然而，发展也指可以不可检测的进展。出于本公开的目的，发展或进展是指症状的生物学过程。“发展”包括发生、复发和发作。

[0303] 如本文所用，疾病的“发作”或“发生”包括初始发作和/或复发。医疗领域普通技术人员已知的常规方法可以用于将分离的多肽或药物组合物施用于受试者，取决于待治疗的疾病的类型或疾病的部位。

[0304] 试剂盒

[0305] 可以将本公开的组合物组装成试剂盒。在一些实施方案中，试剂盒包含用于表达本文所述的核碱基编辑器的核酸载体。在一些实施方案中，试剂盒进一步包含适当的指导核苷酸序列(例如gRNA)或用于表达此类指导核苷酸序列的核酸载体，以将Cas9蛋白或核碱基编辑器靶向所期望的靶序列。

[0306] 本文所述的试剂盒可以包括一个或多个容器，所述容器容纳用于实施本文所述方法的组分和任选的使用说明书。本文所述的任何试剂盒可以进一步包含实施测定方法所需的组分。如果适用，试剂盒的每种组分可以以液体形式(例如，以溶液)或以固体形式(例如，干粉)提供。在某些情况下，一些组分可以是可重构的或以其他方式可加工的(例如，成为活性形式)，例如，通过添加合适的溶剂或其他物质(例如水)，其可以或可以不随试剂盒提供。

[0307] 在一些实施方案中，试剂盒可以任选地包括用于使用所提供组分的说明书和/或广告宣传。如本文所用，“说明书”可以定义说明和/或广告宣传的组分，并且通常涉及针对本公开的包装或与之相关的书面说明书。说明书还可以包括以任何方式提供的任何口头或电子说明书，使得用户将清楚地认识到说明书与该试剂盒相关，例如，视听(例如，录像带、DVD等)，因特网和/或基于网络的通信等。书面说明书可以是以由管理药品或生物制品的制造、使用或销售的政府机构规定的形式，其也可以反映用于动物施用的制造、使用或销售的机构的批准。如本文所用，“宣传的”包括做生意的所有方法，其包括教育的方法、医院和其他临床指导、科学探究、药物发现或开发、学术研究、制药行业活动包括药品销售，以及任何广告或其他广告宣传活动包括与本公开相关的任何形式的书面、口头和电子通信。另外，如本文所述，取决于具体应用，试剂盒可以包括其他组分。

[0308] 试剂盒可以在一个或多个容器中含有本文所述的任何一种或多种组分。这些组分可以无菌制备，包装在注射器中并冷藏运输。或者，其可以容纳在小瓶或其他容器中用于储存。第二容器可以具有无菌制备的其他组分。或者，试剂盒可以包括预混合的活性剂并在小瓶、管或其他容器中运输。

[0309] 试剂盒可以具有多种形式，例如泡罩袋、收缩包装袋、真空可密封袋、可密封的热成型托盘或类似的袋或托盘形式，其中附件松散地包装在袋内，一个或多个管、容器、盒子或袋子。在添加附件后，可以对试剂盒进行灭菌，从而允许容器中的各个附件以其他方式拆开。可以使用任何适当的灭菌技术对试剂盒进行灭菌，例如辐射灭菌、加热灭菌或本领域已知的其他灭菌方法。试剂盒还可以包括其他组分，取决于具体应用，例如容器、细胞培养基、盐、缓冲液、试剂、注射器、针、织物例如纱布，用于施加或去除消毒剂、一次性手套、施用前

对试剂的支持物等。

[0310] 宿主细胞

[0311] 可以含有本文所述任何组合物的细胞包括原核细胞和真核细胞。本文所述的方法用于将Cas9蛋白或核碱基编辑器递送到真核细胞(例如哺乳动物细胞,例如人细胞)中。在一些实施方案中,细胞是体外的(例如,培养的细胞。在一些实施方案中,细胞是体内的(例如,在受试者如人受试者中)。在一些实施方案中,细胞是离体的(例如,从受试者分离的并且可以施用回相同或不同的受试者)。

[0312] 本公开的哺乳动物细胞包括人细胞、灵长类动物细胞(例如vero细胞)、大鼠细胞(例如GH3细胞、OC23细胞)或小鼠细胞(例如MC3T3细胞)。存在多种人细胞系,其包括但不限于,人胚肾(HEK)细胞、HeLa细胞、来自国家癌症研究所(National Cancer Institute)的60癌细胞系(NCI60)的癌症细胞、DU145(前列腺癌)细胞、Lncap(前列腺癌)细胞、MCF-7(乳腺癌)细胞、MDA-MB-438(乳腺癌)细胞、PC3(前列腺癌)细胞、T47D(乳腺癌)细胞、THP-1(急性髓样白血病)细胞、U87(胶质母细胞瘤)细胞、SHSY5Y人神经母细胞瘤细胞(克隆自骨髓瘤)和Saos-2(骨癌)细胞。在一些实施方案中,rAAV载体被递送到人胚肾(HEK)细胞(例如HEK 293或HEK 293T细胞)中。在一些实施方案中,rAAV载体被递送到干细胞(例如人干细胞)中,诸如例如多能干细胞(例如人多能干细胞包括人诱导多能干细胞(hiPSC))。干细胞是指具有在培养中无限期分裂并产生特化细胞的能力的细胞。多能干细胞是指一种类型的干细胞,其能够分化成生物体的所有组织,但不能单独支撑完整的生物体发育。人诱导多能干细胞是指通过被迫表达对维持胚胎干细胞的定义性质重要的基因和因子而重编程为胚胎干细胞样状态的体细胞(例如,成熟或成体)细胞(参见例如,Takahashi and Yamanaka,Cell 126(4):663-76,2006,通过引用本文并入)。人诱导多能干细胞表达干细胞标志物并且能够产生所有三个胚层(外胚层、内胚层、中胚层)的特征性细胞。

[0313] 可以根据本公开使用的细胞系的另外的非限制性实例包括293-T、293-T、3T3、4T1、721、9L、A-549、A172、A20、A253、A2780、A2780ADR、A2780cis、A431、ALC、B16、B35、BCP-1、BEAS-2B、bEnd.3、BHK-21、BR 293、BxPC3、C2C12、C3H-10T1/2、C6、C6/36、Cal-27、CGR8、CHO、CML T1、CMT、COR-L23、COR-L23/5010、COR-L23/CPR、COR-L23/R23、COS-7、COV-434、CT26、D17、DH82、DU145、DuCaP、E14Tg2a、EL4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、Hepa1c1c7、High Five细胞、HL-60、HMEC、HT-29、HUVEC、J558L细胞、Jurkat、JY细胞、K562细胞、KCL22、KG1、Ku812、KYO1、LNCap、Ma-Mel1、2、3....48、MC-38、MCF-10A、MCF-7、MDA-MB-231、MDA-MB-435、MDA-MB-468、MDCK II、MG63、MONO-MAC 6、MOR/0.2R、MRC5、MTD-1A、MyEnd、NALM-1、NCI-H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、NW-145、OPCN/OPCT Peer、PNT-1A/PNT 2、PTK2、Raji、RBL细胞、RenCa、RIN-5F、RMA/RMAS、S2、Saos-2细胞、Sf21、Sf9、SiHa、SKBR3、SKOV-3、T-47D、T2、T84、THP1、U373、U87、U937、VCaP、WM39、WT-49、X63、YAC-1和YAR细胞。

[0314] 无需进一步详细阐述,相信本领域技术人员基于以上描述可以最充分地利用本公开。因此,以下具体实施方案应被解释为仅是说明性的,并且不以任何方式限制本公开的其余部分。出于本文引用的目的或主题,本文引用的所有出版物均通过引用并入。

实施例

[0315] 为了可以更全面地理解本文所述的发明,提出以下实施例。提供本申请中描述的合成实施例以说明本文提供的化合物和方法,并且不以任何方式解释为限制其范围。

[0316] 实施例1:Cas9蛋白和核碱基编辑器的氨基酸序列

[0317] 提供了合适的Cas9蛋白和变体,以及核碱基编辑器和变体的非限制性实例。本公开提供了Cas9变体,例如来自一种或多种生物体的Cas9蛋白,其可以包含一个或多个突变(例如,以产生dCas9或Cas9切口酶)。在一些实施方案中,Cas9蛋白的一个或多个氨基酸残基(下文用星号标识)可以得以突变。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的D10和/或H840残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的突变,得以突变。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的D10残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的突变,得以突变为任何氨基酸残基,除了D。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的D10残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的突变,得以突变为A。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的H840残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的残基,是H。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的H840残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的突变,得以突变为任何氨基酸残基,除了H。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的H840残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的突变,得以突变为A。在一些实施方案中,SEQ ID NO:1中提供的氨基酸序列的D10残基,或SEQ ID NO:2-275和394-397中提供的任何氨基酸序列中相应的残基,是D。

[0318] 对来自各种物种的许多Cas9序列进行比对以确定是否可以在其他Cas9蛋白中鉴定SEQ ID NO:1的D10和H840的相应的同源氨基酸残基,从而允许产生具有同源氨基酸残基的相应的突变的Cas9变体。使用NCBI基于约束的多重比对工具(NCBI Constraint-based Multiple Alignment Tool)(COBALT(可在st-va.ncbi.nlm.nih.gov/tools/cobalt获得)进行比对,具有以下参数。比对参数:空位罚分-11,-1;末端空位罚分-5,-1。CDD参数:使用RPS BLAST开启(on);Blast E值0.003;查找保守列和重新计算开启。查询聚类参数:使用查询聚类开启;词大小4;最大聚类距离0.8;字母常规。

[0319] 酿脓链球菌Cas9野生型(NCBI参考序列:NC_002737.2,Uniprot参考序列:Q99ZW2)

[0320]

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS
DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN
FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 1)

[0321] 酿脓链球菌dCas9(D10A和H840A)

[0322]

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS
DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN
FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL

[0323]

NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI
DLSQLGGD (SEQ ID NO: 2)

[0324] 酿脓链球菌Cas9切口酶(D10A)

[0325]

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS
DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN
FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 3)

[0326] VRER-nCas9(D10A/D1135V/G1218R/R1335E/T1337R)酿脓链球菌Cas9切口酶

[0327]

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS
DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN
FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI

[0328]

NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKEYRSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 4)

[0329] VQR-nCas9(D10A/D1135V/R1335Q/T1337R)酿脓链球菌Cas9切口酶

[0330]

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS
DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN
FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 5)

[0331] EQR-nCas9(D10A/D1135E/R1335Q/T1337R)酿脓链球菌Cas9切口酶

[0332]

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS
DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN
FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK

[0333]

GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFESPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 6)

[0334] KKH-nCas9(D10A/E782K/N968K/R1015H)金黄色葡萄球菌Cas9切口酶

[0335]

MKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR
RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGV
HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYV
KEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMG
HCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKP
TLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILT
IYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIF
NRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELARE
KNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEA
IPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISY
ETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLL
RSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEW
KKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK
KPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQ
TYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDI
TDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEA
KKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMND
KRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 7)

[0336] 嗜热链球菌CRISPR1 Cas9(St1Cas9)切口酶(D9A)

[0337]

MSDLVLGLAIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLTRRKK
HRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYL
DDASDDGNSSIGDYAQIVKENSKQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHR
LINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYYHGPGNEKSRTDYGR
YRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQ
KNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTL
ETLDIEQMDRETLDKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSS
IFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY
NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAA
MLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQF
EVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRES
KTLSNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVS
VVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLL
DIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATR
QAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPI
LENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDIT
PKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKGTGTYKISQEKYNDI
KKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHYVELKPYDKQK

[0338]

FEGGEALIKVLGNVANSGQCKKGLGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF
(SEQ ID NO: 8)

[0339] 嗜热链球菌CRISPR3Cas9(St3Cas9)切口酶(D10A)

[0340]

MTKPYSIGLAIGTNSVGWAVITDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITA
EGRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIF
GNLVEEKVYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSK
NNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIF
SEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKKLY
DAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKN
GYAGYIDGKTNQEDFYVYLKNLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHL
QEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPW
NFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRD
YQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDL
LNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGW
GKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKG
NIKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQ
QRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDID
RLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDFPSLEVVKKRKTFWYQLLKSKL
ISQRKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRAV
RTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVIASALLKKYPKLEPEFVY
GDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKE
SDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAK
EYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLL
EKGYKDIELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHA
KRISNTINENHRKYVENHKKEFEELFYYILEFNENYVGAKKNGKLLNSAFQSWQNHSID
ELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGL
YETRIDLAKLGEG (SEQ ID NO: 9)

[0341] 金黄色葡萄球菌Cas9野生型

[0342]

MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR
RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGV
HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYV
KEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMG
HCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKP
TLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILT
IYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIF
NRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELARE
KNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEA
IPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISY
ETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLL
RSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEW
KKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK
KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQ
TYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDI
TDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEA
KKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMND
KRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 10)

[0343] 金黄色葡萄球菌Cas9切口酶(D10A)

[0344]

MKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR
RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGV
HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYV
KEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMG
HCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKP
TLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILT
IYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIF
NRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELARE
KNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEA
IPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISY
ETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLL
RSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEW
KKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK
KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQ
TYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDI
TDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEA
KKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMND
KRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 11)

[0345] 嗜热链球菌野生型CRISPR3Cas9(St3Cas9)

[0346]

MTKPYSIGLDIGTNSVGWAVITDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITA
EGRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIF
GNLVEEKVYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSK
NNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIF
SEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKKLY
DAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKN
GYAGYIDGKTNQEDFYVYLKNLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHL
QEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPW
NFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRD
YQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDL
LNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGW
GKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKG
NIKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQ
QRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDID
RLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDFPSLEVVKKRKTFWYQLLKSKL
ISQRKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRAV
RTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVIASALLKKYPKLEPEFVY
GDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKE
SDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAK
EYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLL
EKGYKDIELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHA
KRISNTINENHRKYVENHKKEFEELFYYILEFNENYVGAKKNGKLLNSAFQSWQNHSID
ELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGL
YETRIDLAKLGEG (SEQ ID NO: 12)

[0347] 嗜热链球菌CRISPR1 Cas9野生型(St1Cas9)

[0348]

MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLTRRKK
HRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYL
DDASDDGNSSIGDYAQIVKENSKQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHR
LINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYYHGPGNEKSRTDYGR
YRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQ

[0349]

KNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTL
ETLDIEQMDRETLDKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSS
IFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY
NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAA
MLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQF
EVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRES
KTLSNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVS
VVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLL
DIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATR
QAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPI
LENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDIT
PKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKGTGTYKISQEKYNDI
KKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHYVELKPYDKQK
FEGGEALIKVLGNVANSGQCKKGLGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF
(SEQ ID NO: 13)

[0350] 来自冰岛硫化叶菌(Sulfolobus islandicus)(菌株REY15A)的CasX

[0351]

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEDAAAER
RGKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQVKEC
EEVSAPSFVKPEFYKFGRSPGMVERTRRVKLEVEPHYLIMAAAGWVLTRLGKAKVSEG
DYVGVNVFTPTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSVVSIYTIS
DAVGQNPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMRERYIVLANYIYEYL
TGSKRLEDLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG (SEQ ID NO: 14)

[0352] 来自冰岛硫化叶菌(菌株REY15A)的CasY

[0353]

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEDAAAER
RGKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQVKEC
EEVSAPSFVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIIAAAGWVLTRLGKAKVSEGD
YVGVNVFTPTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSVVRIYTISD
AVGQNPTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMRERYIVLANYIYEYLT
GSKRLEDLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG (SEQ ID NO: 15)

[0354] 野生型新凶手弗朗西斯菌(Francisella novicida)Cpf1(D917、E1006和D1255是粗体且加下划线的)

[0355]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**D**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE

[0356]

KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
DANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 16)

[0357] 新凶手弗朗西斯菌Cpf1 D917A(A917、E1006和D1255是粗体且加下划线的)

[0358]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**A**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
DANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 17)

[0359] 新凶手弗朗西斯菌Cpf1 E1006A(D917、A1006和D1255是粗体且加下划线的)

[0360]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**D**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA

[0361]

AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
DANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 18)

[0362] 新凶手弗朗西斯菌Cpf1 D1255A(D917、E1006和A1255是粗体且加下划线的)

[0363]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**D**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 19)

[0364] 新凶手弗朗西斯菌Cpf1 D917A/E1006A(A917、A1006和D1255是粗体且加下划线的)

[0365]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**A**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
DANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 20)

[0366] 新凶手弗朗西斯菌Cpf1 D917A/D1255A(A917、E1006和A1255是粗体且加下划线的)

[0367]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 21)

[0368] 新凶手弗朗西斯菌Cpf1 E1006A/D1255A(D917、A1006和A1255是粗体且加下划线的)

[0369]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**D**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 22)

[0370] 新凶手弗朗西斯菌Cpf1 D917A/E1006A/D1255A(A917、A1006和A1255是粗体且加下划线的)

[0371]

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQF
FIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKN
LFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGF
HENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELT
FDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYI
NLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFK
TVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
QIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFA
AIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIF
HISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENST
LANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKL
LPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFID
FYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGK
LYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKI
THPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINL
LLKEKANDVHILSI**A**RGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIE
KDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKV
EKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYV
PAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKA
AKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIC
GESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 23)

[0372] 野生型格氏嗜盐碱杆菌(Natronobacterium gregoryi)Argonaute

[0373]

MTVIDLDSTTTADELTSGHTYDISVTLTGVYDNTDEQHPRMSLAFEQDNGERRYITLWK
NTTPKDVFTYDYATGSTYIFTNIDYEVKDGYENLTATYQTTVENATAQEVGTTDEDETF
AGGEPLDHHLDDALNETPDDAETESDSGHVMTSFASRDQLPEWTLHTYTLTATDGAKT
DTEYARRTLAYTVRQELYTDHDAAPVATDGLMLLTPEPLGETPLDLDCGVRVEADETR
TLDYTTAKDRLLARELVEEGLKRSLWDDYLVRGIDEVLSKEPVLTCDEFDLHERYDLS
VEVGHSGRAYLHINFRHRFVPKLTLADIDDDNIYPGLRVKTTYRPRRGHIVWGLRDECA
TDSLNTLGNQSVVAYHRNNQTPINTDLLDAIEAADRRVVETRRQGHGDDAVSFPQELL
AVEPNTHQIKQFASDGFHQQARSKTRLSASRCSEKAQAFAERLDPVRLNGSTVEFSSEFF
TGNNEQQLRLLYENGESVLTFRDGARGAHPDETFSKGIVNPPESFEVAVVLPEQQADTC
KAQWDTMADLLNQAGAPPTRSETVQYDAFSSPESISLNVAGAIDPSEVDAAFVVLPPDQ
EGFADLASPTETYDELKKALANMGIYSQMAYFDRFRDAKIFYTRNVALGLLAAAGGV
AFTTEHAMPGDADMFIGIDVSRSYPEDGASGQINIAATATAVYKDGTILGHSSTRPQLGE
KLQSTDVRDIMKNAILGYQQVTGESPTHIVIHRDGFMNEDLDPATEFLNEQGVEYDIVEI
RKQPQTRLLAVSDVQYDTPVKSIAAINQNEPRATVATFGAPEYLATRDGGGLPRPIQIER
VAGETDIETLTRQVYLLSQSHIQVHNSTARLPITTAYADQASTHATKGYLVQTGAFESN
VGFL (SEQ ID NO: 24)

[0374] 具有降低的Cas9和DNA主链之间的静电相互作用的Cas9变体

[0375]

DKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG
NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD
AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNG

[0376]

YAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE
LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF
EEVVDKGASAQSFIERMT**A**FDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK
PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRR
YTGWG**A**LSRKLINGIRDKQSGKTILDFLKSDGFANRNFM**A**LIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR**A**ITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 25)CasY (ncbi.nlm.nih.gov/protein/APG80656.1)

[0377] ＞APG80656.1 CRISPR相关蛋白CasY[未培养的Parcubacteria组细菌]

[0378]

MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDDY
VGLYGLSNFDDLYNAEKRNEEKVYSVLDFWYDCVQYGAVFSYTAPGLLKNVAEVRG
GSYELTKTLKGSHLYDELQIDKVIKFLNKKEISRANGSLDKLKKDIIDCFKAEYRERHKD
QCNKLADDIKNAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTCCLLPFD
TVNNNRNRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSGTAFSNFLGEGFLGRLR
ENKITELKKAMMDITDAWRGQEQEEELEKRLRILAALTIKLREPKFDNHWGGYRSDIN
GKLSSWLQNYINQTVKIKEDLKGHKKDLKKAKEMINRFGESDTKEEAVVSSLLESIEKI
VPDDSADDEKPDIPAIAIYRRFLSDGRLTLNRFVQREDVQEALIKERLEAEKKKKPKKRK
KKSDAEDEKETIDFKELFPHLAKPLKLVPNFYGDSKRELYKKYKNAAIYTDALWKAVE
KIYKSAFSSSLKNSFFDTDFDKDFFIKRLQKIFSVYRRFNTDKWKPIVKNSFAPYCDIVSL
AENEVLYKPKQSRSRKSAAIDKNRVRLPSTENIAKAGIALARELSVAGFDWKDLLKKEE
HEEYIDLIELHKTALALLLAVTETQLDISALDFVENGTVKDFMKTRDGNLVLEGRFLEM
FSQSIVFSELRGLAGLMSRKEFITRSAIQTMNGKQAELLYIPHEFQSAKITTPKEMSRAFL
DLAPAEFATSLEPESLSEKSLLKLKQMRYYPHYFGYELTRTGQGIDGGVAENALRLEKS
PVKKREIKCKQYKTLGRGQNKIVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLI
DEKKVKTRWNYDALTVALEPVSGSERVFVSQPFTIFPEKSAEEEGQRYLGIDIGEYGIAY
TALEITGDSAKILDQNFISDPQLKTLREEVKGLKLDQRRGTFAMPSTKIARIRESLVHSLR
NRIHHLALKHKAKIVYELEVSRFEEGKQKIKKVYATLKKADVYSEIDADKNLQTTVWG
KLAVASEISASYTSQFCGACKKLWRAEMQVDETITTQELIGTVRVIKGGTLIDAIKDFMR
PPIFDENDTPFPKYRDFCDKHHISKKMRGNSCLFICPFCRANADADIQASQTIALLRYVK
EEKKVEDYFERFRKLKNIKVLGQMKKI (SEQ ID NO: 26)

[0379] 高保真Cas9结构域

[0380]

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG
NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD
AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNG
YAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE
LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF

[0381]

EEVVDKGASAQSFIERMTAFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK
PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRR
YTGWGALSRKLINGIRDKQSGKTILDFLKSDGFANRNFMALIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRAITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 394)

[0382] C2c1(uniprot.org/uniprot/T0D7A2#)

[0383] sp|T0D7A2|C2C1_ALIAG CRISPR相关内切核酸酶C2c1 OS＝酸土脂环酸芽孢杆菌(Alicyclobacillus acidoterrestris)(菌株ATCC 49025/DSM 3922/CIP 106132/NCIMB 13137/GD3B)GN＝c2c1 PE＝1 SV＝1

[0384]

MAVKSIKVKLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNGDG
EQECDKTAEECKAELLERLRARQVENGHRGPAGSDDELLQLARQLYELLVPQAIGAKG
DAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEKEKAETRKSA
DRTADVLRALADFGLKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQAIERM
MSWESWNQRVGQEYAKLVEQKNRFEQKNFVGQEHLVHLVNQLQQDMKEASPGLESK
EQTAHYVTGRALRGSDKVFEKWGKLAPDAPFDLYDAEIKNVQRRNTRRFGSHDLFAK
LAEPEYQALWREDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGG
NLHQYTFLFNEFGERRHAIRFHKLLKVENGVAREVDDVTVPISMSEQLDNLLPRDPNEPI
ALYFRDYGAEQHFTGEFGGAKIQCRRDQLAHMHRRRGARDVYLNVSVRVQSQSEARG
ERRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDGKLGSEGLLSGLRVMSVDLGLR
TSASISVFRVARKDELKPNSKGRVPFFFPIKGNDNLVAVHERSQLLKLPGETESKDLRAI
REERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSWAKLIEQPVDAANHMTPDWREA
FENELQKLKSLHGICSDKEWMDAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRG
YAKDVVGGNSIEQIEYLERQYKFLKSWSFFGKVSGQVIRAEKGSRFAITLREHIDHAKE
DRLKKLADRIIMEALGYVYALDERGKGKWVAKYPPCQLILLEELSEYQFNNDRPPSEN
NQLMQWSHRGVFQELINQAQVHDLLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQ
EHNPEPFPWWLNKFVVEHTLDACPLRADDLIPTGEGEIFVSPFSAEEGDFHQIHADLNA
AQNLQQRLWSDFDISQIRLRCDWGEVDGELVLIPRLTGKRTADSYSNKVFYTNTGVTY
YERERGKKRRKVFAQEKLSEEEAELLVEADEAREKSVVLMRDPSGIINRGNWTRQKEF
WSMVNQRIEGYLVKQIRSRVPLQDSACENTGDI (SEQ ID NO: 395)

[0385] C2c2(uniprot.org/uniprot/P0D0C6)

[0386] ＞sp|P0D0C6|C2C2_LEPSD CRISPR相关内切核糖核酸酶C2c2 OS＝Leptotrichia shahii(菌株DSM 19757/CCUG 47503/CIP 107916/JCM 16776/LB37)GN＝c2c2 PE＝1 SV＝1

[0387]

MGNLFGHKRWYEVRDKKDFKIKRKVKVKRNYDGNKYILNINENNNKEKIDNNKFIRK
YINYKKNDNILKEFTRKFHAGNILFKLKGKEGIIRIENNDDFLETEEVVLYIEAYGKSEKL
KALGITKKKIIDEAIRQGITKDDKKIEIKRQENEEEIEIDIRDEYTNKTLNDCSIILRIIENDE
LETKKSIYEIFKNINMSLYKIIEKIIENETEKVFENRYYEEHLREKLLKDDKIDVILTNFME
IREKIKSNLEILGFVKFYLNVGGDKKKSKNKKMLVEKILNINVDLTVEDIADFVIKELEF
WNITKRIEKVKKVNNEFLEKRRNRTYIKSYVLLDKHEKFKIERENKKDKIVKFFVENIK

[0388]

NNSIKEKIEKILAEFKIDELIKKLEKELKKGNCDTEIFGIFKKHYKVNFDSKKFSKKSDEE
KELYKIIYRYLKGRIEKILVNEQKVRLKKMEKIEIEKILNESILSEKILKRVKQYTLEHIMY
LGKLRHNDIDMTTVNTDDFSRLHAKEELDLELITFFASTNMELNKIFSRENINNDENIDF
FGGDREKNYVLDKKILNSKIKIIRDLDFIDNKNNITNNFIRKFTKIGTNERNRILHAISKER
DLQGTQDDYNKVINIIQNLKISDEEVSKALNLDVVFKDKKNIITKINDIKISEENNNDIKY
LPSFSKVLPEILNLYRNNPKNEPFDTIETEKIVLNALIYVNKELYKKLILEDDLEENESKNI
FLQELKKTLGNIDEIDENIIENYYKNAQISASKGNNKAIKKYQKKVIECYIGYLRKNYEE
LFDFSDFKMNIQEIKKQIKDINDNKTYERITVKTSDKTIVINDDFEYIISIFALLNSNAVINK
IRNRFFATSVWLNTSEYQNIIDILDEIMQLNTLRNECITENWNLNLEEFIQKMKEIEKDFD
DFKIQTKKEIFNNYYEDIKNNILTEFKDDINGCDVLEKKLEKIVIFDDETKFEIDKKSNILQ
DEQRKLSNINKKDLKKKVDQYIKDKDQEIKSKILCRIIFNSDFLKKYKKEIDNLIEDMESE
NENKFQEIYYPKERKNELYIYKKNLFLNIGNPNFDKIYGLISNDIKMADAKFLFNIDGKNI
RKNKISEIDAILKNLNDKLNGYSKEYKEKYIKKLKENDDFFAKNIQNKNYKSFEKDYNR
VSEYKKIRDLVEFNYLNKIESYLIDINWKLAIQMARFERDMHYIVNGLRELGIIKLSGYN
TGISRAYPKRNGSDGFYTTTAYYKFFDEESYKKFEKICYGFGIDLSENSEINKPENESIRN
YISHFYIVRNPFADYSIAEQIDRVSNLLSYSTRYNNSTYASVFEVFKKDVNLDYDELKKK
FKLIGNNDILERLMKPKKVSVLELESYNSDYIKNLIIELLTKIENTNDTL (SEQ ID NO:
396)

[0389] C2c3,翻译自＞CEPX01008730.1海洋宏基因组基因组组装TARA_037_MES_0.1-0.22,contig TARA_037_MES_0.1-0.22_支架22115_1,全基因组鸟枪序列。

[0390]

MRSNYHGGRNARQWRKQISGLARRTKETVFTYKFPLETDAAEIDFDKAVQTYGIAEGV
GHGSLIGLVCAFHLSGFRLFSKAGEAMAFRNRSRYPTDAFAEKLSAIMGIQLPTLSPEGL
DLIFQSPPRSRDGIAPVWSENEVRNRLYTNWTGRGPANKPDEHLLEIAGEIAKQVFPKFG
GWDDLASDPDKALAAADKYFQSQGDFPSIASLPAAIMLSPANSTVDFEGDYIAIDPAAE
TLLHQAVSRCAARLGRERPDLDQNKGPFVSSLQDALVSSQNNGLSWLFGVGFQHWKE
KSPKELIDEYKVPADQHGAVTQVKSFVDAIPLNPLFDTTHYGEFRASVAGKVRSWVAN
YWKRLLDLKSLLATTEFTLPESISDPKAVSLFSGLLVDPQGLKKVADSLPARLVSAEEAI
DRLMGVGIPTAADIAQVERVADEIGAFIGQVQQFNNQVKQKLENLQDADDEEFLKGLK
IELPSGDKEPPAINRISGGAPDAAAEISELEEKLQRLLDARSEHFQTISEWAEENAVTLDPI
AAMVELERLRLAERGATGDPEEYALRLLLQRIGRLANRVSPVSAGSIRELLKPVFMEER
EFNLFFHNRLGSLYRSPYSTSRHQPFSIDVGKAKAIDWIAGLDQISSDIEKALSGAGEALG
DQLRDWINLAGFAISQRLRGLPDTVPNALAQVRCPDDVRIPPLLAMLLEEDDIARDVCL
KAFNLYVSAINGCLFGALREGFIVRTRFQRIGTDQIHYVPKDKAWEYPDRLNTAKGPIN
AAVSSDWIEKDGAVIKPVETVRNLSSTGFAGAGVSEYLVQAPHDWYTPLDLRDVAHLV
TGLPVEKNITKLKRLTNRTAFRMVGASSFKTHLDSVLLSDKIKLGDFTIIIDQHYRQSVT
YGGKVKISYEPERLQVEAAVPVVDTRDRTVPEPDTLFDHIVAIDLGERSVGFAVFDIKSC
LRTGEVKPIHDNNGNPVVGTVAVPSIRRLMKAVRSHRRRRQPNQKVNQTYSTALQNYR
ENVIGDVCNRIDTLMERYNAFPVLEFQIKNFQAGAKQLEIVYGS (SEQ ID NO: 397)

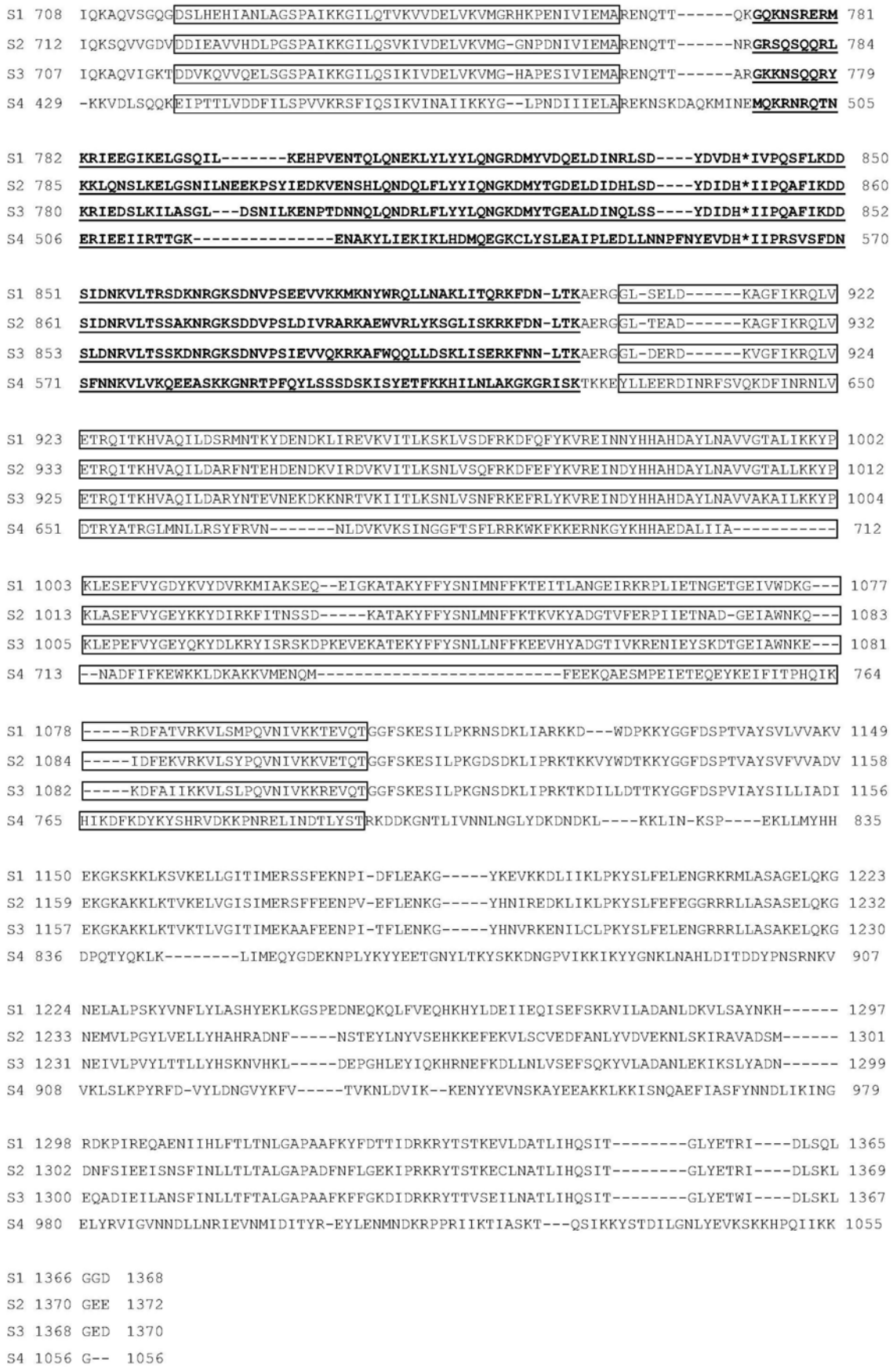
[0391] 以下提供了四个Cas9序列的示例性比对。比对中的Cas9序列是:序列1(S1):SEQ ID NO:1|WP_010922251|gi 499224711|II型CRISPR RNA引导的内切核酸酶Cas9[酿脓链球菌];序列2(S2):SEQ ID NO:27|WP_039695303|gi 746743737|II型CRISPR RNA引导的内切核酸酶Cas9[解没食子酸链球菌(Streptococcus gallolyticus)];序列3(S3):SEQ ID NO:28|WP_045635197|gi 782887988|II型CRISPR RNA引导的内切核酸酶Cas9[缓症链球菌(Streptococcus mitis)];序列4(S4):SEQ ID NO:29|5AXW_A|gi 924443546|金黄色葡萄球菌Cas9。针对四个序列的每一个鉴定HNH结构域(粗体且加下划线)和RuvC结构域(加框)。

S1中的氨基酸残基10和840以及比对序列中的同源氨基酸在相应的氨基酸残基后用星号标识。

[0392]

```
S1 1    --MDKK-YSIGLD*IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI--GALLFDSG--ETAEATRLKRTARRYT  73
S2 1    --MTKKNYSIGLD*IGTNSVGWAVITDDYKVPAKKMKVLGNTDKKYIKKNLL--GALLFDSG--ETAEATRLKRTARRYT  74
S3 1    --M-KKGYSIGLD*IGTNSVGFAVITDDYKVPSKKMKVLGNTDKRFIKKNLI--GALLFDEG--TTAEARRLKRTARRYT  73
S4 1    GSHMKRNYILGLD*IGITSVGYGII--DYET----------------RDVIDAGVRLFKEANVENNEGRRSKRGARRLKR  61

S1 74   RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL 153
S2 75   RRKNRLRYLQEIFANEIAKVDESFFQRLDESFLTDDDKTFDSHPIFGNKAEEDAYHQKFPTIYHLRKHLADSSEKADLRL 154
S3 74   RRKNRLRYLQEIFSEEMSKVDSSFFHRLDDSFLIPEDKRESKYPIFATLTEEKEYHKQFPTIYHLRKQLADSKEKTDLRL 153
S4 62   RRRHRIQRVKKLL--------------FDYNLLTD--------------------HSELSGINPYEARVKGLSQKLSEEE 107

S1 154  IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEK  233
S2 155  VYLALAHMIKFRGHFLIEGELNAENTDVQKIFADFVGVYNRTFDDSHLSEITVDVASILTEKISKSRRLENLIKYYPTEK  234
S3 154  IYLALAHMIKYRGHFLYEEAFDIKNNDIQKIFNEFISIYDNTFEGSSLSGQNAQVEAIFTDKISKSAKRERVLKLFPDEK  233
S4 108  FSAALLHLAKRRG----------------------VHNVNEVEEDT---------------------------------  131

S1 234  KNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT 313
S2 235  KNTLFGNLIALALGLQPNFKTNFKLSEDAKLQFSKDTYEEDLEELLGKIGDDYADLFTSAKNLYDAILLSGILTVDDNST 314
S3 234  STGLFSEFLKLIVGNQADFKKHFDLEDKAPLQFSKDTYDEDLENLLGQIGDDFTDLFVSAKKLYDAILLSGILTVTDPST 313
S4 132  -----GNELS-----------------TKEQISRN---------------------------------------------  144

S1 314  KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM--DGTEELLV  391
S2 315  KAPLSASMIKRYVEHHEDLEKLKEFIKANKSELYHDIFKDKNKNGYAGYIENGVKQDEFYKYLKNILSKIKIDGSDYFLD  394
S3 314  KAPLSASMIERYENHQNDLAALKQFIKNNLPEKYDEVFSDQSKDGYAGYIDGKTTQETFYKYIKNLLSKF--EGTDYFLD  391
S4 145  ----SKALEEKYVAELQ-----------------------------------------------LERLKKDG------  165

S1 392  KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE  471
S2 395  KIEREDFLRKQRTFDNGSIPHQIHLQEMHAILRRQGDYYPFLKEKQDRIEKILTFRIPYYVGPLVRKDSRFAWAEYRSDE  474
S3 392  KIEREDFLRKQRTFDNGSIPHQIHLQEMNAILRRQGEYYPFLKDNKEKIEKILTFRIPYYVGPLARGNRDFAWLTRNSDE  471
S4 166  --EVRGSINRFKTSD--------YVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGP--GEGSPFGW------K  227

S1 472  TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL 551
S2 475  KITPWNFDKVIDKEKSAEKFITRMTLNDLYLPEEKVLPKHSHVYETYAVYNELTKIKYVNEQGKE-SFFDSNMKQEIFDH 553
S3 472  AIRPWNFEEIVDKASSAEDFINKMTNYDLYLPEEKVLPKHSLLYETFAVYNELTKVKFIAEGLRDYQFLDSGQKKQIVNQ 551
S4 228  DIKEW---------------YEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEK---LEYYEKFQIIEN 289

S1 552  LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR---FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED  628
S2 554  VFKENRKVTKEKLLNYLNKEFPEYRIKDLIGLDKENKSFNASLGTYHDLKKIL-DKAFLDDKVNEEVIEDIIKTLTLFED  632
S3 552  LFKENRKVTEKDIIHYLHN-VDGYDGIELKGIEKQ---FNASLSTYHDLLKIIKDKEFMDDAKNEAILENIVHTLTIFED  627
S4 290  VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF---TNLKVYHDIKDITARKEII---ENAELLDQIAKILTIYQS  363

S1 629  REMIEERLKTYAHLFDDKVMKQLKR-RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED 707
S2 633  KDMIHERLQKYSDIFTANQLKKLER-RHYTGWGRLSYKLINGIRNKENNKTILDYLIDDGSANRNFMQLINDDTLPFKQI 711
S3 628  REMIKQRLAQYDSLFDEKVIKALTR-RHYTGWGKLSAKLINGICDKQTGNTILDYLIDDGKINRNFMQLINDDGLSFKEI 706
S4 364  SEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDE------LWHTNDNQIAIFNRLKLVP--------- 428
```

[0393]



[0394] 比对证明,可以通过使用本领域已知的比对程序和算法鉴定与参照序列或参照残

基比对的氨基酸序列或残基，来在Cas9序列变体间鉴定与参照Cas9氨基酸序列或氨基酸残基同源的氨基酸序列和氨基酸残基，所述Cas9序列变体包括但不限于来自不同物种的Cas9序列。本公开提供了Cas9变体，其中如本文所述突变通过SEQ ID NO:1和27-29(例如，分别为S1、S2、S3和S4)中的星号标识的一个或多个氨基酸残基。SEQ ID NO:1的Cas9中的残基D10和H840(其对应于SEQ ID NO:1和27-29中通过星号标识的残基)在本文中称为“同源的”或“相应的”残基。此类同源的残基可以通过序列比对来鉴定，例如，如上所述，并通过鉴定与参照序列或残基比对的序列或残基来鉴定。类似地，对应于本文SEQ ID NO:1中鉴定的突变的Cas9序列中的突变，例如SEQ ID NO:1中的残基10和840的突变，在本文中称为“同源的”或“相应的”突变。例如，对于上述四个比对序列，对应于SEQ ID NO:1(S1)中的D10A突变的突变是S2的D11A、S3的D10A和S4的D13A;SEQ ID NO:1(S1)中H840A的相应的突变是S2的H850A、S3的H842A和S4的H560A。

[0395] 提供了来自不同物种的总共250个Cas9序列(SEQ ID NO:1和27-275)。对应于SEQ ID NO:1的残基10和840的氨基酸残基可以以与上述相同的方式鉴定。可以根据本公开使用所有这些Cas9序列。

[0396]

WP_010922251.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 1

WP_039695303.1 II型CRISPR RNA引导的内切核酸酶Cas9 [解没食子酸链球菌] SEQ ID NO: 27

WP_045635197.1 II型CRISPR RNA引导的内切核酸酶Cas9 [缓症链球菌] SEQ ID NO: 28

5AXW_A Cas9, A链, 晶体结构 [金黄色葡萄球菌] SEQ ID NO: 29

WP_009880683.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 30

WP_010922251.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 31

WP_011054416.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 32

WP_011284745.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 33

WP_011285506.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 34

WP_011527619.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 35

WP_012560673.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 36

WP_014407541.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 37

WP_020905136.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 38

WP_023080005.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 39

WP_023610282.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 40

WP_030125963.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 41

WP_030126706.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 42

WP_031488318.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 43

WP_032460140.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 44

WP_032461047.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 45

WP_032462016.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 46

WP_032462936.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 47

WP_032464890.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 48

[0397]

WP_033888930.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 49

WP_038431314.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 50

WP_038432938.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 51

WP_038434062.1 II型CRISPR RNA引导的内切核酸酶Cas9 [酿脓链球菌] SEQ ID NO: 52

BAQ51233.1 CRISPR相关蛋白，Csn1家族 [酿脓链球菌] SEQ ID NO: 53

KGE60162.1 假设的蛋白质MGAS2111_0903 [酿脓链球菌MGAS2111] SEQ ID NO: 54

KGE60856.1 CRISPR相关内切核酸酶蛋白 [酿脓链球菌SS1447] SEQ ID NO: 55

WP_002989955.1 多物种：II型CRISPR RNA引导的内切核酸酶Cas9 [链球菌] SEQ ID NO: 56

WP_003030002.1 多物种：II型CRISPR RNA引导的内切核酸酶Cas9 [链球菌] SEQ ID NO: 57

WP_003065552.1 多物种：II型CRISPR RNA引导的内切核酸酶Cas9 [链球菌] SEQ ID NO: 58

WP_001040076.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 59

WP_001040078.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 60

WP_001040080.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 61

WP_001040081.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 62

WP_001040083.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 63

WP_001040085.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 64

WP_001040087.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 65

WP_001040088.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 66

WP_001040089.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 67

WP_001040090.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 68

WP_001040091.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 69

WP_001040092.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 70

WP_001040094.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 71

[0398]

WP_001040095.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 72

WP_001040096.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 73

WP_001040097.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 74

WP_001040098.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 75

WP_001040099.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 76

WP_001040100.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 77

WP_001040104.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 78

WP_001040105.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 79

WP_001040106.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 80

WP_001040107.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 81

WP_001040108.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 82

WP_001040109.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 83

WP_001040110.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 84

WP_015058523.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 85

WP_017643650.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 86

WP_017647151.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 87

WP_017648376.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 88

WP_017649527.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 89

WP_017771611.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 90

WP_017771984.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 91

CFQ25032.1 CRISPR相关蛋白 [无乳链球菌] SEQ ID NO: 92

CFV16040.1 CRISPR相关蛋白 [无乳链球菌] SEQ ID NO: 93

KLJ37842.1 CRISPR相关蛋白 Csn1 [无乳链球菌] SEQ ID NO: 94

[0399]

KLJ72361.1 CRISPR相关蛋白 Csn1 [无乳链球菌] SEQ ID NO: 95

KLL20707.1 CRISPR相关蛋白 Csn1 [无乳链球菌] SEQ ID NO: 96

KLL42645.1 CRISPR相关蛋白 Csn1 [无乳链球菌] SEQ ID NO: 97

WP_047207273.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 98

WP_047209694.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 99

WP_050198062.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 100

WP_050201642.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 101

WP_050204027.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 102

WP_050881965.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 103

WP_050886065.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无乳链球菌] SEQ ID NO: 104

AHN30376.1 CRISPR相关蛋白 Csn1 [无乳链球菌 138P] SEQ ID NO: 105

EAO78426.1 网织红细胞结合蛋白 [无乳链球菌 H36B] SEQ ID NO: 106

CCW42055.1 CRISPR相关蛋白, SAG0894家族 [无乳链球菌 ILRI112] SEQ ID NO:107

WP_003041502.1 II型CRISPR RNA引导的内切核酸酶Cas9 [咽峡炎链球菌] SEQ ID NO: 108

WP_037593752.1 II型CRISPR RNA引导的内切核酸酶Cas9 [咽峡炎链球菌] SEQ ID NO: 109

WP_049516684.1 CRISPR相关蛋白 Csn1 [咽峡炎链球菌] SEQ ID NO: 110

GAD46167.1 假设的蛋白质ANG6_0662 [咽峡炎链球菌 T5] SEQ ID NO: 111

WP_018363470.1 II型CRISPR RNA引导的内切核酸酶Cas9 [Streptococcus caballi] SEQ ID NO: 112

WP_003043819.1 II型CRISPR RNA引导的内切核酸酶Cas9 [犬链球菌] SEQ ID NO: 113

WP_006269658.1 II型CRISPR RNA引导的内切核酸酶Cas9 [星座链球菌] SEQ ID NO: 114

WP_048800889.1 II型CRISPR RNA引导的内切核酸酶Cas9 [星座链球菌] SEQ ID NO: 115

WP_012767106.1 II型CRISPR RNA引导的内切核酸酶Cas9 [停乳链球菌] SEQ ID NO: 116

WP_014612333.1 II型CRISPR RNA引导的内切核酸酶Cas9 [停乳链球菌] SEQ ID NO: 117

[0400]

WP_015017095.1 II型CRISPR RNA引导的内切核酸酶Cas9 [停乳链球菌] SEQ ID NO: 118

WP_015057649.1 II型CRISPR RNA引导的内切核酸酶Cas9 [停乳链球菌] SEQ ID NO: 119

WP_048327215.1 II型CRISPR RNA引导的内切核酸酶Cas9 [停乳链球菌] SEQ ID NO: 120

WP_049519324.1 CRISPR相关蛋白 Csn1 [停乳链球菌] SEQ ID NO: 121

WP_012515931.1 II型CRISPR RNA引导的内切核酸酶Cas9 [马链球菌] SEQ ID NO: 122

WP_021320964.1 II型CRISPR RNA引导的内切核酸酶Cas9 [马链球菌] SEQ ID NO: 123

WP_037581760.1 II型CRISPR RNA引导的内切核酸酶Cas9 [马链球菌] SEQ ID NO: 124

WP_004232481.1 II型CRISPR RNA引导的内切核酸酶Cas9 [马肠链球菌] SEQ ID NO: 125

WP_009854540.1 II型CRISPR RNA引导的内切核酸酶Cas9 [解没食子酸链球菌] SEQ ID NO: 126

WP_012962174.1 II型CRISPR RNA引导的内切核酸酶Cas9 [解没食子酸链球菌] SEQ ID NO: 127

WP_039695303.1 II型CRISPR RNA引导的内切核酸酶Cas9 [解没食子酸链球菌] SEQ ID NO: 128

WP_014334983.1 II型CRISPR RNA引导的内切核酸酶Cas9 [婴儿链球菌] SEQ ID NO: 129

WP_003099269.1 II型CRISPR RNA引导的内切核酸酶Cas9 [海豚链球菌] SEQ ID NO: 130

AHY15608.1 CRISPR相关蛋白 Csn1 [海豚链球菌] SEQ ID NO: 131

AHY17476.1 CRISPR相关蛋白 Csn1 [海豚链球菌] SEQ ID NO: 132

ESR09100.1 假设的蛋白质 IUSA1_08595 [海豚链球菌 IUSA1] SEQ ID NO: 133

AGM98575.1 CRISPR相关蛋白 Cas9/Csn1, 亚型 II/NMEMI [海豚链球菌 SF1] SEQ ID NO: 134

ALF27331.1 CRISPR相关蛋白 Csn1 [中间型链球菌] SEQ ID NO: 135

WP_018372492.1 II型CRISPR RNA引导的内切核酸酶Cas9 [马寨链球菌] SEQ ID NO: 136

WP_045618028.1 II型CRISPR RNA引导的内切核酸酶Cas9 [缓症链球菌] SEQ ID NO: 137

WP_045635197.1 II型CRISPR RNA引导的内切核酸酶Cas9 [缓症链球菌] SEQ ID NO: 138

WP_002263549.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 139

WP_002263887.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 140

[0401]

WP_002264920.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 141

WP_002269043.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 142

WP_002269448.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 143

WP_002271977.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 144

WP_002272766.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 145

WP_002273241.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 146

WP_002275430.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 147

WP_002276448.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 148

WP_002277050.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 149

WP_002277364.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 150

WP_002279025.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 151

WP_002279859.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 152

WP_002280230.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 153

WP_002281696.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 154

WP_002282247.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 155

WP_002282906.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 156

WP_002283846.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 157

WP_002287255.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 158

WP_002288990.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 159

WP_002289641.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 160

WP_002290427.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 161

WP_002295753.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 162

WP_002296423.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 163

[0402]

WP_002304487.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 164

WP_002305844.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 165

WP_002307203.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 166

WP_002310390.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 167

WP_002352408.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 168

WP_012997688.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 169

WP_014677909.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 170

WP_019312892.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 171

WP_019313659.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 172

WP_019314093.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 173

WP_019315370.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 174

WP_019803776.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 175

WP_019805234.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 176

WP_024783594.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 177

WP_024784288.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 178

WP_024784666.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 179

WP_024784894.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 180

WP_024786433.1 II型CRISPR RNA引导的内切核酸酶Cas9 [变异链球菌] SEQ ID NO: 181

WP_049473442.1 CRISPR相关蛋白 Csn1 [变异链球菌] SEQ ID NO: 182

WP_049474547.1 CRISPR相关蛋白 Csn1 [变异链球菌] SEQ ID NO: 183

EMC03581.1 假设的蛋白质 SMU69_09359 [变异链球菌 NLML4] SEQ ID NO: 184

WP_000428612.1 II型CRISPR RNA引导的内切核酸酶Cas9 [口腔链球菌] SEQ ID NO: 185

WP_000428613.1 II型CRISPR RNA引导的内切核酸酶Cas9 [口腔链球菌] SEQ ID NO: 186

[0403]

WP_049523028.1 CRISPR相关蛋白 Csn1 [副血链球菌] SEQ ID NO: 187

WP_003107102.1 II型CRISPR RNA引导的内切核酸酶Cas9 [副乳房链球菌] SEQ ID NO: 188

WP_054279288.1 II型CRISPR RNA引导的内切核酸酶Cas9 [海豹链球菌] SEQ ID NO: 189

WP_049531101.1 CRISPR相关蛋白 Csn1 [假肺炎链球菌] SEQ ID NO: 190

WP_049538452.1 CRISPR相关蛋白 Csn1 [假肺炎链球菌] SEQ ID NO: 191

WP_049549711.1 CRISPR相关蛋白 Csn1 [假肺炎链球菌] SEQ ID NO: 192

WP_007896501.1 II型CRISPR RNA引导的内切核酸酶Cas9 [假乖链球菌] SEQ ID NO: 193

EFR44625.1 CRISPR相关蛋白, Csn1家族 [假乖链球菌 SPIN 20026] SEQ ID NO: 194

WP_002897477.1 II型CRISPR RNA引导的内切核酸酶Cas9 [血链球菌] SEQ ID NO: 195

WP_002906454.1 II型CRISPR RNA引导的内切核酸酶Cas9 [血链球菌] SEQ ID NO: 196

WP_009729476.1 II型CRISPR RNA引导的内切核酸酶Cas9 [链球菌 sp. F0441] SEQ ID NO: 197

CQR24647.1 CRISPR相关蛋白 [链球菌 sp. FF10] SEQ ID NO: 198

WP_000066813.1 II型CRISPR RNA引导的内切核酸酶Cas9 [链球菌 sp. M334] SEQ ID NO: 199

WP_009754323.1 II型CRISPR RNA引导的内切核酸酶Cas9 [链球菌 sp. 分类 056] SEQ ID NO: 200

WP_044674937.1 II型CRISPR RNA引导的内切核酸酶Cas9 [猪链球菌] SEQ ID NO: 201

WP_044676715.1 II型CRISPR RNA引导的内切核酸酶Cas9 [猪链球菌] SEQ ID NO: 202

WP_044680361.1 II型CRISPR RNA引导的内切核酸酶Cas9 [猪链球菌] SEQ ID NO: 203

WP_044681799.1 II型CRISPR RNA引导的内切核酸酶Cas9 [猪链球菌] SEQ ID NO: 204

WP_049533112.1 CRISPR相关蛋白 Csn1 [猪链球菌] SEQ ID NO: 205

WP_029090905.1 II型CRISPR RNA引导的内切核酸酶Cas9 [热杀索丝菌] SEQ ID NO: 206

WP_006506696.1 II型CRISPR RNA引导的内切核酸酶Cas9 [Catenibacterium mitsuokai] SEQ ID NO: 207

AIT42264.1 Cas9hc:NLS:HA [克隆载体 pYB196] SEQ ID NO: 208

WP_034440723.1 II型CRISPR内切核酸酶Cas9 [核菌属细菌 S5-A11] SEQ ID NO: 209

[0404]

AKQ21048.1 Cas9 [CRISPR介导的基因靶向载体p(hsp68-Cas9)] SEQ ID NO: 210

WP_004636532.1 II型CRISPR RNA引导的内切核酸酶Cas9 [嗜情疫诈球菌] SEQ ID NO: 211

WP_002364836.1 多物种: II型CRISPR RNA引导的内切核酸酶Cas9 [肠球菌] SEQ ID NO: 212

WP_016631044.1 多物种: II型CRISPR RNA引导的内切核酸酶Cas9 [肠球菌] SEQ ID NO: 213

EMS75795.1 假设的蛋白质H318_06676 [耐久肠球菌 IPLA 655] SEQ ID NO: 214

WP_002373311.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 215

WP_002378009.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 216

WP_002407324.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 217

WP_002413717.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 218

WP_010775580.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 219

WP_010818269.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 220

WP_010824395.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 221

WP_016622645.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 222

WP_033624816.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 223

WP_033625576.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 224

WP_033789179.1 II型CRISPR RNA引导的内切核酸酶Cas9 [粪肠球菌] SEQ ID NO: 225

WP_002310644.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 226

WP_002312694.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 227

WP_002314015.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 228

WP_002320716.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 229

WP_002330729.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 230

WP_002335161.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 231

WP_002345439.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 232

[0405]

WP_034867970.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 233

WP_047937432.1 II型CRISPR RNA引导的内切核酸酶Cas9 [屎肠球菌] SEQ ID NO: 234

WP_010720994.1 II型CRISPR RNA引导的内切核酸酶Cas9 [小肠肠球菌] SEQ ID NO: 235

WP_010737004.1 II型CRISPR RNA引导的内切核酸酶Cas9 [小肠肠球菌] SEQ ID NO: 236

WP_034700478.1 II型CRISPR RNA引导的内切核酸酶Cas9 [小肠肠球菌] SEQ ID NO: 237

WP_007209003.1 II型CRISPR RNA引导的内切核酸酶Cas9 [意大利肠球菌] SEQ ID NO: 238

WP_023519017.1 II型CRISPR RNA引导的内切核酸酶Cas9 [蒙氏肠球菌] SEQ ID NO: 239

WP_010770040.1 II型CRISPR RNA引导的内切核酸酶Cas9 [啄木鸟肠球菌] SEQ ID NO: 240

WP_048604708.1 II型CRISPR RNA引导的内切核酸酶Cas9 [肠球菌 sp. AM1] SEQ ID NO: 241

WP_010750235.1 II型CRISPR RNA引导的内切核酸酶Cas9 [Enterococcus villorum] SEQ ID NO: 242

AII16583.1 Cas9内切核酸酶 [表达载体 pCas9] SEQ ID NO: 243

WP_029073316.1 II型CRISPR RNA引导的内切核酸酶Cas9 [Kandleria vitulina] SEQ ID NO: 244

WP_031589969.1 II型CRISPR RNA引导的内切核酸酶Cas9 [Kandleria vitulina] SEQ ID NO: 245

KDA45870.1 CRISPR相关蛋白 Cas9/Csn1, 亚型 II/NMEMI [动物乳杆菌] SEQ ID NO: 246

WP_039099354.1 II型CRISPR RNA引导的内切核酸酶Cas9 [弯曲乳杆菌] SEQ ID NO: 247

AKP02966.1 假设的蛋白质ABB45_04605 [香肠乳杆菌] SEQ ID NO: 248

WP_010991369.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无害李斯特氏菌] SEQ ID NO: 249

WP_033838504.1 II型CRISPR RNA引导的内切核酸酶Cas9 [无害李斯特氏菌] SEQ ID NO: 250

EHN60060.1 CRISPR相关蛋白, Csn1家族 [无害李斯特氏菌 ATCC 33091] SEQ ID NO: 251

EFR89594.1 crispr相关蛋白, Csn1家族 [无害李斯特氏菌 FSL S4-378] SEQ ID NO: 252

WP_038409211.1 II型CRISPR RNA引导的内切核酸酶Cas9 [伊氏李斯特氏菌] SEQ ID NO: 253

EFR95520.1 crispr相关蛋白 Csn1 [伊氏李斯特氏菌 FSL F6-596] SEQ ID NO: 254

WP_003723650.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特氏菌] SEQ ID NO: 255

[0406]

WP_003727705.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特氏菌] SEQ ID NO: 256

WP_003730785.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 257

WP_003733029.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 258

WP_003739838.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 259

WP_014601172.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 260

WP_023548323.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 261

WP_031665337.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 262

WP_031669209.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 263

WP_033920898.1 II型CRISPR RNA引导的内切核酸酶Cas9 [单核增生李斯特特氏菌] SEQ ID NO: 264

AKI42028.1 CRISPR相关蛋白 [单核增生李斯特特氏菌] SEQ ID NO: 265

AKI50529.1 CRISPR相关蛋白 [单核增生李斯特特氏菌] SEQ ID NO: 266

EFR83390.1 crispr相关蛋白 Csn1 [单核增生李斯特特氏菌 FSL F2-208] SEQ ID NO: 267

WP_046323366.1 II型CRISPR RNA引导的内切核酸酶Cas9 [斯氏李斯特特氏菌] SEQ ID NO: 268

AKE81011.1 Cas9 [植物多重基因组编辑载体 pYLCRISPR/Cas9Pubi-H] SEQ ID NO: 269

CUO82355.1 细菌中保守的未表征的蛋白质 [Roseburia hominis] SEQ ID NO: 270

WP_033162887.1 II型CRISPR RNA引导的内切核酸酶Cas9 [Sharpea azabuensis] SEQ ID NO: 271

AGZ01981.1 Cas9内切核酸酶 [合成构建体] SEQ ID NO: 272

AKA60242.1 核酸酶缺陷Cas9 [合成构建体] SEQ ID NO: 273

AKS40380.1 Cas9 [合成质粒 pFC330] SEQ ID NO: 274

4UN5_B Cas9, B链, 晶体结构 SEQ ID NO: 275

[0407] 提供了合适的胞嘧啶脱氨酶结构域的非限制性实例。

[0408] 人AID

[0409]

MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCH
VELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYF
CEDRKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLS
RQLRRILLPLYEVDDLRDAFRTLGL (SEQ ID NO: 276)

[0410] 小鼠AID

[0411]

MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLDFGHLRNKSGCH
VELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWNPNLSLRIFTARLYF
CEDRKAEPEGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVRLT
RQLRRILLPLYEVDDLRDAFRMLGF (SEQ ID NO: 277)

[0412] 狗AID

[0413]

MDSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGHLRNKSGCH
VELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFAARLYF
CEDRKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGLHENSVRLS
RQLRRILLPLYEVDDLRDAFRTLGL (SEQ ID NO: 278)

[0414] 牛AID

[0415]

MDSLLKKQRQFLYQFKNVRWAKGRHETYLCYVVKRRDSPTSFSLDFGHLRNKAGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFTARLYFC
DKERKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLS
RQLRRILLPLYEVDDLRDAFRTLGL (SEQ ID NO: 279)

[0416] 小鼠APOBEC-3

[0417]

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSL
HHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLAT
HHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRR
FRPWKRLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEGRRMDPL
SEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFL
DKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKG
LCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGL
QDLVNDFGNLQLGPPMS (SEQ ID NO: 280)

[0418] 大鼠APOBEC-3

[0419]

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLRYAIDRKDTFLCYEVTRKDCDSPVSLH
HGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQVLRFLAT
HHNLSLDIFSSRLYNIRDPENQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRR
FRPWKKLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVERRRVHLL
SEEEFYSQFYNQRVKHLCYYHGVKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFL
DKIRSMELSQVIITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGL
CSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESWGL
QDLVNDFGNLQLGPPMS (SEQ ID NO: 281)

[0420] 恒河猴APOBEC-3G

[0421]

MVEPMDPRTFVSNFNNRPILSGLNTVWLCCEVKTKDPSGPPLDAKIFQGKVYSKAKYH
PEMRFLRWFHKWRQLHHDQEYKVTWYVSWSPCTRCANSVATFLAKDPKVTLTIFVAR
LYYFWKPDYQQALRILCQKRGGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLP

[0422]

KHYTLLQATLGELLRHLMDPGTFTSNFNNKPWVSGQHETYLCYKVERLHNDTWVPLN QHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTCFTSWSPCFSCAQEM AKFISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEYCWDTFVDR QGRPFQPWDGLDEHSQALSGRLRAI (SEQ ID NO: 282) (斜体：核酸编辑结构域；下划线：细胞质定位信号)

[0423] 黑猩猩APOBEC-3G

[0424]

MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPLDAKIFRGQV YSKLKYHPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKCTRDVATFLAEDPKVT LTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFE PWNNLPKYYILLHIMLGEILRHSMDPPTFTSNFNNELWVRGRHETYLCYEVERLHNDT WVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLHQDYRVTCFTSWSPCF SCAQEMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLAKAGAKISIMTYSEFKHCWD TFVDHQGCPFQPWDGLEEHSQALSGRLRAILQNQGN (SEQ ID NO: 283)

[0425] 绿猴APOBEC-3G

[0426]

MNPQIRNMVEQMEPDIFVYYFNNRPILSGRNTVWLCYEVKTKDPSGPPLDANIFQGKLY PEAKDHPEMKFLHWFRKWRQLHRDQEYEVTWYVSWSPCTRCANSVATFLAEDPKVT LTIFVARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPF KPRKNLPKHYTLLHATLGELLRHVMDPGTFTSNFNNKPWVSGQRETYLCYKVERSHN DTWVLLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFWKLDDQQYRVTCFTSWSPC FSCAQKMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLHRDGAKIAVMNYSEFEYC WDTFVDRQGRPFQPWDGLDEHSQALSGRLRAI (SEQ ID NO: 284)

[0427] 人APOBEC-3G

[0428]

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQV YSELKYHPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKCTRDMATFLAEDPKVT LTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFE PWNNLPKYYILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDT WVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCF SCAQEMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWD TFVDHQGCPFQPWDGLDEHSQDLSGRLRAILQNQEN (SEQ ID NO: 285)

[0429] 人APOBEC-3F

[0430]

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPRLDAKIFRGQV YSQPEHHAEMCFLSWFCGNQLPAYKCFQITWFVSWTPCPDCVAKLAEFLAEHPNVTLT ISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKF DDNYAFLHRTLKEILRNPMEAMYPHIFYFHFKNLRKAYGRNESWLCFTMEVVKHHSPV SWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSPCPECAGEVAE FLARHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWENFVYND DEPFKPWKGLKYNFLFLDSKLQEILE (SEQ ID NO: 286)

[0431] 人APOBEC-3B

[0432]

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQ
VYFKPQYHAEMCFLSWFCGNQLPAYKCFQITWFVSWTPCPDCVAKLAEFLSEHPNVTL
TISAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWY
KFDENYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEVERLDNGTWVL
MDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWG
CAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWD
TFVYRQGCPFQPWDGLEEHSQALSGRLRAILQNQGN (SEQ ID NO: 287)

[0433] 人APOBEC-3C：

[0434]

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRN
QVDSETHCHAERCFLSWFCDDILSPNTKYQVTWYTSWSPCPDCAGEVAEFLARHSNVN
LTIFTARLYYFQYPCYQEGLRSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFKPWKG
LKTNFRLLKRRLRESLQ (SEQ ID NO: 288)

[0435] 人APOBEC-3A：

[0436]

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQ
AKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQE
NTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPF
QPWDGLDEHSQALSGRLRAILQNQGN (SEQ ID NO: 289)

[0437] 人APOBEC-3H：

[0438]

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICF
INEIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKP
QQKGLRLLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRR
LERIKIPGVRAQGRYMDILCDAEV (SEQ ID NO: 290)

[0439] 人APOBEC-3D

[0440]

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGP
VLPKRQSNHRQEVYFRFENHAEMCFLSWFCGNRLPANRRFQITWFVSWNPCLPCVVKV
TKFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFV
CNEGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKACGRNESWL
CFTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTS
WSPCPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYK
DFVSCWKNFVYSDDEPFKPWKGLQTNFRLLKRRLREILQ (SEQ ID NO: 291)

[0441] 人APOBEC-1

[0442]

MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRSSGKNTT
NHVEVNFIKKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARL
FWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWM
MLYALELHCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR
(SEQ ID NO: 292)

[0443] 小鼠APOBEC-1

[0444]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTS NHVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIARLY HHTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPPSNEAYWPRYPHLWVKLY VLELYCIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK (SEQ ID NO: 293)

[0445] 大鼠APOBEC-1

[0446]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK (SEQ ID NO: 294)

[0447] 海七鳃鳗CDA1(pmCDA1)

[0448]

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQ SGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGH TLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNEN RWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAV (SEQ ID NO: 295)

[0449] 人APOBEC3G D316R_D317R

[0450]

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQV YSELKYHPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKCTRDMATFLAEDPKVT LTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFE PWNNLPKYYILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDT WVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCF SCAQEMAKFISKNKHVSLCIFTARIYRRQGRCQEGLRTLAEAGAKISIMTYSEFKHCWD TFVDHQGCPFQPWDGLDEHSQDLSGRLRAILQNQEN (SEQ ID NO: 296)

[0451] 人APOBEC3G A链

[0452]

MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFL EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTA RIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQD LSGRLRAILQ (SEQ ID NO: 297)

[0453] 人APOBEC3G A链D120R_D121R

[0454]

MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFL EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTA RIYRRQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQD LSGRLRAILQ (SEQ ID NO: 298)

[0455] 提供了非限制性的示例性尿嘧啶糖基化酶抑制剂序列。

[0456] 芽孢杆菌噬菌体PBS2(噬菌体PBS2)尿嘧啶-DNA糖基化酶抑制剂MTNLSDIIEKETGK QLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML(SEQ ID NO:299)

[0457] 塔斯曼尼亚欧文氏菌(Erwinia tasmaniensis)SSB(热稳定性单链DNA结合蛋白)

[0458]

MASRGVNKVILVGNLGQDPEVRYMPNGGAVANITLATSESWRDKQTGETKEKTEWHR
VVLFGKLAEVAGEYLRKGSQVYIEGALQTRKWTDQAGVEKYTTEVVVNVGGTMQML
GGRSQGGGASAGGQNGGSNNGWGQPQQPQGGNQFSGGAQQQARPQQQPQQNNAPA
NNEPPIDFDDDIP (SEQ ID NO: 300)

[0459] UdgX(与DNA中的尿嘧啶结合但不切除)

[0460]

MAGAQDFVPHTADLAELAAAAGECRGCGLYRDATQAVFGAGGRSARIMMIGEQPGD
KEDLAGLPFVGPAGRLLDRALEAADIDRDALYVTNAVKHFKFTRAAGGKRRIHKTPSR
TEVVACRPWLIAEMTSVEPDVVVLLGATAAKALLGNDFRVTQHRGEVLHVDDVPGDP
ALVATVHPSSLLRGPKEERESAFAGLVDDLRVAADVRP (SEQ ID NO: 301)

[0461] UDG(催化无活性的人UDG,与DNA中的尿嘧啶结合但不切除)

[0462]

MIGQKTLYSFFSPSPARKRHAPSPEPAVQGTGVAGVPEESGDAAAIPAKKAPAGQEEPG
TPPSSPLSAEQLDRIQRNKAAALLRLAARNVPVGFGESWKKHLSGEFGKPYFIKLMGFV
AEERKHYTVYPPPHQVFTWTQMCDIKDVKVVILGQEPYHGPNQAHGLCFSVQRPVPPP
PSLENIYKELSTDIEDFVHPGHGDLSGWAKQGVLLLNAVLTVRAHQANSHKERGWEQF

[0463]

TDAVVSWLNQNSNGLVFLLWGSYAQKKGSAIDRKRHHVLQTAHPSPLSVYRGFFGCR
HFSKTNELLQKSGKKPIDWKEL (SEQ ID NO: 302)

[0464] 提供了C至T核碱基编辑器的非限制性实例。

[0465] His_6-rAPOBEC1-XTEN-dCas9用于大肠杆菌表达(SEQ ID NO:303)

[0466]

MGSSHHHHHHMSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHS
IWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPH
VTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWP
RYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK
NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF
LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF
RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN
LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAH
LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD
DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE
NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL
QNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE
VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL
NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK
TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS
KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD
ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSITGLYETRIDLSQLGGDSGGSPKKKRKV

[0467] rAPOBEC1-XTEN-dCas9-NLS用于哺乳动物表达(SEQ ID NO:304)

[0468]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR

[0469]

RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGGSPKKKRKV

[0470] hAPOBEC1-XTEN-dCas9-NLS用于哺乳动物表达(SEQ ID NO:305)

[0471]

MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRSSGKNTT
NHVEVNFIKKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARL
FWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWM
MLYALELHCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR
SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK
NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF
LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF
RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN
LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKP
ENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYY
LQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSE
EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL
NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK
TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS
KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD
ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSITGLYETRIDLSQLGGDSGGSPKKKRKV

[0472] rAPOBEC1-XTEN-dCas9-UGI-NLS(SEQ ID NO:306)

[0473]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH

[0474]

PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTD
ENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0475] rAPOBEC1-XTEN-Cas9切口酶-UGI-NLS(BE3,SEQ ID NO:307)

[0476]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIlLGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIlKDKDFLDNEENEDILEDIVLTITLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIlKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIlEQISEFSKRVILADANLDKVLSAYNKHRDK

[0477]

PIREQAENIlHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGGSTNLSDIlEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTD
ENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0478] pmCDA1-XTEN-dCas9-UGI(细菌)(SEQ ID NO:308)

[0479]

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQ
SGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGH
TLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNEN
RWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAVSGSETPGTSESATPESDKKYSIGLAIG
TNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK
YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQT
YNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNF
KSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE
ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE
EFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYP
FLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF
IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDN
EENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE
EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT
KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD
VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ
KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSMTN
LSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPE
YKPWALVIQDSNGENKIKML

[0480] pmCDA1-XTEN-nCas9-UGI-NLS(哺乳动物构建体)(SEQ ID NO:309)

[0481]

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQ
SGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGH
TLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNEN
RWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAVSGSETPGTSESATPESDKKYSIGLAIG
TNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK
YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQT
YNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNF
KSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE
ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE
EFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYP
FLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF
IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDN
EENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN

[0482]

GIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE
EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT
KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD
VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ
KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNL
SDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEY
KPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0483] huAPOBEC3G-XTEN-dCas9-UGI(细菌)(SEQ ID NO:310)

[0484]

MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFL
EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTA
RIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQD
LSGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVL
GNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI
YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSA
RLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD
DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT
LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVG
PLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH
SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNR
GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQL
VETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN
YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKK
TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK
KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE
FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR
YTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSMTNLSDIIEKETGKQLVIQESILM
LPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKM
L

[0485] huAPOBEC3G-XTEN-nCas9-UGI-NLS(哺乳动物构建体)(SEQ ID NO:311)

[0486]

MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFL
EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTA
RIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQD
LSGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVL

[0487]

GNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI
YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSA
RLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD
DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT
LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVG
PLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH
SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNR
GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQL
VETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN
YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKK
TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK
KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE
FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR
YTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLP
EEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLS
GGSPKKKRKV

[0488] huAPOBEC3G(D316R_D317R)-XTEN-nCas9-UGI-NLS(哺乳动物构建体)(SEQ ID NO：312)

[0489]

MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFL
EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTA
RIYRRQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQD
LSGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVL
GNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI
YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSA
RLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD
DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT
LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVG
PLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH
SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNR
GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQL
VETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN
YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKK
TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK

[0490]

KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE
FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR
YTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLP
EEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLS
GGSPKKKRKV

[0491] 高保真核碱基编辑器(SEQ ID NO:313)

[0492]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTAFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGALSRKLINGIRDKQSGKTILDFLKSDGFANRNFMALIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRAITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGD

[0493] rAPOBEC1-XTEN-SaCas9n-UGI-NLS)

[0494]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLK
RRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRR
GVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSD
YVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEML
MGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQK
KKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIA
KILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDN
QIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIE
LAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLY

[0495]

SLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDS
KISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLM
NLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFK
EWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRV
DKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHD
PQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAH
LDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYE
EAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENM
NDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGSGGSTNLSDIIEKETGK
QLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQD
SNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 399)

[0496] rAPOBEC1-XTEN-SaCas9n-UGI-NLS

[0497]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLK
RRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRR
GVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSD
YVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEML
MGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQK
KKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIA
KILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDN
QIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIE
LAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLY
SLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDS
KISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLM
NLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFK
EWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRV
DKKPNR*K*LINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHD
PQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAH
LDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYE
EAKKLKKISNQAEFIASFY*K*NDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENM
NDKRPP*H*IIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGSGGSTNLSDIIEKETGK
QLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQD
SNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 400)

[0498] 核碱基编辑器4-SSB

[0499]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW

[0500]

NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGGSGGSGGSASRGVNKVILVGNLGQDPEVRYMPNGGAVANITLATSESWR
DKATGEMKEQTEWHRVVLFGKLAEVASEYLRKGSQVYIEGQLRTRKWTDQSGQDRY
TTEVVVNVGGTMQMLGGRQGGGAPAGGNIGGGQPQGGWGQPQQPQGGNQFSGGAQ
SRPQQSAPAAPSNEPPMDFDDDIPFSGGSPKKKRKV (SEQ ID NO: 401)

[0501] 核碱基编辑器4-(GGS)$_3$

[0502]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTA
YDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO:
402)

[0503] 核碱基编辑器4-XTEN

[0504]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGSETPGTSESATPESTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDIL
VHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ
ID NO: 403)

[0505] 核碱基编辑器4-32aa接头

[0506]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGGSSGGSSGSE
TPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLL
AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV
RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL
RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT
YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL
IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK

[0507]

PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY
YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS
EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH
VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF
KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGF
SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL
LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG
NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIG
NKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKK
RKV (SEQ ID NO: 404)

[0508] 核碱基编辑器4-2XUGI

[0509]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV
LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESA
TPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM
RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR
RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE
SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTD
ENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSTNLSDIIEKETGKQLVIQESILML
PEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML
SGGSPKKKRKV (SEQ ID NO: 405)

[0510] 核碱基编辑器4

[0511]

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH
HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYV

[0512]

LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGGSSGGSSGSE
TPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLL
AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV
RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL
RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT
YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL
IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK
PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY
YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS
EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH
VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF
KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGF
SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL
LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG
NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSITGLYETRIDLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEE
VEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGG
SGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENV
MLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 406)

[0513] 提供了接受DNA作为底物的演化的腺苷脱氨酶的非限制性实例。

[0514] ecTadA

[0515]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG
SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID
NO: 314)

[0516] ecTadA(D108N)

[0517]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGAAG
SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID
NO: 315)

[0518] ecTadA(D108G)

[0519]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 316)

[0520] ecTadA(D108V)

[0521]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 317)

[0522] ecTadA(H8Y,D108N,N127S)

[0523]

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGAAG SLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 318)

[0524] ecTadA(H8Y,D108N,N127S,E155D)

[0525]

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGAAG SLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQDIKAQKKAQSSTD (SEQ ID NO: 319)

[0526] ecTadA(H8Y,D108N,N127S,E155G)

[0527]

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGAAG SLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQGIKAQKKAQSSTD (SEQ ID NO: 320)

[0528] ecTadA(H8Y,D108N,N127S,E155V)

[0529]

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGAAG SLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQVIKAQKKAQSSTD (SEQ ID NO: 321)

[0530] ecTadA(A106V,D108N,D147Y和E155V)

[0531]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD (SEQ ID NO:322)

[0532] ecTadA(A106V,D108N,D147Y和E155V)

[0533]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD (SEQ ID NO: 407)

[0534] ecTadA(L84F,A106V,D108N,H123Y,D147Y,E155V,I156F)

[0535]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 408)

[0536] ecTadA(S2A,I49F,A106V,D108N,D147Y,E155V)

[0537]

AEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPFGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAA

[0538]

GSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD (SEQ ID NO: 409)

[0539] ecTadA(H8Y,A106T,D108N,N127S,K160S)

[0540]

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGTRNAKTGAAGS LMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQSKAQSSTD (SEQ ID NO: 410)

[0541] ecTadA(R26G,L84F,A106V,R107H,D108N,H123Y,A142N,A143D,D147Y,E155V, 1156F)

[0542]

SEVEFSHEYWMRHALTLAKRAWDEGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 411)

[0543] ecTadA(E25G,R26G,L84F,A106V,R107H,D108N,H123Y,A142N,A143D,D147Y, E155V,I156F)

[0544]

SEVEFSHEYWMRHALTLAKRAWDGGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTGAA GSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 412)

[0545] ecTadA(E25D,R26G,L84F,A106V,R107K,D108N,H123Y,A142N,A143G,D147Y, E155V,I156F)

[0546]

SEVEFSHEYWMRHALTLAKRAWDDGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVKNAKTGAA GSLMDVLHYPGMNHRVEITEGILADECNGLLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 413)

[0547] ecTadA(R26Q,L84F,A106V,D108N,H123Y,A142N,D147Y,E155V,I156F)

[0548]

SEVEFSHEYWMRHALTLAKRAWDEQEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 414)

[0549] ecTadA(E25M,R26G,L84F,A106V,R107P,D108N,H123Y,A142N,A143D,D147Y, E155V,I156F)

[0550]

SEVEFSHEYWMRHALTLAKRAWDMGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVPNAKTGAA GSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 415)

[0551] ecTadA(R26C,L84F,A106V,R107H,D108N,H123Y,A142N,D147Y,E155V,I156F)

[0552]

SEVEFSHEYWMRHALTLAKRAWDECEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 416)

[0553] ecTadA(L84F,A106V,D108N,H123Y,A142N,A143L,D147Y,E155V,I156F)

[0554]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECNLLLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 417)

[0555] ecTadA(R26G,L84F,A106V,D108N,H123Y,A142N,D147Y,E155V,I156F)

[0556]

SEVEFSHEYWMRHALTLAKRAWDEGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 418)

[0557] ecTadA(R51H,L84F,A106V,D108N,H123Y,D147Y,E155V,I156F,K157N)

[0558]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGHHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD (SEQ ID NO: 419)

[0559] ecTadA(E25A,R26G,L84F,A106V,R107N,D108N,H123Y,A142N,A143E,D147Y, E155V,I156F)

[0560]

SEVEFSHEYWMRHALTLAKRAWDAGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVNNAKTGAA GSLMDVLHYPGMNHRVEITEGILADECNELLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 420)

[0561] ecTadA(L84F,A106V,D108N,H123Y,D147Y,E155V,I156F)

[0562]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 421)

[0563] ecTadA(N37T,P48T,L84F,A106V,D108N,H123Y,D147Y,E155V,I156F)

[0564]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHTNRVIGEGWNRTIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 422)

[0565] ecTadA(N37S,L84F,A106V,D108N,H123Y,D147Y,E155V,I156F)

[0566]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHSNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 423)

[0567] ecTadA(H36L,L84F,A106V,D108N,H123Y,D147Y,E155V,1156F)

[0568]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 424)

[0569] ecTadA(L84F,A106V,D108N,H123Y,S146R,D147Y,E155V,I156F)

[0570]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLRYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 425)

[0571] ecTadA(H36L,P48L,L84F,A106V,D108N,H123Y,D147Y,E155V,I156F)

[0572]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRLIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 426)

[0573] ecTadA(H36L,L84F,A106V,D108N,H123Y,D147Y,E155V,K57N,I156F)

[0574]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGRHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 427)

[0575] ecTadA(H36L,L84F,A106V,D108N,H123Y,S146C,D147Y,E155V,I156F)

[0576]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGRHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 428)

[0577] ecTadA(L84F,A106V,D108N,H123Y,S146R,D147Y,E155V,I156F)

[0578]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLRYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 429)

[0579] ecTadA(N37S,R51H,L84F,A106V,D108N,H123Y,D147Y,E155V,I156F

[0580]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHSNRVIGEGWNRPIGHHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 430)

[0581] ecTadA(R51L,L84F,A106V,D108N,H123Y,D147Y,E155V,I156F,K157N

[0582]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 431)

[0583] saTadA(wt)

[0584]

MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCSGSL
MNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 432)

[0585] saTadA(D108N)

[0586]

GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEHI
AIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADNPKGGCSGSLMN
LLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 433)

[0587] saTadA(D107A_D108N)

[0588]

GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEHI
AIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGSLMN
LLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 434)

[0589] saTadA(G26P_D107A_D108N)

[0590]

GSHMTNDIYFMTLAIEEAKKAAQLPEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEHI AIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGSLMN LLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 435)

[0591] saTadA(G26P_D107A_D108N_S142A)

[0592]

GSHMTNDIYFMTLAIEEAKKAAQLPEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEHI AIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGSLMN LLQQSNFNHRAIVDKGVLKEACATLLTTFFKNLRANKKSTN (SEQ ID NO: 436)

[0593] saTadA(D107A_D108N_S142A)

[0594]

GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEHI AIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGSLMN LLQQSNFNHRAIVDKGVLKEACATLLTTFFKNLRANKKSTN (SEQ ID NO: 437)

[0595] ecTadA(P48S)

[0596]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRSIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 438)

[0597] ecTadA(P48T)

[0598]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRTIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 439)

[0599] ecTadA(P48A)

[0600]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRAIGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAA GSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 440)

[0601] ecTadA(A142N)

[0602]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECNALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 441)

[0603] ecTadA(W23R)

[0604]

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 442)

[0605] ecTadA(W23L)

[0606]

SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 443)

[0607] ecTadA(R152P)

[0608]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMPRQEIKAQKKAQSSTD (SEQ ID NO: 444)

[0609] ecTadA(R152H)

[0610]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMHRQEIKAQKKAQSSTD (SEQ ID NO: 445)

[0611] ecTadA(L84F,A106V,D108N,H123Y,D147Y,E155V,I156F)

[0612]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD (SEQ ID NO: 446)

[0613] ecTadA(H36L,R51L,L84F,A106V,D108N,H123Y,S146C,D147Y,E155V,I156F, K157N)

[0614]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGLHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD (SEQ ID NO: 447)

[0615] ecTadA(H36L,P48S,R51L,L84F,A106V,D108N,H123Y,S146C,D147Y,E155V,I156F, K157N)

[0616]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRSIGLHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD (SEQ ID NO: 448)

[0617] ecTadA(H36L,P48A,R51L,L84F,A106V,D108N,H123Y,S146C,D147Y,E155V,I156F,

K157N)

[0618]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD (SEQ ID NO: 449)

[0619] ecTadA(W23L,H36L,P48A,R51L,L84F,A106V,D108N,H123Y,S146C,D147Y,R152P, E155V,I156F,K157N)

[0620]

SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD (SEQ ID NO: 450)

[0621] ecTadA(W23R,H36L,P48A,R51L,L84F,A106V,D108N,H123Y,S146C,D147Y,R152P, E155V,I156F,K157N)

[0622]

SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD (SEQ ID NO: 479)

[0623] 金黄色葡萄球菌TadA：

[0624]

MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCSGSL MNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 451)

[0625] 枯草芽孢杆菌TadA：

[0626]

MTQDELYMKEAIKEAKKAEEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEMLVI D
EACKALGTWRLEGATLYVTLEPCPMCAGAVVLSRVEKVVFGAFDPKGGCSGTLMNLL QEERFNHQAEVVSGVLEEECGGMLSAFFRELRKKKKAARKNLSE (SEQ ID NO: 452)

[0627] 鼠伤寒沙门氏菌(Salmonella typhimurium)(S.typhimurium)TadA：

[0628]

MPPAFITGVTSLSDVELDHEYWMRHALTLAKRAWDEREVPVGAVLVHNHRVIGEGWN RPIGRHDPTAHAEIMALRQGGLVLQNYRLLDTTLYVTLEPCVMCAGAMVHSRIGRVVF GARDAKTGAAGSLIDVLHHPGMNHRVEIIEGVLRDECATLLSDFFRMRRQEIKALKKA DRAEGAGPAV (SEQ ID NO: 453)

[0629] 腐败希瓦氏菌(Shewanella putrefaciens)(S.putrefaciens)TadA：

[0630]

MDEYWMQVAMQMAEKAEAAGEVPVGAVLVKDGQQIATGYNLSISQHDPTAHAEILC LRSAGKKLENYRLLDATLYITLEPCAMCAGAMVHSRIARVVYGARDEKTGAAGTVVN LLQHPAFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE (SEQ ID NO: 454)

[0631] 流感嗜血杆菌(Haemophilus influenzae)F3031(H.influenzae)TadA：

[0632]

MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQSDP TAHAEIIALRNGAKNIQNYRLLNSTLYVTLEPCTMCAGAILHSRIKRLVFGASDYKTGAI GSRFHFFDDYKMNHTLEITSGVLAEECSQKLSTFFQKRREEKKIEKALLKSLSDK (SEQ ID NO: 455)

[0633] 新月柄杆菌(Caulobacter crescentus)(C.crescentus)TadA:

[0634]

MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAHDP TAHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCAMCAGAISHARIGRVVFGADDPKG GAVVHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI (SEQ ID NO: 456)

[0635] 硫还原地杆菌(Geobacter sulfurreducens)(G.sulfurreducens)TadA:

[0636]

MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGHNLREGSNDP SAHAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAIILARLERVVFGCYDPKGG AAGSLYDLSADPRLNHQVRLSPGVCQEECGTMLSDFFRDLRRRKKAKATPALFIDERK VPPEP (SEQ ID NO: 457)

[0637] 提供了A至G核碱基编辑器的非限制性实例。

[0638] ecTadA(wt)-XTEN-nCas9-NLS(SEQ ID NO:323)

[0639]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDSGGSPKKKRKV

[0640] ecTadA(D108N)-XTEN-nCas9-NLS:(哺乳动物构建体,在DNA上有活性,A至G编辑,SEQ ID NO:324)

[0641]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT

[0642]

ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSPKKKRKV

[0643] ecTadA(D108G)-XTEN-nCas9-NLS:(哺乳动物构建体,在DNA上有活性,A至G编辑,SEQ ID NO:325)

[0644]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSPKKKRKV

[0645] ecTadA(D108V)-XTEN-nCas9-NLS:(哺乳动物构建体,在DNA上有活性,A至G编辑,SEQ ID NO:326)

[0646]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK

[0647]

SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSPKKKRKV

[0648] ecTadA(D108N)-XTEN-nCas9-UGI-NLS(A至G编辑器的BE3类似物,SEQ ID NO:327)

[0649]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0650] ecTadA(D108G)-XTEN-nCas9-UGI-NLS(A至G编辑器的BE3类似物,SEQ ID NO:328):

[0651]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET

[0652]

PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0653] ecTadA(D108V)-XTEN-nCas9-UGI-NLS(A至G编辑器的BE3类似物,SEQ ID NO:329):

[0654]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK

[0655]

YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0656] ecTadA(D108N)-XTEN-dCas9-UGI-NLS(哺乳动物细胞,A至G编辑器的BE2类似物,SEQ ID NO:330):

[0657]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0658] ecTadA(D108G)-XTEN-dCas9-UGI-NLS(哺乳动物细胞,A至G编辑器的BE2类似物,SEQ ID NO:331):

[0659]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR

[0660]

FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0661] ecTadA(D108V)-XTEN-dCas9-UGI-NLS(哺乳动物细胞,A至G编辑器的BE2类似物,SEQ ID NO:332):

[0662]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV

[0663] ecTadA(D108N)-XTEN-nCas9-AAG(E125Q)-NLS-cat.烷基腺苷糖基化酶(SEQ ID

NO:333)

[0664]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA

[0665]

AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELR
GRIVETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDG
ACVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRD
LAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRV
AEQDTQASGGSPKKKRKV

[0666] ecTadA(D108G)-XTEN-nCas9-AAG(E125Q)-NLS-cat.烷基腺苷糖基化酶(SEQ ID NO:334)

[0667]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR

[0668]

MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELR
GRIVETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDG
ACVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRD
LAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRV
AEQDTQASGGSPKKKRKV

[0669] ecTadA(D108V)-XTEN-nCas9-AAG(E125Q)-NLS-cat.烷基腺苷糖基化酶(SEQ ID NO:335)

[0670]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELR
GRIVETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDG
ACVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRD
LAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRV
AEQDTQASGGSPKKKRKV

[0671] ecTadA(D108N)-XTEN-nCas9-EndoV(D35A)-NLS:含有cat.内切核酸酶V(SEQ ID NO:336)

[0672]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA

[0673]

AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSDLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQG
GEVTRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDL
VFVDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQ
LAWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA
FVRYTANQPSGGSPKKKRKV

[0674] ecTadA(D108G)-XTEN-nCas9-EndoV(D35A)-NLS:含有cat.内切核酸酶V(SEQ ID NO:337)

[0675]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR

[0676]

MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSDLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQG
GEVTRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDL
VFVDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQ
LAWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA
FVRYTANQPSGGSPKKKRKV

[0677] ecTadA(D108V)-XTEN-nCas9-EndoV(D35A)-NLS:含有cat.内切核酸酶V(SEQ ID NO:338)

[0678]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDSGGSDLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQG
GEVTRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDL
VFVDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQ
LAWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA
FVRYTANQPSGGSPKKKRKV

[0679] 产生自第一轮演化的变体(细菌中)ecTadA(H8Y_D108N_N127S)-XTEN-dCas9(SEQ ID NO:339)

[0680]

MSEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA

[0681]

AGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGD

[0682] 来自第二轮演化的丰富变体(细菌中)ecTadA(H8Y_D108N_N127S_E155X)-XTEN-dCas9;X＝D,G或V(SEQ ID NO:340)

[0683]

MSEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA
AGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQXIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS

[0684]

SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGD

[0685] pNMG-160:ecTadA(D108N)-XTEN-nCas9-GGS-AAG*(E125Q)-GGS-NLS(SEQ ID NO: 341)

[0686]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDGGSKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRG
RIVETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGA
CVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRDL
AQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRVA
EQDTQAGGSPKKKRKV

[0687] pNMG-161:ecTadA(D108N)-XTEN-nCas9-GGS-EndoV*(D35A)-GGS-NLS(SEQ ID NO:342)

[0688]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN

[0689]

GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM
ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS
SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK
YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDGGSDLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQGG
EVTRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVF
VDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLA
WVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFV
RYTANQPGGSPKKKRKV

[0690] pNMG-371：

[0691] ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)-SGGS-SGGS-XTEN-SGGS-SGGS-

[0692] ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)-SGGS-SGGS-XTEN-SGGS-SGGS-nCas9-SGGS-NLS(SEQ ID NO:458)

[0693]

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTDSGGSSG
GSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVL
VHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFF
RMRRQVFKAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAI
GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR
RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE
KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV
QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLT
PNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA
SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED
FYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA
QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK
AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV

[0694]

YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0695] pNMG-616氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:459)

[0696]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0697] pNMG-624氨基酸序列:ecTadA(野生型)-32a.a.接头-ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-24a.a.接头_nCas9_SGGS_NLS(SEQ ID NO:460)

[0698]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS

[0699]

SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDKKYSIGLAIGTNSVG
WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY
HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQL
FEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK
ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSPKKKR
KV

[0700] pNMG-476氨基酸序列(演化#3异源二聚体,wt TadA+TadA evo#3突变):ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:461)

[0701]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLSY
FFRMRRQVFKAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI

[0702]

ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0703] pNMG-477氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:462)

[0704]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYF
FRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLA
IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR
RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV
QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLT
PNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA
SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED
FYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA
QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK
AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0705] pNMG-558氨基酸序列:ecTadA(野生型)-32a.a.接头-ecTadA(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)-24a.a.接头_nCas9_SGGS_NLS(SEQ ID NO:463)

[0706]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYF
FRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDKKYSIGLAIGTNSVG
WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY
HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQL
FEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK
ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSPKKKR
KV

[0707] pNMG-576氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:464)

[0708]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYF
FRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLA
IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR
RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV
QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLT

[0709]

PNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA
SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED
FYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA
QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK
AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0710] pNMG-577氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_1156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:465)

[0711]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCYF
FRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLA
IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR
RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV
QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLT
PNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA
SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED
FYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA
QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK
AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG

[0712]

FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0713] pNMG-586氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:466)

[0714]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0715] pNMG-588氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:467)

[0716]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS

[0717]

SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCY
FFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0718] pNMG-620氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:468)

[0719]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR

[0720]

KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0721] pNMG-617氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:469)

[0722]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCY
FFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0723] pNMG-618氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:470)

[0724]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0725] pNMG-620氨基酸序列:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS(SEQ ID NO:471)

[0726]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY

[0727]

HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV

[0728] pNMG-621氨基酸序列:ecTadA(野生型)-32a.a.接头-ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-24a.a.接头_nCas9_GGS_NLS(SEQ ID NO:472)

[0729]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDKKYSIGLAIGTNSVG
WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY
HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQL
FEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK
ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK

[0730]

MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSPKKKR
KV

[0731] pNMG-622氨基酸序列:ecTadA(野生型)-32a.a.接头-ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P_E155V_I156F_K157N)-24a.a.接头_nCas9_GGS_NLS(SEQ ID NO:473)

[0732]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDKKYSIGLAIGTNSVG
WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY
HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQL
FEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK
ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSPKKKR
KV

[0733] pNMG-623氨基酸序列:ecTadA(野生型)-32a.a.接头-ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-24a.a.接头_nCas9_GGS_NLS(SEQ ID NO:474)

[0734]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA

[0735]

AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDKKYSIGLAIGTNSVG
WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY
HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQL
FEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK
ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSPKKKR
KV

[0736] ABE6.3ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:475)

[0737]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPVGA
VLVLNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYF
FRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLA
IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR
RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV
QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLT
PNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA
SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED
FYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA
QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK
AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF

[0738]

LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV*

[0739] ABE7.8ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:476)

[0740]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCY
FFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV*

[0741] ABE7.9ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P¬_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:477)

[0742]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV*

[0743] ABE7.10ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P¬_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

[0744]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGS
SGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY
FFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY

[0745]

HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSP
KKKRKV*

[0746] ABE6.4:ecTadA(野生型)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS(SEQ ID NO:480)

[0747]

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHD
PTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSG
GSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDEREVPV
GAVLVLNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECNALL
CYFFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSI
GLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEV
AYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLF
IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL
GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR
RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG
EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK
DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR
ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY
DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ
RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYG

[0748]

DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGE
IVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYK
EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG
SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAE
NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
SGGSPKKKRKV

[0749] 实施例2:分裂的核碱基编辑器的AAV递送

[0750] 设计该研究以显示核碱基编辑器可以通过重组AAV(rAAV)以两个部分递送,其可以经由蛋白质剪接在细胞中接合以形成完整且有活性的核碱基编辑器。测试rAAV构建体的不同元件用于优化的核碱基编辑器表达和活性。

[0751] 重组AAV(rAAV)广泛用于转基因递送。将转基因插入反向末端重复(ITR)序列之间的AAV基因组中并包装到AAV病毒颗粒中,其用于转导宿主细胞(例如哺乳动物细胞,人细胞)。然而,对可以包装到rAAV中的转基因的大小存在限制,通常约为4.9千碱基。编码核碱基编辑器(例如,胞嘧啶脱氨酶-dCas9-UGI)的核酸通常超过rAAV的包装限制。如本文所述,编码核碱基编辑器的核酸得以分裂(参见图1A),并将每个部分包装到分开的rAAV颗粒中。将核碱基编辑器的两个部分递送至细胞,并且可以经由蛋白质剪接(例如,由内含肽如DnaE内含肽介导;参见图1C)连接以形成完整的核碱基编辑器。经连接的完整核碱基编辑器在编辑靶碱基中有活性(见图1B)。编码分裂的核碱基编辑器的rAAV构建体在不同的细胞系(例如U118和HEK293T)中进行测试,并且在编辑靶碱基中有活性(参见图3A-3B和图5A-5B)。

[0752] 在rAAV构建体中测试了不同的转录终止子和核定位信号(NLS)以优化核碱基编辑器的表达和活性(参见图4、6和7)。

[0753] 实施例3:使用AAV编码的分裂的核碱基编辑器编辑小鼠神经元中的DNMT1基因

[0754] 设计该研究以测试体内AAV编码的分裂的核碱基编辑器的碱基编辑活性。使用如图1A中所示的分裂的核碱基编辑器。dCas9结构域和脱氨酶结构域之间的接头的氨基酸序列是SGGSSGGSSGSETPGTSESATPESSGGSSGGS(SEQ ID NO:384)。选择靶向DNMT1基因中充分表征的位点的指导RNA。预计细胞能够耐受编辑。这些实验旨在确定AAV编码的分裂的碱基编辑器是否可以在包括初级神经元在内的几种细胞类型中体外或体内编辑基因座。

[0755] 在一个实验中,在分离和培养皮质神经元两天后,使用编码分裂的核碱基编辑器的AAV载体和靶向DNMT1的指导RNA转导解离的小鼠皮质神经元。转导后16天收获神经元并对DNMT1基因进行测序(图8A)以确定编辑效率以及脱靶效应。检测到17.34%的编辑效率(C至T编辑,图8B中的深灰色),而仅检测到0.82%的不期望的编辑(C至G或C至A变化,图8B中的浅灰色)。

[0756] 在另一个实验中,培养的小鼠Neuro-2细胞用编码分裂的核碱基编辑器的AAV载体和靶向DNMT1的指导RNA转导,或者用脂质包裹的编码核碱基编辑器的DNA和指导RNA转染,从而允许直接比较使用核碱基编辑器的不同递送方法的编辑效率(图9A)。对于AAV编码的分裂核碱基编辑器,观察到5.96%的编辑效率(C至T编辑,图9B中的深灰色),而对于脂质转染的DNA编码的核碱基编辑器,观察到27.3%的编辑效率(C至T编辑,图9B中的深灰色)。对于AAV编码的分裂的核碱基编辑器,不期望的产物的量为0.15%,而对于脂质转染的DNA编码的核碱基编辑器,不期望的产物的量为1.3%(C至G或C至A变化,图9B中的浅灰色)。

[0757] 等同实施方案和范围

[0758] 在权利要求中,诸如“一种”、“一个”和“该”的冠词可以意指一个或超出一个,除非相反地指出或者从上下文中以其他方式显而易见。若一个、超出一个或所有组成员在给定产物或过程中存在、采用或以其他方式相关,则认为在组中的一个或多个成员之间包括“或”的权利要求或说明书是满足的,除非另有说明或从上下文中以其他方式显而易见。本发明包括实施方案,其中组的恰好一个成员在给定产物或过程中存在、采用或以其他方式相关。本发明包括实施方案,其中超出一个或所有组成员在给定产物或过程中存在、采用或以其他方式相关。

[0759] 此外,本发明涵盖所有变型、组合和置换,其中来自一个或多个列出的权利要求的一个或多个限制、元素、条款、描述性术语被引入到另一个权利要求中。例如,可以修改依赖于另一个权利要求的任何权利要求以包括在依赖于相同基本权利要求的任何其他权利要求中找到的一个或多个限制。在将元素呈现为列表(例如,以马库什群组格式)的情况下,还公开了元素的每个子群,并且可以从群组中移除任何元素。应当理解,通常,在本发明或本发明的方面称为包含特定元素和/或特征的情况下,本发明或本发明的方面的某些实施方案由此类元素和/或特征组成,或基本上由之组成。出于简化的目的,那些实施方案未在本文中具体阐述。

[0760] 还应注意,术语“包含”和“含有”旨在是开放的并且允许包括另外的元素或步骤。在给出范围的情况下,端点包括在内。此外,除非另有说明或从上下文和本领域普通技术人员的理解中以其他方式明显看出,否则表达为范围的值可以假定在本发明的不同实施方案

中的所述范围内的任何特定值，至该范围的下限的单位的十分之一，除非上下文另有明确规定。

[0761] 本申请涉及各种已授权的专利、公布的专利申请，期刊文章和其他出版物，其所有都通过引用并入本文。如果任何并入的参考文献与本说明书之间存在冲突，则以本说明书为准。另外，落入现有技术内的本发明的任何特定实施方案可以明确地从任何一个或多个权利要求中排除。因为此类实施方案认为是本领域普通技术人员已知的，所以即使本文未明确阐述排除，也可以排除它们。无论是否与现有技术的存在相关，本发明的任何特定实施方案可以出于任何原因排除在任何权利要求之外。

[0762] 本领域技术人员将认识到或能够仅仅使用常规的实验确定本文所述的具体实施方案的许多等同实施方案。本文所述的本实施方案的范围不意图限于以上说明书，而是如所附权利要求中所述。本领域普通技术人员将理解，在不脱离如以下权利要求所限定的本发明的精神或范围的情况下，可以对本说明书进行各种改变和修改。

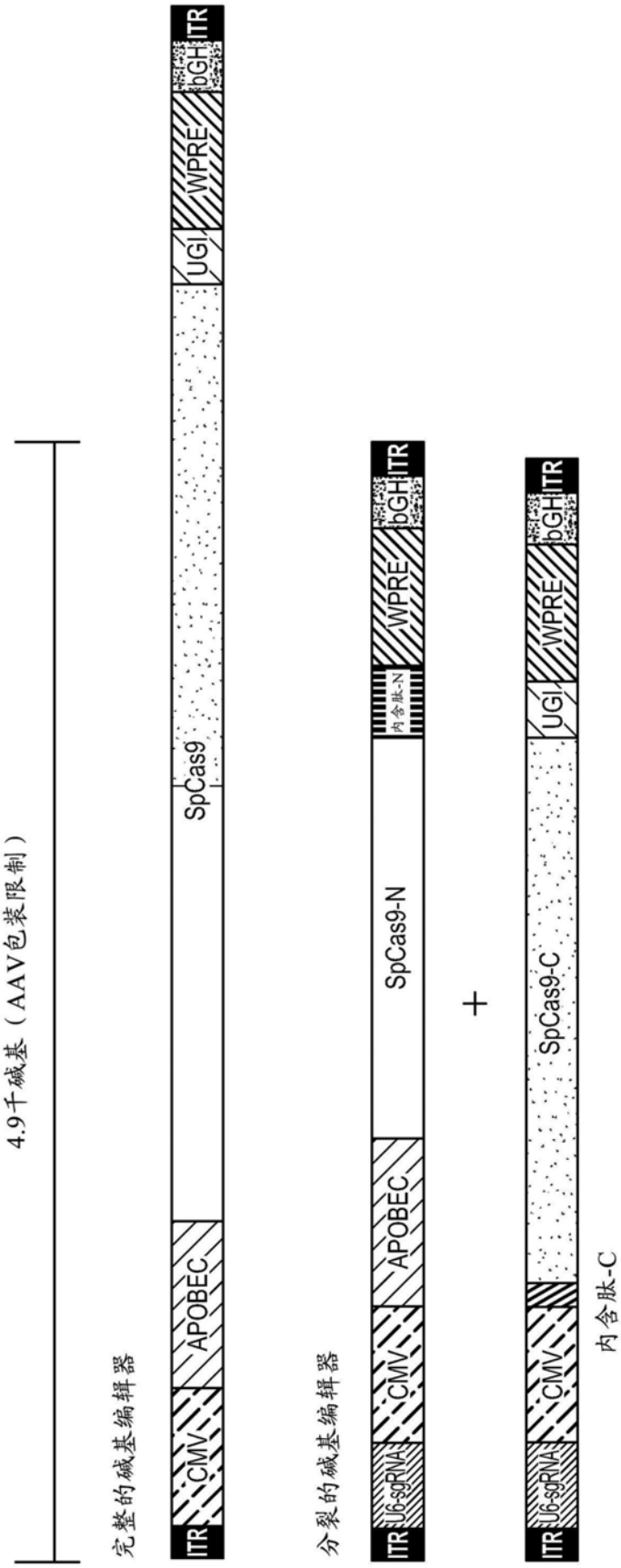

4.9千碱基（AAV包装限制）
完整的碱基编辑器
ITR
CMV
APOBEC
SpCas9
UGI
WPRE
bGH
ITR
分裂的碱基编辑器
ITR
U6-sgRNA
CMV
APOBEC
SpCas9-N
内含肽-N
WPRE
bGH
ITR
+
ITR
U6-sgRNA
CMV
SpCas9-C
UGI
WPRE
bGH
ITR
内含肽-C

图1A

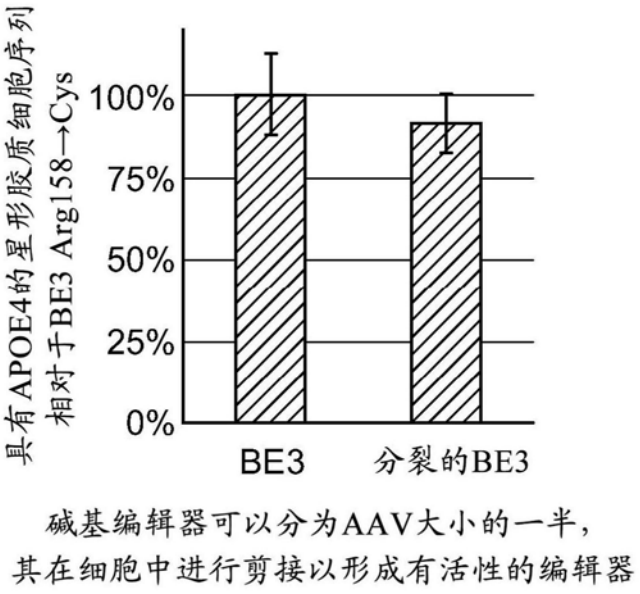

具有APOE4的星形胶质细胞序列
相对于BE3 Arg158→Cys
100%
75%
50%
25%
0%
BE3
分裂的BE3
碱基编辑器可以分为AAV大小的一半，
其在细胞中进行剪接以形成有活性的编辑器

图1B

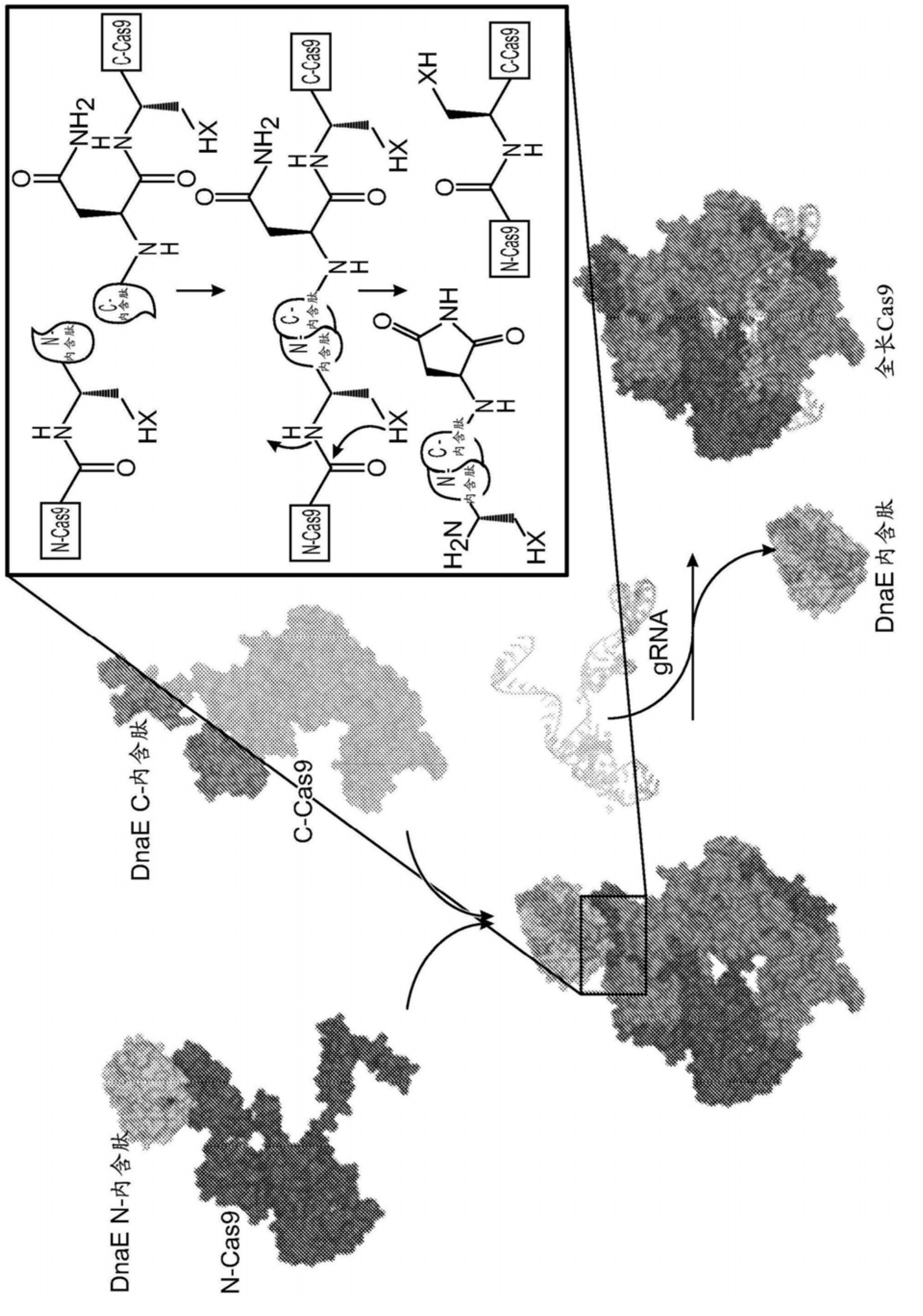

DnaE N-内含肽
N-Cas9
DnaE C-内含肽
C-Cas9
gRNA
DnaE 内含肽
全长Cas9
N-Cas9
C-Cas9
N-内含肽
C-内含肽
HX
XH
NH2
H2N

图1C

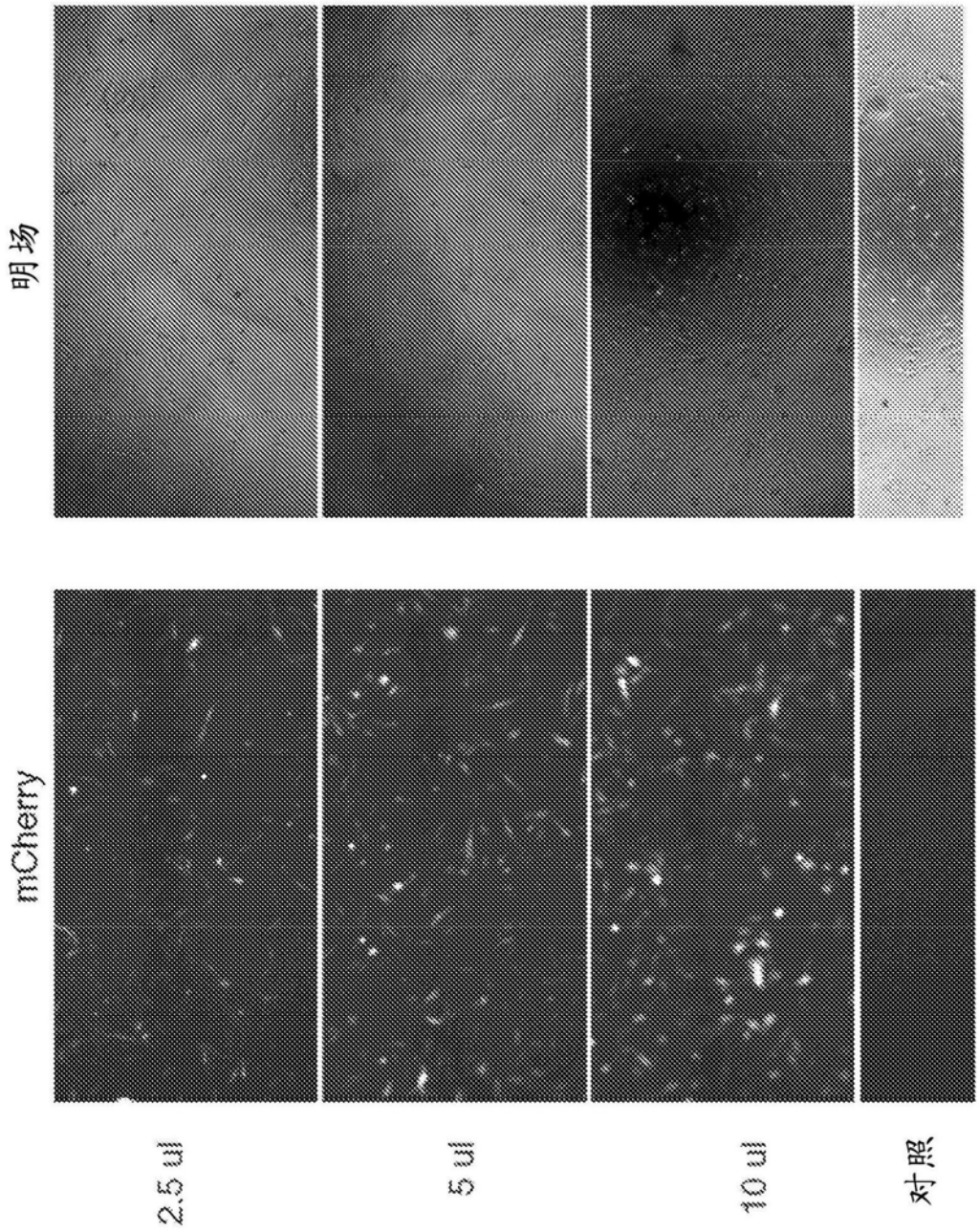


图2

病毒收获在
24小时
72小时:
C1
N1
U118
HEK
C7 C8
对照-用BE3和PRNP指导进行脂质转染（U118顶行；HEK底行）
Gly Ser -> Ser Arg -> stop Tyr Pro Gly Glu
G G C4 A G C7 C8 G A T A C C C G G G C A G G G

图3A

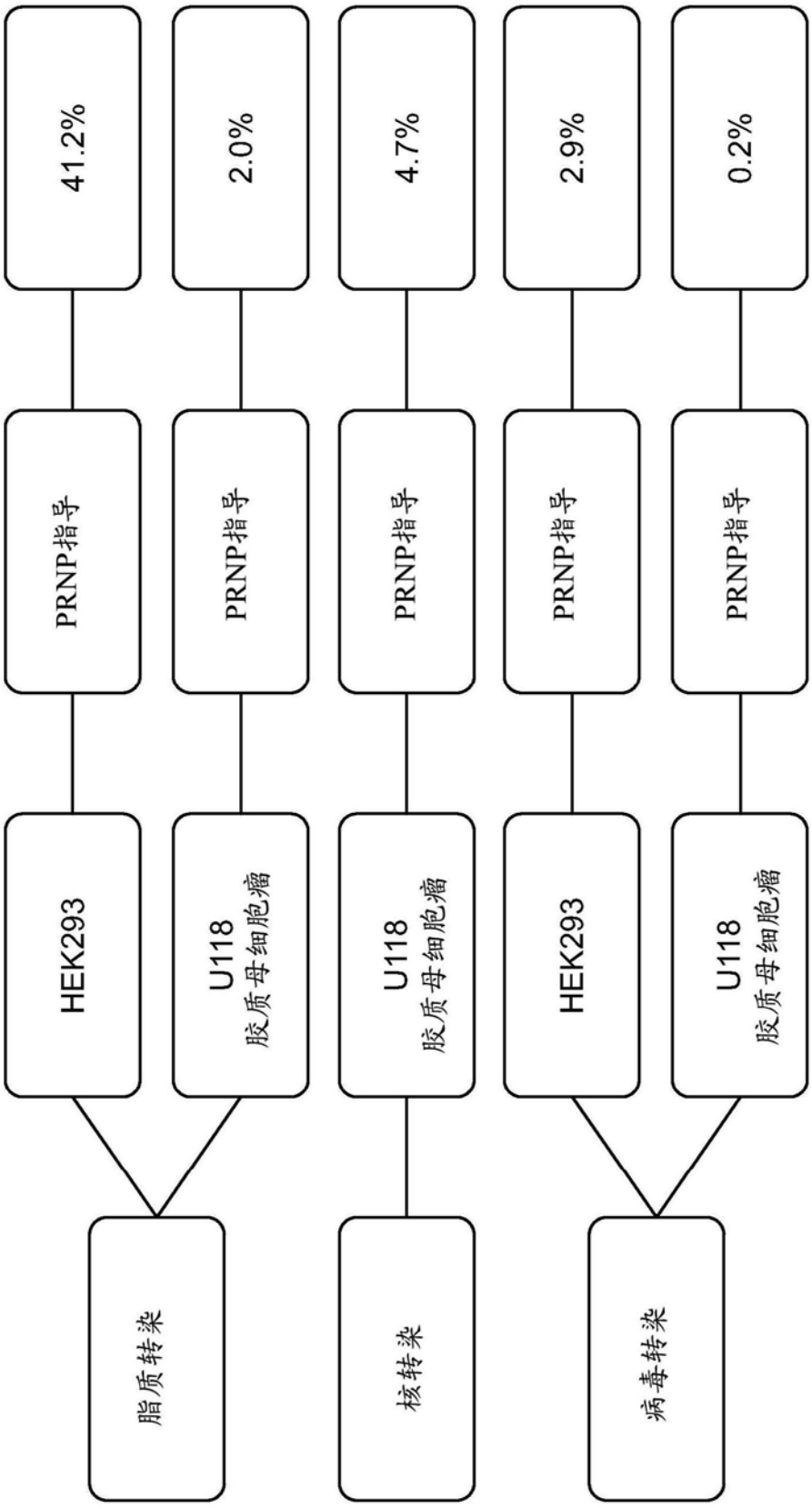

41.2%
2.0%
4.7%
2.9%
0.2%
PRNP指导
PRNP指导
PRNP指导
PRNP指导
PRNP指导
HEK293
U118
胶质母细胞瘤
U118
胶质母细胞瘤
HEK293
U118
胶质母细胞瘤
脂质转染
核转染
病毒转染

图3B

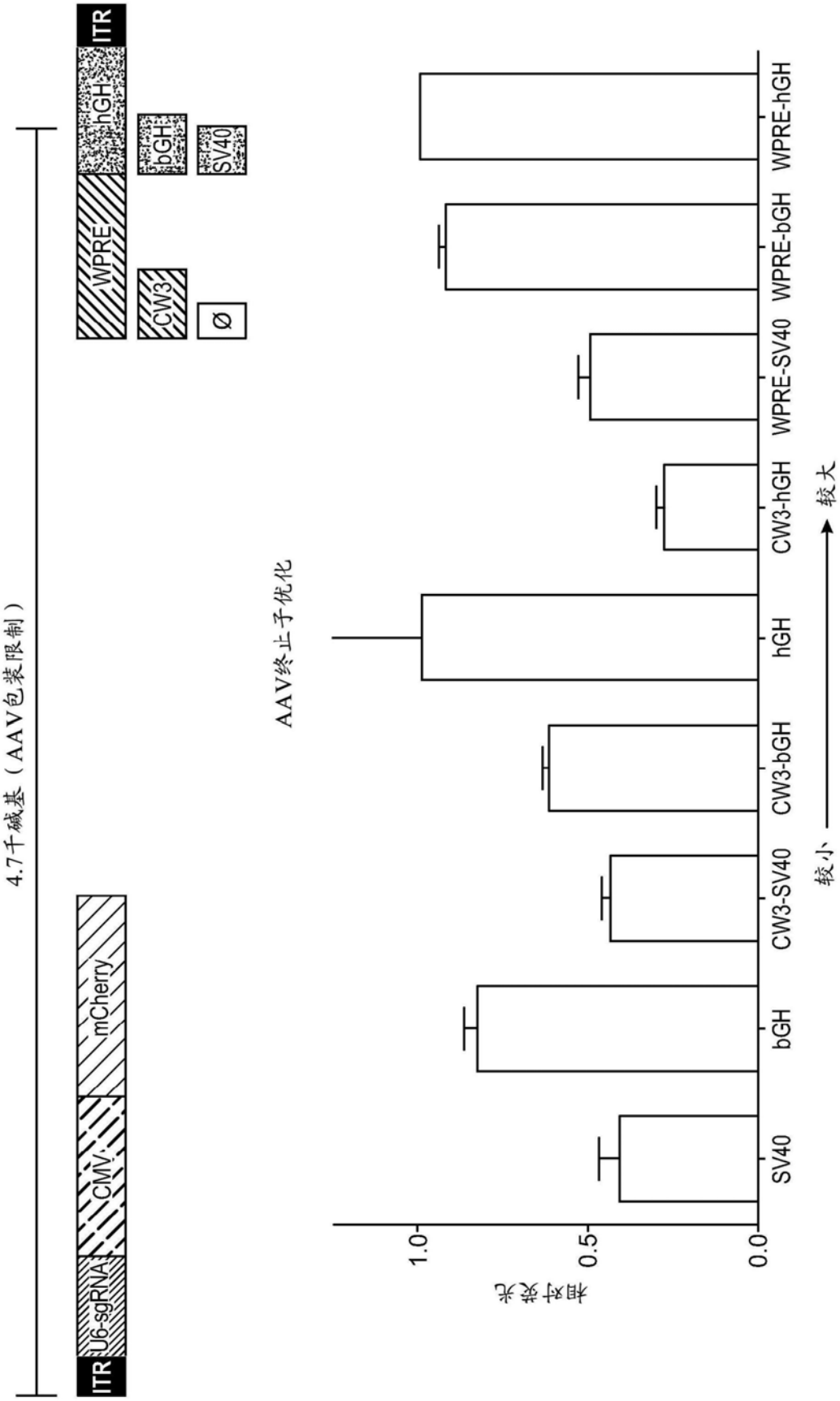

4.7千碱基（AAV包装限制）
ITR
U6-sgRNA
CMV
mCherry
WPRE
hGH
CW3
bGH
Ø
SV40
ITR
AAV终止子优化
相对表达
1.0
0.5
0.0
SV40
bGH
CW3-SV40
CW3-bGH
hGH
CW3-hGH
WPRE-SV40
WPRE-bGH
WPRE-hGH
较小
较大

图4

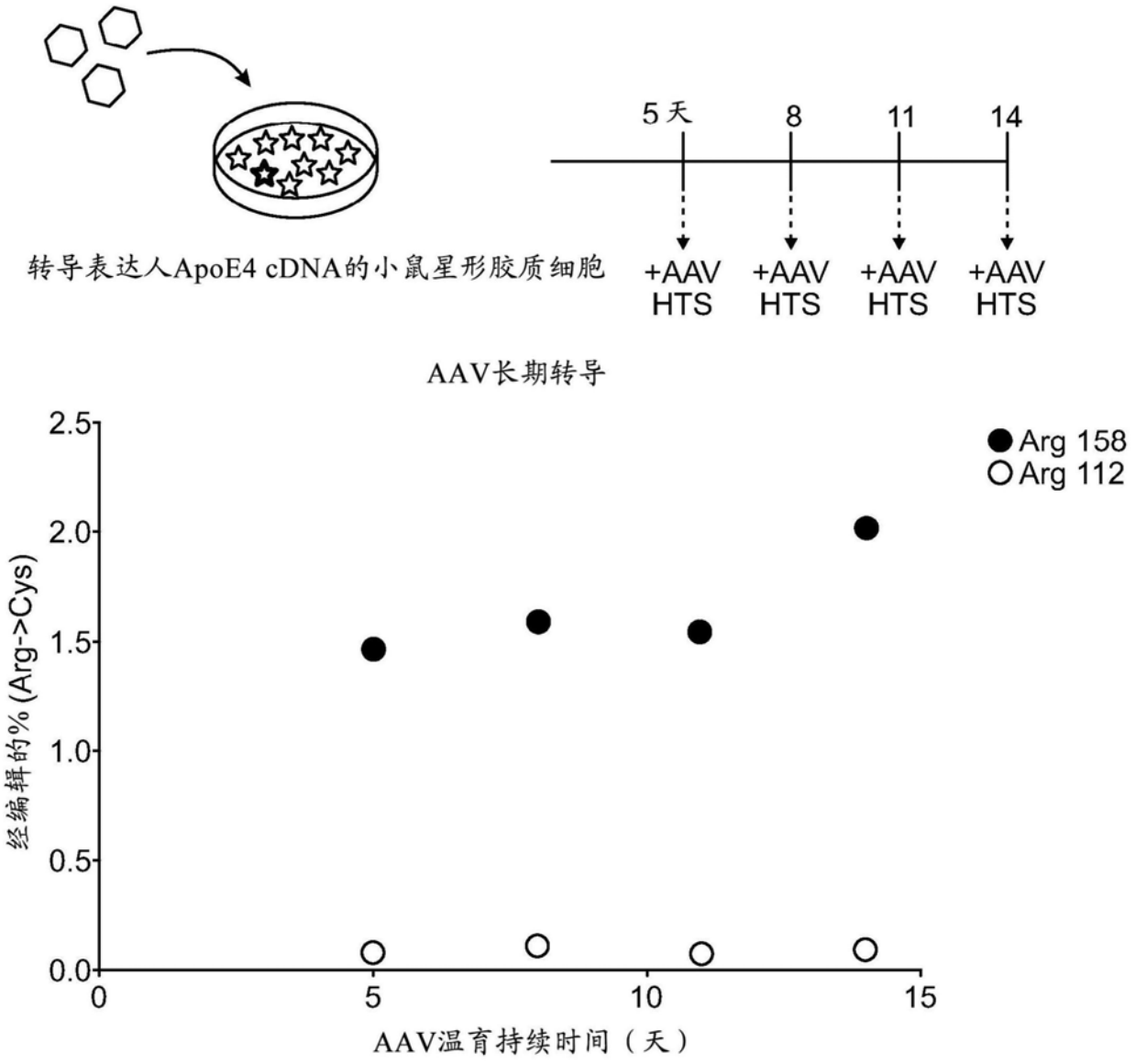

转导表达人ApoE4 cDNA的小鼠星形胶质细胞
5天
8
11
14
+AAV HTS
+AAV HTS
+AAV HTS
+AAV HTS
AAV长期转导
Arg 158
Arg 112
经编辑的% (Arg->Cys)
2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
AAV温育持续时间（天）

图5A

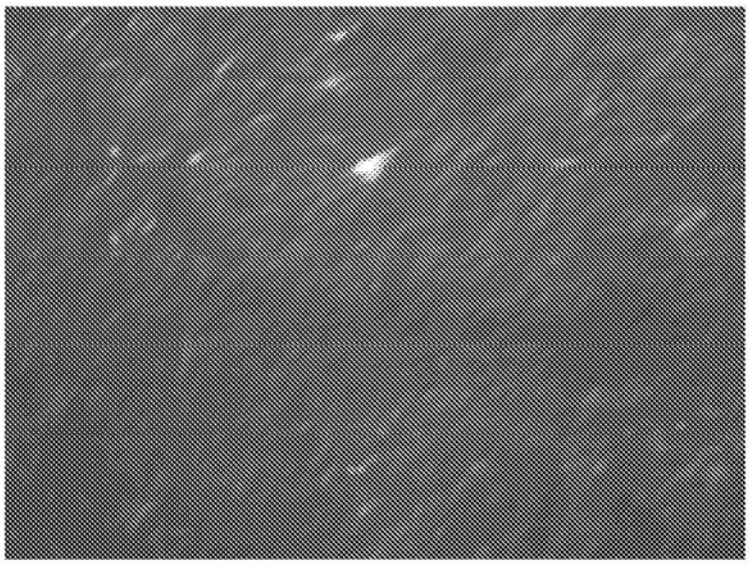


图5B

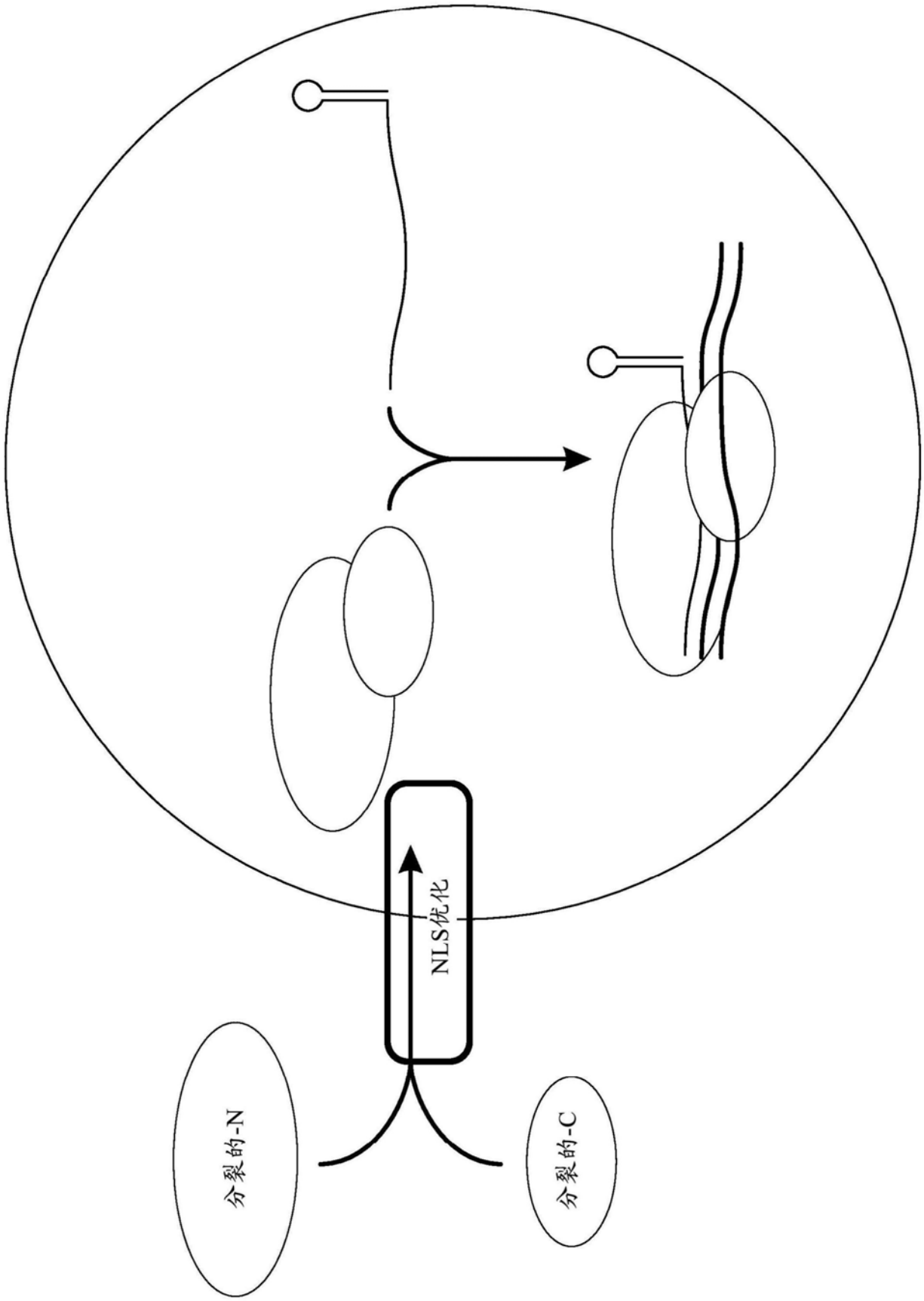


图6

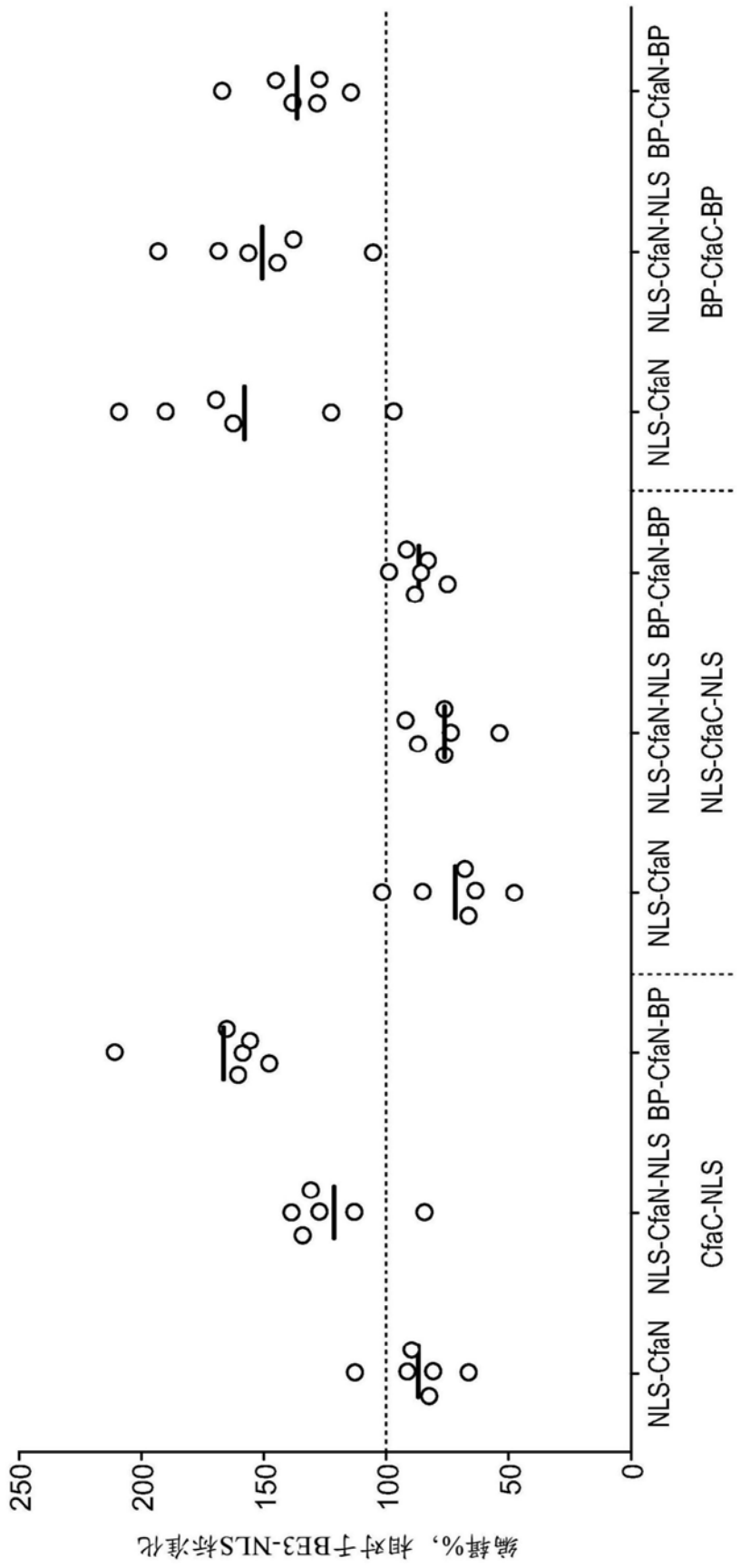

编辑%，相对于BE3-NLS标准化
250
200
150
100
50
0
NLS-CfaN
NLS-CfaN-NLS
BP-CfaN-BP
CfaC-NLS
NLS-CfaN
NLS-CfaN-NLS
BP-CfaN-BP
NLS-CfaC-NLS
NLS-CfaN
NLS-CfaN-NLS
BP-CfaN-BP
BP-CfaC-BP

图7

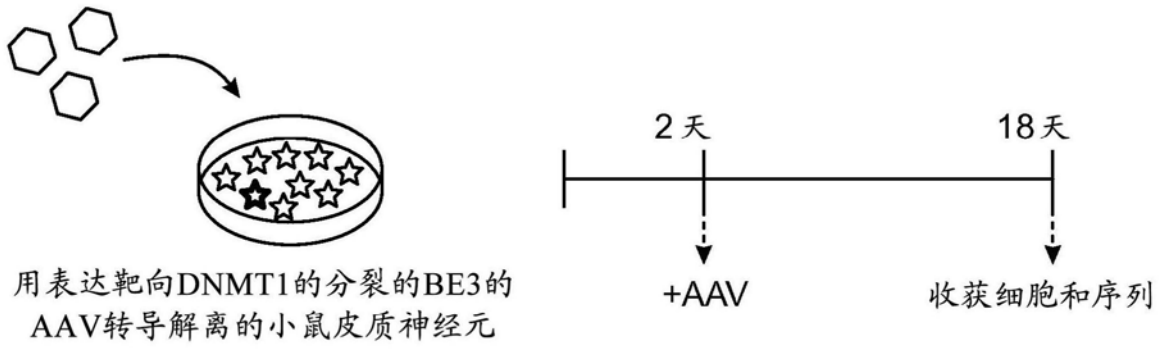

用表达靶向DNMT1的分裂的BE3的
AAV转导解离的小鼠皮质神经元
2天
18天
+AAV
收获细胞和序列

图8A

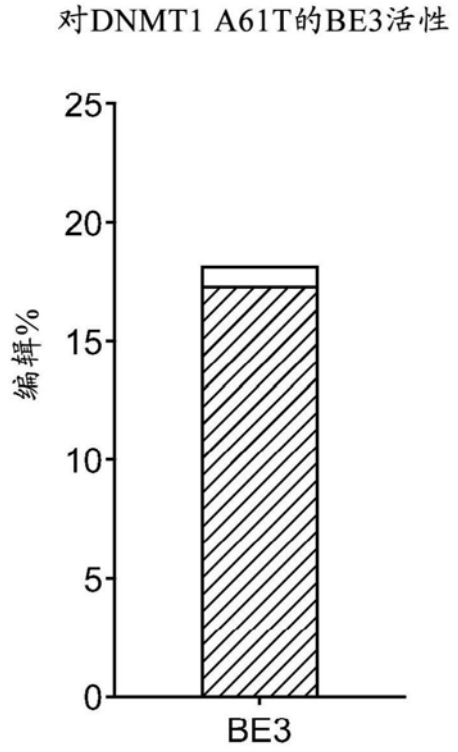

对DNMT1 A61T的BE3活性
编辑%
25
20
15
10
5
0
BE3

图8B

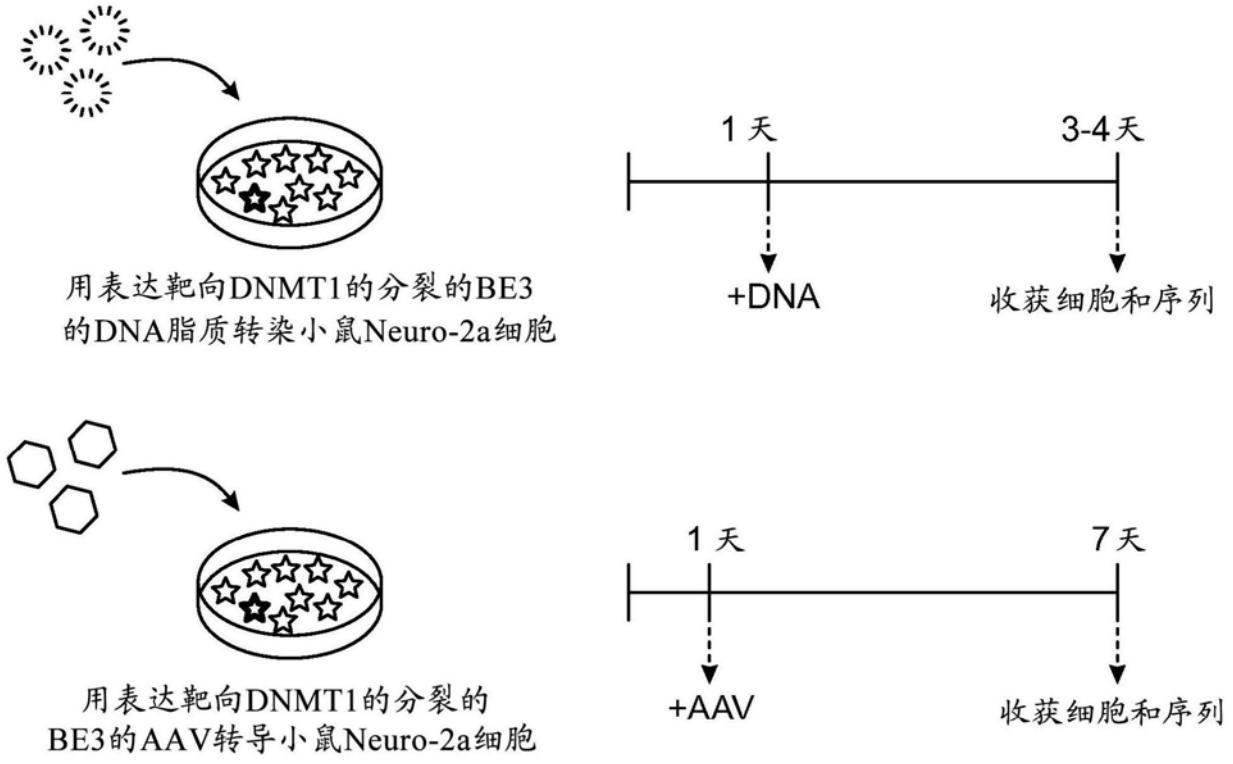

用表达靶向DNMT1的分裂的BE3
的DNA脂质转染小鼠Neuro-2a细胞
1天
3-4天
+DNA
收获细胞和序列
用表达靶向DNMT1的分裂的
BE3的AAV转导小鼠Neuro-2a细胞
1天
7天
+AAV
收获细胞和序列

图9A

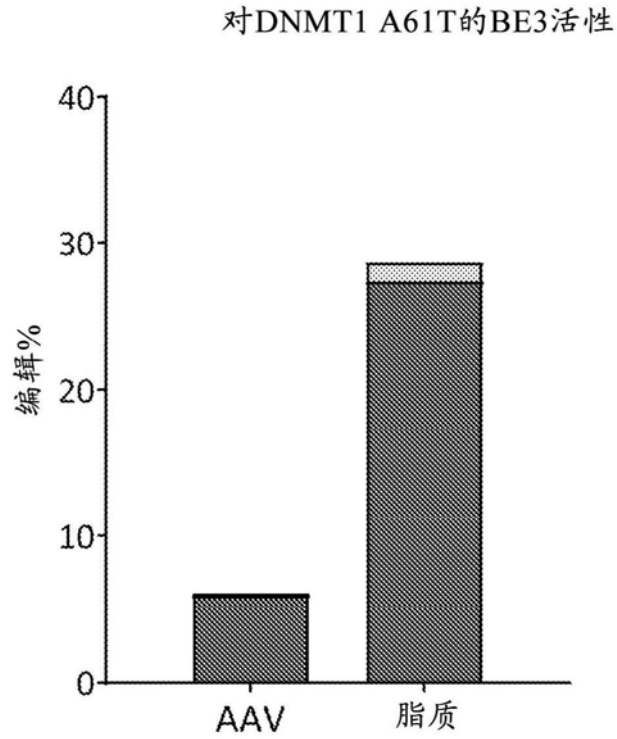

对DNMT1 A61T的BE3活性
编辑%
40
30
20
10
0
AAV
脂质

图9B