(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 111868240 A
(43)申请公布日 2020.10.30

(21)申请号 201880082218.2

(22)申请日 2018.11.09

(30)优先权数据
62/584,310 2017.11.10 US
62/596,375 2017.12.08 US
62/667,084 2018.05.04 US

(85)PCT国际申请进入国家阶段日
2020.06.18

(86)PCT国际申请的申请数据
PCT/US2018/060126 2018.11.09

(87)PCT国际申请的公布数据
WO2019/094791 EN 2019.05.16

(71)申请人 马萨诸塞大学
地址 美国马萨诸塞州

(72)发明人 E·J·索尼默 R·伊波拉赫姆 W·薛 A·米尔 A·艾德拉克 G·伊尔达尔

(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
代理人 郑天松

(51)Int.Cl.
C12N 9/22(2006.01)
C12N 15/79(2006.01)
C12N 15/86(2006.01)

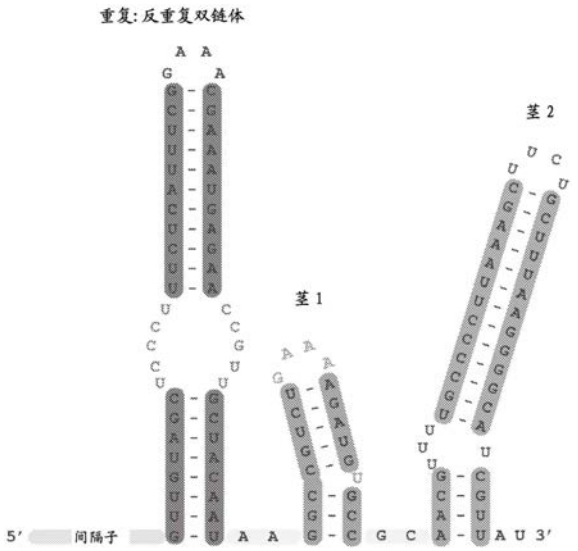
权利要求书3页 说明书50页 附图74页

(54)发明名称

靶向CRISPR递送平台

(57)摘要

本发明涉及用于基因治疗的组合物和方法。本文所述几种方法利用脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9系统,其提供了超精确的CRISPR基因编辑平台。此外,本发明掺入了全长和截短的单引导RNA序列,其允许将完整的sgRNA-Nme1Cas9载体插入与体内施用相容的腺相关病毒质粒中。此外,已经鉴定出II-C型Cas9直系同原物,其靶向原间隔子相邻基序序列限于1～4个所需核苷酸之间。



CN 111868240 A

1.单引导核糖核酸(sgRNA)序列,其包含截短的重复:反重复区。

2.权利要求1的sgRNA序列,其还包含截短的茎2区。

3.权利要求2的sgRNA序列,其还包含截短的间隔子区。

4.权利要求1的sgRNA序列,其中所述sgRNA序列具有121个核苷酸的长度。

5.权利要求2的sgRNA序列,其中所述sgRNA序列长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。

6.权利要求3的sgRNA序列,其中所述sgRNA序列具有100个核苷酸的长度。

7.权利要求1的sgRNA序列,其中所述sgRNA序列是Nme1Cas9单引导核糖核酸序列或Nme2Cas9单引导核糖核酸序列。

8.单引导核糖核酸(sgRNA)序列,其包含截短的茎2区。

9.权利要求8的sgRNA序列,其还包含截短的重复:反重复区。

10.权利要求9的sgRNA序列,其还包含截短的间隔子区。

11.权利要求9的sgRNA序列,其中所述sgRNA序列长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。

12.权利要求10的sgRNA序列,其中所述sgRNA序列具有100个核苷酸的长度。

13.腺相关病毒(AAV)质粒,其包含单引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9核酸载体。

14.权利要求13的AAV质粒,其中所述单引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9核酸载体包含至少一个启动子。

15.权利要求14的AAV质粒,其中所述至少一个启动子选自:U6启动子和U1a启动子。

16.权利要求13的AAV质粒,其中所述单引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9核酸载体包含Kozak序列。

17.权利要求13的AAV质粒,其中所述sgRNA包含与目的基因序列互补的核酸序列。

18.权利要求17的AAV质粒,其中所述目的基因序列选自:PCSK9序列和ROSA26序列。

19.权利要求13的AAV质粒,其中所述sgRNA包含截短的重复:反重复序列。

20.权利要求19的AAV质粒,其中所述sgRNA还包含截短的茎2区。

21.权利要求20的AAV质粒,其中所述sgRNA还包含截短的间隔子区。

22.权利要求19的AAV质粒,其中所述sgRNA序列具有121个核苷酸的长度。

23.权利要求20的AAV质粒,其中所述sgRNA序列的长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。

24.权利要求21的AAV质粒,其中所述sgRNA序列具有100个核苷酸的长度。

25.权利要求13的AAV质粒,其中所述sgRNA包含截短的茎2区。

26.权利要求25的AAV质粒,其中所述sgRNA还包含截短的重复:反重复区。

27.权利要求26的AAV质粒,其中所述sgRNA还包含截短的间隔子区。

28.权利要求26的AAV质粒,其中所述sgRNA序列的长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。

29.权利要求27的AAV质粒,其中所述sgRNA序列具有100个核苷酸的长度。

30.方法,其包括:

(a)提供;

(i)表现出至少一种医学病症的症状的患者,其中所述患者包含与所述医学病症有关的多个基因;

(ii)腺相关病毒(AAV)质粒,其包含单引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9核酸载体,其中所述sgRNA包含与所述多个基因中至少一个的一部分互补的核酸序列,和

(b)在减轻所述医学病症的所述至少一种症状的条件下,向所述患者施用所述AAV质粒。

31.权利要求30的方法,其中所述医学病症包括高胆固醇血症。

32.权利要求30的方法,其中所述多个基因中的所述至少一个是PCSK9基因。

33.权利要求32的方法,其中所述sgRNA核酸与所述PCSK9基因的一部分互补。

34.权利要求30的方法,其中所述sgRNA包含截短的重复:反重复序列。

35.权利要求34的方法,其中所述sgRNA还包含截短的茎2区。

36.权利要求35的方法,其中所述sgRNA还包含截短的间隔子区。

37.权利要求34的方法,其中所述sgRNA序列的长度为121个核苷酸。

38.权利要求35的方法,其中所述sgRNA序列的长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。

39.权利要求21的方法,其中所述sgRNA序列的长度为100个核苷酸。

40.权利要求30的方法,其中所述sgRNA包含截短的茎2区。

41.权利要求40的方法,其中所述sgRNA还包含截短的重复:反重复区。

42.权利要求41的方法,其中所述sgRNA还包含截短的间隔子区。

43.权利要求41的方法,其中所述sgRNA序列的长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。

44.权利要求42的方法,其中所述sgRNA序列具有100个核苷酸的长度。

45.编码II-C型Cas9核酸酶蛋白的腺相关病毒(AAV)质粒,其中所述蛋白包含配置有与原间隔子相邻基序序列的结合位点的原间隔子相邻基序识别结构域,所述原间隔子相邻基序序列包含1～4个所需核苷酸。

46.权利要求45的腺相关病毒质粒,其中所述II-C型Cas9核酸酶蛋白选自:脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株De10444 Nme2Cas9核酸酶蛋白,副流感嗜血杆菌(Haemophilus parainfluenzae)HpaCas9核酸酶蛋白和米氏西蒙斯菌(Simonsiella muelleri)SmuCas9核酸酶蛋白。

47.权利要求46的腺相关病毒质粒,其中所述包含1～4个所需核苷酸的原间隔子相邻基序序列选自:N_4CN_3、N_4CT、N_4CCN、N_4CCA和N_4GNT_3。

48.权利要求45的腺相关病毒质粒,其中所述1～4个所需核苷酸选自:C、CC、CT、CCN、CCA、CN_3和GNT_2。

49.权利要求45的腺相关病毒质粒,其中所述II-C型Cas9核酸酶蛋白与截短的sgRNA结合。

50.方法,其包括:

(a)提供;

(i)表现出至少一种医学病症的症状的患者,其中所述患者包含与所述医学病症相关

的多个基因,其中所述多个基因包含原间隔子相邻基序,所述原间隔子相邻基序包含2～4个所需核苷酸;

(ii)包含至少一种编码II-C型Cas9核酸酶蛋白的核酸的递送平台,其中所述蛋白包含配置有与原间隔子相邻基序序列的结合位点的原间隔子相邻基序识别结构域,所述原间隔子相邻基序序列包含2～4个所需核苷酸;和

(b)在使所述医学病症的所述至少一种症状减轻的条件下,将所述递送平台施用于所述患者。

51.权利要求50的方法,其中所述递送平台包含腺相关病毒质粒。

52.权利要求50的方法,其中所述递送平台包括微粒。

53.权利要求50的方法,其中所述II-C型Cas9核酸酶蛋白选自:脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株De10444 Nme2Cas9核酸酶蛋白,副流感嗜血杆菌(Haemophilus parainfluenzae)HpaCas9核酸酶蛋白和米氏西蒙斯菌(Simonsiella muelleri)SmuCas9核酸酶蛋白。

54.权利要求50的方法,其中所述包含1～4个所需核苷酸的原间隔子相邻基序序列选自:N_4C_3、N_4CT、N_4CCN、N_4CCA和N_4GNT_3。

55.权利要求50的腺相关病毒质粒,其中所述1～4个所需核苷酸选自:C、CC、CT、CCN、CCA、CN_3和GNT_2。

56.权利要求50的方法,其中所述II-C型Cas9核酸酶蛋白与截短的sgRNA结合。

57.权利要求50的方法,其中所述医学病症选自高脂血症和酪氨酸血症。

靶向CRISPR递送平台

【发明领域】

[0001] 本发明涉及用于基因治疗的组合物和方法。本文描述的几种方法利用脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9系统,其提供超精确的CRISPR基因编辑平台。此外,本发明结合了该Cas9系统的改进:例如截短单引导RNA序列,以及将Nme1Cas9或Nme2Ca9与其引导RNA一起包装在可与体内给药相容的腺相关病毒载体中。此外,已经鉴定出II-C型Cas9直系同原物,其靶向原间隔子相邻基序序列限于1～4个所需核苷酸之间。

[0002] 【发明背景】

[0003] 成簇的规则间隔子的短回文重复序列(CRISPR)-CRISPR相关的(Cas)是在古细菌和细菌中发现的独特的RNA引导的适应性免疫系统。这些系统通过靶向和灭活源自外源遗传元件的核酸来提供免疫力。迄今为止,已经确定了许多不同类型的CRISPR-Cas系统,并将其分为两类。

[0004] 在II类CRISPR系统中,II型CRISPR-Cas系统的特征是称为Cas9的单一效应蛋白,该蛋白与CRISPR RNA(crRNA)和反式激活RNA(tracrRNA)形成核糖核蛋白(RNP)复合物,以靶向和靶向切割DNA。crRNA包含一个可编程的引导序列,该序列可将Cas9引导至生物体中几乎任何DNA序列。

[0005] Cas9 RNP复合物的这种可编程性已被许多研究人员用于真核系统中的基因组编辑。它已被用于编辑哺乳动物细胞,人胚胎,植物,啮齿动物和其他生物的基因组。Cas9 RNP已用于精确的(带有供体模板)和不精确的基因组编辑,这两种方法都已在基因治疗,农业和其他领域得到应用。此外,Cas9直系同原物的核酸酶死型也用于转录调节,位点特异性DNA标记和特定基因组位点的蛋白组图谱分析。几种不同的Cas9已用于这些应用程序。Cas9可编程性及其应用的核心是在crRNA中引入任何引导序列的能力。可以将crRNA和tracrRNA融合在一起以形成单引导RNA(sgRNA),它更稳定并提供增强的基因组编辑能力。

[0006] 本领域需要改进的Cas9s和sgRNA序列,其可以提供更广泛范围的靶位点的特异性和准确的编辑,特别是当与可靠的核酸递送平台组合时。

[0007] 【发明概述】

[0008] 本发明涉及基因治疗的组合物和方法。本文描述的几种方法利用脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9系统,其提供超精确的CRISPR基因编辑平台。此外,本发明结合了该Cas9系统的改进:例如,截短单引导RNA序列,以及将Nme1Cas9或Nme2Cas9与其引导RNA一起包装在可与体内施用相容的腺相关病毒载体中。此外,已经鉴定出II-C型Cas9直系同原物,其靶向限于1～4个所需核苷酸之间的原间隔子相邻基序序列。

[0009] 在一个实施方式中,本发明涉及包含截短的重复:反重复区的单引导核糖核酸(sgRNA)序列。在一个实施方式中,所述sgRNA序列还包含截短的茎2区。在一个实施方式中,所述sgRNA序列还包含截短的间隔子区。在一个实施方式中,所述sgRNA序列具有121个核苷酸的长度。在一个实施方式中,所述sgRNA序列长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。在一个实施方式中,所述sgRNA序列具有100个核苷酸的长度。在一个实施方式中,所述sgRNA序列是Nme1Cas9单引导

核糖核酸序列。在一个实施方式中，所述sgRNA序列是Nme2Cas9单引导核糖核酸序列。在一个实施方式中，所述sgRNA序列是Nme1Cas9单引导核糖核酸序列或Nme2Cas9单引导核糖核酸序列。

[0010] 在一个实施方式中，本发明涉及包含截短的茎2区的单引导核糖核酸(sgRNA)序列。在一个实施方式中，所述sgRNA序列还包含截短的重复:反重复区。在一个实施方式中，所述sgRNA序列还包含截短的间隔子区。在一个实施方式中，所述sgRNA序列的长度选自：111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。在一个实施方式中，所述sgRNA序列具有100个核苷酸的长度。

[0011] 在一个实施方式中，本发明涉及包含单引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9(sgRNA-Nme1Cas9或sgRNA-Nme2Cas9)核酸载体的腺相关病毒(AAV)载体。在一个实施方式中，所述单一引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9核酸载体包含至少一个启动子。在一个实施方式中，所述至少一种启动子选自:U6启动子和U1a启动子。在一个实施方式中，所述单一引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9核酸载体包含Kozak序列。在一个实施方式中，所述sgRNA包含与目的基因序列互补的核酸序列。在一个实施方式中，所述目的基因序列选自PCSK9序列和ROSA26序列。在一个实施方式中，所述sgRNA包含长度为145个核苷酸的截短的序列。在一个实施方式中，所述sgRNA包含截短的重复:反重复序列。在一个实施方式中，所述sgRNA还包含截短的茎2区。在一个实施方式中，所述sgRNA还包含截短的间隔子区。在一个实施方式中，所述sgRNA序列具有121个核苷酸的长度。在一个实施方式中，所述sgRNA序列的长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。在一个实施方式中，所述sgRNA序列具有100个核苷酸的长度。在一个实施方式中，所述sgRNA包含截短的茎2区。在一个实施方式中，所述sgRNA还包含截短的重复:反重复区。在一个实施方式中，所述sgRNA还包含截短的间隔子区。在一个实施方式中，所述sgRNA序列的长度选自:111个核苷酸，107个核苷酸，105个核苷酸，103个核苷酸，101个核苷酸和99个核苷酸。在一个实施方式中，所述sgRNA序列具有100个核苷酸的长度。在一个实施方式中，所述sgRNA包含长度为145个核苷酸的截短的序列。

[0012] 在一个实施例中，本发明构想了一种方法，其包括:(a)提供;(i)表现出至少一种医学病症的症状的患者，其中所述患者包含与所述医学病症有关的多个基因;(ii)包含单引导核糖核酸-脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9(sgRNA-Nme1Cas9或sgRNA-Nme2Cas9)核酸载体的递送平台，其中所述sgRNA包含与所述多个核糖核酸中的至少一个的一部分互补的核酸序列;及(b)在减轻所述医学病症的所述至少一种症状的条件下向所述患者施用所述AAV质粒。在一个实施方式中，递送平台包含腺相关病毒(AAV)载体。在一个实施方式中，递送平台包含微粒。在一个实施方式中，所述医学病症包括高胆固醇血症。在一个实施方式中，所述医学病症包括酪氨酸血症。在一个实施方式中，所述多个基因中的所述至少一个是PCSK9基因。在一个实施方式中，所述sgRNA核酸与所述PCSK9基因的一部分互补。在一个实施方式中，所述多个基因中的至少一个是FAH基因。在一个实施方式中，所述sgRNA核酸与所述FAH基因的一部分互补。在一个实施方式中，所述sgRNA包含截短的重复:反重复序列。在一个实施方式中，所述sgRNA还包含截短的茎2区。在一个实施方式中，所述sgRNA还包含截短的间隔子区。在一个实施方式中，所述sgRNA序列具有121个核苷酸的长

度。在一个实施方式中,所述sgRNA序列的长度选自:111个核苷酸,107个核苷酸,105个核苷酸,103个核苷酸,101个核苷酸和99个核苷酸。在一个实施方式中,所述sgRNA序列具有100个核苷酸的长度。在一个实施方式中,所述sgRNA包含截短的茎2区。在一个实施方式中,所述sgRNA还包含截短的重复:反重复区。在一个实施方式中,所述sgRNA还包含截短的间隔子区。在一个实施方式中,所述sgRNA序列的长度选自:111个核苷酸、107个核苷酸、105个核苷酸、103个核苷酸、102个核苷酸、101个核苷酸和99个核苷酸。在一个实施方式中,所述sgRNA序列具有100个核苷酸的长度。在一个实施方式中,所述sgRNA包含长度为145个核苷酸的截短的序列。

[0013] 在一个实施方式中,本发明涉及编码II-C型Cas9核酸酶蛋白的腺相关病毒(AAV)质粒,其中所述蛋白包含配置有与原间隔子相邻基序序列的结合位点的原间隔子相邻基序识别结构域,所述原间隔子相邻基序序列包含1～4个所需核苷酸。在一个实施方式中,所述II-C型Cas9核酸酶蛋白选自:脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株De10444Nme2Cas9核酸酶蛋白,副流感嗜血杆菌(Haemophilus parainfluenzae)HpaCas9核酸酶蛋白和米氏西蒙斯菌(Simonsiella muelleri)SmuCas9核酸酶蛋白。在一个实施方式中,所述包含1～4个所需核苷酸的原间隔子相邻基序序列选自:N_4CN_3、N_4CT、N_4CCN、N_4CCA和N_4GNT_3。在一个实施方式中,一个至4个所需核苷酸选自:C,CT,CCN,CCA,CN_3和GNT_2。在一个实施方式中,所述II-C型Cas9核酸酶蛋白与截短的sgRNA结合。在一个实施方式中,腺相关病毒质粒编码2个sgRNA序列。在一个实施方式中,腺相关病毒质粒编码多腺苷序列。在一个实施方式中,腺相关病毒质粒编码同源性引导的修复供体核苷酸模板。在一个实施方式中,腺相关病毒质粒是多合一腺相关病毒质粒。

[0014] 在一个实施例中,本发明考虑了一种方法,其包括:(a)提供;(i)表现出至少一种医学病症的症状的患者,其中所述患者包含与所述医学病症有关的多个基因,其中所述多个基因包含原间隔子相邻基序,所述原间隔子相邻基序包含1～4个所需核苷酸;(ii)一种递送平台,其包含至少一个编码II-C型Cas9核酸酶蛋白的核酸,其中所述蛋白包含配置有与原间隔子相邻基序序列的结合位点的原生原间隔子相邻基序识别结构域,所述原间隔子相邻基序序列包含2～4个所需核苷酸;(b)在使所述医学病症的所述至少一种症状减轻的条件下,将所述递送平台给予所述患者。在一个实施方式中,所述医学病症包括高胆固醇血症。在一个实施方式中,所述医学病症包括酪氨酸血症。在一个实施方式中,所述多个基因中的所述至少一个是PCSK9基因。在一个实施方式中,所述sgRNA核酸与所述PCSK9基因的一部分互补。在一个实施方式中,所述多个基因中的至少一个是FAH基因。在一个实施方式中,所述sgRNA核酸与所述FAH基因的一部分互补。在一个实施方式中,所述递送平台包含腺相关病毒质粒。在一个实施方式中,所述递送平台包含微粒。在一个实施方式中,所述II-C型Cas9核酸酶蛋白选自:脑膜炎奈瑟氏球菌(Neisseria meningitidis)De10444Nme2Cas9核酸酶蛋白,副流感嗜血杆菌(Haemophilus parainfluenzae)HpaCas9核酸酶蛋白和米氏西蒙斯菌(Simonsiella muelleri)SmuCas9核酸酶蛋白。在一个实施方式中,包含1～4个所需核苷酸的所述原间隔子相邻基序序列选自N_4CN_3,N_4CT,N_4CCN,N_4CCA和N_4GNT_3。在一个实施方式中,1～4个所需核苷酸选自:C,CT,CCN,CCA,CN_3和GNT_2。在一个实施方式中,所述II-C型Cas9核酸酶蛋白与截短的sgRNA结合。在一个实施方式中,腺相关病毒质粒编码2个sgRNA序列。在一个实施方式中,腺相关病毒质粒编码多腺苷序列。在一个实施方式中,腺相关病毒

质粒编码同源性引导的修复供体核苷酸模板。在一个实施方式中，腺相关病毒质粒是多合一腺相关病毒质粒。

[0015] 在一个实施方式中，本发明涉及编码II-C型Cas9核酸酶蛋白的腺相关病毒(AAV)质粒，其中所述蛋白包含被配置为与原间隔子相邻基序(PAM)序列结合的原间隔子相邻基序识别结构域(例如，PAM-相互作用结构域;PID)，所述PAM序列包含相邻的胞嘧啶二核苷酸对。在一个实施方式中，相邻的胞嘧啶二核苷酸对在PAM的位置五(5)和六(6)。在一个实施方式中，所述II-C型Cas9核酸酶蛋白衍生自脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株。在一个实施方式中，脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株是De10444。在一个实施方式中，II-C型Cas9核酸酶蛋白是Nme2Cas9核酸酶蛋白。在一个实施方式中，脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株为98002。在一个实施方式中，II-C型Cas9核酸酶蛋白是Nme3Cas9核酸酶蛋白。在一个实施方式中，所述PAM序列选自：N_4CC，N_4CCN_3，N_4CCA，$N_4CC(X)$，N_4CA_3和N_{10}。在一个实施方式中，PAM序列是N_3CC。在一个实施方式中，II-C型Cas9核酸酶蛋白还包含sgRNA序列。在一个实施方式中，sgRNA序列包含间隔子，间隔子的长度在大约十七(17)～二十四(24)个核苷酸之间。

[0016] 在一个实施例中，本发明考虑了一种方法，其包括：(a)提供;(i)表现出至少一种医学病症的症状的患者，其中所述患者包含与所述医学病症相关的多个基因，其中所述多个基因包含与原胞间隔子相邻的基序，所述基序包括相邻的胞嘧啶二核苷酸对;(ii)递送平台，其包含至少一种编码II-C型Cas9核酸酶蛋白的核酸，其中所述蛋白包含被配置为与所述原间隔子相邻基序序列结合的原间隔子相邻基序识别结构域(例如，PAM相互作用结构域;PID)，所述原间隔子相邻基序序列包含相邻的胞嘧啶二核苷酸对;(b)在使所述医学病症的所述至少一种症状减轻的条件下，将所述递送平台给予所述患者。在一个实施方式中，所述递送平台包含腺相关病毒载体。在一个实施方式中，腺相关病毒载体是腺相关病毒载体八(AAV8)。在一个实施方式中，所述医学病症包括高胆固醇血症。在一个实施方式中，所述医学病症包括酪氨酸血症。在一个实施方式中，医学病症是x连锁慢性肉芽肿病。在一个实施方式中，医学病症是天冬氨酰糖尿症。在一个实施方式中，所述多个基因中的所述至少一个是PCSK9基因。在一个实施方式中，所述sgRNA核酸与所述PCSK9基因的一部分互补。在一个实施方式中，所述多个基因中的至少一个是FAH基因。在一个实施方式中，所述sgRNA核酸与所述FAH基因的一部分互补。在一个实施方式中，腺相关病毒质粒编码至少一个sgRNA序列。在一个实施方式中，腺相关病毒质粒编码2个sgRNA序列。在一个实施方式中，腺相关病毒质粒编码多腺苷序列。在一个实施方式中，腺相关病毒质粒编码同源性引导的修复供体核苷酸模板。在一个实施方式中，腺相关病毒质粒是多合一腺相关病毒质粒。在一个实施方式中，所述递送平台包含微粒。在一个实施方式中，相邻的胞嘧啶二核苷酸对在PAM的位置五(5)和六(6)。在一个实施方式中，所述II-C型Cas9核酸酶蛋白衍生自脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株。在一个实施方式中，脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株是De10444。在一个实施方式中，II-C型Cas9核酸酶蛋白是Nme2Cas9核酸酶蛋白。在一个实施方式中，脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株是98002。在一个实施方式中，II-C型Cas9核酸酶蛋白是Nme3Cas9核酸酶蛋白。在一个实施方式中，所述PAM序列选自：N_4CC，N_4CCN_3，N_4CCA，$N_4CC(X)$，N_4CA_3和N_{10}。在一个实施方式中，PAM序列是N_3CC。在一个实施方式中，II-C型Cas9核酸酶蛋白还包含sgRNA序列。在一个实施方式

中,sgRNA序列包含间隔子,间隔子的长度在大约十七(17)-二十四(24)个核苷酸之间。

[0017] 【定义】

[0018] 为了促进对本发明的理解,下面定义了许多术语。本文所定义的术语具有与本发明有关的领域的普通技术人员通常理解的含义。诸如“一个”,“一个”和“该”的术语不旨在仅指单数实体,还指复数实体,并且还包括一般类别,其具体示例可用于说明。本文中的术语用于描述本发明的特定实施例,但是除了权利要求中概述的以外,它们的用法不限制本发明。

[0019] 在任何测定测量值的任何情况下,如本文所用,术语“约”或“近似”是指给定测量值的+/-5%。

[0020] 如本文所用,“Rosa26基因”或“Rosa26基因”是指广泛用于在小鼠中实现广义表达的人或小鼠(分别)基因座。靶向Rosa26基因座可通过将所需基因引入基因座的第一内含子中,在原始基因捕获系上游约248bp的唯一XbaI位点处实现。可以使用腺病毒剪接受体,随后是目的基因和在独特的XbaI位点插入的聚腺苷酸化位点来构建构建体。新霉素抗性盒也可以包括在靶向载体中。

[0021] 如本文所用,“PCSK9基因”或“Pcsk9基因”是指编码PCSK9蛋白的人或小鼠(分别)的基因座。PCSK9基因位于1号染色体1p32.3处,包括13个外显子。该基因可以通过选择性剪接产生至少两种同工型。

[0022] 术语“蛋白原转化酶舒替兰酶/kexin 9型”和“PCSK9”是指由调节低密度脂蛋白水平的基因编码的蛋白。蛋白原转化酶舒替兰酶/kexin 9型,也称为PCSK9,是一种在人中由PCSK9基因编码的酶。Seidah et al.,"The secretory proprotein convertase neural apoptosis-regulated convertase 1(NARC-1):liver regeneration and neuronal differentiation"Proc.Natl.Acad.Sci.U.S.A.100(3):928-933(2003)。在许多物种中发现了相似的基因(直系同源基因)。许多酶,包括PSCK9,在首次合成时都是无活性的,因为它们的一部分肽链会阻断其活性;蛋白原转化酶去除该部分以激活该酶。据信PSCK9在胆固醇稳态中起调节作用。例如,PCSK9可以与低密度脂蛋白受体(LDL-R)的表皮生长因子样重复序列A(EGF-A)域结合,从而导致LDL-R内在化和降解。显然,可以预期的是,降低的LDL-R水平会导致LDL-C的代谢减少,这可能导致高胆固醇血症。

[0023] 如本文所用,术语“高胆固醇血症”是指血液胆固醇水平升高到超过临床推荐水平的任何医学病症。例如,如果使用低密度脂蛋白(LDL)测量胆固醇,则如果测量的LDL水平高于例如大约70mg/dl,则可能存在高胆固醇血症。或者,如果使用游离血浆胆固醇测量胆固醇,如果测得的游离胆固醇水平高于例如约200-220mg/dl,则可能存在高胆固醇血症。

[0024] 如本文所用,术语“CRISPRs”或“聚簇的规则间隔子的短回文重复序列”是指DNA基因座的首字母缩写,其包含碱基序列的多个,短的,直接的重复。每个重复序列包含一系列碱基,后接30个左右的碱基对,称为“间隔子”序列。间隔子是病毒DNA的短片段,可以用作过去暴露的“记忆”,以促进针对未来入侵的适应性防御。Doudna et al.,"Genome editing.The new frontier of genome engineering with CRISPR-Cas9"Science,346(6213):1258096(2014)。

[0025] 如本文所用,术语“Cas”或“CRISPR相关(cas)”是指经常与CRISPR重复间隔子阵列相关的基因。

[0026] 如本文所用,术语“Cas9”是指来自II型CRISPR系统的核酸酶,该酶专门用于在DNA中产生双链断裂,具有2个活性切割位点(HNH和RuvC结构域),每个酶切位点双螺旋链tracrRNA和间隔子RNA可以组合成一个“单引导RNA”(sgRNA)分子,与Cas9混合,可以通过sgRNA内的引导序列与靶DNA序列之间的Watson-Crick配对找到并切割DNA靶。Jinek et al.,"A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity"Science 337(6096):816-821(2012)。

[0027] 如本文所用,术语“催化活性的Cas9”是指包括完整核酸酶活性的未修饰的Cas9核酸酶。

[0028] 如本文所用,术语“切口酶”是指仅切割单条DNA链的核酸酶,这归因于其天然功能或因为其已被设计成仅切割单条DNA链。具有RuvC或HNH结构域突变的Cas9切口酶变体可控制DNA链被切割并保持完整。Jinek et al.,"A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity",Science 337(6096):816-821(2012)和Cong et al.,"Multiplex genome engineering using CRISPR/Cas systems",Science,339(6121):819-823(2013)。

[0029] 如本文所用,术语“反式激活crRNA”,“tracrRNA”是指小的反式编码的RNA。例如,CRISPR/Cas(聚簇的规则间隔子的短回文重复序列/CRISPR相关蛋白)构成了RNA介导的防御系统,可防御病毒和质粒。该防御路径包括3个步骤。首先,将入侵核酸的一个拷贝整合到CRISPR基因座中。接下来,从该CRISPR基因座转录CRISPR RNA(crRNA)。然后将crRNA掺入效应物复合物中,其中crRNA将复合物引导至侵入的核酸,并且Cas蛋白降解该核酸。CRISPR激活有多种途径,其中一种需要tracrRNA,而tracrRNA在crRNA的成熟中起作用。TracrRNA与pre-crRNA的重复序列互补,形成RNA双链体。这被RNA特异性核糖核酸酶RNase III切割,形成crRNA/tracrRNA杂交体。该杂交体充当核酸内切酶Cas9的引导,该酶切割入侵的核酸。

[0030] 如本文所用,术语“原间隔子相邻基序(或PAM)”是Cas9/sgRNA形成R环以通过其引导RNA与基因组的Watson-Crick配对询问特定DNA序列可能需要的DNA序列。PAM特异性可以是Cas9蛋白的DNA结合特异性的函数(例如,Cas9的C末端的“原间隔子相邻基序识别结构域”)。

[0031] 如本文所用,术语“原间隔子相邻基序识别结构域”,“PAM相互作用结构域”或“PID”是指Cas9氨基酸序列,其包含与DNA靶PAM序列的结合位点。

[0032] 如本文所用,术语“结合位点”是指具有特定三级和/或四级结构的任何分子排列,所述三级和/或四级结构与结合成分发生物理连接或紧密结合。例如,分子排列可以包含氨基酸序列。备选地,分子排列可以包含核酸序列。此外,分子排列可以包含脂质双层或其他生物材料。

[0033] 如本文所用,术语“sgRNA”是指与CRISPR相关系统(Cas)结合使用的单引导RNA。sgRNA是crRNA和tracrRNA的融合体,包含与所需靶位点互补的核苷酸序列。Jinek et al.,"A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity"Science 337(6096):816-821(2012)。sgRNA与靶位点的Watson-Crick配对可形成R环,与功能性PAM结合使用可进行DNA切割,或者在核酸酶缺陷的情况下,Cas9可使该位点与DNA结合。

[0034] 如本文所用,术语“正交”是指不重叠,不相关或独立的靶。例如,如果利用2个正交

的Cas9同工型,它们将采用仅对Cas9同工型之一编程以进行DNA识别和切割的正交sgRNA。Esvelt et al.,"Orthogonal Cas9 proteins for RNA-guided gene regulation and editing"Nat Methods10(11):1116-1121(2013)。例如,这将允许一个Cas9同工型(例如酿脓链球菌(S.pyogenes)Cas9或SpyCas9)充当由可能对其特异的sgRNA编程的核酸酶,而另一个Cas9同工型(例如脑膜炎奈瑟氏球菌(N.meningitidis)Cas9或NmeCas9)作为核酸酶死亡的Cas9,通过其PAM特异性和正交sgRNA提供靶向结合位点的DNA。其他Cas9包括金黄色葡萄球菌(S.aureus)Cas9或SauCas9和内生链霉菌(A.naeslundii)Cas9或AnaCas9。

[0035] 当用于多核苷酸序列或氨基酸序列时,如本文所用,术语"截短的"是指野生型序列的至少一部分可以不存在。在某些情况下,sgRNA或crRNA中的截短的引导序列可以提高Cas9的编辑精度。Fu,et al."Improving CRISPR-Cas nuclease specificity using truncated guide RNAs"Nat Biotechnol.2014Mar;32(3):279-284(2014)。

[0036] 如本文所用,术语"碱基对"是指特定的核碱基(也称为含氮碱基),其是形成DNA和RNA的一级结构的核苷酸序列的结构单元。双链DNA的特征可以在于特定的氢键模式。碱基对可包括但不限于鸟嘌呤-胞嘧啶和腺嘌呤-胸腺嘧啶碱基对。

[0037] 如本文所用,术语"特异性基因组靶"是指能够结合本文考虑的Cas9蛋白的任何预定的核苷酸序列。靶可包括但不限于与可编程DNA结合域互补的核苷酸序列或用其自身的引导RNA编程的正交Cas9蛋白,与单引导RNA互补的核苷酸序列,原间隔子相邻基序识别序列,中靶结合序列和脱靶结合序列。

[0038] 如本文所用,术语"中靶结合序列"是指特定基因组靶的亚序列,其可以与可编程DNA结合域和/或单引导RNA序列完全互补。

[0039] 如本文所用,术语"脱靶结合序列"是指特定基因组靶的亚序列,其可以与可编程DNA结合域和/或单引导RNA序列部分互补。

[0040] 如本文所用,术语"结合失败"是指任何核苷酸-核苷酸相互作用或核苷酸-氨基酸相互作用,其表现出部分互补性,但互补性不足以识别以触发Cas9核酸酶切割靶位点。这样的结合失败可能导致2个分子的弱或部分结合,使得预期的生物学功能(例如,核酸酶活性)失败。

[0041] 如本文所用,术语"切割"可以定义为DNA中断裂的产生。取决于可以使用的核酸酶的类型,这可以是单链断裂或双链断裂。

[0042] 如本文所用,术语"编辑"或"编辑"是指通过使用外源提供的DNA模板选择性删除特定的基因组靶或特定包含新序列来改变多核苷酸的核酸序列(例如,野生型天然存在的核酸序列或天然突变的序列)的方法。这样的特定基因组靶包括但不限于染色体区,线粒体DNA,基因,启动子,开放阅读框或任何核酸序列。

[0043] 如本文所用,术语"删除","删除","删除"或"删除"可定义为核苷酸或氨基酸序列的变化,其中一个或多个核苷酸或氨基酸残基分别为不存在或变得不存在。

[0044] 如本文所用,术语"目的基因"是指可能需要缺失的任何预定基因。

[0045] 如本文所用,术语"等位基因"是指相同基因或相同遗传基因座的多种替代形式中的任一种。

[0046] 如本文所用,术语"有效量"是指包含达到临床有益结果(即例如减轻症状)的治疗剂的药物组合物的特定量。此类组合物的毒性和治疗功效可通过细胞培养或实验动物中的

标准制药程序确定，例如用于确定LD_{50}(致死人群的50％的剂量)和ED_{50}(以50为治疗有效剂量)人口百分率)。毒性和治疗效果之间的剂量比是治疗指数，可以表示为LD_{50}/ED_{50}之比。表现出大治疗指数的化合物是优选的。从这些细胞培养测定法和其他动物研究中获得的数据可用于配制供人使用的剂量范围。此类化合物的剂量优选在循环浓度范围内，包括几乎没有毒性或无毒性的ED_{50}。剂量在此范围内变化，这取决于所采用的剂型，患者的敏感性和给药途径。

[0047] 如本文所用，术语“症状”是指患者观察到的疾病或身体不适的任何主观或客观证据。例如，主观证据通常基于患者的自我报告，并且可以包括但不限于疼痛，头痛，视力障碍，恶心和/或呕吐。备选地，客观证据通常是医学测试的结果，包括但不限于体温，全血细胞计数，脂质板，甲状腺板，血压，心率，心电图，组织和/或身体成像扫描。

[0048] 如本文所用，术语“疾病”或“医学病症”是指对动物或植物体或其部分之一的正常状态的任何破坏，其中断或改变了生命功能的执行。通常表现为明显的体征和症状，通常是对以下方面的反应：(i)环境因素(如营养不良，工业危害或气候)；(ii)特定的感染因子(如蠕虫，细菌或病毒)；(iii)有机体的固有缺陷(作为遗传异常)；和/或(iv)这些因素的组合。

[0049] 当提及未经治疗的受试者相对于治疗的受试者的任何症状时，“减少”，“抑制”，“减少”，“抑制”，“减少”，“防止”和语法等价形式(包括“降低”，“减小”等)意味着经治疗的受试者中症状的数量和/或程度要比未经治疗的受试者低任何医学训练人员认为与临床相关的量。在一个实施方式中，所治疗的受试者中症状的数量和/或严重程度比低于，至少降低25％，低于至少50％，低于至少75％和/或低于至少10％，和/或低于比未经治疗的受试者的症状的数量和/或严重程度低至少90％。

[0050] 如本文所用，术语“附着”是指介质(或载体)与药物之间的任何相互作用。附件可能是可逆的或不可逆的。这种附接包括但不限于共价键，离子键，范德华力或摩擦力等。如果将药物浸渍，掺入，包被，悬浮，与溶液混合或与之混合，则将药物附着到介质(或载体)上。

[0051] 如本文所用，术语“药物”或“化合物”是指能够给药以实现所需效果的任何药理活性物质。药物或化合物可以是合成的或天然存在的非肽，蛋白或肽，寡核苷酸或核苷酸，多糖或糖。

[0052] 如本文所用，术语“给药”是指向患者提供组合物以使该组合物对患者具有预期作用的任何方法。施用的示例性方法是通过直接机制，例如局部组织施用(即，血管外放置)，口服，透皮贴剂，局部，吸入，栓剂等。

[0053] 如本文所用，术语“患者”或“受试者”是人或动物，不需要住院。例如，门诊病人，疗养院中的人是“病人”。病人可以包括任何年龄的人或非人动物，因此包括成年和未成年人(即儿童)。术语“患者”并不意味着需要治疗，因此，无论是临床还是支持基础科学研究，患者都可以自愿或非自愿地参与实验。

[0054] 如本文所用，术语“亲和力”是指物质或颗粒之间使它们进入并保持化学结合的任何吸引力。例如，与具有低亲和力的抑制剂相比，对受体具有高亲和力的抑制剂化合物将在防止受体与其天然配体相互作用方面提供更大的功效。

[0055] 如本文所用，术语“源自”是指化合物或序列的来源。一方面，化合物或序列可以源自生物或特定物种。在另一方面，化合物或序列可以衍生自更大的复合物或序列。

[0056] 如本文所用，术语“蛋白”是指由肽键连接的氨基酸残基组成的多种天然存在的极其复杂的物质(作为酶或抗体)中的任何一种，包含碳，氢，氮，氧元素，通常是硫。通常，蛋白包含数百个数量级内的氨基酸。

[0057] 如本文所用，术语“肽”是指通过一个酸的氨基与另一个酸的羧基结合而衍生自2个或更多个氨基酸的各种酰胺中的任何酰胺，并且通常是通过蛋白的部分水解而获得的。通常，肽包含具有数十个数量级的氨基酸。

[0058] 术语“多肽”是指通过一个酸的氨基与另一个酸的羧基结合而衍生自2个或更多个氨基酸的各种酰胺中的任何酰胺，并且通常是通过蛋白的部分水解获得的。通常，肽包含具有数十或更大数量级的氨基酸。

[0059] 如本文所用，术语“药学上”或“药理学上可接受的”是指当施用于动物或人时不产生不利的，过敏的或其他不良反应的分子实体和组合物。

[0060] 如本文所用，术语“药学上可接受的载体”包括任何和所有溶剂，或分散介质，其包括但不限于水，乙醇，多元醇(例如，甘油，丙二醇和液体聚乙烯)二醇等)，其合适的混合物以及植物油，涂料，等渗和吸收延迟剂，脂质体，可商购的清洁剂等。补充生物活性成分也可掺入此类载体中。

[0061] 如本文所用，“核酸序列”和“核苷酸序列”是指寡核苷酸或多核苷酸及其片段或部分，并且是指可以是单链或双链的基因组或合成来源的DNA或RNA，和代表有义或反义链。

[0062] 如本文所用，术语“分离的核酸”是指已经从其天然状态去除(例如，从细胞去除并且在优选的实施方式中，不含其他基因组核酸)的任何核酸分子。

[0063] 本文所用的术语“氨基酸序列”和“多肽序列”是可互换的，并且是指氨基酸序列。

[0064] 如本文所用，术语“部分”在涉及蛋白时(如在“给定蛋白的一部分”中)是指该蛋白的片段。片段的大小范围可以从4个氨基酸残基到整个氨基酸序列减去一个氨基酸。

[0065] 当用于核苷酸序列时，术语“部分”是指该核苷酸序列的片段。片段的大小范围可以从5个核苷酸残基到整个核苷酸序列减去一个核酸残基。

[0066] 术语“生物活性的”是指具有结构，调节或生化功能的任何分子。例如，生物学活性可以例如通过在缺乏蛋白活性的细胞中恢复野生型生长来确定。缺乏蛋白活性的细胞可以通过许多方法产生(例如，点突变和移码突变)。通过用表达蛋白，其衍生物或其部分的表达载体转染缺乏蛋白活性的细胞来实现互补。

[0067] 如本文所用，术语“互补的”或“互补性”用于指与碱基配对规则相关的“多核苷酸”和“寡核苷酸”(它们是可互换的术语，指的是核苷酸序列)。例如，序列“CAGT”与序列“GTCA”互补。互补可以是“部分的”或“全部的”。“部分”互补是指一个或多个核酸碱基根据碱基配对规则错配，核酸之间的“全部”或“完全”互补是指每个核酸碱基在碱基配对下均与另一个碱基匹配规则。核酸链之间的互补程度对核酸链之间的杂交效率和强度具有重要影响。这在扩增反应以及取决于核酸之间结合的检测方法中特别重要。

[0068] 如本文所用，术语“杂交”是指使用核酸链通过碱基配对与互补链结合以形成杂交复合物的任何过程来配对互补核酸。杂交和杂交强度(即核酸之间缔合的强度)受以下因素影响，例如核酸之间的互补程度，所涉及条件的严格性，形成的杂交体的T_m以及核酸内的G：C比。

[0069] 如本文所用，术语“杂交复合物”是指由于互补的G和C碱基之间以及互补的A和T碱

基之间的氢键的形成而在2个核酸序列之间形成的复合物;这些氢键可通过碱基堆积相互作用进一步稳定。2个互补核酸序列以反平行构型氢键键合。杂交复合物可以在溶液中形成(例如,C_0 t或R_0 t分析)或溶液中存在的一种核酸序列与固定在固相支持物上的另一种核酸序列之间(例如,Southern和Northern印迹,斑点印迹或载玻片中使用的尼龙膜或硝酸纤维素滤膜)原位杂交,包括FISH(荧光原位杂交)。

[0070] 真核生物中的转录控制信号包含“启动子”和“增强子”元件。启动子和增强子由短序列的DNA序列组成,这些序列与参与转录的细胞蛋白发生特异性相互作用。Maniatis,T.et al.,Science 236:1237(1987)。已经从多种真核来源分离了启动子和增强子元件,包括植物,酵母,昆虫,哺乳动物细胞和病毒中的基因(在原核生物中也发现了类似的控制元件,即启动子)。具体启动子和增强子的选择取决于用于表达目的蛋白的细胞类型。

[0071] 如本文所用,术语“聚A位点”或“聚A序列”表示引导新生RNA转录物的终止和聚腺苷酸化的DNA序列。重组转录物的有效聚腺苷酸化是理想的,因为缺乏聚A尾巴的转录物是不稳定的并且迅速降解。表达载体中利用的多聚A信号可以是“异源的”或“内源的”。内源性多聚A信号是在基因组中给定基因的编码区的3’末端天然发现的信号。异源多聚A信号是从一个基因分离并置于另一个基因的3’端的信号。重组DNA序列在真核细胞中的有效表达涉及信号的表达,该信号指导所得转录物的有效终止和聚腺苷酸化。转录终止信号通常在聚腺苷酸化信号的下游发现,并且长度为数百个核苷酸。

[0072] 术语“转染”或“转染的”是指将外源DNA引入细胞。

[0073] 如本文所用,术语“核酸分子编码”,“DNA序列编码”和“DNA编码”是指沿着脱氧核糖核酸链的脱氧核糖核苷酸的顺序或序列。这些脱氧核糖核苷酸的顺序决定了沿着多肽(蛋白)链的氨基酸顺序。因此,DNA序列编码氨基酸序列。

[0074] 如本文所用,术语“编码区”在用于结构基因时是指编码由于mRNA分子翻译而在新生多肽中发现的氨基酸的核苷酸序列。在真核生物中,编码区在5’侧由编码引发剂蛋氨酸的核苷酸三联体“ATG”界定,在3’侧由指定终止密码子的3个三联体之一(即TAA,TAG,TGA)界定)。

[0075] 如本文所用,术语“结构基因”是指编码RNA或蛋白的DNA序列。相反,“调节基因”是编码控制其他基因(例如,转录因子)表达的产物的结构基因。

[0076] 如本文所用,术语“基因”是指脱氧核糖核苷酸序列,其包含结构基因的编码区,并且包括在5’和3’末端上与编码区相邻定位的序列,在DNA上约1kb的距离,使该基因的两端与全长mRNA的长度相对应。位于编码区5’并且存在于mRNA上的序列被称为5’非翻译序列。位于编码区3’或下游并存在于mRNA上的序列称为3’非翻译序列。术语“基因”涵盖基因的cDNA和基因组形式。基因的基因组形式或克隆包含被非编码序列打断的编码区,这些非编码序列称为“内含子”或“中间区”或“中间序列”。内含子是基因的片段,被转录成异质核RNA(hnRNA);内含子可能包含调控元件,例如增强子。内含子从核或初级转录本中去除或“剪接”;因此,信使RNA(mRNA)转录本中没有内含子。mRNA在翻译过程中起作用,以指定新生多肽中氨基酸的序列或顺序。

[0077] 除了含有内含子外,基因的基因组形式还可以包括位于存在于RNA转录物上的序列的5’和3’端的序列。这些序列被称为“侧翼”序列或区(这些侧翼序列位于mRNA转录本上存在的非翻译序列的5’或3’)。5′侧翼区可以含有调控序列,例如控制或影响基因转录的启

动子和增强子。3′侧翼区可包含引导转录终止,转录后切割和聚腺苷酸化的序列。

[0078] 术语“病毒载体”涵盖源自病毒基因组的任何核酸构建体,其能够掺入异源核酸序列以在宿主生物中表达。例如,这种病毒载体可以包括但不限于腺相关病毒载体,慢病毒载体,SV40病毒载体,逆转录病毒载体,腺病毒载体。尽管病毒载体有时是由病原性病毒产生的,但可以通过对其进行修饰的方法来最大程度地降低其整体健康风险。这通常涉及删除与病毒复制有关的一部分病毒基因组。这种病毒可以有效地感染细胞,但是一旦感染发生,该病毒可能需要辅助病毒来提供缺失的蛋白以生产新的病毒体。最好是病毒载体应该对其感染的细胞的生理影响最小,并表现出遗传稳定的特性(例如,不进行自发的基因组重排)。大多数病毒载体被工程化以感染尽可能广泛的细胞类型。即使这样,也可以修饰病毒受体以将病毒靶向特定种类的细胞。以这种方式修饰的病毒被称为伪型。病毒载体通常被工程化以掺入有助于识别哪些细胞吸收了病毒基因的某些基因。这些基因称为标志物基因。例如,共同的标志物基因赋予抗生素对某种抗生素的抗性。

[0079] 【附图简述】

[0080] 该专利的文件包含至少一个彩色附图。专利和商标局将根据要求并支付必要的费用,提供带有彩色附图的本专利的副本。

[0081] 图1显示了常规的全长145nt Nme1Cas9和Nme2Cas9 sgRNA的代表性序列。

[0082] 图2显示了示例性Nme1Cas9 sgRNA序列和相关的基因编辑活性,其具有截短的重复:反重复区或截短的茎2区。Nme1Cas9 sgRNA的删除/截短系列。上图:比对序列,颜色编码如图1所示。下图:使用指示的sgRNA作为引导,在Nme1Cas9靶位点7(NTS7)上进行编辑的T7E1分析。

[0083] 图3呈现了示例性Nme1Cas9 sgRNA序列和具有截短的重复:反重复区或截短的茎2区的相关基因编辑活性。最短的Nme1Cas9 sgRNA(#10-101nt;24nt引导序列;#11-100nt;23nt引导序列)有效地编辑了人基因组中的3个不同靶位点(NTS7,NTS27和NTS55)。上图:野生型和最小化sgRNA的序列,使用与之前图中相同的配色方案。下图:HEK293T细胞中3个靶位点的编辑效率的T7E1分析。

[0084] 图4呈现了Nme1Cas9 wt sgRNA,截短的sgRNA 11和12的示例性序列(作为二级结构)以及通过Nme1Cas9和sgRNA的RNP递送的相关基因编辑。使用HEK293T细胞在人基因组中靶向了3个基因组位点(N-TS72,N-TS55和N-TS40)和一个交通灯报告子位点。上图:显示为野生型和最小化sgRNA二级结构的序列。下图:通过T7E1分析或流式细胞仪测量的编辑效率以条形图表示。

[0085] 图5显示了使用最小化的sgRNA 11和体外转录的wt sgRNA在PLB985细胞中的基因编辑。用Nme1Cas9和sgRNA的RNP复合物转染细胞,并通过TIDE在基因组位点N-TS72上进行基因编辑。

[0086] 图6给出了包含完整的CRISPR/Cas9基因编辑复合体的AAV载体的一个实施方式的示意图。各种AAV载体区的代表序列在附录1中用颜色编码。

[0087] 图7显示了Nme单引导RNA和启动子的颜色编码序列的一个实施方式,如图4所示,其中使用SapI将骨架线性化以插入24-nt靶间隔子区。

[0088] U6启动子:蓝绿色。

[0089] Nme单引导RNA:紫色

[0090] SapI限制性位点:粗体

[0091] 图8呈现了如图4所示的Nme1Cas9和启动子的颜色编码序列的一个实施方式,其中起始和终止密码子用粗体加下划线。

[0092] U1a启动子:蓝色

[0093] Kozak序列:灰色

[0094] 人性化的Nme1Cas9:红色

[0095] SV40 NLS:绿色

[0096] 核纤溶蛋白(NP)NLS:黄色

[0097] HA标签(3倍):橙红色

[0098] 合成NLS:蓝绿色

[0099] β-球蛋白聚腺苷酸化信号:青绿色

[0100] 图9显示了示例性数据,显示了使用带有sgRNA-Nme1Cas9构建体的AAV质粒引导Pcsk9基因或Rosa26基因(对照)的各种靶位点的编辑效率。

[0101] 图10显示了针对Pcsk9基因或Rosa26基因(对照)的sgRNA-Nme1Cas9构建体的颜色编码的靶位点序列的一个实施方式。

[0102] 24-nt Nme1Cas9靶间隔子,蓝色加粗

[0103] Nme1Cas9 PAM带有下划线(NNNNGATT)

[0104] T7E1引物结合位点:绿色斜体

[0105] TIDE引物结合位点:紫色斜体

[0106] 图11显示了示例性数据,其显示了通过小鼠尾静脉体内流体动力注射30μg靶向Pcsk9的无内毒素的sgRNA-Nme1Cas9-AAV质粒后的基因编辑效率。

[0107] 图12A呈现了示例性数据,其显示了包装在肝细胞特异性AAV8血清型中的Nme1-Cas9载体在肝细胞中对Pcsk9基因和Rosa26基因的基因编辑效率,每剂量4×10^{11}个基因组拷贝(gc)载体施用后14天的小鼠。图12B显示了示例性数据,显示了包装在肝细胞特异性AAV8血清型中的Nme1-Cas9载体在Pcsk9基因和Rosa26基因在肝脏中的基因编辑效率,每只小鼠50天剂量为4×10^{11}个基因组拷贝(gc)后载体管理。

[0108] 图13显示了示例性数据,其显示了在0、25和50天注射靶向Pcsk9基因,Rosa26基因和PBS对照组的sgRNA-Cas9-AAV载体后,小鼠胆固醇水平降低。

[0109] 图14A和14B显示了示例性数据,该数据显示了通过测序(例如GUIDE-Seq)测定实现的双链断裂(DSB)全基因组范围无偏鉴定,该测定为Pcsk9-sgRNA-Cas9-AAV(A)和Rosa26-sgRNA-Cas9-AAV(B)。

[0110] 图15呈现了示例性数据,其显示了在注射Pcsk9-sgRNA-Cas9-AAV和Rosa26-sgRNA-Cas9-AAV两者后14天的小鼠中的靶向TIDE分析,其揭示了最小的切割。OnT,中靶位点;OT1,OT2等:脱靶位点。

[0111] 图16显示了示例性数据,其显示了在注射靶向Pcsk9基因和Rosa26基因的载体后第14天处死的小鼠的肝脏切片中的苏木精和曙红染色测定。没有观察到宿主免疫反应的证据。

[0112] 图17示出了体外PAM库识别工作流程的一个实施方式。NGS,下一代测序。

[0113] 图18显示了来自图17中所述体外PAM发现测定的推定序列。将来自每种细菌的重

组纯化的Cas9与sgRNA和具有随机PAM的靶一起温育。Nme1Cas9用作对照。

[0114] 图19提供了示例性数据，其显示了HEK293T细胞中人基因组中单个位点(上图)的基因组编辑百分率。百分率显示基于整合到HEK293T细胞基因组中的“交通灯”报告子，使用T7E1核酸内切酶测定(Nme2Cas9，HpaCas9)或荧光测定(对于SmuCas9)估计的插入删除形成。

[0115] 图20呈现了示例性数据，其显示了在具有Nme2Cas9的整合交通灯报告子的HEK293T细胞中的基因组编辑，所述Nme2Cas9靶向具有各种PAM的各种原间隔子(X轴)。结果表明人细胞中Nme2Cas9的优选NNNNCC PAM。

[0116] 图21呈现了示例性数据，其显示了在存在各种抗CRISPR(Acr)蛋白的情况下HEK293T细胞中的基因组编辑。T7E1消化显示质粒转染(表达Nme2Cas9及其sgRNA)或RNA/蛋白递送(HpaCas9及其sgRNA)后的基因组编辑。Nme2Cas9被2个Acr蛋白(AcrIIC3$_{Nme}$和AcrIIC4$_{Hpa}$)强烈抑制，而HpaCas9被4个先前报道的II-C型Acr抑制。这些结果表明，这2个Cas9蛋白受反CRISPR的开关控制。

[0117] 图22展示了使用“CC”二核苷酸PAM上的Nme2Cas9-sgRNA复合体编辑交通灯报告子(TLR)基因的示例性数据。图22A。蓝色条表示显示荧光的细胞百分率，而红色条表示基于测序(“TIDE分析”)更准确地进行编辑的百分率。

[0118] 图23显示了在AAVS1，染色体14NTS4，VEGF和CFTR基因座处使用T7E1测定法通过Nme2Cas9进行基因编辑的示例性数据。

[0119] 图24给出了野生型Nme2Cas9细菌开放阅读框DNA序列的一个实施方式。

[0120] 图25显示了野生型Nme2Cas9细菌蛋白序列的一个实施方式。

[0121] 图26显示了Nme2Cas9人密码子优化的开放阅读框DNA序列的一个实施方式。黄色SV40 NLS;绿色3X-HA-标签;蓝色:类似于cMyc的NLS。

[0122] 图27显示了Nme2Cas9人源化蛋白序列的一个实施方式。黄色SV40 NLS;绿色3X-HA-标签;蓝色;类似于cMyc的NLS。

[0123] 图28显示了HpaCas9细菌蛋白序列的一个实施方式。

[0124] 图29显示了SmuCas9天然细菌开放阅读框DNA序列的一个实施方式。

[0125] 图30显示了SmuCas9细菌蛋白序列的一个实施方式。

[0126] 图31显示了SmuCas9人密码子优化的开放阅读框DNA序列的一个实施方式。黄色SV40 NLS;绿色3X-HA-标签;蓝色:类似于cMyc的NLS。

[0127] 图32显示了SmuCas9人源化蛋白序列的一个实施方式。黄色SV40 NLS;绿色3X-HA-标签;蓝色:类似于cMyc的NLS。

[0128] 图33显示了与短的富含C的PAM兼容的示例性II-C型Cas9直系同原物单引导RNA序列。黄色-crRNA;灰色-接头;紫色-tracrRNA。

[0129] 图34说明了3个密切相关的脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9直系同源物，它们具有不同的PAM。

[0130] 图34A:显示Nme2Cas9(左)和Nme3Cas9(右)之间的突变残基(橙色球)的示意图，映射到Nme1Cas9的预测结构上，揭示了PID中的突变簇(黑色)。

[0131] 图34B:体外PAM发现测定的实验工作流程，其中原间隔子下游具有10nt的随机PAM序列。将适配体连接至裂解产物并测序。

[0132] 图34C:体外PAM发现测定的序列标志,证明了Nme1Cas9的N_4GATT PAM,如先前在细胞中所示。

[0133] 图34D:序列标识显示Nme1Cas9(其PID与Nme2Cas9(左)和Nme3Cas9(右)的PID互换)在位置5处识别为C。由于蛋白嵌合体的适度切割效率,其余核苷酸的可信度较低(图35C)。

[0134] 图34E:序列标识,说明全长Nme2Cas9基于PAM发现测定法识别N_4CC PAM,在位置5具有固定的C,并且对PAM nts 1-4和6-8进行了随机分配。

[0135] 图35显示了根据图34具有快速发展的PID的脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9直系同源物的特征。

[0136] 图35A:NmeCas9直系同源物的无根系统发育树,与Nme1Cas9>80%相同。出现了3个不同的分支,大多数突变聚集在PID中。与Nme1Cas9的同源性>98%的第1组(蓝色)PID,与Nme1Cas9的约52%相同的PID的第2组(橙色),与Nme1Cas9的约86%相同的PID的第3组(绿色)。每个组的3个代表性Cas9直系同源物(Nme1Cas9,Nme2Cas9和Nme3Cas9)被标记。

[0137] 图35B:显示编码来自(A)的3个Cas9直系同原物(Nme1Cas9,Nme2Cas9和Nme3Cas9)的菌株的CRISPR位点的示意图,每个CRISPR-Cas组分与脑膜炎奈瑟氏球菌(Neisseria meningitidis)8013(编码Nme1Cas9)的百分率同一性是如图所示。

[0138] 图35C:从体外测定的完整Nme1Cas9以及与Nme1Cas9的PID交换为Nme2Cas9和Nme3Cas9的嵌合体的嵌合体中裂解的DNA读长的数目。减少的阅读计数表明嵌合体中较低的切割效率。

[0139] 图35D;来自Nme1Cas9在NNNNCNNN随机PAM上进行体外PAM发现测定的序列标识,其PID与Nme2Cas9(左)或Nme3Cas9(右)的PID互换。

[0140] 图36显示,Nme2Cas9使用22-24个核苷酸的间隔子来识别和编辑与N_4CC PAM相邻的位点。所有实验均一式三份进行,误差条代表平均值的标准误差(sem)。

[0141] 图36A:显示HEK293T TLR2.0细胞上的瞬时转染工作流程的示意图。转染Nme2Cas9和sgRNA质粒,转染72小时后检测到mCherry+细胞。

[0142] 图36B:使用Nme2Cas9定位TLR2.0中的PAM阵列。具有N_4CC PAM的所有位点均以不同程度的效率靶向,而在N_4GATT PAM或不存在sgRNA时未观察到Nme2Cas9靶向。SpyCas9(靶向NGG)和Nme1Cas9(靶向N_4GATT)用作阳性对照。

[0143] 图36C:间隔子长度对Nme2Cas9编辑效率的影响。靶向TLR2.0位点(带有N_4CCA PAM)的sgRNA,间隔子长度从24到20nts不等(包括5'端G),显示了22-24个核苷酸间隔子的最高编辑效率。

[0144] 图36D:可以串联使用^{Nme2Cas9}切口酶(HNH切口酶=Nme2Cas9^{D16A};RuvC切口酶=Nme2Cas9^{H588A})以在TLR2.0中产生插入删除。使用任一种切口酶靶向切割位点分别为32个碱基对和64个碱基对的靶以产生插入删除。HNH切口酶显示有效的编辑,尤其是当切割位点接近(32bp)时。野生型Nme2Cas9用作对照。绿色是GFP(HDR),红色是mCherry(NHEJ)。

[0145] 根据图36,图37展示了针对哺乳动物细胞中Nme2Cas9靶向的PAM,间隔子和种子元素的示例性数据。所有实验均一式三份进行,误差条代表sem。

[0146] 图37A:Nme2Cas9靶向TLR2.0中的N_4CD位点。检查每个非C核苷酸在测试位置(N_4CA,N_4CT和N_4CG)的4个位点,并将N_4CC位点用作阳性对照。

[0147] 图37B:Nme2Cas9靶向TLR2.0中N_4DC位点[类似于(A)]。

[0148] 图37C:在另一个TLR2.0位点上的引导截短,显示出与图36C中观察到的长度要求相似的长度要求。

[0149] 图37D:Nme2Cas9靶向效率对种子序列中的单核苷酸错配敏感。数据显示,沿着TLR靶位点中23-nt间隔子区的sgRNA中行走的单核苷酸错配的影响。

[0150] 图38显示了示例性数据,其显示了通过多种递送方法在哺乳动物细胞中的基因组位点处的Nme2Cas9基因组编辑效率。所有结果代表3个独立的生物学重复,误差条代表sem。

[0151] 图38A:使用瞬时转染,用靶向人HEK293T细胞中整个基因组的基因座的sgRNA进行Nme2Cas9基因组编辑,基于38个位点的初步筛选选择了14个位点,以证明不同位点的插入/缺失范围(通过TIDE检测)由Nme2Cas9诱导。Nme1Cas9靶位点(具有N_4GATT PAM)用作阴性对照。

[0152] 图38B:左图:通过TIDE检测到的针对Hepa1-6小鼠细胞中Pcsk9和Rosa26基因座的多合一质粒(Nme2Cas9+sgRNA)的瞬时转染。右图:将sgRNA质粒电穿孔到稳定表达来自慢病毒载体的Nme2Cas9的K562细胞中,可在预期的基因座上形成有效的插入删除。

[0153] 图38C:Nme2Cas9可以作为RNP复合物进行电穿孔,以进行有效的基因组编辑。将40皮摩尔的Cas9以及50皮摩尔的靶向3个不同基因座的体外转录sgRNAs电穿孔到HEK293T细胞中。72小时后使用TIDE测量插入删除。

[0154] 图39呈现了示例性数据,其显示了根据图38的Nme2Cas9对剂量的依赖性和阻断缺失。

[0155] 图39A:增加电穿孔的Nme2Cas9质粒的剂量(500ng,图3A中为200ng)提高了2个位点(TS16和TS6)的编辑效率。

[0156] 图39B:Nme2Cas9可用于创建区块删除。与Nme2Cas9同时靶向2个裂解位点相距32bp的TLR2.0靶。产生的大多数病变恰好是32bp缺失(绿色)。

[0157] 图40呈现示例性数据,其显示II-C型抗CRISPR蛋白可用于在体外和体内抑制Nme2Cas9基因编辑活性(例如,作为开关)。所有实验均一式三份进行,误差条代表sem。

[0158] 图40A:在五个先前表征的抗CRISPR蛋白(Acr:Cas9的比例为10:1)的存在下,Nme1Cas9和Nme2Cas9的体外切割测定。上图:在没有Acr的情况下或在存在对照Acr(AcrE2)的情况下,Nme1Cas9可以有效地裂解含有带有N_4GATT PAM的原间隔子的片段。如预期的那样,所有其他先前表征的Acrs抑制Nme1Cas9。下图:Nme2Cas9有效地切割含有protospacer具有N的靶$_4$在AcrE2和AcrIIC5的存在CC PAM_{Smu},表明$AcrIIC5_{Smu}$不能抑制Nme2Cas9以10:1的摩尔比。

[0159] 图40B:存在五个先前描述的抗CRISPR蛋白时的基因组编辑。将表达Nme2Cas9,sgRNA和各个Acr的质粒(200ng Cas9、100ng sgRNA,200ng Acr)共转染到HEK293T细胞中,并在转染后72小时使用TIDE测量基因组编辑。除了AcrE2和$AcrIIC5_{Smu}$,其他所有Acrs都抑制基因组编辑,尽管效率不同。

[0160] 图40C:对Nme2Cas9的Acr抑制是剂量依赖性的,具有明显的表观效力。$AcrIIC1_{Nme}$和$AcrIIC4_{Hpa}$分别以共转染质粒的2:1和1:1比例完全抑制Nme2Cas9。

[0161] 图41显示的示例性数据显示,Nme2Cas9 PID交换使Nme1Cas9对$AcrIIC5_{Smu}$抑制不敏感,根据图40。Nme1Cas9-Nme2Cas9PID嵌合体的体外切割是在先前鉴定的Acr蛋白(10μM

Cas9-sgRNA+100μM Acr)存在下进行的。

[0162] 图42呈现了示例性数据，其显示了Nme2Cas9在哺乳动物细胞中没有可检测的脱靶。

[0163] 图42A：示意图，显示了由于SpyCas9和Nme2Cas9由于其非重叠的PAM而可靶向的双位点(DS)。Nme2Cas9 PAM(橙色)和SpyCas9PAM(蓝色)突出显示。

[0164] 图42B：Nme2Cas9和SpyCas9在双位点诱导插入删除。选择带有GN_3GN_{19}NGGNCC序列的VEGFA中的六个双位点用于2个直系同源物之间的直接比较。在HEK293T细胞中，将表达每个Cas9的质粒(具有相同的启动子和NLS)与每个直系同源物的同源引导一起转染。转染后72小时通过TIDE确定插入删除率。在所有六个位点上都可以检测到Nme2Cas9编辑，并且在2个位点(DS2和6)上比SpyCas9更有效。SpyCas9编辑了六个位点中的4个(DS1、2、4和6)，其中2个位点的编辑率明显高于Nme2Cas9(DS1和4)。选择DS2、4和6进行GUIDE-Seq分析，因为在这些位置，Nme2Cas9分别比SpyCas9效率更高，效率更低。

[0165] 图42C：Nme2Cas9在人细胞中具有干净的脱靶概况。显示了由GUIDE-Seq在单个靶位点针对每种核酸酶检测到的离靶位点数。SpyCas9离靶编号显示为黑色。除了双位点以外，TS6(由于其高效性和潜在的脱靶位点)和2个鼠标位点(以测试另一种细胞类型的准确性)也每个引导显示零个或一个脱靶位点。

[0166] 图42D：靶向深度测序证实了GUIDE-seq指示的高Nme2Cas9准确性。通过GUIDE-seq检测到的最高脱靶基因座被扩增并进行了深度测序。SpyCas9最多显示出脱靶位，而对于Nme2Cas9，只有一个(Rosa26位点)显示出脱靶基因座的编辑水平相对较低(脱靶为~40%对比脱靶为~1%)。注意y轴上的对数刻度。

[0167] 图42E：Nme2Cas9和SpyCas9效率因基因座和靶位点而异。选择整个基因组中的位点(具有GN_3GN_{19}NGGNCC序列)以直接比较2个直系同源物的编辑。将表达每个Cas9的质粒(具有相同的启动子，接头，标签和NLS)和其同源引导转染到HEK293T细胞中。转染后72小时通过TIDE测定插入/缺失效率。箱线图显示了Nme2Cas9和SpyCas9在左二十八(28)个双位点的编辑效率。分析中未显示未编辑的位点。Nme2Cas9和SpyCas9的相对效率表明，Nme2Cas9的效率平均低于SpyCas9(右)。在右侧面板中，相对于所有效率的分析都包括了在所有二十八(28)个位点上通过2个Cas9直系同源物编辑的效率。

[0168] 图42F显示了经验证的Rosa26指南的脱靶位点的核酸序列，显示了PAM区(带下划线)，共有CC PAM二核苷酸(加粗)和PAM远端的3个错配间隔子的一部分(红色)。

[0169] 图43显示了示例性数据，该数据显示了根据图42在2个靶位点处Nme2Cas9和SpyCas9的正交性和相对准确性。

[0170] 图43A：Nme2Cas9和SpyCas9引导是正交的。TIDE结果显示了由靶向DS12的两种核酸酶与其同源sgRNA或其他直系同源物的sgRNA产生的插入删除的频率。

[0171] 图43B：Nme2Cas9和SpyCas9在GUIDE-seq期间具有可比的按靶编辑效率。条形图表示每个直系同源物靶向的3个双位点从GUIDE-Seq进行的靶读长计数。橙色条代表Nme2Cas9，黑色条代表SpyCas9。

[0172] 图43C：每个位点的SpyCas9的中靶读长与脱靶读长。橙色条表示中靶的读长，而黑色条表示脱靶的读长。

[0173] 图43D：每个位点的Nme2Cas9的中靶读长与脱靶读长。

[0174] 图43E:条形图,显示基于CRISPRseek的预期脱靶位点的TIDE,未检测到脱靶位点的插入删除。

[0175] 图44呈现示例性数据,其显示了通过多合一AAV递送在体内编辑Nme2Cas9基因组。

[0176] 图44A:递送AAV8.Nme2Cas9+sgRNA的工作流程,以通过靶向Pcsk9降低小鼠的胆固醇水平。上图:表达Nme2Cas9和sgRNA的多合一AAV载体示意图。下图:尾静脉注射AAV8.Nme2Cas9+sgRNA的时间表,然后在第14天进行胆固醇测量,在第28天进行插入删除,组织学和胆固醇分析。

[0177] 图44B:深度测序分析,以测量从注射有靶向Pcsk9和Rosa26(对照)基因座的AAV8.Nme2Cas9+sgRNA的小鼠肝脏中提取的DNA插入删除。

[0178] 图44C:与靶向Rosa26的对照组相比,注射靶向Pcsk9的指南的小鼠血清胆固醇水平降低。通过不成对的T检验计算P值。

[0179] 图44D:注射有AAV8.Nme2Cas9+sgRosa26(左)或AAV8.Nme2Cas9+sgPcsk9(右)的小鼠肝脏的H&E染色。比例尺,25μm。

[0180] 图45示出了最小化的AAV主链和与常规大小的AAV主链的示例性比较TLR 2.0数据的一个实施方式。

[0181] 图46显示了sgRNA 11截短和sgRNA 12截短的Nme2Cas9结构的比较。

[0182] 图47示出了具有短多聚A信号的最小化多合一AAV的一个实施方式。

[0183] 图48示出了最小化的多合一AAV主链的2个实施方式。串联的双重sgRNA(上图)。同源性定向修复的供体模板(下图)。

[0184] 图49显示了多合一AAV-sgRNA-hNme1Cas9构建体的验证。

[0185] 图49A:表达人密码子优化的Nme1Cas9及其sgRNA的单个rAAV载体的示意图。主链侧接AAV反向末端重复序列(ITR)。多聚(a)信号来自兔β-珠蛋白(BGH)。

[0186] 图49B:Pcsk9(上图)和Rosa26(下图)小鼠基因的示意图。红色条代表外显子。放大视图显示了原间隔子符序列(红色),而Nme1Cas9 PAM序列以绿色突出显示。标出了双链断裂位置位点(黑色箭头)。

[0187] 图49C:堆叠的直方图,显示了在靶向Pcsk9(sgPcsk9)和Rosa26(sgRosa26)基因的Hepa1-6细胞中,TIDE在AAV-sgRNA-hNme1Cas9质粒转染后通过TIDE获得的插入-缺失(插入删除)的代表性百分率分布。数据表示为来自3个生物学重复的平均值±SD。

[0188] 图49D:堆叠的直方图,显示了在通过流体动力学注射AAV-sgRNA-hNme1Cas9质粒后,由TIDE获得的C57BL/6小鼠肝脏中Pcsk9的插入删除的代表性百分率分布。

[0189] 图50提供了示例性数据,其显示了许多N_4GN_3 PAM是无活性的,并且没有显示出在小鼠基因组中具有少于4个错配的脱靶位点。

[0190] 图51提供了示例性数据,其显示了Nme1Cas9介导的Hpd的敲除可以挽救I型遗传性酪氨酸血症小鼠的致死表型。

[0191] 图51A:Hpd小鼠基因的示意图。红色条代表外显子。放大视图显示了针对外显子8(sgHpd1)和外显子11(sgHpd2)的原间隔子序列(红色)。Nme1Cas9 PAM序列显示为绿色,并标出了双链断裂位置(黑色箭头)。

[0192] 图51B:实验设计。向三组I型遗传性酪氨酸血症$Fah^{-/-}$小鼠注射PBS或多合一AAV-sgRNA-hNme1Cas9质粒sgHpd1或sgHpd2。

[0193] 图51C:在NTBC撤除后监测了用PBS(绿色),动力注射PBS(绿色),靶向Hpd外显子8的AAV-sgRNA-hNme1Cas9质粒sgHpd1(红色)或靶向sgHpd2的Hpd外显子11(蓝色)的小鼠的体重。误差条代表PBS和sgHpd1组的三只小鼠和sgHpd2组的两只小鼠。数据表示为平均值±SD。

[0194] 图51D:堆叠的直方图,显示了在水力注射PBS或sgHpd1和sgHpd2质粒后,由TIDE获得的$Fah^{-/-}$小鼠肝脏中Hpd的插入删除的代表性百分率分布。NTBC撤药结束时(第43天)收获肝脏。

[0195] 图52呈现了示例性数据,该数据显示了图51中所示指南的平均插入删除效率。

[0196] 图53显示了示例性的组织学显微照片,显示与注射PBS的$Fah^{mut/mut}$小鼠相比,经sgHpd1和sgHpd2处理的小鼠的肝损伤严重程度较轻,与PBS-注射的小鼠相比,多核肝细胞数量较少。

[0197] 图54显示了用于体内基因组编辑的Nme1Cas9的AAV递送。

[0198] 图54A:在C57BL/6小鼠中,AAV8-sgRNA-hNme1Cas9载体尾静脉注射靶向Pcsk9(sgPcsk9)和Rosa26(sgRosa26)的实验概述。在注射后4(n=1)或50天(n=5)处死小鼠,并收集肝组织。在注射后第0、25和50天收集血液血清以测量胆固醇水平。

[0199] 图54B:血清胆固醇水平。p值通过不配对t检验计算。

[0200] 图54C:堆叠的直方图,显示了通过靶向深度测序分析测量的小鼠肝脏中Pcsk9或Rosa26的插入删除的代表性百分率分布。数据表示为每组五只小鼠的平均值±SD。

[0201] 图54D:使用从50天小鼠肝匀浆中收集的总蛋白的代表性抗PCSK9蛋白印迹。总共包含2ng的重组小鼠PCSK9(r-PCSK9)作为迁移标准。星号表示比对照重组蛋白大的交叉反应蛋白。

[0202] 图55提供了示例性数据,其显示注射有AAV8-sgRNA-hNme1Cas9的小鼠产生抗Nme1Cas9抗体。

[0203] 图56呈现了示例性数据,其显示了Nme1Cas9的GUIDE-seq全基因组特异性。数据表示为平均值±SD。

[0204] 图56A:中靶(OnT)和脱靶(OT)位点的GUIDE-seq读长数。

[0205] 图56B:靶向深度测序以测量Hepa1-6细胞中每个OT位点的损伤率。每个OT位点与OnT原间隔子的错配都会突出显示(蓝色)。数据表示为来自3个生物学重复的平均值±SD。

[0206] 图56C:使用从注射了多合一AAV8-sgRNA-hNme1Cas9 sgPcsk9和sgRosa26并在第14天(D14)处死的小鼠获得的基因组DNA进行靶向深度测序,以测量每个OT部位的病变率注射后第50天(D50)。

[0207] 图57显示了Nme2Cas9在小鼠合子中离体编辑酪氨酸酶(Tyr)基因的示例性数据,与图58相关。

[0208] 图57A:测试了Tyr基因中的2个位点,每个位点具有N_4CC PAM,在Hepa1-6细胞中进行编辑。sgTyr2指南显示出更高的编辑效率,因此被选择进行进一步测试。

[0209] 图57B:通过TIDE测定,7只小鼠在出生后发育中存活,并且每只表现出毛色表型以及按靶编辑。

[0210] 图57C:如TIDE分析所示,来自图57B的每只小鼠和未编辑的C57BL/6NJ小鼠的尾部DNA的插入删除特征。指出了各种大小的插入(阳性)和缺失(阴性)的效率。

[0211] 图58显示了使用多合一AAV递送的离体Nme2Cas9基因组编辑的示例性数据。

[0212] 图58A:通过靶向Tyr基因离体编辑以产生白化病C57BL/6NJ小鼠的单AAV Nme2Cas9的工作流程。合子在含有AAV6.Nme2Cas9:sgTyr的KSOM中培养5-6小时,在M2中漂洗,并培养一天,然后转移到假妊娠输卵管。

[0213] 图58B:由3×10^{9}个GC产生的白化病小鼠(左)和毛丝鼠或杂种小鼠(中),由AAV6.Nme2Cas9:sgTyr的合子由3×10^{8}个GC产生的毛丝鼠或杂种小鼠(右)。

[0214] 图58C:两种AAV剂量的Nme2Cas9.sgTyr单AAV离体Tyr编辑实验的总结。

[0215] 图59显示了Nme1Cas9和Nme2Cas9核苷酸序列的比对。图例:非PID aa差异(蓝绿色阴影);PID aa差异(黄色阴影);活性位点残基(红色字母)。

[0216] 图60显示了Nme1Cas9和Nme3Cas9核苷酸序列的比对。图例:非PID aa差异(蓝绿色阴影);PID aa差异(黄色阴影);活性位点残基(红色字母)。

[0217] 图61显示了Nme2Cas9氨基酸序列的一个实施方式。图例:SV40 NLS(黄色阴影);3X-HA-标签(绿色阴影);类似于cMyc的NLS(蓝绿色底纹);接头(紫色底纹)。

[0218] 图62显示了Nme2Cas9氨基酸序列的一个实施方式。图例:SV40 NLS(黄色阴影);3X-HA-标签(绿色阴影);类核于核质蛋白的NLS(红色阴影);c-myc NLS(蓝绿色底纹);接头(紫色底纹)。

[0219] 图63显示了重组Nme2Cas9(rNme2Cas9)氨基酸序列的一个实施方式。图例:SV40 NLS(黄色阴影);类核于核质蛋白的NLS(红色阴影);接头(紫色底纹)。

[0220] 图64显示了多合一AAV-sgRNA-hNmeCas9质粒核苷酸序列的一个实施方式。图例:sgRNA支架(棕色字母);GUIDE序列(黑色字母);U6启动子(蓝色字母);U1a启动子(紫色字母):NLS NLS(绿色字母);hNmeCas9(红色字母);NLS 3X-HA和NLS BGH-pA(绿色/黑色字母交替)。

[0221] 【发明详述】

[0222] 本发明涉及用于基因治疗的组合物和方法。本文所述几种方法利用脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9系统,其提供了超精确的CRISPR基因编辑平台。此外,本发明结合了该Cas9系统的改进:例如,截短单引导RNA序列,以及将-Nme1Cas9或Nme2Cas9与其引导RNA一起包装在与体内施用相容的腺相关病毒载体中。此外,已经鉴定出II-C型Cas9直系同原物,其靶向原间隔子相邻基序序列限于1～4个所需核苷酸之间。

[0223] 【I.脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9(Nme1Cas9)/CRISPR基因编辑准确性】

[0224] 以前,报道了称为脑膜炎奈瑟氏球菌(Neisseria meningitidis)Cas9(Nme1Cas9)的II-C型CRISPR-Cas9系统的超高精度版本。除了非常精确外,Nme1Cas9还比广泛使用的酿脓链球菌(S.pyogenes)Cas9(SpyCas9)小,从而可以通过基于病毒和信使RNA(mRNA)的方法更轻松地递送Nme1Cas9。Nme1Cas9基因组编辑通常使用质粒转染完成。Zhang et al.,"Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis"Mol Cell 50:488-503(2013);Hou et al.,"Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis"Procd Natl Acad Sci U.S.A.110:15644-15649(2013);Esvelt et al.,"Orthogonal Cas9 proteins for RNA-guided gene regulation and editing"Nature

Methods 10:1116-1121(2013);Zhang et al.,"DNase H activity of Neisseria meningitidis Cas9"Mol Cell 60:242-255(2015);Lee et al.,"The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells"Molecular Therapy 24:645-654(2016);Pawluk et al.,"Naturally occurring off-switches for CRISPR-Cas9"Cell 167:1829-1838(2016);and Amrani et al.,"Nme1Cas9 is an intrinsically high-fidelity genome editing platform"biorxiv.org/content/early/2017/08/04/172650(2017)。

[0225] 然而,尚未广泛探索基于Nme1Cas9病毒,RNA和核糖核蛋白(RNP)的递送。与其他递送方法相比,用于基因组工程的Cas9直系同原物基于RNA和RNP的递送具有多个优势。由于它们绕过了与基于DNA的Cas9及其sgRNA递送有关的表达问题,因此它们不仅可以加快编辑速度,而且还可以减少与基于Cas9的编辑相关的脱靶效应。脱靶活性的降低归因于对Cas9 RNA和RNP浓度的更好控制,以及细胞中相对较快的Cas9 RNA和RNP降解。已经证明细胞中活性Cas9的延长存在与更高的脱靶效应有关。由于Cas9 RNA和RNP在细胞内降解更快,因此作为RNA或RNP递送的Cas9不会长时间持续存在,因此降低了脱靶效应。

[0226] 常规使用的全长145nt Nme1Cas9sgRNA包括48个核苷酸(nt)crRNA,4nt接头和93nt tracrRNA。sgRNA的crRNA区由第一个24nt间隔子区序列和第二个24nt重复序列组成,第二个24nt重复序列与24nt tracrRNA反重复5'区配对,从而形成重复:反重复区。其余的69nt tracrRNA区包括茎1区和茎2区。图1。

[0227] 该全长Nme1Cas9 sgRNA已成功地用于基于质粒的方法中的基因组编辑。此外,可以将体外转录的Nme1Cas9 sgRNA与纯化的Nme1Cas9复合,并用于人细胞中的基因组编辑。尽管在体外转录的sgRNA中成功完成了人细胞的基因组编辑,但在难以转染的人细胞系(如PLB985)中,Nme1Cas9 RNP的编辑效率却降低了。

[0228] 以前已经显示,Cas9 RNP的编辑效率与其sgRNA的化学稳定性成正比。尽管不必了解本发明的机制,但据信采用了几种细胞机制来快速降解RNA。因此,Cas9 sgRNA通常通过化学方法进行修饰。赋予sgRNA更高的稳定性的一些化学修饰包括但不限于核糖2'-O-甲基化和/或硫代磷酸酯键。尽管化学修饰的RNA是通过Cas9 RNP改进基因组编辑的选择,但其有效性受到以下事实的限制:随着RNA长度的增加,RNA的化学合成变得越来越困难且昂贵。在145nt,Nme1Cas9sgRNA的合成对于采用化学合成sgRNA的常规基因组编辑应用来说是遥不可及的。

[0229] 【II.截短的Nme1Cas9 sgRNA序列】

[0230] 由于上述确定的局限性,即全长145nt Nme1Cas9 sgRNA太大而无法常规化学合成sgRNA进行基因组编辑,因此,本发明的一个实施方式考虑了截短的Nme1Cas9 sgRNA。尽管没有必要理解本发明的机制,但是据信截短的Nme1Cas-sgRNA不会损害Nme1Cas9 RNP的功能。此外,Nme1Cas9和Nme2Cas9的sgRNA是相同且可互换的(图35B),因此sgRNA截短同样适用于Nme1Cas9和Nme2Cas9。截短的sgRNA和相关的靶位点的示例性序列在下面公开,其中在引导区中的可变sgRNA nts以"N"残基给出。在靶序列中,由sgRNA引导区识别的24个核苷酸带有下划线,原间隔子相邻基序(PAM)区以粗体显示。表1。

[0231] 【表1:示例性截短的sgRNA序列和相关的基因组靶】

[0232]

描述	序列
wt sgRNA	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCUUUC UCAUUUCGGA AACGAAAUGA GAACCGUUGC UACAAUAAGG CCGUCUGAAA AGAUGUGCCG CAACGCUCUG CCCCUUAAAG CUUCUGCUUU AAGGGGCAUC GUUUA
sgRNA #1	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCCC UUAAAGCUUC UGCUUUAAGG GGCAUCGUUU A
sgRNA #2	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCCC UUUUCUAAGG GGCAUCGUUU A
sgRNA #3	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCCC UUUUCUAAGG GGCAU
sgRNA #4	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCUUC UGCAUCGUU
sgRNA #5	NNNNNNNNNN NNNNNNNNNN NNNGUUGUAG CUCCCGAAAC GUUGCUACAA UAAGGCCGUC UGAAAAGAUG UGCCGCAACG CUCUGCUUCU GCAUCGUUUA
sgRNA #6	NNNNNNNNNN NNNNNNNNNN NGUUGUAGCU CCCGAAACGU UGCUACAAUA AGGCCGUCUG AAAAGAUGUG CCGCAACGCU CUGCCCUUCU GGGCAUCGUU
sgRNA #7	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCCC UUUCUAGGGG CAUCGUU
sgRNA #8	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCCC UUCUGGGGCA UCGUU
sgRNA #9	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCCU UCUGGGCAUC GUU
sgRNA #10	NNNNNNNNNN NNNNNNNNNN NNNNGUUGUA GCUCCCGAAA CGUUGCUACA AUAAGGCCGU CUGAAAAGAU GUGCCGCAAC GCUCUGCCUU CUGGCAUCGU U
sgRNA #11	NNNNNNNNNN NNNNNNNNNN NNNGUUGUAG CUCCCGAAAC GUUGCUACAA

[0233]

	UAAGGCCGUC UGAAAAGAUG UGCCGCAACG CUCUGCCUUC UGGCAUCGUU
N-TS7 间隔子（24nt）	GAGGGAGAGA GGUGAGCGGA UGAA
N-TS7 间隔子（23nt）	GGGGAGAGAG GUGAGCGGAU GAA
N-TS27 间隔子（24nt）	GUUCUCCAAG CCCUCGGACC UCGU
N-TS27 间隔子（23nt）	GUCUCCAAGC CCUCGGACCU CGU
N-TS55 间隔子（24nt）	GCUGGAUUAC UGUGUGGUAG AGGG
N-TS55 间隔子（23nt）	GUGGAUUACU GUGUGGUAGA GGG
N-TS7 基因组靶位点	AGCTTGAGCA AAGGGAGAGA GGTGAGCGGA TGAA**GGGAGA TT**GGTGAGTA TC
N-TS27 基因组靶位点	CGCTTCGCGG CTTCTCCAAG CCCTCGGACC TCGT**GGGCGT CT**TCTCCTGC GT
N-TS55 基因组靶位点	GAATTCACTA GCTGGATTAC TGTGTGGTAG AGGG**AGGTGA TT**AGCACCTG TG

[0234] 如本文所预期的，截短的Nme1Cas9sgRNA将不仅允许以合理的成本进行合成，而且还促进了基于病毒的递送方法(例如，腺相关病毒递送平台)的使用，其中基于DNA的允许长度为有限。在一个实施方式中，截短的sgRNA降低脱靶的Nme1Cas9编辑作用。在一个实施方

式中，截短的Nme1Cas9 sgRNA包含至少一种化学修饰，其增加Nme1Cas9编辑效率。

[0235] 如上所述，Nme1Cas9的全长145nt的sgRNA包括引导区，重复:反重复双链体区，茎1区和茎2区。图1。但是，由于sgRNA的长度对于常规基因组编辑是有问题的，因此非常需要开发用于Nme1Cas9的截短的sgRNA。当前，可商购的RNA合成方法要求RNA终产物不超过~100nt。

[0236] 在一个实施方式中，本发明涉及一种Nme1Cas9sgRNA，其包含截短的重复:反重复双链体。在一个实施方式中，本发明涉及包含截短的茎2的Nme1Cas9 sgRNA。图2。此外，先前已显示Nme1Cas9的5'可变引导crRNA区(例如间隔子区)也可以被截短几个核苷酸而不会失去功能。Amrani et al.,"Nme1Cas9 is an intrinsically high-fidelity genome editing platform"biorxiv.org/content/early/2017/08/04/172650(2017);和Lee et al.,"The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells"Molecular Therapy 24:645-654(2016)。

[0237] 在一个实施方式中，本发明涉及一种100nt Nme1Cas9截短的sgRNA。图3，构建体#11。通过在HEK293T细胞中进行瞬时转染，在3个不同的人基因组位点上测试了这种100nt Nme1Cas9截短的sgRNA构建体#11，在这3个位点上它们均以与全长Nme1Cas9 sgRNA相同甚至更高的水平支持Nme1Cas9功能。图3，下部图。此外，还使用RNP递送在多个基因组靶位点测试了sgRNA 11和sgRNA 13，编辑效率与wt sgRNA相似或更高。图4。还在PLB985细胞中测试了构建体#11的合成形式，相对于体外转录的wt sgRNA，其编辑效率更高。图5。

[0238] 【III.伴随腺病毒CRISPR递送平台】

[0239] 与转录激活物样效应子核酸酶(TALENs)和锌指核酸酶(ZFNs)相比，Cas9s以其灵活性和多功能性而著称。Komor et al.,"CRISPR-based technologies for the manipulation of eukaryotic genomes"Cell 2017;168:20-36。这些特性使其成为推动基因组工程领域前进的理想选择。在过去的几年中，除了在基因疗法和个性化医学中的有希望的应用之外，CRISPR-Cas9还被用于增强农业，食品和工业领域的产品。Barrangou et al.,"Applications of CRISPR technologies in research and beyond"Nat Biotechnol.2016;34:933-41。尽管已描述了2类CRISPR系统的多样性，但仅开发了其中几个系统并对其进行了体内基因组编辑的验证。NmeCas9是一种紧凑的高保真Cas9，可以考虑将其用于将来使用多合一rAAV的体内基因组编辑应用程序。NmeCas9独特的PAM能够在其他靶进行编辑，而其他2个紧凑的一体式rAAV验证直系同原物(SauCas9和CjeCas9)则无法访问这些靶。

[0240] 使用细菌CRISPR系统进行基因组编辑开辟了人基因治疗的新途径。CRISPR复合物以捕获细菌中入侵核酸片段的成簇规则间隔子短回文重复序列命名，包含引导核酸酶Cas9(CRISPR相关蛋白9)切割互补双链DNA的引导RNA(例如sgRNA)。Cas9诱导的DNA断裂的非同源修复导致灭活靶基因的小插入或缺失(插入删除)，但也可以通过同源DNA模板修复断裂，从而导致基因替换。Nelson et al.,"In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy"Science 351:403-407(2016);和Ran et al.,"In vivo genome editing using Staphylococcus aureus Cas9"Nature 520:186-191(2015);和Yin et al.,"Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype"Nature Biotechnology 32:551-553

(2014)。

[0241] 作为一种灵活的基因组编辑工具,目前广泛使用的II-A型酿脓链球菌(S.pyogenes)(Spy)Cas9表现出几个缺点:(i)递送效率低;(ii)脱靶裂解;(iii)不受管制的活动。这些缺点严格限制了SpyCas9作为潜在的基因治疗工具。如本文所述,高度准确和精确的Nme1Cas9或Nme2Cas9复合物可以克服这些SpyCas9限制。

[0242] 本文已显示Nme1Cas9和Nme2Cas9是哺乳动物细胞中的有效基因组编辑平台,并且作为比SpyCas9更小的蛋白,更容易工程化用于体内递送的病毒载体。此外,与SpyCas9相比,Nme1Cas9和Nme2Cas9的脱靶编辑显著降低,并且已经确定了可以控制Nme1Cas9和Nme2Cas9活性的抗CRISPR蛋白。Esvelt et al.,"Orthogonal Cas9 proteins for RNA-guided gene regulation and editing"Nature Methods 10:1116-1121(2013);Amrani et al.,"Nme1Cas9 is an intrinsically high-fidelity genome editing platform" biorxiv.org/content/early/2017/08/04/172650(2017);Lee et al.,"The Neisseria meningitidis CRISPR-Cas9 System Enables Specific Genome Editing in Mammalian Cells"Molecular Therapy 24:645-654(2016);Hou et al.,"Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis"Procd Natl Acad.Sci USA 110:15644-15649(2013);和Pawluk et al.," Naturally Occurring Off-Switches for CRISPR-Cas9"Cell 167:1829-38e9(2016);和图21。

[0243] 腺相关病毒(AAV)已被证明是在临床前和临床环境中具有最小致病性的递送穿梭机,但其包装能力有限。Nme1Cas9由一个约3.3kb的开放阅读框编码,其引导RNA在AAV的包装范围内。Nme2Cas9具有类似的优势。与需要通过sgRNA和Cas9的单独载体递送的SpyCas9不同,Nme1Cas9,Nme2Cas9及其sgRNA足够小,可以通过单个AAV载体递送。

[0244] AAV已经成功地在体内递送了其他Cas9直系同原物,例如空肠弯曲杆菌Cas9(CjeCas9)和金黄色葡萄球菌(S.aureus)(SauCas9)。Kim et al.,"In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni Nat Commun 8:14500(2017);和Ran et al.,"In vivo genome editing using Staphylococcus aureus Cas9"Nature 520:186-191(2015)。Nme1Cas9通常与N_4GATT PAM相关,这不同于CjeCas9 PAM(例如N_4RYAC)或SauCas9 PAM(例如NNGRRT)(R=嘌呤(A或G),Y=嘧啶(C或T))。

[0245] Nme1Cas9已作为核糖核蛋白(RNP)复合体成功交付到人细胞中。图2和图3。此外,本文提供的数据表明,在尾静脉注射后,可以使用多合一sgRNA-Nme1Cas9-AAV载体在小鼠体内将Nme1Cas9核酸序列表达至靶基因。

[0246] 本文提供的数据证明靶向小鼠蛋白原转化酶舒替兰酶/Kexin 9型(Pcsk9)基因。PCSK9充当低密度脂蛋白(LDL)受体的拮抗剂,并限制LDL胆固醇的摄取。血清中胆固醇水平降低的检测可以使用PCSK9引导的Cas9平台直接提供有效Nme1Cas9编辑的功能读长。

[0247] 在一个实施方式中,本发明涉及包含Nme1Cas9-sgRNA复合物或Nme2Cas9-sgRNA复合物的腺相关病毒载体。尽管不必理解本发明的机制,但是据信AAV/Nme1Cas9-sgRNA复合物或AAV/Nme2Cas9-sgRNA复合物与体内递送途径相容以便提供基因编辑。

[0248] 在一个实施方式中,本发明涉及一种sgRNA-Nme1Cas9-AAV载体,其包含sgRNA序

列,RNA聚合酶III U6启动子序列,人密码子优化的Nme1Cas9序列和RNA聚合酶II U1a启动子序列。图6。U1a是普遍存在的启动子,可在各种目的组织中广泛表达Cas9。可以通过使用常规限制性位点(例如Sap1)将与靶基因匹配的间隔子区序列插入sgRNA盒中来靶向待编辑的特定基因。sgRNA-Nme1Cas9-AAV各种元件的代表性序列以彩色注释显示。图7和8。

[0249] 使用Pcsk9-sgRNA-Nme1Cas9-AAV质粒和Rosa26-sgRNA-Nme1Cas9-AAV质粒在几个靶位点的编辑效率通过瞬时转染到小鼠Hepa1-6肝癌细胞中后的T7E1分析进行评估。图9。与Pcsk9-sgRNA-Nme1Cas9-AAV质粒和Rosa26-sgRNA-Nme1Cas9-AAV质粒互补的Pcsk9基因和Rosa26基因内的代表性靶位点序列以彩色注释显示。图10。

[0250] 通过尾静脉静脉内动态注射30μg靶向Pcsk9的无内毒素的sgRNA-Nme1Cas9-AAV质粒,在小鼠体内验证了质粒设计。注射后10天在小鼠肝脏中检测到显著的基因编辑,这是通过DEcomposition(TIDE)跟踪插入删除进行的,DEcomposition是一种基于序列的评估插入删除效率的方法。图11。

[0251] 靶向Pcsk9基因和Rosa26基因的质粒骨架被包装在肝细胞特异性AAV8血清型中,并且通过尾静脉注射每只小鼠4×10^{10}个基因组拷贝(gc)的剂量。初步数据显示,注射后14天处死的小鼠的插入删除值在肝脏Pcsk9和Rosa26基因中具有明显的插入删除水平。图12A。注射后第50天也已收集了深度测序数据。

[0252] 在注射后第50天处死3个小鼠组,并使用TIDE使用肝gDNA测量Pcsk9和Rosa26的插入删除值。图12B。还进行了深度测序分析,以记录对插入删除值的准确测量。

[0253] PCSK9蛋白的“敲低”可能导致小鼠胆固醇水平显著降低。通过Infinity™比色终点分析(Thermo-Scientific)在注射了靶向Pcsk9基因,Rosa26基因和PBS对照组的载体的3个小鼠组中测量血清胆固醇水平。结果表明,Nme1Cas9诱导的插入删除形成导致Pcsk9基因正常阅读框的中断,这在注射后25天和50天时血清胆固醇值显著降低即可证明。图13。在第50天时,还进行了蛋白印迹分析以测量小鼠肝脏中PCSK9蛋白的水平。

[0254] 在注射靶向Pcsk9基因和Rosa26基因的载体后,通过测序测定法(例如,Illumina的GUIDE-Seq®)实现了全基因组范围内的双链断裂(DSB)的无偏性鉴定。数据显示Pcsk9有四(4)个潜在的脱靶位点,而Rosa26有六(6)个潜在的脱靶位点。图14A和14B。

[0255] 靶向TIDE分析揭示了在注射靶向&Pcsk9基因和Rosa26基因的AAV载体后第14天,细胞和小鼠中的靶基因组编辑。图15。在注射后50天,还对这些位点的脱靶切割进行了深度测序分析。

[0256] 在注射靶向Pcsk9基因和Rosa26基因的载体后第14天处死的小鼠肝脏中,苏木精和曙红染色法未显示大量免疫细胞浸润的迹象。图16。注射后50天将进行特异性免疫反应测定。

[0257] 在一个实施方式中,本发明设想了一种通过对Nme2Cas9进行多合一AAV递送来治疗性体内基因组编辑的方法。尽管不必理解本发明的机理,但是据信紧凑,小PAM和高保真度使Nme2Cas9成为使用AAV进行体内基因组编辑的理想工具。为此,将Nme2Cas9及其同源sgRNA及其各自的启动子克隆到单个AAV载体主链中。图44A;最上。这种多合一的AAV.sgRNA.Nme2Cas9包装在肝细胞选择性AAV8衣壳中。靶向2个基因:(i)Rosa26,一种常用位点作为阴性对照;(ii)9型蛋白原转化酶舒替兰酶/kexin(Pcsk9),这是循环胆固醇稳态的主要调节剂。研究表明,使用Cas9敲除Pcsk9会降低胆固醇水平(Ran等)。

[0258] 给两组小鼠(n＝5)注射包装的靶向Pcsk9或Rosa26的AAV8.sgNA.Nme2Cas9。在载体注射后0、14和28天收集血清用于胆固醇水平测量。注射后28天处死小鼠，并收集肝组织。图44A下。深度测序分析显示Pcsk9和Rosa26的插入删除水平显著高。图44B。这些插入删除值伴随着在14天和28天后注射sgPcsk9的小鼠的胆固醇水平显著降低在整个研究中，用sgRosa26维持胆固醇的正常水平。图44C。H&E分析显示Nme2Cas9表达后两组的毒性和组织损伤均没有迹象。图44D。这些数据证实Nme2Cas9在体内具有很高的功能。

[0259] 在一个实施方式中，本发明考虑了最小化的AAV.hNmeCas9构建体。见图44A。如上所述，本发明考虑了工程化的多合一AAV.sgRNA.hNme1Cas9构建体，其被包装在成功编辑小鼠肝脏中的Pcsk9和Rosa26基因的AAV8病毒体中。

[0260] 在一个实施方式中，本发明涉及一种包括Nme2Cas9盒的AAV8骨架。与Nme1Cas9类似，Nme2Cas9在小鼠的Pcsk9和Rosa26也显示出强大的编辑能力(见下文)。本文提供的数据表明，向小鼠体内施用AAV8-NmeCas9伴随载体注射后28天后循环胆固醇水平显著降低。

[0261] 为了增加这种多合一AAV平台的实用性，引入了各种截断法以最小化货物尺寸，从而为AAV衣壳中的其他特征(例如双sgRNA或供体DNA片段)留出空间。

[0262] 为了最小化多合一AAV主链的负载，在不损害Cas9的核酸酶活性的情况下，系统地去除了额外的特征(3×HA标签和2×NLS序列)。Nme1Cas9使用交通灯报告子(TLR)系统显示，此最小化的多合一AAV.sgRNA.hNme1Cas9(4.468kb)与以前的带有4个NLS序列的较长版本一样有效。见图45。使用新的sgRNA12构建截短的sgRNA，以释放更多空间，这与sgRNA11版本相似，但在3'末端添加了UA。见图46。

[0263] 以前，已经报道了短的多聚A序列对于Cas9构建体可能是有用的。Platt等(2015)。在一个实施方式中，本发明涉及包含BGH多聚A的AAV-Nme2Cas9构建体。见图47。尽管不必理解本发明的机理，但据信该多聚A序列进一步减小了多合一AAV主链的大小。

[0264] 进一步认为，这种最小化的(4.4kb)多合一AAV主链通过包括用于双基因敲除或DNA片段切除的另一种sgRNA而增加了Nme1Cas9和Nme2Cas9的效用。见图48上。此配置还提供了AAV衣壳中的自由空间，以包括用于同源性定向修复应用的供体模板(约600个碱基对)。见图48下。在一些实施方式中，双重sgRNA AAV构建体包装在单个AAV载体内。

[0265] 相对紧凑的Nme1Cas9在一系列细胞类型的基因组编辑中活跃。为了利用此Cas9直系同源基因的小尺寸，在小鼠U1a启动子的表达下以及由U6启动子驱动的sgRNA的作用下，通过人密码子优化的Nme1Cas9生成了多合一AAV构建体。见图49A。首先在小鼠基因组中选择了2个位点，以在体内测试Nme1Cas9的核酸酶活性:Rosa26“安全港”基因(由sgRosa26靶向);以及原蛋白转化酶舒替兰酶/kexin 9型(Pcsk9)基因(以sgPcsk9为靶)，这是降低循环胆固醇和降低心血管疾病风险的常见治疗靶。图49B。使用Bioconductor软件包CRISPRseek 1.9.1和N将计算出这些指南的全基因组脱靶预测$_4$GN$_3$PAM和最多六个错配。Zhu et al.,"CRISPRseek:a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems"PLoS One 2014;9:e108424。许多N$_4$GN$_3$PAMS处于非活动状态，因此这些检索参数几乎可以肯定比真正的脱靶配置文件具有更大的覆盖范围。尽管该检索具有广阔的性质，但分析并未发现在小鼠基因组中错配少于4个的脱靶位点。见图50。通过质粒转染在小鼠Hepa1-6肝癌细胞中评估了这些靶位点的靶编辑效率，并使用序列追踪分解(Indes by Decomposition,TIDE)网络工具通过序列迹线分解进行了插入删除

定量。Brinkman et al.,"Easy quantitative assessment of genome editing by sequence trace decomposition"Nucleic Acids Res.2014;42:e168。数据显示所选指南的插入删除值＞25％,其中大多数为缺失。见图49C。

[0266] 为了评估构建的多合一AAV-sgRNA-hNme1Cas9载体的初步功效,通过尾静脉注射将无内毒素的sgPcsk9质粒流体动力学地施用到C57BL/6小鼠中。这种方法可以将质粒DNA输送到约40％的肝细胞中进行瞬时表达。Liu et al.,"Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA"Gene Ther.1999;6:1258-66。通过TIDE使用从肝脏组织中提取的DNA进行的插入删除分析显示,在载体施用后10天,插入删除占5-9％,与SpyCas9相似测试获得的编辑效率相当。见图49D和Xue et al.,"CRISPR-mediated direct mutation of cancer genes in the mouse liver"Nature 2014;514:380-4。这些结果表明,Nme1Cas9能够在体内编辑肝细胞。

[0267] I型遗传性酪氨酸血症(HT-1)是由FAH基因的常染色体隐性突变引起的致命遗传疾病,该基因编码富马酰乙酰乙酸羟化酶(FAH)。FAH减少的患者的酪氨酸分解代谢途径受阻,酪氨酸分解代谢途径受阻,导致有毒的富马酸乙酰乙酸酯和琥珀酰乙酰乙酸积累,从而引起肝和肾损害。Grompe M.,"The pathophysiology and treatment of hereditary tyrosinemia type 1"Semin Liver Dis.2001;21:563-71。在过去的二十年中,该疾病一直由2-(2-硝基-4-三氟甲基苯甲酰基)-1,3-环己二酮(NTBC)控制,该疾病抑制酪氨酸降解途径上游的4-羟基苯基丙酮酸双加氧酶,从而阻止了积累的有毒代谢产物。Lindstedt et al.,"Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate Dioxygenase"Lancet 1992;340:813-7。但是,这种治疗需要终生控制饮食和药物,最终可能需要进行肝移植。Das,AM.,"Clinical utility of nitisinone for the treatment of hereditary tyrosinemia type-1(HT-1)"Appl Clin Genet.2017;10:43-8。

[0268] 已经测试了几种基因治疗策略,以利用CRISPR-Cas9的定点诱变或同源性修复来纠正有缺陷的FAH基因。Paulk et al.,"Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo"Hepatology 2010;51:1200-8;Yin et al.,"Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo"Nat Biotechnol.2016;34:328-33;和Yin et al.,"Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype"Nat Biotechnol.2014;32:551-3。据报道,仅修饰肝中1/10,000的肝细胞就足以挽救$Fah^{mut/mut}$的表型。老鼠。近来,已经提出了一种代谢途径重编程方法,其中肝脏中Hpd基因的外显子3和4的缺失破坏了羟苯基丙酮酸双加氧酶(HPD)酶的功能。Pankowicz et al.,"Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia"Nat Commun.2016;7:12642。这提供了一种背景,其中例如通过靶向Hpd并评估Fah突变小鼠中疾病表型的拯救来测试Nme1Cas9编辑的功效。Grompe et al.,"Loss of fumarylacetoacetate hydrolase is responsible for the neonatal hepatic dysfunction phenotype of lethal albino mice"Genes Dev.1993;7:2298-307。以此目的,在Hpd的开放阅读框中筛选并鉴定了2个靶位点(外显子8[sgHpd1]和外显子11[sgHpd2]中各一个)。见图51A。这些指南(例如sgRNA)通过在Hepa1-6细胞中进

行质粒转染，分别促进了Nme1Cas9诱导的平均插入删除效率为10.8％和9.1％。图52。

[0269] 通过水动力注射磷酸盐缓冲盐水(PBS)或2个sgHpd1和sgHpd2多合一AAV-sgRNA-hNme1Cas9质粒之一对三组小鼠进行水力注射。sgHpd1组的一只小鼠和sgHpd2组的两只小鼠由于尾静脉注射失败而被排除在随访研究之外。注射后7天，将小鼠从含NTBC的水中取出，并在注射后43天监测其重量。见图51B。注射了PBS的小鼠体重严重减轻(HT-1的特征)，并且在体重减轻20％后被处死。总体而言，所有sgHpd1和sgHpd2小鼠总体上成功维持其体重43天，而没有NTBC则至少维持21天。见图51C。

[0270] 对于两只接受sgHpd1的小鼠和另一只接受sgHpd2的小鼠，NTBC治疗必须恢复2-3天，以使它们在质粒注射后的第三周内恢复体重，这可能是由于初始编辑效率低，肝损伤由于流体动力注入，或两者兼而有之。相反，所有其他sgHpd1和sgHpd2处理的小鼠均获得了频率在35-60％范围内的插入删除。见图51D。这种基因失活的水平可能不仅反映了最初的编辑事件，而且反映了已编辑细胞谱系的竞争性扩张(在NTBC撤离后)，但其代价是未编辑的对应基因。肝脏组织学研究显示，与$Fah^{mut/mut}$相比，经sgHpd1和sgHpd2处理的小鼠的肝损伤严重程度较轻注射PBS的小鼠，与注射PBS的小鼠相比，多核肝细胞数量较少。见图53。

[0271] AAV载体最近已用于生成基因组编辑的小鼠，而无需显微注射或电穿孔，只需将合子浸入含有AAV载体的培养基中，然后再植入假妊娠雌性动物中即可。以前使用双AAV系统获得了编辑，在该系统中SpyCas9及其sgRNA分别递送到单独的载体中。Yoon et al.,"Streamlined ex vivo and in vivo genome editing in mouse embryos using recombinant adeno-associated viruses"Nat.Commun.9:412(2018)。为了测试Nme2Cas9是否可以使用多合一AAV递送系统在小鼠合子中进行准确，高效的编辑，我们以酪氨酸酶基因(Tyr)为靶，该基因的双等位基因失活会破坏黑色素的产生，从而导致白化病幼崽。Yokoyama et al.,"Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice"Nucleic Acids Res.18:7293-7298(1990)。

[0272] 在Hepa1-6细胞中通过瞬时转染验证了有效的Tyr sgRNA(从经典的白化突变位点切割Tyr位点仅17bp)。见图57，接下来，将C57BL/6NJ合子在含有3×10^9或3×10^8的培养基中温浴5-6小时。表达Nme2Cas9和Tyr sgRNA的多合一AAV6载体的GC，在新鲜培养基中过夜培养后，将那些发展到两细胞阶段的合子转移到假孕受体的输卵管中，并使其发育为术语。见图58A。幼犬的毛色分析显示，小鼠为白化病，浅灰色(暗示Tyr亚型等位基因)，或杂色的毛色由白化病和浅灰色斑点组成，但缺乏黑色色素沉着。见图58B和58C。这些结果表明，双等位基因突变的频率很高，因为单个野生型Tyr等位基因的存在会导致黑色素沉着。从3×10^9出生了总共五只幼仔(10％)GC实验。在表型上，他们中的全部都带有插入删除，在表型上，2个是白化病，一个是浅灰色，另外2个具有斑驳的色素沉着，表明有马赛克。从3×10^8GC实验中，获得了四(4)只幼仔(14％)，其中有两只在出生时死亡，从而阻止了毛色或基因组分析。其余两只幼犬的外套颜色分析显示，一只浅灰色和一只马赛克幼犬。这些结果表明，Nme2Cas9及其sgRNA的单AAV递送可用于在小鼠合子中产生突变，而无需显微注射或电穿孔。

[0273] 为了测量Tyr基因中中靶插入删除的形成，从每只小鼠的尾巴中分离DNA，扩增基因座，并进行TIDE分析。数据显示，所有小鼠的Nme2Cas9均具有高水平的按靶编辑能力，范

围从84％到100％不等。见图57B和5C。白化病小鼠9-1中的大多数病变均为1或4bp缺失，表明是镶嵌性或反杂合性。白化病小鼠9-2表现出统一的2-bp删除。见图58C。对来自浅灰色小鼠的尾巴DNA的分析显示，存在框内突变的存在，这可能是造成浅灰色外套颜色的原因。有限的突变复杂性表明编辑在这些小鼠的胚胎发育早期发生。一只雌性(9-2小鼠)与一头经典的白化病雄性交配，所有六只幼崽均为白化病，证明通过合子Ame递送Nme2Cas9+sgRNA产生的突变可以通过种系传播。这些结果通过应用单个AAV载体(在这种情况下同时递送Nme2Cas9及其sgRNA)，为哺乳动物诱变提供了一条简化的途径。

[0274] Hpd基因突变的患者被认为患有III型酪氨酸血症，并在血液中表现出高水平的酪氨酸，但在其他方面似乎无症状。Szymanska et al.,"Tyrosinemia type III in an asymptomatic girl.Mol Genet Metab Rep.2015;5:48-50;和Nakamura et al.,"Animal models of tyrosinemia"J Nutr.2007;137:1556S-60S。HPD在酪氨酸分解代谢途径中作用于FAH的上游，Hpd的破坏可通过防止因FAH丧失而导致的有毒代谢物积累来改善HT-1症状。HPD的结构分析表明，HPD酶的催化结构域位于酶的C末端，由外显子13和14编码。Huang et al.,"The different catalytic roles of the metal-binding ligands in human 4-hydroxyphenylpyruvate dioxygenase"Biochem J.2016;473:1179-89。从而，外显子13上游的移码诱导插入删除将使酶失活。此上下文用于证明通过水动力注射Nme1Cas9质粒使Hpd失活是拯救HT-1小鼠的可行方法。Nme1Cas9可以编辑带有多个不同PAM的位点(N_4GATT[共识]，N_4GCTT，N_4GTTT，N_4GACT，N_4GATA，N_4GTCT和N_4GACA)。Hpd编辑实验使用sgHpd2指南证实了其中一种变体PAM，该指南针对的是N_4GACT PAM的位点。

[0275] 尽管质粒流体动力注射可以产生插入删除，但是治疗的发展可能需要侵入性较小的递送策略，例如通过使用rAAV。为此，将多合一AAV-sgRNA-hNme1Cas9质粒包装在嗜肝细胞AAV8衣壳中，以靶向Pcsk9(sgPcsk9)和Rosa26(sgRosa26)。见图49B;Gao et al,"Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy"Proc Natl Acad Sci USA 2002;99:11854-9,和Nakai et al."Unrestricted hepatocyte transduction with adeno-associated virus serotype 8vectors in mice"J Virol 2005,79:214-24。Pcsk9和Rosa26被部分用来使Nme1Cas9 AAV递送能够与其他Cas9直系同源物的递送类似，并针对相同的基因座。Ran et al.,"In vivo genome editing using Staphylococcus aureus Cas9"Nature 2015;520:186-91。通过尾静脉将载体给予C57BL/6小鼠。见图54A。在注射后25和50天，监测血清中的胆固醇水平并测量肝组织中的PCSK9蛋白和插入删除频率。

[0276] 使用比色终点分析，确定在注射后25天和50天与PBS和Nme1Cas9/sgRosa26小鼠相比，施用Nme1Cas9/sgPcsk9的小鼠中的循环血清胆固醇水平显著降低($p<0.001$)。。见图54B。在施用载体后50天，对Pcsk9和Rosa26靶位点进行的靶向深度测序分析显示非常有效的插入删除，分别为35％和55％。图54C。另外，在注射后第14天对每组的一只小鼠实施安乐死，并显示在Pcsk9和Rosa26上的靶点插入删除效率分别为37％和46％。正如预期的那样，与注射PBS和Nme1Cas9/sgRosa26的小鼠相比，Nme1Cas9/sgPcsk9治疗的小鼠的肝脏中PCSK9蛋白水平大大降低。见图54D。有效的编辑，PCSK9的减少和血清胆固醇的降低表明Nme1Cas9在Pcsk9基因座处的成功递送和活性。

[0277] 已知由病毒载体递送的SpyCas9引起宿主免疫应答。Chew et al.,"A

multifunctional AAV-CRISPR-Cas9 and its host response"Nat Methods 2016;13:868-74;和Wang et al.,"Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses"Hum Gene Ther.2015;26:432-42。为了研究注射AAV8-sgRNA-hNme1Cas9的小鼠是否产生抗Nme1Cas9抗体,将被治疗动物的血清用于IgG1 ELISA。这些结果表明,Nme1Cas9在这些动物中引起体液反应。见图55。尽管存在免疫应答,但rAAV递送的Nme1Cas9在体内具有很高的功能,没有明显的异常或肝损伤迹象。见图16。

[0278] 在治疗性CRISPR/Cas9基因组编辑中,一个重要的问题是脱靶编辑时活性的可能性。例如,已经发现野生型Nme1Cas9是培养的哺乳动物细胞中的天然高精度基因组编辑平台。Lee et al.,"The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells"Mol Ther.2016;24:645-54。为了确定Nme1Cas9是否在小鼠细胞和体内保持最小的脱靶谱,使用测序技术(GUIDE-seq)对全基因组进行了无偏性的DSB鉴定,在小鼠基因组中筛选了脱靶位点。Tsai et al.,"Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases"Nat Rev Genet.2016;17:300-12。用sgPcsk9,sgRosa26,sgHpd1,并且将sgHpd2多合一AAV-sgRNA-hNme1Cas9质粒和所得的基因组DNA进行GUIDE-seq分析。与人细胞中的观察结果一致(数据未显示),GUIDE-seq显示出小鼠基因组中极少的脱靶(OT)位点。sgPcsk9鉴定出4个潜在的OT位点,sgRosa26鉴定出另外六个。未检测到使用sgHpd1和sgHpd2的脱靶编辑。见图56A。这些数据进一步验证了Nme1Cas9本质上是超精确的。

[0279] sgPcsk9和sgRosa26的一些可能的OT位点缺乏Nme1Cas9 PAM偏好(即N_4GATT,N_4GCTT,N_4GTTT,N_4GACT,N_4GATA,N_4GTCT和N_4GACA)。见图56B。为了验证这些OT位点,使用来自Hepa1-6细胞的基因组DNA进行了靶向的深度测序。通过这种更敏感的读长,在所有这些OT位点上都无法检测到高于背景的插入删除,但Pcsk9的OT1除外,插入删除频率<2%。见图56B。为了验证Nme1Cas9在体内的高保真度,在来自AAV8-Nme1Cas9处理,以sgPcsk9为靶和以sgRosa26为靶的小鼠的肝脏基因组DNA的这些OT位点测量了插入删除形成。在第14天,除sgPcsk9 OT1以外,所有部位处死的小鼠肝脏中几乎没有或没有可检测到的脱靶编辑,其病变效率<2%。更重要的是,即使在50天后,此OT编辑水平仍保持在<2%以下,并且对于所有其他候选OT位点而言,仍不可检测或非常低。这些结果表明,Nme1Cas9在体内的延长表达(50天)不会影响其靶向保真度。见图56C。

[0280] 为了实现Nme1Cas9到体内各种组织的靶向递送,由于Nme1Cas9转基因的紧凑大小,rAAV载体是一个有前途的递送平台,它允许以多合一的形式递送Nme1Cas9及其指南。本文提供的数据验证了这种靶向成年小鼠中Pcsk9和Rosa26基因的方法,甚至在注射后14天也观察到有效编辑。Nme1Cas9本质上是准确的,即使没有为减少SpyCas9脱靶所需的广泛工程设计也是如此。Lee et al.,"The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells"Mol Ther.2016;24:645-54;Bolukbasi et al.,"Creating and evaluating accurate CRISPR/Cas9 scalpels for genomic surgery"Nat Methods 2016;13:41-50,Tsai et al.,"Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases"Nat Rev Genet.2016;17:300-12;和Tycko et al.,"Methods for optimizing CRISPR-Cas9 genome editing

specificity"Mol Cell.2016;63:355-70。

[0281] Nme1Cas9 OT编辑的并排比较是在培养的细胞和体内通过靶向深度测序进行的，发现在两种情况下脱靶都是最小的。在sgPcsk9 OT1位点(在未注释的基因座内)的编辑可检测到的最高值约为2%。

[0282] 【IV.富含胞嘧啶的PAM的小型Cas9直系同源物】

[0283] 如上所述，CRISPR系统可分为至少六(6)个不同类型。通常，II型系统按Cas9核酸酶蛋白的存在进行分类。例如，Cas9核酸酶蛋白被认为是一种RNA引导的核酸酶，可以在包括人在内的几乎所有生物体中重新用作基因组编辑平台。报告表明，Cas9基因组编辑已用于医学，农业，人基因治疗和许多其他应用。

[0284] 通常，可以通过与单引导RNA(sgRNA)结合的Cas9核酸酶蛋白实现人基因组中特定基因座的靶向，该单引导RNA通过与特定核酸序列(例如，例如，，是与原间隔子符相邻的主题;PAM)。sgRNA通常包含与靶核酸序列互补的20-24个核苷酸区段，随后是与Cas9蛋白相互作用(例如，结合)的恒定区。为了使Cas9核酸酶蛋白执行基因组编辑，Cas9:sgRNA复合体首先识别通常在靶位点序列下游的原间隔子相邻基序(PAM)序列。尽管不必了解本发明的机制，但可以相信每种Cas9核酸酶蛋白都对特定的PAM具有亲和力(即(例如，由原间隔子相邻基序识别结构域介导)。在缺乏结合至下游PAM靶核酸序列的PAM识别结构域的情况下，Cas9核酸酶不能切割双链DNA(dsDNA)。

[0285] 报告表明，只有少数Cas9直系同原物已被验证可用于人基因组编辑。报道的CRISPR-Cas9类型中的三种包括II-A，II-B和II-C。II-A型Cas9(例如化脓链球菌(SpyCas9))是迄今为止最常用的Cas9。但是，SpyCas9(以及大多数其他II-A型直系同源物)具有一些特性，可能使其不适用于某些应用。首先，SpyCas9相对较大，使得该Cas9不适合有效包装到病毒载体中。其次，尽管已经设计了更高特异性的变体，但SpyCas9的脱靶活性率很高(即，它在人基因组中的非预期基因座上切割DNA)。最后，SpyCas9的PAM(例如NGG)在人基因组的某些位点中的使用受到限制，或用于在编辑过程中识别特定核苷酸的应用。为了克服这些缺点，几个小组将其他Cas9直系同源物重新定位为在人和其他生物体中起作用。如上所述，II-C型Cas9直系同源物(例如Nme1Cas9)足够小，可用于多合一病毒包装(例如，腺相关病毒(AAV)载体)，从而在哺乳动物细胞中具有更高的保真度。野生型Cas9 II-C PAM的长度通常约为四(4)个核苷酸，而SpyCas9PAM的长度通常为两(2)个核苷酸。此额外的PAM长度可能会限制野生型Cas9 II-C PAM可以定位的基因座数量。这在本领域中需要鉴定更多用于基因组编辑的Cas9直系同原物。

[0286] 尽管在NCBI数据库中有成千上万的Cas9直系同源物可供选择，但需要经验性的过程来开发小型II-C Cas9直系同源物，其限制性更小的PAM能够在哺乳动物细胞中提供改善的功能。在一个实施方式中，本发明设想了一种改进的II-C Cas9直系同原物，其能够以更广泛的靶位点进行精确的基因组编辑。在一个实施方式中，改进的II-C型Cas9直系同原物具有能够有效地递送病毒的紧凑尺寸。在一个实施方式中，改良的II-C型Cas9直系同原物包括但不限于副流感嗜血杆菌(Haemophilus parainfluenzae)(HpaCas9)，穆氏西门氏菌(SmuCas9)和脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株De10444(Nme2Cas9)。

[0287] 【A.与II-C型Cas9直系同源物相关的短PAM】

[0288] 本文提供的数据显示了几种II-C型Cas9直系同源物的短PAM靶的表征。图17。例

如,II-C型Cas9直系同原物可以与包含1～4个所需核苷酸之间的短PAM相互作用。尽管不必理解本发明的机制,但据信这些短的富含C的PAM可提供对靶位点的改进的Cas9基因组编辑,这些靶位点以前甚至无法通过更紧凑的Cas9直系同源物(例如Nme1Cas9)访问。在一个实施方式中,Nme2Cas9 PAM具有NNNNCc的序列,其中"c"是唯一的部分偏好。在一个实施方式中,SmuCas9 PAM具有NNNNCT序列。图18。

[0289] 目前认为,尚未验证具有短的富含C的PAM的Cas9直系同原物用于基因组编辑,并且Nme2Cas9作为在人细胞中高效基因编辑活性的潜在候选者特别引人注目。在一个实施方式中,本发明涉及与野生型Nme1Cas9 sgRNA(例如,脑膜炎奈瑟氏球菌(Neisseria meningitidis)8013Cas9,先前称为NmeCas9)结合的Nme2Cas9核酸酶。Nme1Cas9先前已经描述过。Sontheimer et al.,"RNA-Directed DNA Cleavage and Gene Editing by Cas9 Enzyme From Neisseria Meningitidis",美国专利申请公开号2014/0349,405(在此引入作为参考)。尽管Nme1Cas9可用于基因组编辑,但其主要限制是其相对较长的PAM,这限制了任何给定基因组基因座中可编辑位点的数量。

[0290] 在一些实施方式中,本发明考虑了II-C型Cas9直系同源物的较短和较不严格的PAM,包括但不限于Nme2Cas9。尽管没有必要了解发明的机制,但人们认为,较短而不太严格的PAM可以部分缓解靶限制的局限性,同时仍然保留了Nme1Cas9的许多(即使不是大多数)优点,包括但不限于小的优点尺寸(例如紧凑性)以实现高效的多合一AAV交付并提高靶精确度(例如减少脱靶切割)。此外,上面讨论的Nme1Cas9的最小化sgRNA也与Nme2Cas9构建体兼容。因此,这种截短的引导RNA也可能被Nme2Cas9用于基因组编辑。

[0291] 在一个实施方式中,本发明考虑了一种具有NNNNGNTTT序列的HpaCas9 PAM。尽管事实上长的PAM限制了人基因组中可靶向位点的数量,但人们认为HpaCas9 PAM可以以非常高的准确性靶向位点,这与极度准确的Nme1Cas9类似(同上)。

[0292] 本文提供的数据证明了靶向短富CPAM的II-C型Cas9核酸酶在人(HEK293T)细胞中进行基因组编辑的能力。Certo et al.,"Tracking genome engineering outcome at individual DNA breakpoints"Nature Methods 8:671-676(2011)。例如,HpaCas9和Nme2Cas9在特定基因座上提供了有效的基因组编辑,表明它们在哺乳动物中具有活性。图19和表2。

[0293] 【表2:人基因组中代表性的II-C型Cas9直系同源物靶序列】

[0294]

as	间隔子序列	PAM	染色体
me2	GAATATCAGG AGACTAGGAA GGAG	GAGGCCTA	19
pa	GGACAGGAGT CGCCAGAGGC CGGT	GGTGGATTT	4
mn	GCACCTGCCT CGTGGAATAC GGT	AAACCTAC	交通灯报告子

[0295] 这些数据表明,Nme2Cas9和HpaCas9在相同的基因组位点以与先前验证的Nme1Cas9相当的水平进行了基因组编辑。对于SmuCas9,编辑效率相对较低,尽管很重要的一点是活动不为零,并且有望提高效率。然后,将Nme2Cas9用于测试整合到HEK293T细胞基因组中的交通灯报告子(TLR)中的另外十四(14)个位点。在这些测定中,每个位点均符合PAM模板,即"C"是PAM区的第五个核苷酸(即NNNNCNNN)。值得注意的是,所有14个位点均由Nme2Cas9编辑,表明该酶在多种酶中均具有活性哺乳动物细胞中的指南。最成功的引导RNA符合NNNNCCN PAM共识。图20。

[0296] 测试了II-C型Cas9直系同原物对抗CRISPR蛋白的敏感性。抗CRISPR蛋白是天然存在的蛋白，当不再需要Cas9活性时，可以关闭Cas9。数据显示所有三种II-C型Cas9直系同原物均被某些抗CRISPR抑制。图21。抗CRISPRs对这些Cas9直系同源物的可控性可以增加其在基因组编辑中的潜在效用。

[0297] 【B.Nme2Cas9基因编辑】

[0298] 本文提供的数据显示了使用Nme2Cas9-sgRNA复合物的基因编辑。该数据采用交通灯报告子(TLR)系统来证明，在NNNNCC序列的背景下，基因靶序列中的任何CC二核苷酸均可充当PAM。图22。蓝色条形是显示荧光的细胞百分率，而红色条形表示基于测序(“TIDE分析”)更准确地进行编辑的百分率。这些数据证实二核苷酸足以用于Nme2Cas9 PAM结合，而不是三核苷酸序列的要求(例如序列NNNNCCX中的“X”)。尽管不必了解发明的机制，但可以认为这意味着Nme2Cas9可编辑基因组靶位点的频率至少与SpyCas9可编辑位点相同，并且比SauCas9、Nme1Cas9或CjeCas9和其他当前替代方法更频繁。

[0299] 此外，T7E1测定法被用于分析天然基因组位点的编辑(例如，非整合的人工荧光报告子)。这些数据表明，在某些情况下，甚至可能不需要第二个“C”。见图23。请注意，靶位置DeTS1和DeTS4都位于AAVS1位置，可分别在靶位置使用NNNNCA和NNNNCG候选PAM进行编辑。本文公开了这些Nme2Cas9靶位点中的几个。见表3。

[0300] 【表3:Nme2Cas9的代表性PAM靶位点】

[0301]

靶位点名称	靶基因座	靶序列（间隔子-PAM）
Nme2TS1	AAVS1	ATGTGGCTCT GGTTCTGGGT ACTTTTATCT GTCCCCTCCA CCCCACAGTG GG
Nme2TS4	AAVS1	CAGATAAGGA ATCTGCCTAA CAGGAGGTGG GGGTTAGACG AATATCAGGA GA
Nme2TS5	AAVS1	GGGGTTAGAC GAATATCAGG AGACTAGGAA GGAGGAGGCC TAAGGATGGG GG
Nme2TS6	Chr.14	CCCACCCGGC GGCGCCTCCC TGCAGGGCTG CTCCCCAGCC CAAACCGCCG CG
Nme2TS10	AAVS1	TCCGAGAGCT CAGCTAGTCT TCTTCCTCCA ACCCGGGCCC TATGTCCACT TC
Nme2TS11	AAVS1	TGGGTACTTT TATCTGTCCC CTCCACCCCA CAGTGGGGCC ACTAGGGACA GG
Nme2TS12	AAVS1	GTAGGGGAGC TGCCCAAATG AAAGGAGTGA GAGGTGACCC GAATCCACAG GA
Nme2TS13	AAVS1	TAGCACCTCT CCATCCTCTT GCTTTCTTTG CCTGGACACC CCGTTCTCCT GT
Nme2TS14	AAVS1	GTCTCCCTTG CGTCCCGCCT CCCCTTCTTG TAGGCCTGCA TCATCACCGT TT
Nme2TS15	AAVS1	CCTCACCCAA CCCCATGCCG TGTTCACTCG CTGGGTTCCC TTTTCCTTCT CCT
Nme2TS16	Chr.14	GCGCAGGACA GGAGTCGCCA GAGGCCGGTG GTGGATTTCC

[0302]

		TCCCCGCATC TC
Nme2TS17	Chr.14	CGCGGGGACG CCCAGCGGCC GGATATCAGC TGCCACGCCC GCGTGGGCGG A
Nme2TS22	VEGF	GATTCCAATA GATCTGTGTG TCCCTCTCCC CACCCGTCCC TGTCCGGCTC TC
Nme2TS23	VEGF	TGACCCCTGG CCTTCTCCCC GCTCCAACGC CCTCAACCCC ACACGCACAC AC
Nme2TS24	VEGF	TCCCTCTCCC CACCCGTCCC TGTCCGGCTC TCCGCCTTCC CCTGCCCCCT TC
Nme2TS25	VEGF	ACACGCACAC ACTCACTCAC CCACACAGAC ACACACGTCC TCACTCTCGA AG
Nme2TS26	Chr.7 (CFTR)	TAAGCACAGT GGAAGAATTT CATTCTGTTC TCAGTTTTCC TGGATTATGC CT
Nme2TS27	Chr.7 (CFTR)	TTCATTCTGT TCTCAGTTTT CCTGGATTAT GCCTGGCACC ATTAAAGAAA AT

[0303] 尽管没有必要理解发明的机制,但据信这些数据表明基因组中平均每4-8个碱基对可能存在候选编辑位点。这些数据还表明大多数Cas9 sgRNA具有某些功能,因此,本领域可能过分强调了对sgRNA筛选的需求。

[0304] 【C.迅速发展的PAM交互域】

[0305] CRISPR-Cas9的体内应用具有改变生物技术和治疗学许多领域的潜力。自然界中有成千上万的Cas9直系同原物,其中只有少数经过了体内基因组编辑的验证。酿脓链球菌(S.pyogenes)的Cas9(SpyCas9)由于其高效和非限制性NGG原间隔子相邻基序(PAM)而被广泛使用。但是,相对较大的SpyCas9大小限制了其在体内治疗中的使用,这种方法使用的是包装能力有限的输送梭,例如腺相关病毒(AAV)。已知几个较小的Cas9直系同源物在哺乳动物细胞中具有活性,但它们具有限制性更强的PAM,从而限制了靶位点密度。可以利用密切相关的Cas9直系同源物的PAM相互作用结构域(PID)中的自然变异来鉴定克服这些限制的基因组编辑酶。在一些实施方式中,本发明考虑使用Nme2Cas9复合物,其是具有N_4CC PAM的紧凑,天然超准确的Cas9。本文提供的数据表明Nme2Cas9是一种高保真哺乳动物基因组编辑平台,可提供与SpyCas9相同的靶位点密度。通过多合一AAV载体递送Nme2Cas9及其引导RNA,可在成年小鼠中进行有效的基因组编辑,在肝脏中以Pcsk9基因为靶可诱导血清胆固醇降低,而没有明显的脱靶(见下文)。Nme2Cas9还为哺乳动物体内基因组编辑提供了多合一AAV兼容性,自然超高精度和高靶位密度的独特组合。

[0306] 除了靶密度之外,使Cas9用作安全治疗剂非常需要使脱靶活性(例如,在不需要的基因座处的裂解)最小化。野生型(wt)SpyCas9由于其独特的杂交动力学而具有高度的脱靶活性。(Klein等人,2018)。特别地,仍然存在关于它们的按靶编辑效率的问题,并且这些变体没有克服上述关于整体大小的限制。相反,本文已显示Nme1Cas9和CjeCas9的实施方式包括天然准确的基因编辑活性。尽管不必了解本发明的机理,但据信以前没有报道Cas9直系同原物:(i)在人细胞中有活性;(ii)表现出SpyCas9极高的靶位密度;(iii)足够紧凑以实现多合一的AAV交付;(iv)自然是超高精度的。在一个实施方式中,本发明考虑了Nme2Cas9作为基因组编辑平台,其包括上述所有特征。例如,Nme2Cas9包含对N_4CC PAM具有高亲和力的结合位点是超高精度的,在哺乳动物细胞中有效发挥作用。在一个实施方式中,将Nme2Cas9包装在用于治疗性基因组编辑的多合一AAV递送平台中。

[0307] 【1.具有快速发展的PID的密切相关的Nme1Cas9直系同源物】

[0308] 先前已经报道,Nme1Cas9(来自脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株8013)是用于体内基因组编辑的小型,超高精度Cas9(Amrani等,2018)。但是,Nme1Cas9绑定到一个较长的PAM(N_4GMTT),这限制了它在某些可以锁定小窗口的情况下的使用。Cas9对PAM的识别主要通过Cas9的PAM相互作用结构域(PID)和与PAM相邻的核苷酸之间的蛋白-DNA相互作用进行。PID受到噬菌体和其他移动遗传元件(MGE)的高选择压力。例如,已经显示抗CRISPR蛋白与PID相互作用以抑制Cas9(下文)。这可能会导致具有密切相关的Cas9直系同源物,这些直系同源物的PID识别出完全不同的PAM。

[0309] 最近,使用两种土芽孢杆菌属(Geobacillus)突出了这一原理。确定嗜热脂肪土芽孢杆菌(G.stearothermophilus)包含N_4CRAA PAM特有的PID,但当将其交换为LC300 PID菌株时,其亲和力变为N_4GMAA PAM(Harrington等,2017)。假设假设脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株高度测序,则可以发现与快速识别的PAM迅速进化的PID密切相关的Cas9直系同源物。研究了与NmeCas9菌株8013具有高度序列同一性(>80%)的Cas9直系同源物,因为该Cas9已为基因组编辑充分表征,体积小且准确性高。鉴定了几个Cas9直系同原物,其与菌株8013相比,其PID氨基酸序列不同。图34A。

[0310] 发现三组不同的Cas9直系同源物具有明显不同的PID,图35A。从每个PID组中选择一个菌株,例如,第2组的De11444和第3组的98002。这2个CRISPR基因座具有完整的Cas9开放阅读框和带有几个间隔子的CRISPR阵列,表明它们是活性基因座。有趣的是,这些CRISPR基因座的crRNA和tracrRNA与8013相同,可以利用相同的sgRNA。图35B。

[0311] 为了测试这些Cas9直系同源物是否确实具有对不同PAMs具有亲和力的PID,由于这些直系同源物的蛋白的其余部分具有高序列同一性,所以将8013PID与98002PID和De11444 PID互换。为了鉴定PAM,将这些蛋白“嵌合体”重组表达,纯化并用于体外PAM鉴定,如前所述。简而言之,使用重组Cas9和靶向该原间隔子的sgRNA在体外切割包含原间隔子和下游十(10)个核苷酸随机序列的DNA片段。图34B。G23核苷酸间隔子区长度用于sgRNA,与Nme1Cas9 8013和其他II-C型系统一致。PAM鉴定分析表明,这些不同的Cas9嵌合体具有识别不同PAM的PID。例如,通过识别第5位的C残基而不是Nme1Cas9 8013用其N_4GATT PAM识别的G。图34C。

[0312] 但是,由于嵌合蛋白的低切割效率,不能可靠地表征剩余的核苷酸,这表明PID之外的少数残基可能涉及有效的活性。图35C。为了进一步解析PAM,在具有7个核苷酸的随机PAM(位置为5的C)的库中进行了体外测定(例如,NNNNCNNN)。结果表明NmeCas9-De11444和NmeCas9-98002分别识别NNNNCC(A)和NNNNCAAA PAM。图35D。NmeCas9-De11444对位置5的C有很强的偏好,但对核苷酸6和7的偏好较弱。如本文所用,Cas9 De11444直系同源物被称为“Nme2Cas9”,而Cas9 98002直系同源物被称为“Nme3Cas9”。

[0313] 我们还使用全长(例如,未进行PID交换的)Nme2Cas9进行了此测定,并观察到相似的结果。图34E。这些结果表明,Nme2Cas9和Nme3Cas9具有识别NAM1Cas9的PAM完全不同的PID。

[0314] 【2.人细胞中的Nme2Cas9】

[0315] 因为Nme2Cas9 PID与一个小的PAM序列结合,所以该直系同原物可用于人基因组编辑,特别是在涉及高靶密度时。为了表征Nme2Cas9,将全长(未进行PID交换的)人源化Nme2Cas9与NLS一起克隆到CMV驱动的质粒中,以进行哺乳动物表达。为了表征人细胞,使用

了Traffic Light Reporter系统,类似于先前所述系统(Certo等,2011)。

[0316] 通过TLR 2.0基因座中的非同源末端连接(NHEJ)的不完美修复,产生+1移码插入/缺失的诱导。在没有供体DNA的情况下,产生了框内mCherry蛋白,其可以通过流式细胞术定量。图36A。作为初始测试,将Nme2Cas9质粒与十五个(15)sgRNA质粒一起转染,其中的间隔子靶向N_4CCX PAM的原间隔子。作为对照,使用了SpyCas9和Nme1Cas9及其针对NGG和N_4的同源sgRNA。GATT原间隔子。七十二(72)小时后收获细胞,并对每个靶位点的mCherry阳性细胞进行定量。SpyCas9和Nme1Cas9在其各自的靶(分别约为28%和10%mCherry)上显示了有效的编辑效果(图36B)。对于Nme2Cas9,具有N_4CCX PAM的所有十五(15)个靶均具有不同程度的功能(范围从4%到20%mCherry),而没有伴随sgRNA和/或N_4GATT对照的NmeCas9处理没有产生mCherry细胞。图36B。这些数据表明,Nme2Cas9识别人细胞中的N_4CC PAM。

[0317] 为了进一步解析Nme2Cas9 PAM,还在TLR报道细胞中用N_5CX和N_4CD(D=A,T,G)测试了靶位点。在具有N_5CX和N_4CD PAM的靶位点处未观察到可编辑的编辑,表明基于TLR 2.0报告子,Nme2Cas9的活性需要5和6位的2个C核苷酸。图37A和37B。这些结果表明,Nme2Cas9包含与N_4CC PAM结合的PID,并且在TLR 2.0基因座的哺乳动物细胞中始终发挥功能。

[0318] 在不同的Cas9直系同原物之间,crRNA间隔子区的长度是不同的。SpyCas9的最佳间隔子区长度是二十(20)个核苷酸,但是,可以截短到十七(17)个核苷酸的截短。Fu et al.,Nature Biotechnology 32,279(2014)。相比之下,Nme1Cas9包含具有二十四(24)个核苷酸间隔子区的sgRNA,并能耐受低至18(18)个核苷酸的截短。(Amrani et al.,2018)。为了测试Nme2Cas9的间隔子区长度,创建了靶向相同基因座但间隔子区长度不同的sgRNA质粒。图36C和图37B。当使用G23,G22和G21间隔子时,观察到了可比的活性,当将引导切成G20和G19时,活性明显降低。图36C。这些结果表明,Nme2Cas9'最佳间隔子区长度在22-24个核苷酸之间,与Nme1Cas9,GeoCas9和CjeCas9相似。因此,以下描述的所有实验均使用23-24个核苷酸间隔子进行。

[0319] Cas9直系同原物被认为使用它们的HNH和RuvC结构域分别在靶DNA的互补链和非互补链中诱导双链断裂。或者,Cas9切口酶已用于通过产生突出端来改善基因组编辑特异性和同源性定向修复(HDR)。(Ran等,2013)。但是,由于SpyCas9具有较高的靶密度,因此只能通过使用SpyCas9来获得成功。要将Nme2Cas9用作昵称,请使用Nme2Cas9^{D16A}和Nme2Cas9^{H588A}产生分别在RuvC和HNH结构域的催化残基中提供突变的α-淀粉酶。由于TLR 2.0还可以用于研究HDR的效率,当提供供体时,修复后的基因座会表达GFP,因此包含供体DNA序列可通过这些Nme2Cas9切口酶测试HDR。在TLR2.0基因中选择靶位点,以使用靶向间隔子为32bp和64bp的切割位点的引导RNA来测试每种切口酶的功能。作为对照,靶向单个位点的野生型Nme2Cas9显示有效的编辑,并伴随有NHEJ和HDR修复途径的诱导。对于切口酶,使用Nme2Cas9^{D16A}(HNH切口酶)显示间隔子为32bp和64bp的切割位点可编辑,但使用Nme2Cas9^{H588A}均未^{切割}任何靶^{标}。图36D。

[0320] Cas9直系同原物包含通常与PAM近端的八至十二(8-12)个核苷酸之间的靶序列杂交的种子序列。种子序列和PAM之间的错配(例如,非互补性)可以降低Cas9核酸酶活性。通过沿着二十三(23)个核苷酸间隔子区走单核苷酸错配,进行了一系列针对TLR 2.0基因中相同基因座的瞬时转染。图37C。与其他Cas9直系同源物相似,数据表明Nme2Cas9在前八至九(8-9)个核苷酸中具有一个"种子序列",该序列与PAM近端的靶序列杂交,这可以从数量

的减少中得出mCherry阳性细胞。即使对错配的耐受性高度依赖于sgRNA的序列和靶基因座,但这些结果表明Nme2Cas9对错配的耐受性非常低,尤其是在其种子序列中。

[0321] 【3.Nme2Cas9基因组编辑效率】

[0322] 使用瞬时转染,Nme2Cas9被用于靶向HEK293T细胞中整个人基因组中的四十(40)个不同的靶位点。表4。

[0323] 【表4:代表性的HEK293T细胞Nme2Cas9靶位点】

[0324]

编号	位点名	间隔点子序列	PAM	基因座	150ng Cas9 p	TIDE引物名	FW TIDE引物	RV TIDE引物
1	TS1	GGTTCTGGGTACTTTTATCTGTCC	CCTCCACC	AAVS1	0.2	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCAACTTATTCTG
2	TS4	GTCTGCCTAACAGGAGGTGGGGGT	TAGACGAA	AAVS1	11	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCAACTTATTCTG
3	TS5	GAATATCAGGAGACTAGGAAGGAG	GAGGCCTA	AAVS1	15	AAVS1_TIDE1	TGGCTTAGCACCTCTCCCAT	AGAACTCAGGACCAACTTATTCTG
4	TS6	GCCTCCCTGCAGGGCTGCTCCC	CAGCCCAA	LINC01588	20	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
5	TS10	GAGCTAGTCTTCTTCCTCCAACCC	GGGCCCTA	AAVS1	3.5	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCAACTTATTCTG
6	TS11	GATCTGTCCCCTCCACCCCACAGT	GGGGCCAC	AAVS1	9	AAVS1_TIDE1	TGGCTTTAGCACCTCTCCAT	AGAACTCAGGACCAACTTATTCTG
7	TS12	GGCCCAAATGAAAGGAGTGAGAGG	TGACCCGA	AAVS1	10	AAVS1_TIDE2	TCCGTCTTCCTCCACTCC	TAGGAAGGAGGAGGCCTAAG
8	TS13	GCATCCTCTTGCTTTCTTTGCCTG	GACACCCC	AAVS1	2	AAVS1_TIDE2	TCCGTCTTCCTCCACTCC	TAGGAAGGAGGAGGCCTAAG
9	TS16	GGAGTCGCCAGAGGCCGGTGGTGG	ATTTCCTC	LINC01588	28	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
10	TS17	GCCCAGCGGCCGGATATCAGCTGC	CACGCCCG	LINC01588	0.2	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
11	TS18	GGAAGGGAACATATTACTATTGC	TTTCCCTC	CYBB	1	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
12	TS19	GTGGAGTGGCCTGCTATCAGCTAC	CTATCCAA	CYBB	6	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
13	TS20	GAGGAAGGGAACATATTACTATTG	CTTTCCCT	CYBB	11.2	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
14	TS21	GTGAATTCTCATCAGCTAAAATGC	CAAGCCTT	CYBB	1	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
15	TS25	GCTCACTCACCCACACAGACACAC	ACGTCVTC	VEGFA	15.6	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
16	TS26	GGAAGAATTTCATTCTGTTCTCAG	TTTTCCTG	CFTR	2	hCFTR_TIDE1	TGGTGATTATGGGAGAACTGGAGC	ACCATTGAGGACGTTTGTCTCAC
17	TS27	GCTCAGTTTTCCTGGATTATGCCT	GGCACCAT	CFTR	4	hCFTR_TIDE1	TGGTGATTATGGGAGAACTGGAGC	ACCATTGAGGACGTTTGTCTCAC

18	TS31	GCGTTGGAGCGGGGAGAAGGCCAG	GGGTCACT	VEGFA	9	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
19	TS34	GGGCCGCGGAGATAGCTGCAGGGC	GGGGCCCC	LINC01588	0	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
20	TS35	GCCCACCCGGCGGCGCCTCCCTGC	AGGGCTGC	LINC01588	0	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
21	TS36	GCGTGGCAGCTGATATCCGGCCGC	TGGGCGTC	LINC01588	0	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
22	TS37	GCCGCGGCGCGACGTGGAGCCAGC	CCCGCAAA	LINC01588	0.5	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
23	TS38	GTGCTCCCCAGCCCAAACCGCCGC	GGCGCGAC	LINC01588	2	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAGAGGGCA
24	TS41	GTCAGATTGGCTTGCTCGGAATTG	CCAGCCAA	AGA	3	AGA_TIDE1	GGCATAAGGAAATCGAAGGTC	CATGTCCTCAAGTCAAGAACAAG
25	TS44	GCTGGGTGAATGGAGCGAGCAGCG	TCTTCGAG	VEGFA	3	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
26	TS45	GTCCTGGAGTGACCCCTGGCCTTC	TCCCCGCT	VEGFA	7.4	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
27	TS46	GATCCTGGAGTGACCCCTGGCCTT	CTCCCCGC	VEGFA	6	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
28	TS47	GTGTGTCCCTCTCCCCACCCGTCC	CTGTCCGG	VEGFA	23.1	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
29	TS48	GTTGGAGCGGGGAGAAGGCCAGGG	GTCACTCC	VEGFA	2	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
30	TS49	GCGTTGGAGCGGGGAGAAGGCCAG	GGGTCACT	VEGFA	4	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAATTCCAGCACCGAGCGC
31	TS50	GTACCCTCCAATAATTTTGGCTGGC	AATTCCGA	AGA	6	AGA_TIDE1	GGCATAAGGAAATCGAAGGTC	CATGTCCTCAAGTCAAGAACAAG
32	TS51	GATAATTTGGCTGGCAATTCCGAG	CAAGCCAA	AGA	4.5	AGA_TIDE1	GGCATAAGGAAATCGAAGGTC	CATGTCCTCAAGTCAAGAACAAG
33	TS58 (DS11)	GCAGGGGCCAGGTGTCCTTCTCTG	GGGGCCTC	VEGFA	5	VEGF_TIDE4	ACACGGGCAGCATGGGAATAGTC	GCTAGGGGAGAGTCCCACTGTCCA
34	TS59 (DS12)	GAATGGCAGGCGGAGGTTGTACTG	GGGGCCAG	VEGFA	11.5	VEGF_TIDE5	CCTGTGTGGCTTTGCTTTGGTCG	GTAGGGTGTGATGGGAGGCTAAGC
35	TS60 (DS13)	GAGTGAGAGAGTGAGAGAGACA	CGGGCCAG	VEGFA	3	VEGF_TIDE5	CCTGTGTGGCTTTGCTTTGGTCG	GTAGGGTGTGATGGGAGGCTAAGC
36	TS61 (DS14)	GTGAGCAGGCACCTGTGCCAACAT	GGGCCCGC	VEGFA	3.5	VEGF_TIDE5	CCTGTGTGGCTTTGCTTTGGTCG	GTAGGGTGTGATGGGAGGCTAAGC
37	TS62 (DS15)	GCGTGGGGGCTCCGTGCCCCACGC	GGGTCCAT	VEGFA	3.4	VEGF_TIDE6	GGAGGAAGAGTAGCTCGCCGAGG	AGACCGAGTGGCAGTGACAGCAAG
38	TS63 (DS16)	GCATGGGCAGGGGCTGGGGTGCAC	AGGCCCAG	VEGFA	16	VEGF_TIDE7	AGGGAGAGGGAAGTGTGGGGAAGG	GTCTTCCTGCTCTGTGCGCACGAC
39	TS64	GAAAATTGTGATTTCCAGATCCAC	AAGCCAA	FANCJ	7	Fancl)TIDE5	GTTGGGGGCTCTAAGTTATGTAT	CTTCATCTGTATCTTCAGGATCA

40	TS65	AGCAGAAAAAATTG TGATTTCC	AGAT CCAC	FANCJ	0	Fancl_ TIDE5	GTTGGGGGCTCTAA GTTATGTAT	CTTCATCTGTATC TTCAGGATCA

[0325] 转染后72小时,收获细胞,然后进行gDNA提取和靶基因座的选择性扩增。通过分解追踪插入删除(TIDE)分析用于测量每个基因座的插入删除率。观察到Nme2Cas9进行的有效编辑,即使插入删除率根据靶序列和基因位点而有显著变化。图38A。此外,Nme2Cas9对与治疗相关位点(如CYBB(突变引起X连锁慢性肉芽肿病)和AGA(突变引起天冬氨酰糖尿症)的靶点附近/附近的靶位点的亲和力表明Nme2Cas9具有治疗潜力。此外,可以通过增加Nme2Cas9质粒的数量来提高编辑效率。图39A。综上所述,这些结果表明,可以构建Nme2Cas9来选择性编辑HEK293T细胞中的特定靶基因组位点。

[0326] 除了HEK293T细胞外,还在其他几种哺乳动物细胞中测定了Nme2Cas9的基因编辑效率,所述其他哺乳动物细胞包括人白血病K562细胞,人骨肉瘤U2OS细胞和小鼠肝肝癌Hepa1-6细胞。创建表达Nme2Cas9的慢病毒构建体,并在SFFV启动子的控制下转导K562细胞以稳定表达Nme2Cas9。与未经处理的细胞相比,这种稳定的细胞系在生长和形态上没有显示任何显著差异,这表明Nme2Cas9在稳定表达时没有毒性。这些细胞用表达靶向几个靶位点的sgRNA的质粒瞬时电穿孔,并在七十二(72)小时后通过TIDE分析插入/缺失速率。在测试的3个位点观察到有效的编辑,证明了Nme2Cas9在K562细胞中发挥功能的能力。对于Hepa1-6细胞,使用类似于上述HEK293T转导的技术共转染编码Nme2Cas9和sgRNA的质粒。这些数据还表明,Nme2Cas9有效编辑了该小鼠细胞系中的Pcsk9和Rosa26位点。图38B。

[0327] 先前的工作表明,Cas9s的核糖核蛋白(RNP)递送而不是质粒转染可能是某些基因组编辑应用程序的替代选择。例如,与质粒递送相比,RNP电穿孔可以显著降低SpyCas9的脱靶作用。Kim et al.,Genome Research 24:1012-1019(2014)。为了测试Nme2Cas9是否通过RNP递送发挥功能,将带有His标记的Nme2Cas9与三(3)个核定位信号(NLSs)和纯化的重组蛋白一起克隆到细菌表达构建物中。通过T7体外转录产生了靶向多个经过验证的靶位点的sgRNA。Nme2Cas9:sgRNA复合物的电穿孔可诱导靶位点成功编辑,如TIDE所检测。图38C。这些结果表明,Nme2Cas9可以作为质粒递送,或作为RNP复合体。总体而言,这些结果表明Nme2Cas9在具有不同递送方式的各种细胞类型中均具有功能。

[0328] 【4.抗CRISPR蛋白抑制】

[0329] 据报道,针对来自多种细菌物种的Nme1Cas9的五(5)个抗CRISPR(Acr)蛋白家族在体外和在人细胞中均抑制Nme1Cas9。(Pawluk et al.,2016,Lee et al.,mBio,印刷中)。考虑到Nme1Cas9和Nme2Cas9之间的高度序列同一性,似乎这些Acr家族中的至少某些物种也可能会抑制Nme2Cas9。全部五个Acr家族均进行重组表达,纯化,并测试了Nme2Cas9体外切割靶序列的能力(10:1Acr:Cas9摩尔比)。作为阴性对照,使用了大肠杆菌(E.coli)中IE CRISPR系统的抑制剂(AcrE2)。不出所料,所有Arc家庭都抑制了Nme1Cas9,而AcrE2却没有。特别是Acrs-IIC1$_{Nme}$,-IIC2$_{Nme}$,-IIC3$_{Nme}$-IIC4$_{Hpa}$抑制Nme2Cas9基因编辑活性。图40A上。

[0330] 引人注目的是,AcrIIC5$_{Smu}$即使在10倍过量的情况下也无法在体外抑制Nme2Cas9,这表明它可能通过与PID相互作用抑制Nme1Cas9。为了进一步证实这一点,使用NmeCas9的杂交版本(例如,具有Nme2Cas9的PID的Nme1Cas9)进行了相同的体外裂解测定。由于该杂交体的活性降低,因此使用了较高浓度(~30倍)的Cas9,以实现类似的裂解,同时保持了10:1的Cas9:Acr摩尔比。与最初的结果一致,未观察到AcrIIC5$_{Smu}$对该蛋白嵌合体的抑制。

图41。AcrIIC5$_{Smu}$无法抑制杂合蛋白进一步表明AcrIIC5$_{Smu}$可能与Nme1Cas9的PID相互作用。

[0331] 上述体外数据表明，Acrs-IIC1$_{Nme}$，-IIC2$_{Nme}$，-IIC3$_{Nme}$和-IIC4$_{Hpa}$可用作Nme2Cas9基因组编辑的开关。为了测试这一点，如上所述在有或没有编码由哺乳动物启动子驱动的Acrs的质粒的情况下进行转染。大约转染了150ng的每种质粒(例如sgRNA:Cas9:Acr的比例为1:1:1)，因为据报道大多数ACR以这些比例抑制Nme1Cas9。(Pawluk等人，2016)。如体外实验所预期的，Acr-IIC1$_{Nme}$，-IIC2$_{Nme}$，-IIC3$_{Nme}$和-IIC4$_{Hpa}$抑制Nme2Cas9基因组编辑，而AcrIIC5$_{Smu}$没有这样做。图40B。此外，观察到Acr3Nme和Acr4Hpa的完全抑制作用低于检测水平，表明它们与AcrsIIC1$_{Nme}$和AcrIIC2$_{Nme}$相比具有较高的效价。为了进一步比较AcrIIC1$_{Nme}$和AcrIIC4$_{Hpa}$的效价，在各种条件下进行了实验。图40C，因此，AcrIIC4$_{Hpa}$是抗Nme2Cas9的高效抑制剂，浓度低至25ng:100ng Acr:Cas9抑制Nme2Cas9达4倍。总之，这些数据表明，Acr蛋白可以用作关闭基于Nme2Cas9的应用程序的开关。

[0332] 【5.Nme2Cas9超高精度】

[0333] 脱靶效应可能通过产生意想不到的突变而在离体和体内人基因治疗期间混淆治疗应用。由于野生型SpyCas9在人细胞中具有相对较高的脱靶位点，因此人们已经进行了许多努力来设计出高保真度的SpyCas9变体。相比之下，Nme1Cas9自然是超高精度的，在细胞和小鼠模型中显示出显著的保真度。先前的工作表明，杂交动力学(不是由PID决定的)可以确定Cas9的保真度，因此表明Nme2Cas9也可能是非常精确的。

[0334] 为了从经验上评估NmeCas9脱靶位点，通过测序(GUIDE-Seq)技术进行的全基因组，双链断裂的无偏鉴定用于以无偏方式确定潜在的脱靶位点。GUIDE-Seq依赖于将双链寡聚脱氧核苷酸(dsODN)整合到整个基因组的DNA双链断裂位点中。这些切割位点通过扩增和高通量测序来检测。

[0335] 野生型SpyCas9用作GUIDE-Seq的基准。特别是，SpyCas9和Nme2Cas9能够被克隆到由相同启动子驱动的相同主链中，并且由于其不重叠的PAM而用于靶向相同位点。该技术允许并排比较两种核酸酶。六(6)个双位点(DS)以NGGNCCN序列靶向VEGFA。图42A。转染后七十二(72)小时，对靶位点进行TIDE分析。Nme2Cas9在所有六个(6)位点诱导插入删除，尽管其中2个效率低，而SpyCas9在4/6个位点诱导插入删除。图42B。在这4个位点中的2个位点(DS1和DS4)上，SpyCas9诱导的插入删除比Nme2Cas9多7倍，而由N-3诱导的Nme2Cas9在DS6插入删除增加了3倍。对于GUIDE-seq，分别选择了靶DS2，DS4和DS6来确定Nme2Cas9分别比SpyCas9效率相当，效率更低或效率更高的位点处的脱靶裂解。

[0336] 除了3个双重靶位点之外，将具有30-50％插入删除率(取决于细胞类型)的TS6靶位点与小鼠Pcsk9和Rosa26基因一起进行GUIDE-Seq分析。人们认为脱靶谱将更为突出，因为已知TS6靶经过高效的基因编辑。此外，对小鼠Pcsk9和Rosa26位点的测试将揭示Nme2Cas9在不同细胞系中的保真度，以及体内基因组编辑的候选基因座。因此，对每个Cas9及其相关的sgRNA进行了转染，并制备了dsODN和GUIDE-Seq库。GUIDE-Seq分析显示了2个Cas9直系同源物的高效按靶点编辑，TIDE观察到的模式相似。对于脱靶识别，分析显示，尽管3个SpyCas9位点具有预期的高脱靶数量，但这些位点(例如，范围介于10-1000之间)。Nme2Cas9的靶配置非常醒目。具体来说，针对同一双位点的Nme2Cas9最多显示一个脱靶位点。见图42C。

[0337] 为了验证由GUIDE-seq检测到的脱靶位点，在GUIDE-seq-独立编辑后(即，没有

dsODN的共转染)，进行靶向深度测序以测量在上部脱靶位点的插入删除形成。尽管SpyCas9在测试的大多数脱靶位点都显示出可观的编辑能力(在某些情况下，其效率比相应的脱靶位点更有效)，但Nme2Cas9在单独的DS2和DS6候选脱靶位点上没有可检测到的插入删除。使用Rosa26 sgRNA，Nme2Cas9会在Hepa1-6细胞的Rosa26-OT1位点诱导∽1%的编辑，而中靶的编辑则为∽30%。图42D。

[0338] 接下来，由于SpyCas9和Nme2Cas9具有不重叠的PAM，因此要启用SpyCas9作为GUIDE-seq的基准，因此它们可以潜在地编辑侧接5'-NGGNCC的任何双位点(DS)-3'序列，可同时满足2个Cas9的绑定属性的PAM要求。这使得脱靶与结合完全相同的靶位点的sgRNA能够并排比较。使用表达每个Cas9及其各自sgRNA的匹配质粒，将28(28)个DS靶向整个人基因组中的多个基因座。质粒递送后七十二(72)小时，对每种核酸酶靶向的位点进行了TIDE分析。Nme2Cas9诱导了十九(19)个靶位点的插入删除，尽管其中4个位点的效率低(＜5%)，而SpyCas9则在二十三(23)个靶位点诱导插入删除，其中一种情况的效率＜5%。3个双重靶位点都难以通过两种核酸酶进行编辑。虽然SpyCas9总体上显然更有效，但两种酶在许多位点上具有相似的效率，并且在由两种核酸酶编辑的17个位点中的2个位点上，Nme2Cas9在这些条件下更有效。见图42E。

[0339] 值得注意的是，这个脱靶位点具有共有的Nme2Cas9 PAM(ACTCCCT)，在引导互补区的PAM远端(即种子外部)只有3个错配。见图42F。这些数据支持并加强了我们的GUIDE-seq结果，表明Nme2Cas9基因组编辑在哺乳动物细胞中具有很高的准确性。

[0340] 通过靶向扩增每个基因座，然后进行TIDE分析，比较这些方面的靶与靶的偏离。图43A。有趣的是，通过TIDE，在任一sgRNA的脱靶位点均未检测到插入删除，而观察到有效的中靶编辑。此外，与在SpyCas9的情况下观察到的结果相比，这些脱靶的读长计数可以忽略不计，这表明Nme2Cas9是高度特异性的。(图43C，分别为左与右)。为了进一步证实这些GUIDE-Seq结果，CRISPRseek被用于计算预测基因组中高度相似的2个活性最高的sgRNA的潜在脱靶位点。(Zhu等人，2014)。这些用$N_{4进行}$CX PAM和2-5错配，主要发生在PAM远端区。图43D。综上所述，这些数据表明-Nme2Cas9是哺乳动物细胞中的高保真核酸酶。

[0341] 【6.临床应用】

[0342] 在一个实施方式中，本发明考虑了Nme2Cas9复合物作为第一紧凑，超精确的Cas9，其具有小的非限制性PAM，用于通过AAV递送进行治疗性基因组编辑。尽管较小的，先前报道的超高精度Cas9直系同源物比本文公开的PAM具有更长的PAM，但由于给定基因中的靶位点有限(在SauCas9情况下为脱靶谱)，因而限制了它们的治疗用途。在仅可以指定特定窗口为靶或需要精确的块删除的基因座中，这一缺点更加严重。

[0343] 本文建立的多合一AAV递送平台可用于靶向任何组织中的任何基因。此外，Nme2Cas9的超高精度可精确编辑靶基因，因此可减轻由于先前观察到的脱靶活性引起的安全隐患。为此，Nme2Cas9不仅有潜力补充现有工具，而且有可能成为通过病毒递送进行治疗性基因组编辑的优选。

[0344] 此外，各种Acrs对Nme2Cas9的抑制作用表明施加于Cas9的可能的进化压力，以迅速进化特定域。具体地说，缺乏$AcrIIC5_{Smu}$对Nme2Cas9的抑制作用增加了其抑制机制是通过PID的可能性。考虑到$AcrIIC5_{Smu}$是迄今为止最有效的Nme1Cas9抑制剂，本文考虑了其中$AcrIIC5_{Smu}$可用于稳健关闭Nme1Cas9而不是Nme2Cas9的作用。这在蜂窝环境中特别有意义，

在蜂窝环境中，通过控制特定直系同源物的能力将增强多路复用。

[0345] 最后，尽管在公共数据库中有成千上万的Cas9直系同源物，但只有少数是特征。本文考虑的一些实施方式利用紧密相关的Cas9直系同原物的天然变异来分别产生2个新颖的Cas9核酸酶，即Nme2Cas9和Nme3Cas9，分别具有N4CC和N4CAAA PAM。本文提供的数据表明，即使密切相关的直系同源物也可以具有截然不同的性质。例如，这些直系同源物使用与Nme1Cas9完全相同的sgRNA，从而避免了预测tracrRNA以及确定每个直系同源物正确的间隔子区长度的困难。此外，很可能可以设计出更短，更稳定的sgRNA(例如化学修饰)，以扩展到所有3个核酸酶。这些特性可以简化基因组编辑工作，并降低与蛋白质和RNA工程相关的成本。

[0346] 对于本领域技术人员显而易见的是，本文描述的实施方式不限于Cas9，并且可以应用于其他Cas蛋白，例如Cas12和Cas13。还应当理解，Cas9的超可变性不限于PID。本文认为存在与给定的Cas9具有高度同源性但由于其他类型的选择性压力而在其他域中不同的菌株。综上所述，Nme2Cas9是一种新颖的核酸酶，可改善用于治疗性基因组编辑的当前CRISPR平台。

[0347] 【V.核苷酸递送平台】

[0348] 除了上述的AAV核苷酸递送系统，本发明还考虑了与核酸相容的几种递送系统，其提供了大致均匀的分布并且具有可控的释放速率。本发明的一些实施方式考虑了编码如本文所述II-C型Cas9-sgRNA复合物的核酸递送系统。

[0349] 以下描述了可用于创建核酸递送系统的多种不同介质。并不意图任何一种介质或载体限制本发明。注意，任何介质或载体都可以与另一种介质或载体结合。例如，在一个实施方式中，可以将附着于化合物的聚合物微粒载体与凝胶介质混合。

[0350] 本发明考虑的载体或介质包括选自明胶，胶原蛋白，纤维素酯，硫酸葡聚糖，戊聚糖多硫酸盐，几丁质，糖类，白蛋白，纤维蛋白密封剂，合成聚乙烯吡咯烷酮，聚环氧乙烷，聚环氧丙烷，聚环氧乙烷和聚环氧丙烷的嵌段聚合物，聚乙二醇，丙烯酸酯，丙烯酰胺，甲基丙烯酸酯，包括但不限于甲基丙烯酸2-羟乙酯，聚原酸酯，氰基丙烯酸酯，明胶-间苯二酚-醛类生物粘合剂，聚丙烯酸及其共聚物和其嵌段共聚物。

[0351] 【微粒子】

[0352] 本发明的一个实施方式涉及一种包含微粒的核酸递送系统。优选地，微粒包括脂质体，纳米颗粒，微球，纳米球，微胶囊和纳米胶囊。优选地，本发明考虑的一些微粒包括聚(丙交酯-共-乙交酯)，脂族聚酯，包括但不限于聚乙醇酸和聚乳酸，透明质酸，改性多糖，壳聚糖，纤维素，葡聚糖，聚氨酯，聚丙烯酸，拟聚氨基酸，与聚羟基丁酸酯有关的共聚物，聚酸酐，聚甲基丙烯酸甲酯，聚环氧乙烷，卵磷脂和磷脂。

[0353] 【脂质体】

[0354] 本发明的一个实施方式涉及能够如本文所述附着和释放核酸的脂质体。脂质体是围绕水核心的微观球形脂质双层，由亲水性分子(例如磷脂)制成。例如，脂质体可以将核酸捕获在磷脂胶束的疏水尾之间。水溶性试剂可以包埋在核心中，脂溶性试剂可以溶解在壳状双层中。脂质体具有一个特殊的特征，它们使水溶性和水不溶性化学物质可以在介质中一起使用，而无需使用表面活性剂或其他乳化剂。通过在水性介质中强力混合磷脂，脂质体可以自发形成。将水溶性化合物溶解在能够水合磷脂的水溶液中。因此，在脂质体形成时，

这些化合物被截留在脂质体中心水溶液中。脂质体壁是磷脂膜,可容纳脂溶性物质,例如油。脂质体可控制掺入化合物的释放。另外,脂质体可以用水溶性聚合物如聚乙二醇包被以增加药代动力学半衰期。本发明的一个实施方式设想了一种超高剪切技术以改善脂质体的生产,从而得到具有专门设计的结构特征的稳定的单层(单层)脂质体。脂质体的这些独特性质允许同时储存通常不混溶的化合物及其控释能力。

[0355] 在一些实施方式中,本发明涵盖阳离子和阴离子脂质体,以及具有中性脂质的脂质体。优选地,阳离子脂质体通过将材料和脂肪酸脂质体组分混合并使它们带电缔合而包含带负电荷的材料。显然,阳离子或阴离子脂质体的选择取决于最终脂质体混合物的所需pH。阳离子脂质体的实例包括脂质转染素,脂质转染胺和脂质转染剂。

[0356] 本发明的一个实施方式涉及一种包含脂质体的核酸递送系统,所述脂质体提供至少一种核酸的受控释放。优选地,能够控制释放的脂质体:(i)是可生物降解的且无毒的;(ii)携带水溶性和油溶性化合物;(iii)溶解顽固性化合物;(iv)防止化合物氧化;v)促进蛋白稳定;vi)控制水合作用;vii)通过改变双层组成来控制化合物的释放,所述双层组成例如但不限于脂肪酸链长,脂肪酸脂质组成,饱和和不饱和脂肪酸的相对量以及物理构型;viii)具有溶剂依赖性;(iv)具有pH依赖性,并且(v)具有温度依赖性。

[0357] 脂质体的组成大致分为两类。常规脂质体通常是稳定的天然卵磷脂(PC)的混合物,其可以包含合成的相同链磷脂,该磷脂可以包含或不包含糖脂。特殊脂质体可包含:(i)双极性脂肪酸,(ii)结合抗体以进行组织靶向治疗的能力;(iii)涂有诸如但不限于脂蛋白和碳水化合物的材料,(iv)多次包封和v)乳液相容性。

[0358] 脂质体可以在实验室中通过诸如但不限于超声和振动的方法容易地制备。或者,化合物递送脂质体可商购获得。例如,已知Collaborative Laboratories,Inc.制造用于特定递送要求的定制设计的脂质体。

[0359] 【微球,微粒和微胶囊】

[0360] 微球和微胶囊是有用的,因为它们具有维持大体上均匀分布,提供稳定的受控化合物释放并且生产和分配经济的能力。优选地,相关的递送凝胶或浸渍有化合物的凝胶是透明的,或者所述凝胶被着色以易于医务人员观察。

[0361] 微球是可商购的(Alkerme's:Cambridge,MA)。例如,将包含至少一种治疗剂的冻干介质在合适的溶剂中均化并喷雾以制造20至90μm范围内的微球。然后采用在纯化,封装和储存阶段保持持续释放完整性的技术。Scott et al.,"Improving Protein Therapeutics With Sustained Release Formulations",Nature Biotechnology,Volume16:153-157(1998)。

[0362] 通过使用可生物降解的聚合物对微球组合物的修饰可以提供控制核酸释放速率的能力。Miller et al.,"Degradation Rates of Oral Resorbable Implants Polylactates and Polyglycolates:Rate Modification and Changes in PLA/PGA Copolymer Ratios",J.Biomed.Mater.Res.,Vol.II:711-719(1977)。

[0363] 可替代地,使用水内干燥方法制备缓释或控释微球制剂,其中首先制备可生物降解的聚合物金属盐的有机溶剂溶液。随后,将核酸的溶解或分散介质添加到可生物降解的聚合物金属盐溶液中。核酸与可生物降解的聚合物金属盐的重量比可为例如约1:100000至约1:1,优选约1:20000至约1:500,更优选约1:10000至约1:500。接下来,将包含可生物降解

的聚合物金属盐和核酸的有机溶剂溶液倒入水相中以制备油/水乳液。然后将油相中的溶剂蒸发掉以提供微球。最后,这些微球随后被回收,洗涤和冻干。此后,可以在减压下加热微球以除去残留的水和有机溶剂。

[0364] 可用于生产与可生物降解的聚合物金属盐和核酸混合物相容的微球的其他方法是:(i)在逐渐加入凝聚剂的过程中进行相分离;(ii)水内干燥法或相分离法,其中加入了抗絮凝剂以防止颗粒团聚;以及(iii)通过喷雾干燥法。

[0365] 在一个实施方式中,本发明考虑了一种包含微球或微囊的介质,该微球或微囊能够在约1天至6个月的时间内递送核酸的受控释放。在一个实施方式中,微球或微粒可以被着色以允许执业医师能够在分配介质时清楚地看到介质。在另一个实施方式中,微球或微胶囊可以是透明的。在另一个实施方式中,微球或微粒用不透射线的荧光镜染料浸渍。

[0366] 可通过使用已知的包囊技术如离心挤出,锅包衣和空气悬浮来生产控释微胶囊。可以对此类微球和/或微胶囊进行改造以实现所需的释放速率。例如,Oliosphere® (Macromed)是控释微球系统。这些特定的微球尺寸范围在5-500μm之间,并且由可生物相容和可生物降解的聚合物组成。微球的特定聚合物组合物可以控制核酸的释放速率,从而使定制设计的微球成为可能,包括有效控制爆发效应。ProMaxx® (Epic Therapeutics,Inc.)是一种蛋白基质递送系统。该系统本质上是水性的,并且适用于标准药物递送模型。特别是,ProMaxx®是可生物侵蚀的蛋白微球,既可输送小分子药物又可输送大分子药物,并且可以根据微球大小和所需的释放特性进行定制。

[0367] 在一个实施方式中,微球或微粒包含对pH敏感的包封材料,其在小于内部肠系膜的pH的pH下稳定。内部肠系膜的典型范围是pH7.6至pH 7.2。因此,应将微胶囊的pH维持在小于7的水平。但是,如果预期pH会发生变化,则可以根据溶解微胶囊所需的不同pH标准选择pH敏感材料。因此,将针对需要溶解的pH环境选择封装的核酸,并将其保存在预先选择的pH中以保持稳定性。可用作密封剂的pH敏感材料的例子是L-100或S-100(Rohm GMBH),羟丙基甲基纤维素邻苯二甲酸酯,羟丙基甲基纤维素乙酸琥珀酸酯,邻苯二甲酸乙酸乙烯酯,邻苯二甲酸乙酸纤维素和偏苯三酸纤维素。在一个实施方式中,脂质构成微胶囊的内涂层。在这些组合物中,这些脂质可以是但不限于脂肪酸和己糖醇酐的部分酯,以及可食用的脂肪例如甘油三酸酯。Lew C.W.,Controlled-Release pH Sensitive Capsule And Adhesive System And Method。美国专利号5,364,634(通过引用并入本文)。

[0368] 在一个实施方式中,本发明涉及包含明胶或其他具有与明胶相似的电荷密度的聚合阳离子的微粒(即,聚-L-赖氨酸),并用作复合物以形成初级微粒。产生的初级微粒是具有以下成分的混合物:(i)明胶(60块,A型来自猪皮肤),(ii)4-硫酸软骨素(0.005%-0.1%),(iii)戊二醛(25%,等级1)和(iv)1-乙基-3-(3-二甲基氨基丙基)-碳化二亚胺盐酸盐(EDC盐酸盐)和超纯蔗糖(Sigma Chemical Co.,圣路易斯,密苏里州)。明胶的来源并不重要。它可以来自牛,猪,人或其他动物。通常,聚合阳离子为19,000-30,000道尔顿。然后将硫酸软骨素与硫酸钠或乙醇作为凝聚剂一起加入到复合物中。

[0369] 在形成微粒之后,将核酸直接结合至微粒的表面,或使用“桥”或“间隔子”间接地连接核酸。明胶赖氨酸基团的氨基容易被衍生以提供用于化合物直接偶联的位点。或者,诸如亲和素-生物素的间隔子(即,连接分子和靶向配体上的衍生部分)也可用于将靶向配体

间接偶联至微粒。通过由EDC盐酸盐诱导的戊二醛-间隔子交联的量来控制微粒的稳定性。控释介质还由戊二醛-间隔子交联的最终密度凭经验确定。

[0370] 在一个实施方式中，本发明涉及通过将包含纤维蛋白原或凝血酶的组合物与核酸喷雾干燥而形成的微粒。优选地，这些微粒是可溶的，并且所选择的蛋白(即纤维蛋白原或凝血酶)产生微粒的壁。因此，核酸被掺入到微粒的蛋白壁之内和之间。Heath et al., Microparticles And Their Use In Wound Therapy，美国专利号6,113,948(通过引用并入本文)。在将微粒施加到活组织上之后，纤维蛋白原和凝血酶之间的后续反应产生了组织密封剂，从而将掺入的化合物释放到紧邻的周围区。

[0371] 本领域技术人员将理解，微球的形状不必精确地为球形，而只是能够喷射或散布到手术部位内或之上(即，开放或封闭)的非常小的颗粒。在一个实施方式中，微粒由选自聚丙交酯，聚乙交酯和丙交酯/乙交酯(PLGA)，透明质酸，改性多糖和任何其他公知材料的共聚物的生物相容性和/或可生物降解的材料组成。

【实施例】

[0372] 【实施例I:构建多合一sgRNA-Nme1Cas9-AAV载体质粒】

[0373] 细菌Nme1Cas9基因已在人中表达进行了密码子优化，并在U1a普遍存在的启动子下克隆到AAV2质粒中。引导RNA在U6启动子之下。cas9基因串联包含4个核定位信号和3个HA标签序列。通过使用退火的合成寡核苷酸用SapI限制酶消化质粒，将间隔子序列插入crRNA盒中，以产生具有与SapI消化的骨架所产生的突出端相容的突出端的双链体。

[0374] 在U1a启动子的控制下的人密码子优化的Nme1Cas9基因和由U6启动子驱动的sgRNA盒被克隆到AAV2质粒主链中。NmeCas9 ORF除了三重HA表位标签外，还带有4个核定位信号-每个末端2个。该质粒可通过Addgene获得(质粒ID 112139)。见图64。通过连接到SapI克隆位点，将具有靶向Hpd，Pcsk9和Rosa26的间隔子序列的寡核苷酸插入sgRNA盒。

[0375] 在马萨诸塞大学医学院的Horae基因治疗中心进行了AAV载体的生产。简而言之，通过在HEK293细胞中的三质粒转染将质粒包装在AAV8衣壳中，并如前所述通过沉降纯化。Gao et al.,"Introducing genes into mammalian cells:viral vectors"In:Green MR, Sambrook J,editors.Molecular cloning:a laboratory manual.Volume 2.4th ed.New York:Cold Spring Harbor Laboratory Press;2012.p.1209-13。这些间隔子的脱靶特征是使用Bioconductor软件包CRISPRseek计算得出的。检索参数适用于Nme1Cas9设置:gRNA.尺寸＝24,PAM＝“NNNNGATT”,PAM.尺寸＝8,RNA.PAM.样式＝“NNNNGNNN$”,权重＝c(0、0、0、0、0、0、0.014、0、0、0.395、0.317、0、0.389、0.079、0.445、0,508、0.613、0.851、0.732、0.828、0,615、0.804、0.685、0.583)，最大错配＝6，允许错配.PAM＝7，topN＝10,000，最小分数＝0。

[0376] 【实施例II:细胞培养和转染】

[0377] 在具有10％的FBS和1％青霉素/链霉素(Gibco)的DMEM中，在具有5％的CO_2的37℃培养箱中，培养小鼠Hepa1-6肝癌细胞。将人HEK293T细胞和PLB985细胞分别在DMEM和RPMI培养基中培养。两者均补充有10％的FBS和1％青霉素/链霉素(Gibco)。使用Lipofectamine LTX进行Hepa 1-6细胞的瞬时转染，而将Polyfect转染试剂(Qiagen)用于HEK293T细胞。对于瞬时转染，转染前24小时在24孔板中每孔培养大约1×10^5个细胞。根据制造商的规程，使

用带有Plus Reagent(ThermoFisher)的Lipofectamine LTX,用500ng多合一sgRNA-Nme1Cas9-AAV质粒转染每个孔。根据生产商的指南,在24孔板中用400ng表达Nme1Cas9和sgRNA的多合一质粒转染HEK293T细胞(例如,Psck9和Rosa26)。

[0378] 将所有细胞系维持在具有5%的CO_2的37℃培养箱中。在含有10%的FBS和1%青霉素/链霉素(Gibco)的DMEM中培养小鼠Hepa1-6肝癌和HEK293T细胞。K562细胞在相同条件下生长,但使用IMDM。IMR-90细胞在EMEM和10%的FBS中培养。最后,HDFa细胞在DMEM和20%的FBS中生长。

[0379] 【实施例III:Nme1Cas9的表达和纯化】

[0380] 将Nme1Cas9克隆到含有T7启动子,随后是6×His-标签,然后是烟草蚀纹病毒(TEV)蛋白酶切割位点的pMCSG7载体中。将该构建体转化到大肠杆菌(E.coli)的Rosetta2 DE3菌株中,并表达了Nme1Cas9。简而言之,细菌培养物在37℃下生长直至OD_{600}为0.6。此时,温度降低至18℃,然后加入1mM异丙基β-D-1-硫代半乳糖吡喃糖苷(IPTG)。使细胞生长过夜,然后收获以进行纯化。

[0381] Nme1Cas9的纯化分3个步骤进行:镍亲和层析法,阳离子交换层析法和尺寸排阻层析法。这些的详细流程可以在以前的出版物中找到(Jinek等,Science,337,816-821,2012)。

[0382] 【实施例IV:Nme1Cas9的核糖核蛋白(RNP)递送】

[0383] 使用Neon转染系统(ThermoFisher)进行Nme1Cas9的RNP递送。在缓冲液R中混合约20皮摩尔的Nme1Cas9和25皮摩尔的sgRNA,并在室温下温浴20～30分钟。然后将这种预组装的复合物与50,000～100,000个细胞混合,并使用10μl的Neon端进行电穿孔。电穿孔后,将细胞铺板于含有适当培养基但不含抗生素的24孔板中。

[0384] 【实施例V:从细胞和组织中分离DNA】

[0385] 在转染后72小时,根据制造商的方案由DNeasy®血液和组织试剂盒(Qiagen)从细胞分离基因组DNA。处死小鼠,肝组织收获后10天,高压注射或50天后尾部静脉给予载体,和用DNeasy®血液和组织试剂盒(Qiagen)根据制造商的方案分离基因组DNA。

[0386] 【实施例VI:插入删除分析】

[0387] 使用50ng基因组DNA,用基因组位点特异性引物和High Fidelity® 2×PCR Master Mix(New England Biolabs)进行PCR扩增。对于TIDE分析,使用QIAquick® PCR纯化试剂盒(Qiagen)纯化30μl的PCR产物,并进行Sanger测序。如前所述,使用TIDE网络工具(tide-calculator.nki.nl/)获得插入删除值。Brinkman et al.,Nucl.Acids Res.(2014)。

[0388] 对于T7核酸内切酶I(T7EI)测定,将10μl的PCR产物杂交并在1×NEB缓冲液2中用0.5μl的T7核酸内切酶I(New England Biolabs)处理1小时。将样品在2.5%琼脂糖凝胶上运行,并用ImageMaster-TotalLab®程序定量。如前所述计算插入删除百分率。Guschin et al.,Engineered Zinc Finger Proteins:Methods and Protocols(2010)。

[0389] 【实施例VII:脱靶分析的GUIDE-Seq】

[0390] 如前所述进行GUIDE-seq分析。Tsai et al.,Nature Biotechnology(2014),Bolukbasi et al.,Nature Methods(2015a);Amrani et al.,biorxiv.org/content/

early/2017/08/04/172650(2017)。

[0391] 简言之,对于靶向Pcsk9和Rosa26基因的2个间隔子,使用Lipofectamine LTX®Plus®试剂(ThermoFisher)向Hepa1-6细胞转染500ng多合一sgRNA-Nme1Cas9-AAV质粒和7.5皮摩尔退火GUIDE-SEQ寡核苷酸。转染后72小时,按照制造商规程,使用DNeasy®血液和组织试剂盒(Qiagen)提取基因组DNA。如前所述进行库制备,深度测序和读长分析。Tsai et al.,Nature Biotechnology(2014),Bolukbasi et al.,Nature Methods(2015a);Amrani et al.,biorxiv.org/content/early/2017/08/04/172650(2017)。

[0392] 【实施例IX:AAV载体生产】

[0393] 如先前所述,在马萨诸塞州大学医学院的Horae基因治疗中心,通过在HEK 293细胞中的三质粒转染将质粒包装在AAV8中,并通过沉降纯化。Gao GP,Sena-Esteves M.Introducing Genes into Mammalian Cells:Viral Vectors.In:Green MR,Sambrook J,eds.Molecular Cloning,Volume 2:A Laboratory Manual.New York:Cold Spring Harbor Laboratory Press,2012:1209-1313。

[0394] 【实施例X:动物,AAV载体注射和肝组织加工】

[0395] 所有动物实验均在马萨诸塞州大学医学院动物护理和使用委员会的指南下获得批准,对于水动力注射,将2.5mL的30μg不含内毒素的靶向Pcsk9的sgRNA-Nme1Cas9-AAV质粒,或PBS作为对照,通过尾静脉将其注射到9-18周龄的雌性C57BL/6小鼠中。对于AAV8载体注射,每只小鼠通过尾静脉向9-18周龄的雌性C57BL/6小鼠注射4×10^{11}个基因组拷贝。将8周龄的雌性C57BL/6NJ小鼠用于体内基因组编辑实验。对于离体实验,将发育至两细胞阶段的胚胎转移到E0.5假孕雌性小鼠的输卵管中。

[0396] 通过CO_2使小鼠安乐死并收集肝脏。将组织在4%多聚甲醛中固定过夜,然后包埋在石蜡中,切片并用苏木精和曙红(H&E)染色。

[0397] 【实施例XI:血清分析】

[0398] 在载体施用后第0、25和50天从面静脉抽取血液(~200μl)。用血清分离器(BD,目录号365967)分离血清,并在-80℃下保存直至测定。遵循制造商的规程,通过Infinity™比色终点试验(Thermo-Scientific)测量血清胆固醇水平。简而言之,在PBS中制备Data-Cal™Chemistry Calibrator的系列稀释液。在96孔板中,将2μl小鼠血清或校准品稀释液与200μl Infinity™胆固醇液体试剂混合,然后在37℃下温浴5分钟。使用Biotek Synergy®HT酶标仪在500nm测量吸光度。

[0399] 【实施例XII:发现具有超进化PID的Cas9直系同原物】

[0400] Nme1Cas9序列被BLAST比对以发现奈瑟氏球菌属(Neisseria)物种中的所有Cas9直系同源物。选择与Nme1Cas9同源性>80%的直系同源物。然后使用ClustalW2将每个PID与Nme1Cas9的PID(第820位氨基酸到第1082位)进行比对,然后选择PID中有突变簇的那些。

[0401] 在BLAST检索中使用Nme1Cas9肽序列作为查询以查找脑膜炎奈瑟氏球菌(Neisseria meningitidis)菌株中的所有Cas9直系同原物。选择与Nme1Cas9有>80%同一性的直系同源物进行研究。然后使用ClustalW2®,将所述PID与Nme1Cas9(残基820-

1082)进行比对,并选择在PID中有突变簇的那些用于进一步的分析。使用FigTree(tree.bio.ed.ac.uk/software/figtree)构建了NmeCas9直系同源物的无根系统发育树。

[0402] 【实施例XIII:Nme2和Nme3 Cas9和Acr直系同源物的克隆和纯化】

[0403] Nme2Cas9和Nme3Cas9的PID被订购为gBlocks(IDT),以使用具有6×His-标签的细菌表达质粒pMSCG7中的Gibson Assembly(NEB)替换Nme1Cas9的PID。如前所述,将构建体转化到大肠杆菌(E.coli)中,进行表达和纯化。

[0404] 简而言之,将包含各个Cas9质粒的Rosetta(DE3)细胞在37℃下生长至0.6的光密度,并通过1mM IPTG在18℃下诱导蛋白表达16小时。收获细胞,并在补充了溶菌酶和蛋白酶抑制剂混合物(Sigma)的裂解缓冲液(50mM Tris pH 7.5、500mM NaCl,5mM咪唑,1mM DTT)中通过超声处理进行裂解。

[0405] 然后使裂解物通过Ni-NTA琼脂糖柱(Qiagen),结合的蛋白用300mM咪唑洗脱,并透析到存储缓冲液(20mM HEPES pH 7.5,250mM NaCl,1mM DTT)中。对于Acr蛋白,在大肠杆菌(E.coli)BL21 Rosetta(DE3)菌株中表达了6×His标记的蛋白。在摇动的培养箱中将细胞在37℃下生长至光密度(OD_{600nm})为0.6。将细菌培养物冷却至18℃,并通过添加1mM IPTG诱导蛋白表达过夜。第二天,收获细胞并重悬于补充有1mg/mL溶菌酶和蛋白酶抑制剂混合物(Sigma)的裂解缓冲液(50mM Tris pH 7.5、500mM NaCl,5mM咪唑,1mM DTT)中,并使用与Cas9相同的流程进行纯化。通过与烟草蚀纹病毒(TEV)蛋白酶在4℃温浴过夜,去除6×His标签,以分离成功裂解的,未标记的Acrs。

[0406] 【实施例IVX:体外PAM发现测定】

[0407] 使用重叠PCR生成带有随机PAM序列的原间隔子库,其中正向引物包含10个核苷酸的随机PAM。

[0408] 将库进行凝胶纯化,并通过纯化的Cas9与体外转录的sgRNA进行体外切割反应。300nM Cas9:sgRNA复合物用于在37℃的1xNE Buffer 3.1(NEB)中裂解300nM的靶片段1小时。然后在50℃下用蛋白酶K将反应物处理10分钟,并在具有1×TAE的4%琼脂糖凝胶上运行。将切割产物纯化并进行库制备。使用Illumina NextSeq500®测序测序平台和分析。序列标识是使用R生成的。

[0409] 【实施例XV:转染和哺乳动物基因组编辑】

[0410] 使用Gibson Assembly将人源化的Nme2Cas9克隆到先前用于Nme1Cas9和SpyCas9表达的pCDest2质粒中。如先前所述(Amrani等人2018)进行HEK293T和HEK293T-TLR细胞的转染。对于Hepa1-6转染,使用Lipofectamine LTX在转染前24小时在24孔板中培养的每孔约1×10^5个细胞中转染500ng的AAV.sgRNA.Nme2Cas9质粒。对于稳定表达Nme2Cas9的K562细胞,使用10μl的Neon端用500sgRNA质粒电穿孔50,000-150,000个细胞。

[0411] 为了测量在所有细胞中插入删除,在转染后72小时收获细胞,并且使用DNaesy®血液和组织试剂盒(Qiagen)提取基因组DNA。靶向基因座通过PCR扩增,Sanger测序(GENEWIZ®)和TIDE(Brinkman等2014)分析。

[0412] 【实施例XVI:慢病毒转导K562细胞以稳定表达Nme2Cas9】

[0413] 如先前所述产生稳定表达Nme2Cas9的K562细胞。对于慢病毒生产,按照制造商的建议,使用TransIT-LT1转染试剂(Mirus Bio)将慢病毒载体与包装质粒(Addgene 12260和

12259)一起共转染6孔板中的HEK293T细胞。24小时后,从转染的细胞中吸出培养基,并换进新鲜的1mL新鲜的DMEM培养基用。

[0414] 第二天,收集含有来自转染细胞的病毒的上清液,并通过0.45μm过滤器过滤。将10μl未稀释的上清液与2.5μg聚凝胺一起用于转导6孔板中约100万个K562细胞。使用含2.5μg/mL嘌呤霉素的培养基选择转导的细胞。

[0415] 【实施例XVII:用于哺乳动物基因组编辑的RNP递送】

[0416] 对于RNP实验,使用了Neon电穿孔系统。将40皮摩尔的3×NLS Nme2Cas9以及50皮摩尔的体外转录的sgRNA在缓冲液R中组装,并使用10μl的Neon端进行电穿孔。电穿孔后,将细胞铺板在预热的24孔板中,该板包含适当的不含抗生素的培养基。电穿孔参数(电压,宽度,脉冲数)为,HEK293T细胞为1150v,20ms,2脉冲;K562细胞为1000v,50ms,1个脉冲。

[0417] 【实施例XVIII:GUIDE-Seq】

[0418] 如前所述进行了GUIDE-Seq实验,但进行了小修改(Amrani等,2018)。

[0419] 简而言之,使用Polyfect(Qiagen)用SpyCas9或Nme2Cas9靶向双重位点,用200ng Cas9、200ng sgRNA和7.5pmol退火的GUIDE-seq寡核苷酸转染HEK293T细胞。如上所述转染Hepa1-6细胞。

[0420] 在转染后72小时,用DNeasy®血液和组织试剂盒(Qiagen)根据制造商的方案提取基因组DNA。完全按照先前的描述进行库制备和测序。

[0421] 为进行分析,将与靶位点有10个错配的序列相匹配的位点视为潜在的脱靶位点。使用Bioconductor软件包GUIDEseq版本1.1.17(Zhu等人,2017)分析数据。

[0422] 【实施例XIX:靶向深度测序和分析】

[0423] 靶向深度测序用于确认GUIDE-Seq的结果,并更定量地测量插入删除率。使用两步PCR扩增为每个中靶和脱靶位点产生DNA片段。对于SpyCas9,选择了上部脱靶位置。

[0424] 在第一步骤中,基因座特异性的引物与轴承互补末端到适配体悬突普遍用2×Phusion® PCR master mix(NEB)混合,以产生片段轴承突出端。第二步,用通用正向引物和索引反向引物扩增纯化的PCR产物。

[0425] 凝胶提取全尺寸产物(长度约250bp),并使用配对末端的MiSeq进行测序。MiSeq数据分析的执行完全如先前所述(Amrani 2018)。

[0426] 【实施例XX:使用CRISPRseek进行脱靶分析】

[0427] 使用Bioconductor程序包CRISPRseek对TS25和TS47进行了脱靶分析。

[0428] 进行了小更改,以适应未与SpyCas9共享的Nme2Cas9的特征。具体而言,使用以下变化:gRNA.尺寸=24,PAM="NNNNCC",PAM.尺寸=6,RNA.PAM.样式="NNNNCN",收集了错配少于6个的脱靶位点。根据错配的数量和位置选择了最可能出现的脱靶位点。来自每个相应sgRNA靶向细胞的gDNA用于扩增每个脱靶基因座,并通过TIDE分析。

[0429] 【实施例XXI:体内AAV8.Nme2Cas9递送和肝组织加工】

[0430] 所有动物手术程序均由马萨诸塞州大学医学院的机构动物护理和使用委员会(IACUC)审查和批准。

[0431] 对于AAV8载体注射,通过尾静脉向8周龄的雌性C57BL/6小鼠注射每只小鼠4×10^{11}个基因组拷贝(靶向Pcsk9或Rosa26)。施用载体28天后处死小鼠,并收集肝组织用于分析。

将肝脏组织在4％福尔马林中固定过夜，然后包埋在石蜡中，切片并用苏木精和曙红(H&E)染色。在注射后0、14和28天从面静脉抽血，并使用血清分离器(BD，目录号365967)分离血清，并在-80℃下保存直至测定。按照制造商的规程并按照先前的描述(Ibraheim et al，2018)使用Infinity™比色终点测定法(Thermo-Scientific)测量血清胆固醇水平。

[0432] 【实施例XXII:动物和肝组织的加工】

[0433] 对于流体动力注射，通过尾静脉将2.5mL的30μg靶向Pcsk9的无内毒素的AAV-sgRNA-hNme1Cas9质粒或2.5mL PBS注射到9至18周龄的雌性C57BL/6小鼠中。10天后对小鼠实施安乐死，并收集肝组织。对于AAV8载体注射，使用靶向Pcsk9或Rosa26的载体，通过尾静脉给12至16周龄的C57BL/6小鼠每只小鼠注射4×10^{11}个基因组拷贝。在施用载体后14和50天处死小鼠，并收集肝组织用于分析。

[0434] 对于Hpd靶向，通过尾静脉将2mL PBS或2mL 30μg无内毒素的AAV-sgRNA-hNme1Cas9质粒给药至15至21周龄的1型酪氨酸血症FAH基因敲除小鼠(Fahneo)。编码的sgRNA靶向外显子8(sgHpd1)或外显子11(sgHpd2)上的位点。当指示时，给HT1小鼠喂食10mg/L NTBC(2-(2-硝基-4-三氟甲基苯甲酰基)-1,3-环己二酮)(Sigma-Aldrich，目录号PHR1731-1G)。在这些实验中使用了2个性别。注射后将小鼠在NTBC水上维持7天，然后换成普通水。每1-3天监测一次体重。当从NTBC处理中去除了20％的体重后，注射PBS的对照小鼠垂死时被处死。

[0435] 根据我们的方案对小鼠实施安乐死，将肝组织切成薄片，并将碎片保存在-80℃。将一些肝脏组织在4％福尔马林中固定过夜，包埋在石蜡中，切片并用苏木精和曙红(H&E)染色。

[0436] 【实施例XXIII:蛋白印迹】

[0437] 将肝组织级分研磨并重悬于150μl RIPA裂解缓冲液中。按照制造商的方案，通过Pierce™ BCA蛋白检测试剂盒(Thermo-Scientific)估算总蛋白含量。将总共20μg来自组织的蛋白或2ng重组小鼠蛋白原的转化酶9/PCSK9蛋白(R&d系统，9258-SE-020)加载到4-20％的Mini-Rotean® TGX™预制凝胶(Bio-Rad)。将分离的条带转移至PVDF膜上，并在室温下用5％的Blocking-Grade Blocker溶液(Bio-Rad)封闭2h。将膜与兔抗GAPDH(Abcam ab9485，1:2000)或山羊抗PCSK9(R&D Systems AF3985，1:400)抗体在4℃温浴过夜。将膜在TBST中洗涤五次，并与辣根过氧化物酶(HRP)偶联的山羊抗兔(Bio-Rad 1,706,515，1:4000)和驴抗山羊(R&D Systems HAF109，1:2000)二抗在室温下温浴2h。将膜在TBST中洗涤五次，并使用M35A X-OMAT处理器(Kodak)用Clarity™蛋白印迹ECL底物(Bio-Rad)显影。

[0438] 【实施例XXIV:体液免疫反应】

[0439] 按照制造商的方案并作一些修改，通过ELISA(Bethyl；小鼠IgG1ELISA试剂盒，E99-105)测量对Nme1Cas9的体液IgG1免疫应答。简而言之，进行Nme1Cas9和SpyCas9的表达和三步纯化。悬浮在1×包被缓冲液(Bethyl)中的总共0.5μg重组Nme1Cas9或SpyCas9蛋白用于包被96孔板(Corning)，并在4℃振荡温浴12小时。使用1×洗涤缓冲液将孔洗涤3次，同时摇动5分钟。在室温下将板用1×BSA封闭溶液(Bethyl)封闭2h，然后洗涤3次。用PBS以1:40稀释血清样品，并一式两份地加到每个孔中。将样品在4℃温浴5小时后，将板洗涤3次，每次5分钟，并将100μl生物素化的抗小鼠IgG1抗体(Bethyl；在1×BSA封闭溶液中1:100,000)添加到每个孔中。在室温下温浴1小时后，将板洗涤4次，然后向每个孔中加入100μl的TMB底

物。使板在室温下在黑暗中显影20分钟，然后每孔添加100μl的ELISA终止溶液。黄色溶液显影后，使用BioTek Synergy® HT酶标仪在450nm处记录吸光度。

[0440] 【实施例XXV：合子温浴和转染】

[0441] 【小鼠品系和胚胎收集】

[0442] 所有动物实验都是在马萨诸塞大学医学院的机构动物护理和使用委员会(IACUC)的引导下进行的。C57BL/6NJ(货号005304)小鼠获自The Jackson Laboratory。所有动物保持12小时光照周期。在观察到交配栓的那一天的光周期的中间被认为是妊娠的第0.5天(E0.5)。通过用镊子撕开壶腹并在含有透明质酸酶的M2培养基中温育以除去卵丘细胞，收集E0.5合子。

[0443] 【体内AAV8.Nme2Cas9+sgRNA递送和肝组织加工】

[0444] 对于AAV8载体注射，通过尾静脉向8周龄的雌性C57BL/6NJ小鼠每只小鼠注射了4×10^{11}个基因组拷贝，将sgRNA靶向Pcsk9或Rosa26中经过验证的位点。在施用载体28天后处死小鼠并收集肝组织用于分析。将肝脏组织在4％福尔马林中固定过夜，包埋在石蜡中，切片并用苏木精和曙红(H&E)染色。在注射后0、14和28天从面静脉抽血，并使用血清分离器(BD，目录号365967)分离血清，并在-80℃下保存直至测定。按照制造商的规程并如前所述，使用Infinity™比色终点测定法(Thermo-Scientific)测量血清胆固醇水平。Ibraheim et al.,"All-in-One Adeno-associated Virus Delivery and Genome Editing by Neisseria meningitidis Cas9 in vivo"Genome Biology 19:137(2018)。

[0445] 对于抗PCSK9蛋白印迹，将来自组织的40μg蛋白或2ng重组小鼠PCSK9蛋白(R&D Systems，9258-SE-020)加载到MiniProtean® TGX™预制凝胶(Bio-Rad)上。将分离的条带转移至PVDF膜上，并用5％的Blocking-Grade Blocker®溶液(Bio-Rad)在室温下放置2小时。接下来，将膜与兔抗GAPDH(Abcam ab9485，1:2,000)或山羊抗PCSK9(R&D Systems AF3985，1:400)抗体温浴过夜。将膜在TBST中洗涤，并与缀合辣根过氧化物酶(HRP)的山羊抗兔(Bio-Rad 1706515，1:4,000)和驴抗山羊(R&D Systems HAF109，1:2,000)二抗在室温温育2小时。再次在TBST中清洗膜，并使用M35A XOMAT处理器(Kodak)使用Clarity™蛋白印迹ECL底物(Bio-Rad)显影。

[0446] 【小鼠合子中的离体AAV6.Nme2Cas9递送】

[0447] 合子在含有AAV6.Nme2Cas9.sgTyr载体的3×10^{9}或3×10^{8}个GC的KSOM(补充钾的单纯形优化培养基，Millipore，目录号MR-106-D)中温浴5-6小时(每滴4个合子)。温浴后，将合子在M2中冲洗，然后转移到新鲜的KSOM中过夜培养。第二天，将发育至2细胞期的胚胎转移到假怀孕受体的输卵管中，使其发育至足月。

[0448] 【实施例XXVI：量化和统计分析】

[0449] 使用R.GraphPad Prism 6®对体外PAM发现数据进行了分析，以进行所有统计分析。对于使用Nme2Cas9进行的哺乳动物细胞实验，进行了3个独立的重复实验，并使用TIDE软件计算了插入删除百分率，误差条表示sem。设置TIDE参数以量化所有数字的插入删除＜20个核苷酸。对于Nme2Cas9和SpyCas9的并排比较，使用Microsoft Excel计算了平均插入删除百分率。对于小鼠体内实验，对照组和测试对象的n＝5。通过不成对的两尾t检验计算P值。

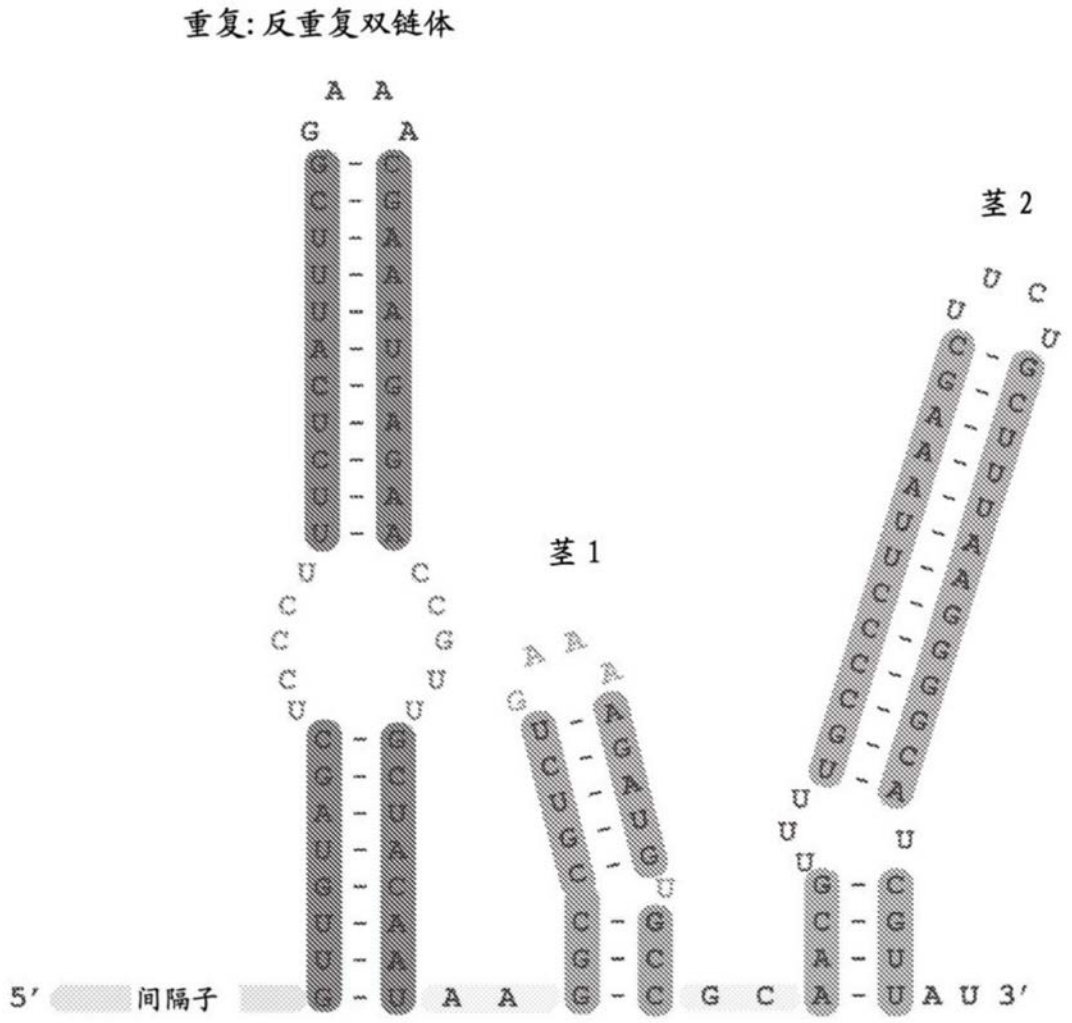


图1

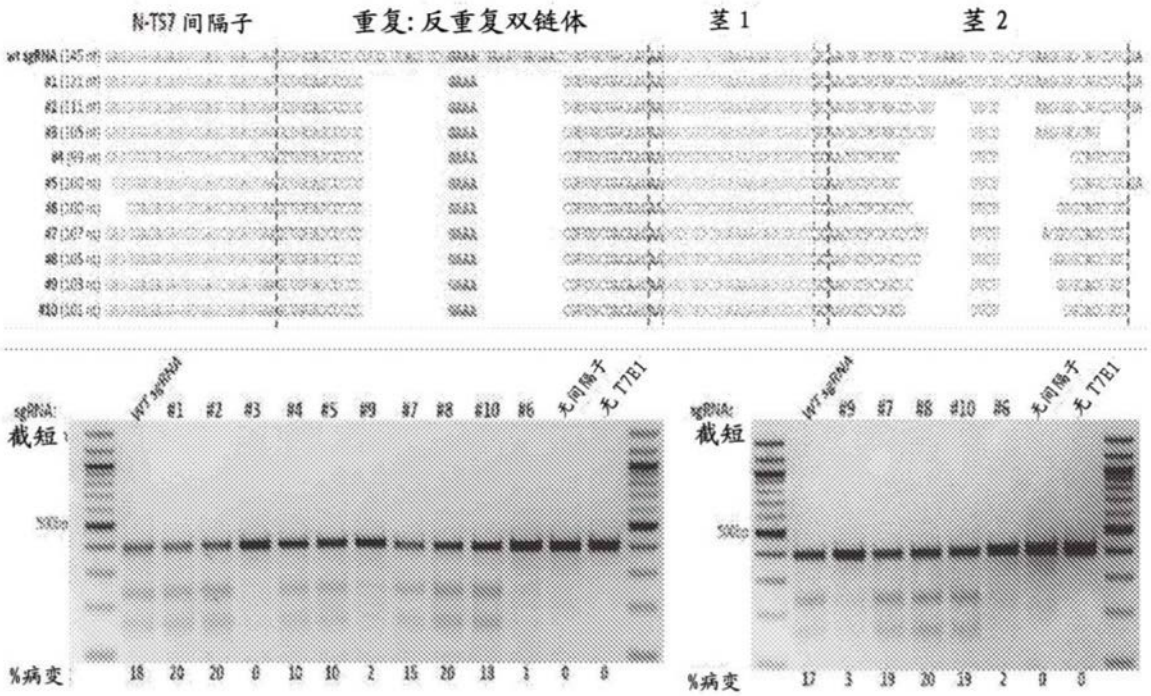

N-TS7 间隔子
重复:反重复双链体
茎 1
茎 2
截短
%病变
18 20 20 8 12 10 2 15 20 13 1 0 0
17 3 19 20 19 2 0 0

图2

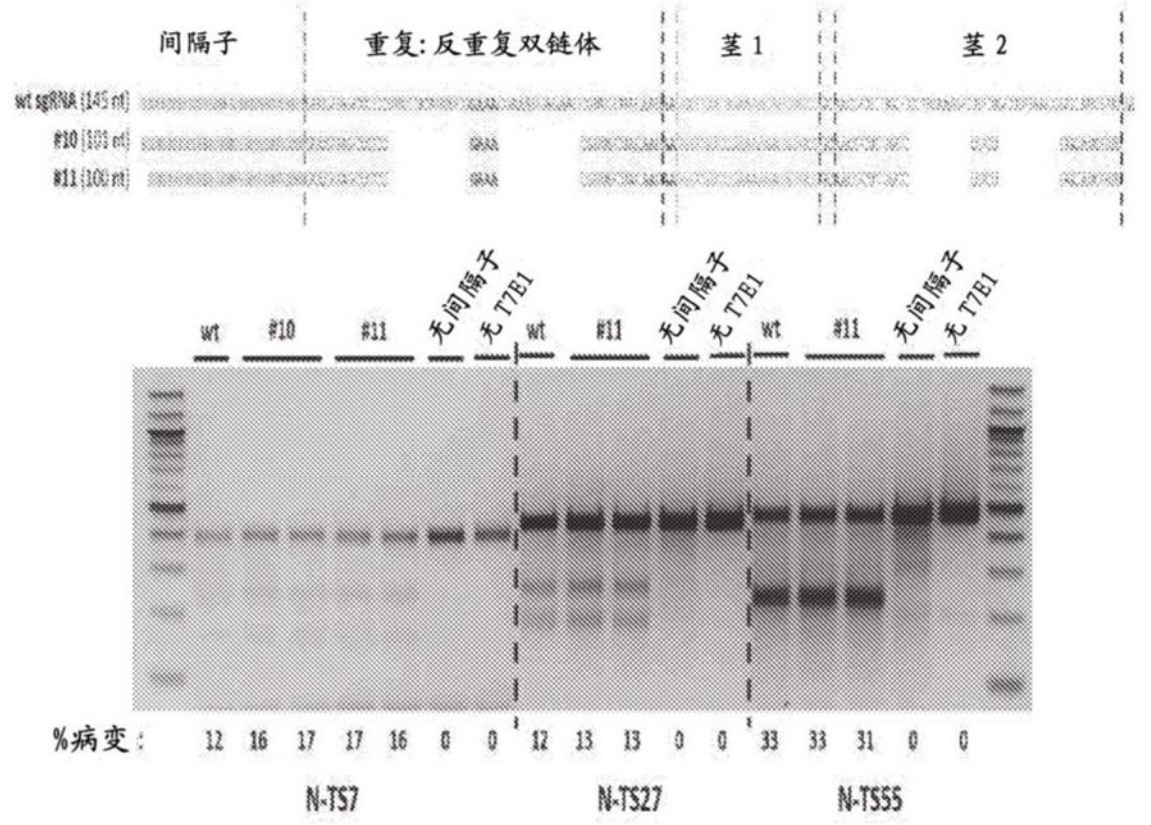

间隔子
重复:反重复双链体
茎 1
茎 2
wt
#10
#11
无间隔子
无 T7E1
%病变:
12 16 17 17 16 0 0 12 13 13 0 0 33 33 31 0 0
N-TS7
N-TS27
N-TS55

图3

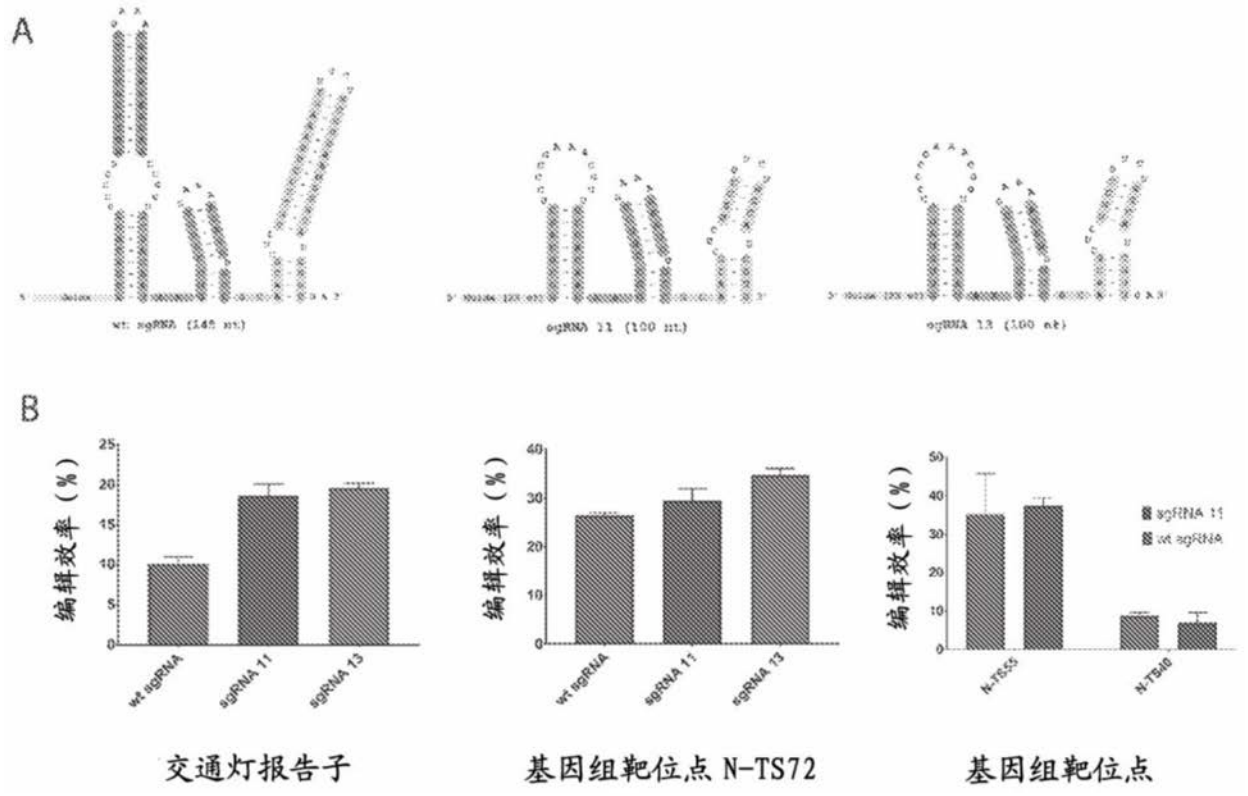
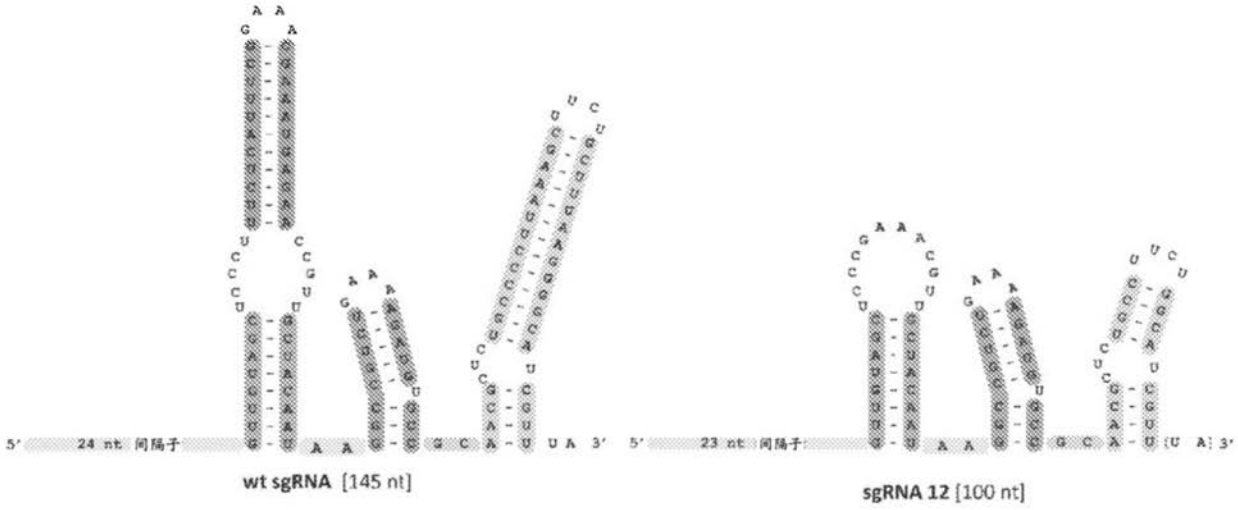

A
wt sgRNA (148 nt)
sgRNA 11 (100 nt)
sgRNA 13 (100 nt)
B
编辑效率（%）
wt sgRNA
sgRNA 11
sgRNA 13
交通灯报告子
编辑效率（%）
wt sgRNA
sgRNA 11
sgRNA 13
基因组靶位点 N-TS72
编辑效率（%）
sgRNA 11
wt sgRNA
N-TS55
N-TS40
基因组靶位点

图4

C


5' 24 nt 间隔子
wt sgRNA [145 nt]
5' 23 nt 间隔子
sgRNA 12 [100 nt]

D

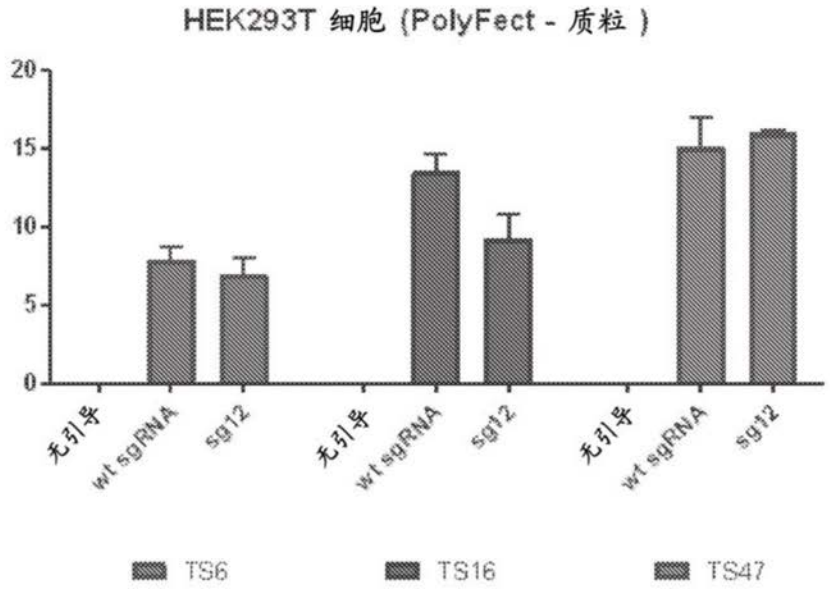

HEK293T 细胞 (PolyFect - 质粒)
20
15
10
5
0
无引导
wt sgRNA
sg12
TS6
TS16
TS47

图4续

E

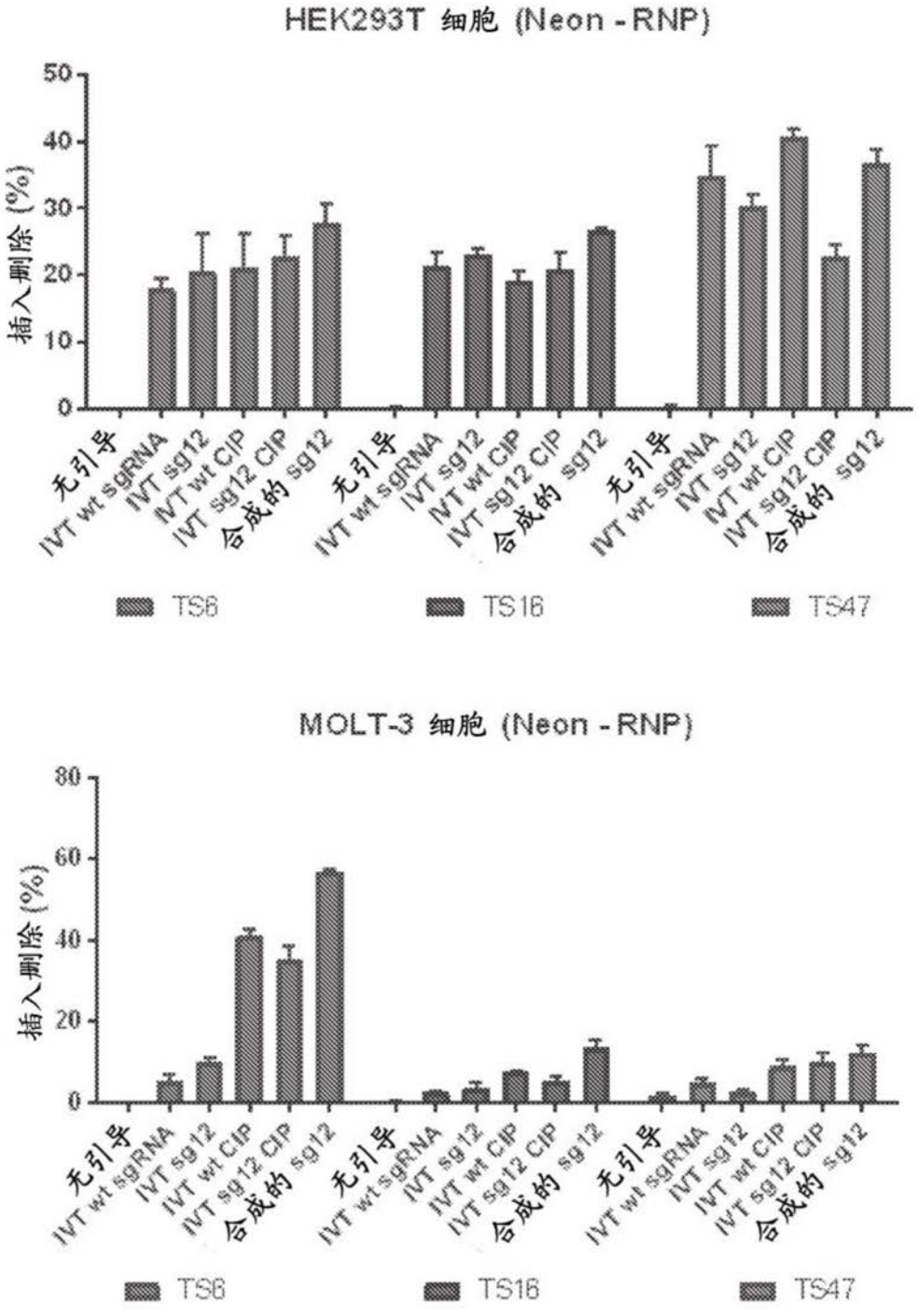

HEK293T 细胞 (Neon - RNP)
插入删除 (%)
50
40
30
20
10
0
无引导
IVT wt sgRNA
IVT sg12
IVT wt CIP
IVT sg12 CIP
合成的 sg12
TS6
TS16
TS47
MOLT-3 细胞 (Neon - RNP)
插入删除 (%)
80
60
40
20
0
无引导
IVT wt sgRNA
IVT sg12
IVT wt CIP
IVT sg12 CIP
合成的 sg12
TS6
TS16
TS47

图4续

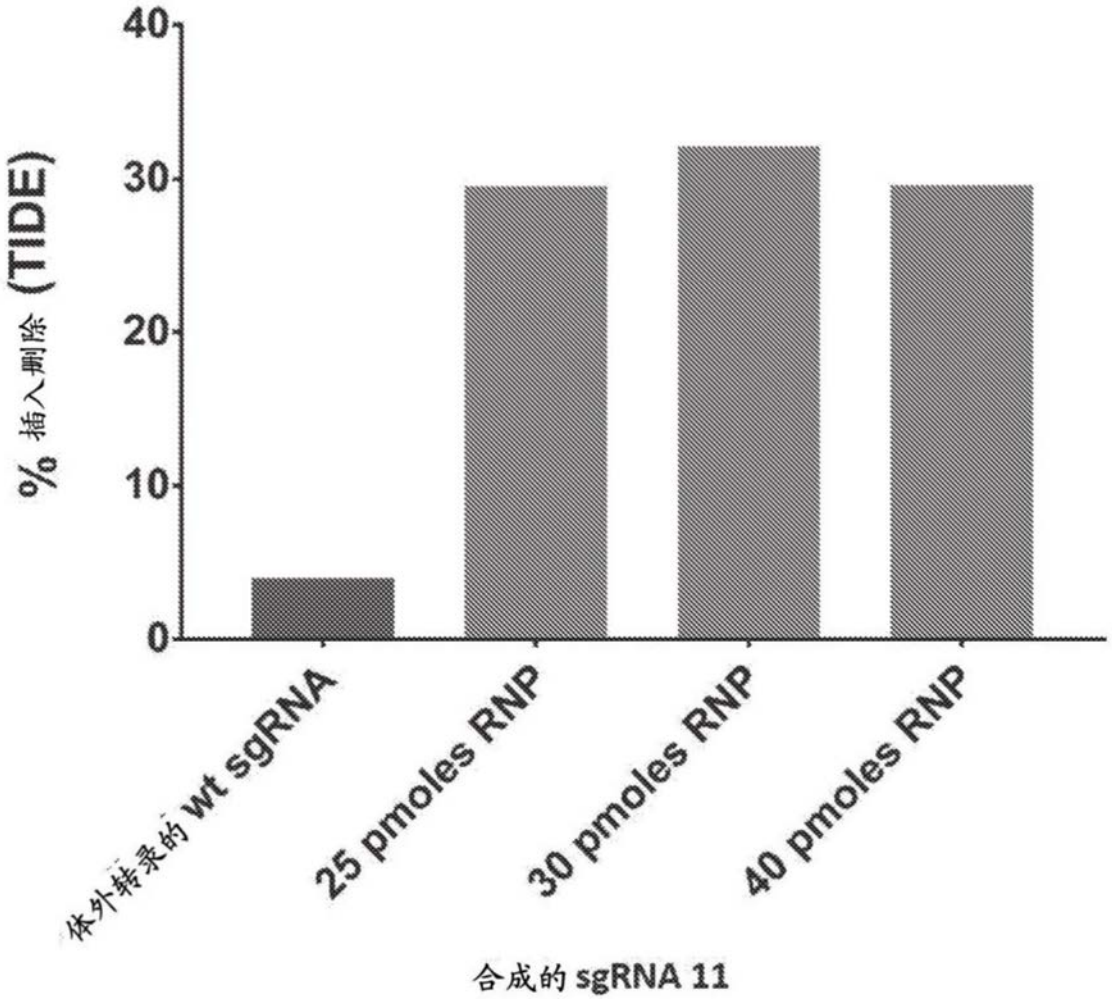

40
30
20
10
0
% 插入删除 (TIDE)
体外转录的 wt sgRNA
25 pmoles RNP
30 pmoles RNP
40 pmoles RNP
合成的 sgRNA 11
基因组靶位点 N-TS72 [在 CYBB 内含子 2 中]

图5

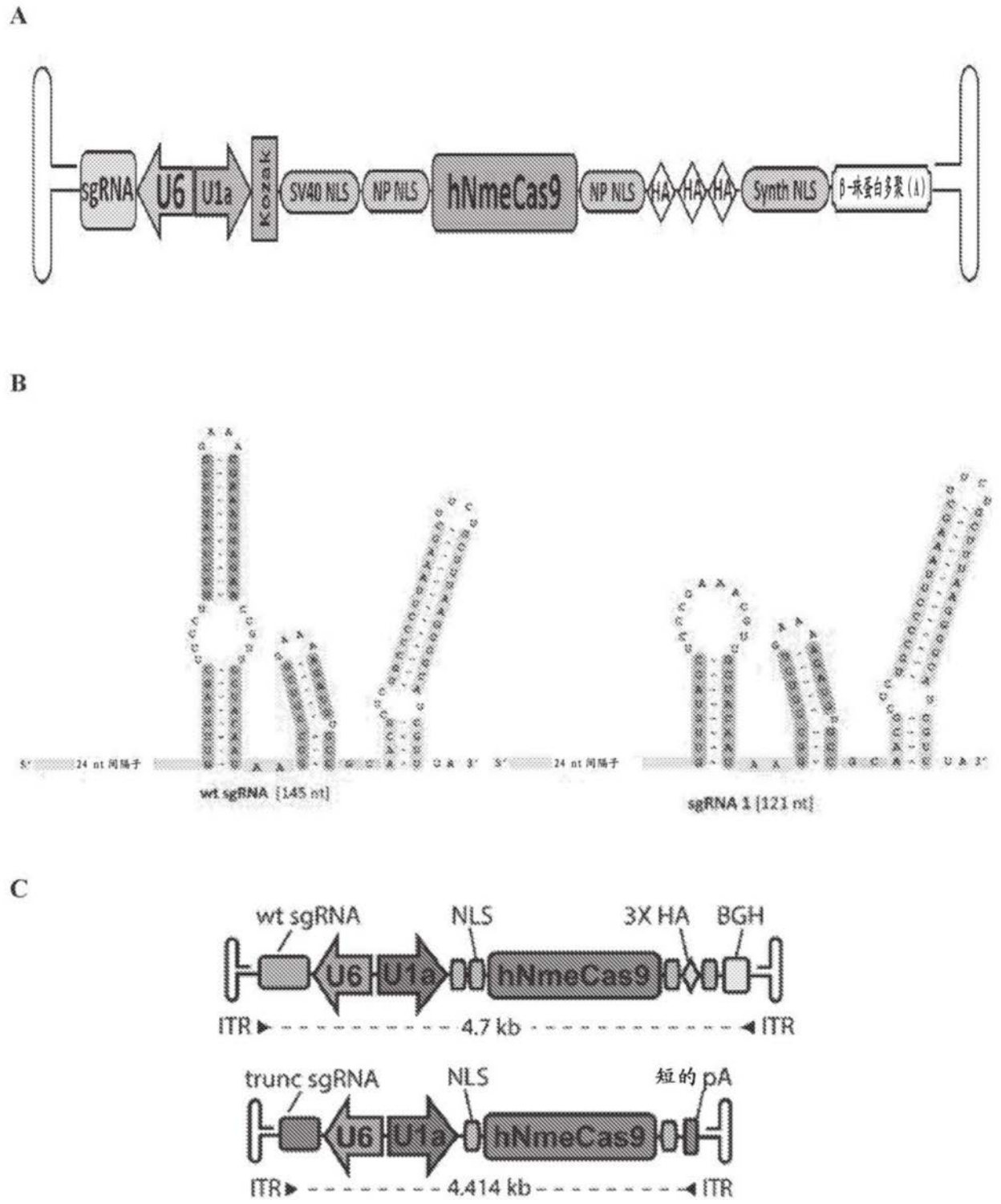

A
sgRNA
U6
U1a
Kozak
SV40 NLS
NP NLS
hNmeCas9
NP NLS
HA
HA
HA
Synth NLS
β-球蛋白多聚(A)
B
24 nt 间隔子
wt sgRNA (145 nt)
24 nt 间隔子
sgRNA 1 (121 nt)
C
wt sgRNA
NLS
3X HA
BGH
U6
U1a
hNmeCas9
ITR
4.7 kb
ITR
trunc sgRNA
NLS
短的 pA
U6
U1a
hNmeCas9
ITR
4.414 kb
ITR

图6

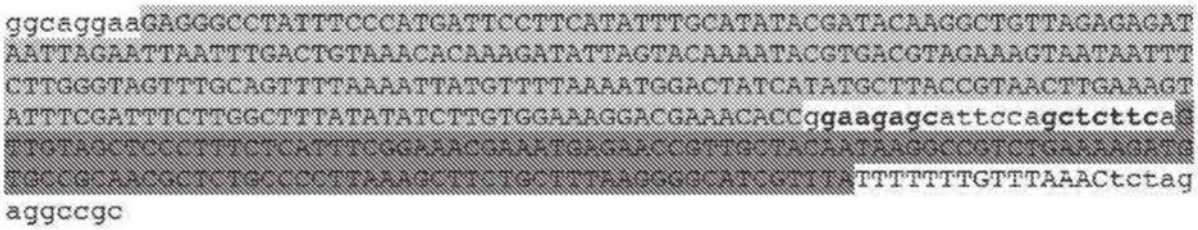

ggcaggaaGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGAT
AATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTT
CTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGT
ATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCggaagagcattccagctcttcaG
TTGTAGCTCCCTTTCTCATTTCGGAAACGAAATGAGAACCGTTGCTACAATAAGGCCGTCTGAAAAGATG
TGCCGCAACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGCATCGTTTATTTTTTTGTTTAAACtctag
aggccgc

图7

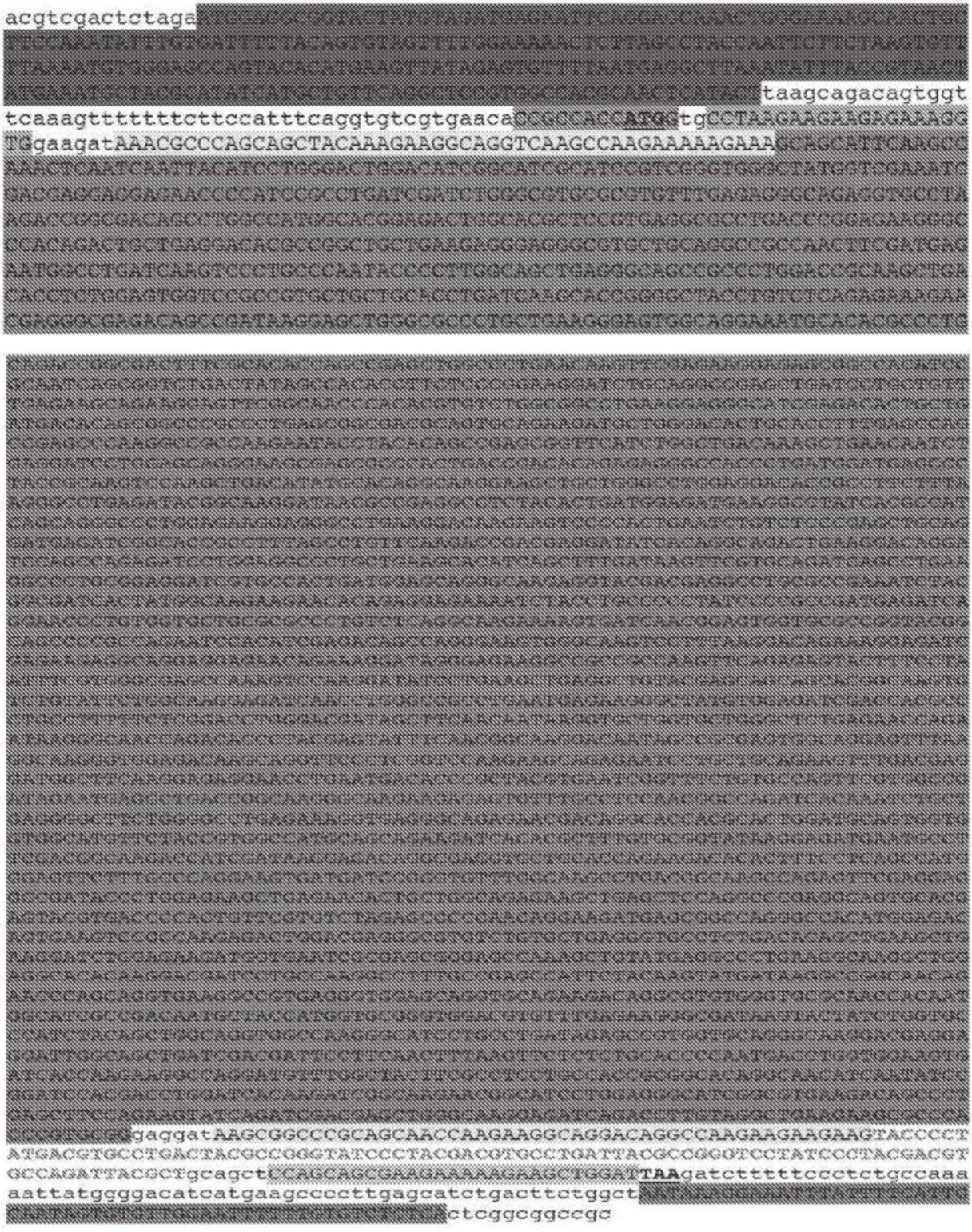


图8

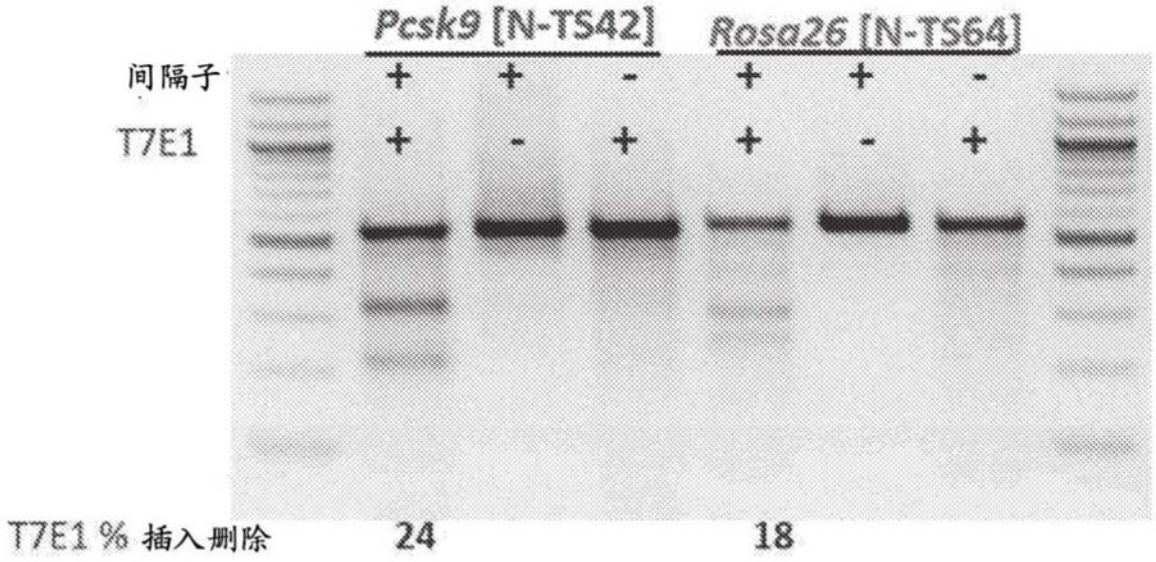


图9

Pcsk9

TGGCGTGATCTCCCGGCCCCCAGGCGTC*CAGTACCCACACCCCAGAA*GGCTTCCACCTTCACGTGGACGC
GCAGGCTGCCGGTG*GGCTCCCGTTCTCTCTCTCT*TTCTGAGGCTAGAGGACTGAGCCAGTCCTTGGCTCC
CCAGAGACATCACGGCCCGCAGCCCCGGAGCCAAGTGCCCCGAGTCCCAGGCGTCCATGTCCTTCCCGAG
GCCGCGCGCACCTCTCCTCGCCCCGATGGGCACCCACTGCTCTGCGTGGCTGCGGTGGCCGCTGTTGCCG
CTGTTGCCGCCGCTGCTGCTGCTGTTGCTGCTACTGTGCCCCACCGGCGCTGGTGCCCAGGACGAGGATG
GAGATTATGAAGAGCTGATGCTCGCCCTCCCGTCCCAGGAGGATGGCCTGGCTGATGAGGCCGCACATGT
GGCCACCGCCACCTTCCGCCGTTGCTCCAAGGTATGGGTGCCAGGCAACGGCGTTGCTTTGGGGTTGGGG
TGATGCTCTTCGGGGGTCTTCTCTGCTCATCTAGCCGTCTGGTGGTCTC*TAAGTGCAGCCCTGAGGTG*CG
GGAGGCGAGGGCAAGACTTAGTGCTCAGCTGCACCTTGTGGCACAGAGTGATGGGGGAGGCCACGTGCTA
AAGGCACTGCGGGGCTTGGTTCCAAAAGTGTGAGGCGGGGAGCGGGCTACCAGTGTGGTCATGCAGAAAA
CGTGTCCTCCGAAGTAAAGTGGCATCGGGAGGCTGAGAACTCTAGTGGCACATCTTTCTCAACTGGTCAT
CCAGCAGTCATCCTGGGTGCCTGTTGGGGTT*CTGTAGGCCTCGATTTAGCG*GAAGGGTGTCAGGG

Rosa26

GGAAAGGGAAAATGCCAATGCTCTGTC*TAGGGGTTGGATAAGCCAGT*ATAATAAATGAAAATGGGGCTAA
AATGAGTGTTCTAAAATACCTTTTGATAAGGCTGCAGAAGGAGCGGGAGAAATGGATATGAAGTACTGGG
CTCTTTAAAAATGATTAAAATTCTGCTTACATAGTCTAACTCGCGACACTGTAATTTCATACTGTAGTAA
GGATCTCAAGCAGGAGAGTATAAAACTCGGGTGAGCATGTCTTTAATCTACCTCGATGGAAAATACTCCG
AGGCGGATCACAA*GCAATAATAACCTCTAGTTTTGCTGC*ATAAAACCCCAGATGACTACCTATCCTCCCA
TTTTCCTTATTTGCCCCTATTAAAAAACTTCCCGACAAAACCGAAAATCTGTGGGAAGTCTTGTCCCTCC
AATTTTACACCTGTTCAATTCCCCTGCAGGACAACGCCCACACACCAGGTTAGCCTTTAAGCCTGCCCAG
AAGACTCCCGCCCATCTTCTAGAAAGACTGGAGTTGCAGATCACGAGGGAAGAGGGGGAAGGGATTCTCC
CAGGCCCAGGGCGGTCCTCAGAAGCCAGGAGGCAGCAGAGAACTCCCAGAAAGGTATTGCAACACTCCCC
TCCCCCCTCCGGAGAAGGGTGCGGCCTTCTCCCCGCCTACTCCACTGCAGCTCCCTTACTGATAACAACT
CAGAGCGACTTTGGGAGAGCAAGTGCTTCCTGC*CTCCAAAACAGCCCAACTGAG*CCCTCGCTCCTTCCCT
CCACTCCCCGGAGTGCGCGATGGAGGTCTGGCTCAGCACGCCCCTCTTGAGGCAACT*CAAGTCGGAAACG*
*TGCTTG*CACCCGCCCCGCAGCCG

图10

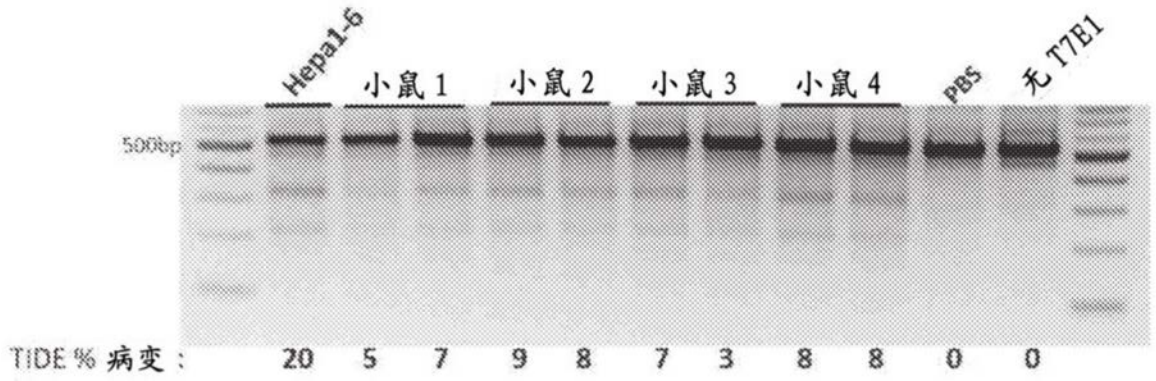


图11

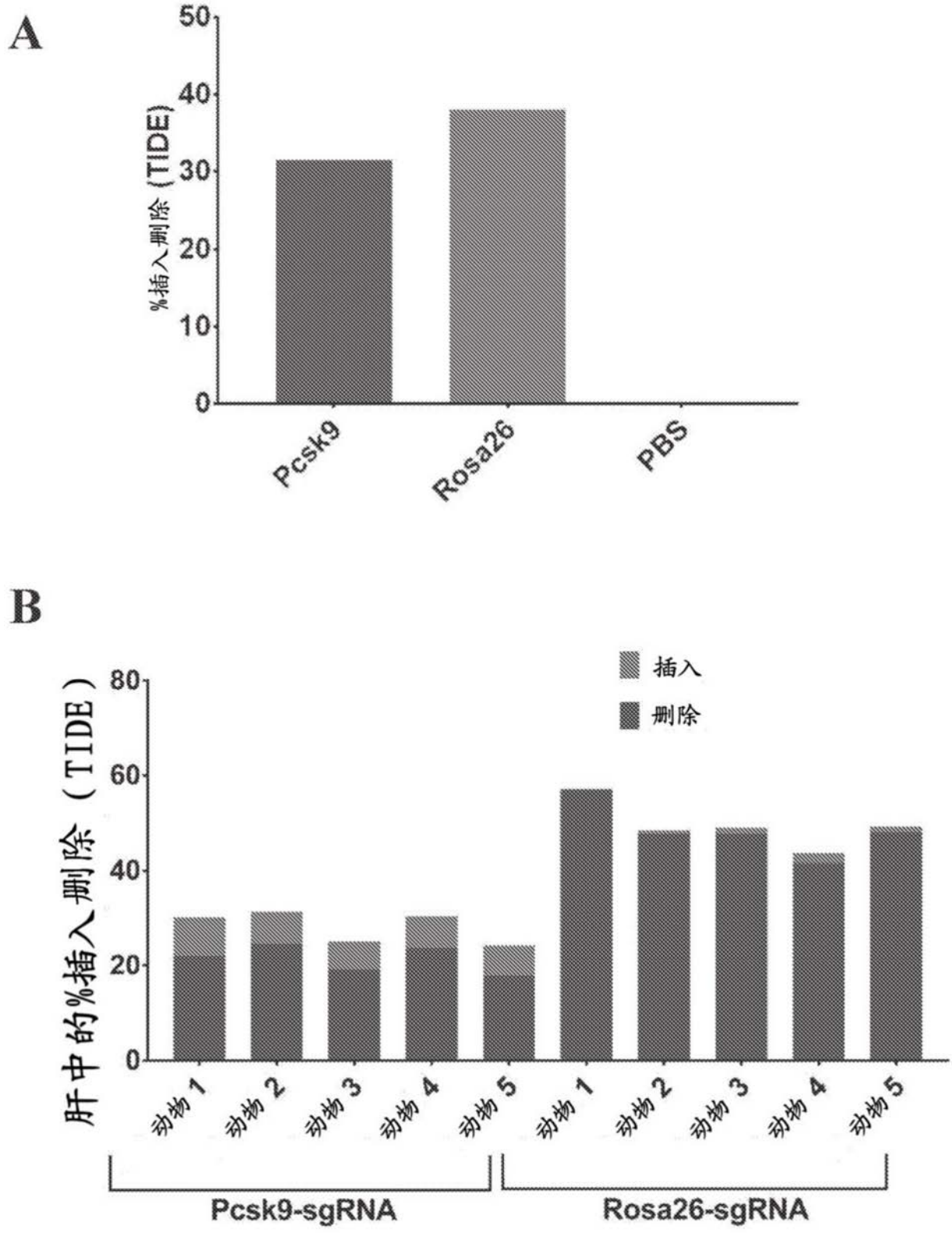

A
%插入删除 (TIDE)
50
40
30
20
10
0
Pcsk9
Rosa26
PBS
B
插入
删除
肝中的%插入删除（TIDE）
80
60
40
20
0
动物 1
动物 2
动物 3
动物 4
动物 5
动物 1
动物 2
动物 3
动物 4
动物 5
Pcsk9-sgRNA
Rosa26-sgRNA

图12

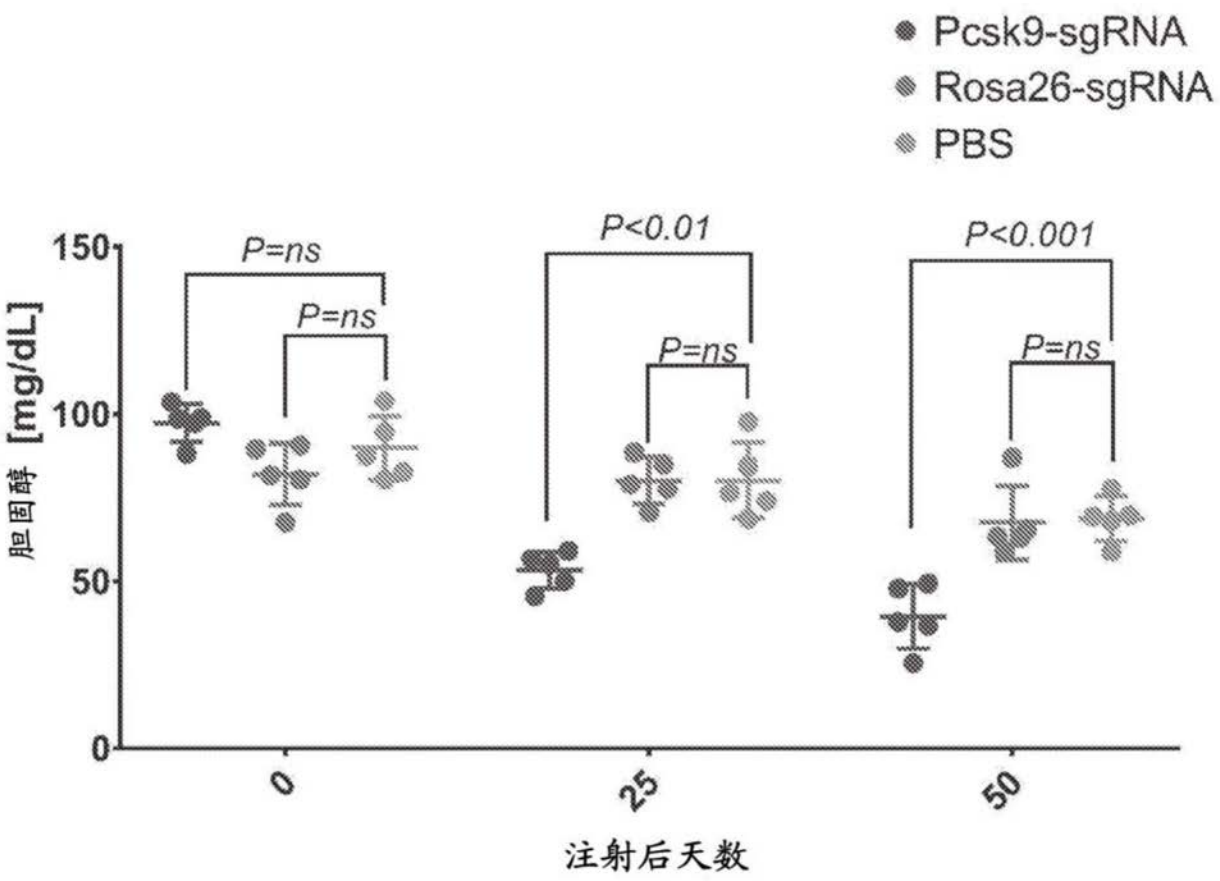
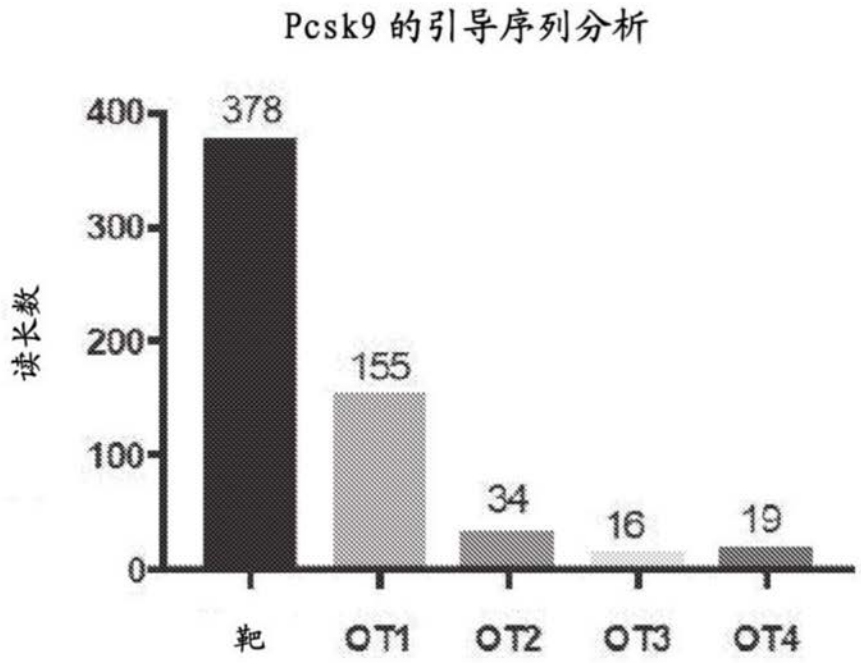

Pcsk9-sgRNA
Rosa26-sgRNA
PBS
胆固醇 [mg/dL]
150
100
50
0
P=ns
P=ns
P<0.01
P=ns
P<0.001
P=ns
0
25
50
注射后天数

图13

A


Pcsk9 的引导序列分析
读长数
400
300
200
100
0
378
155
34
16
19
靶
OT1
OT2
OT3
OT4

B

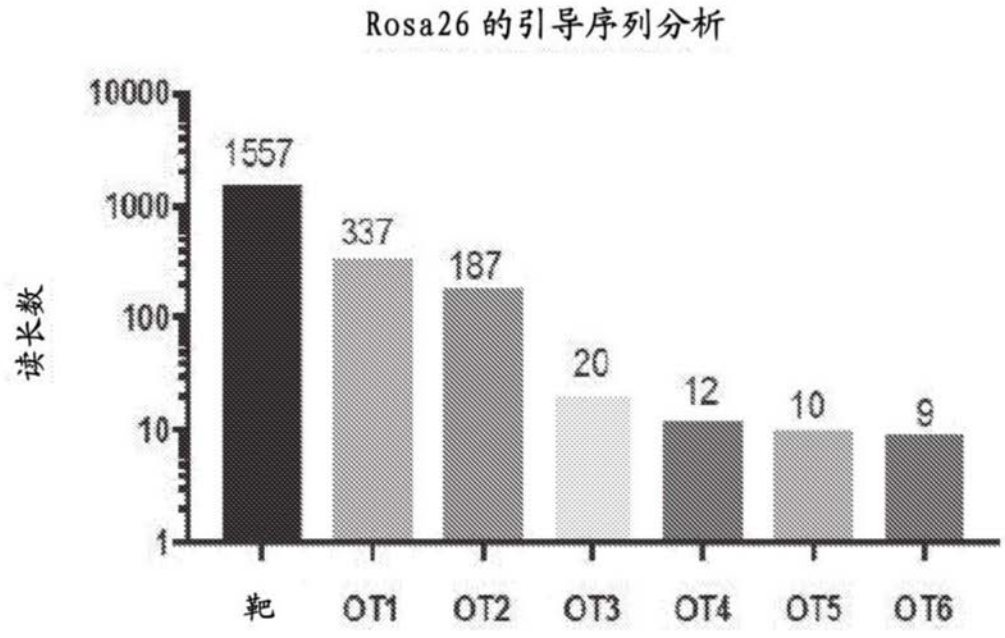

Rosa26 的引导序列分析
读长数
10000
1000
100
10
1
1557
337
187
20
12
10
9
靶
OT1
OT2
OT3
OT4
OT5
OT6

图14

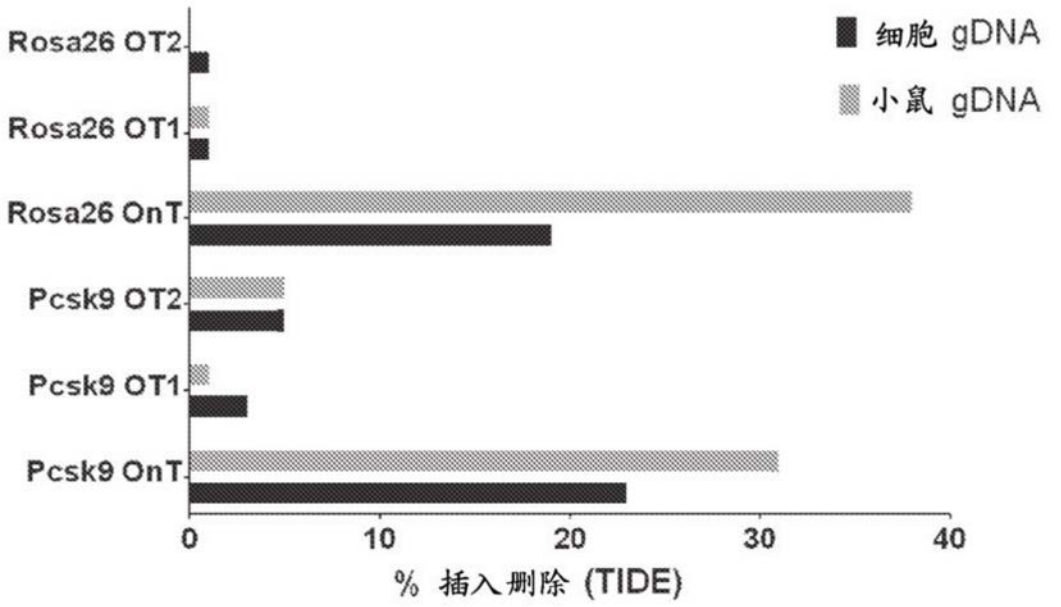

Rosa26 OT2
Rosa26 OT1
Rosa26 OnT
Pcsk9 OT2
Pcsk9 OT1
Pcsk9 OnT
细胞 gDNA
小鼠 gDNA
0
10
20
30
40
% 插入删除 (TIDE)

图15

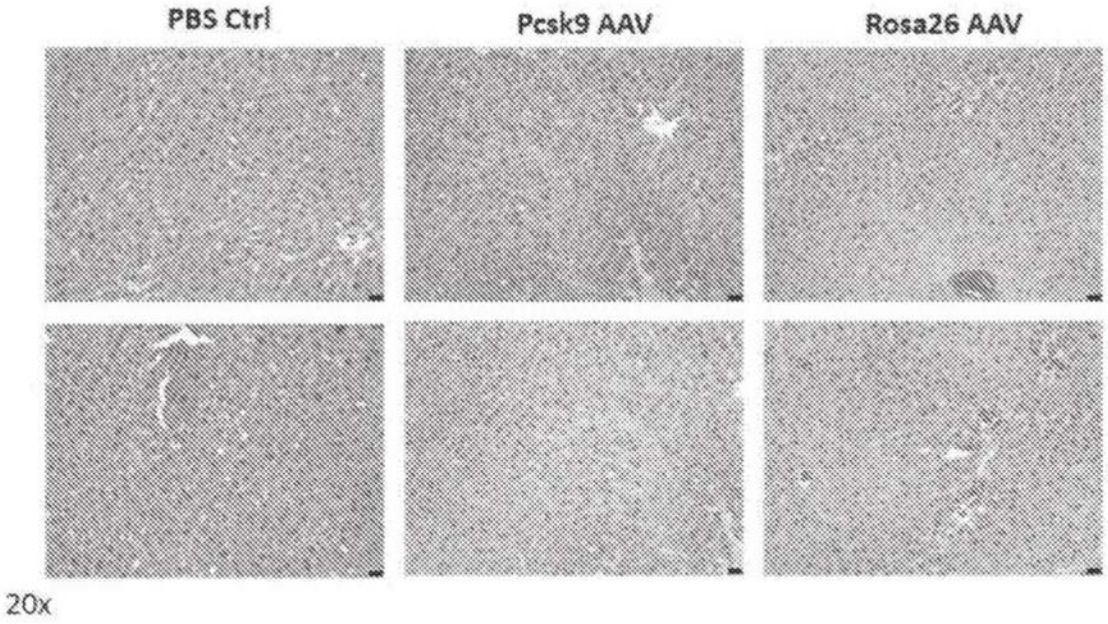

PBS Ctrl
Pcsk9 AAV
Rosa26 AAV
20x

图16

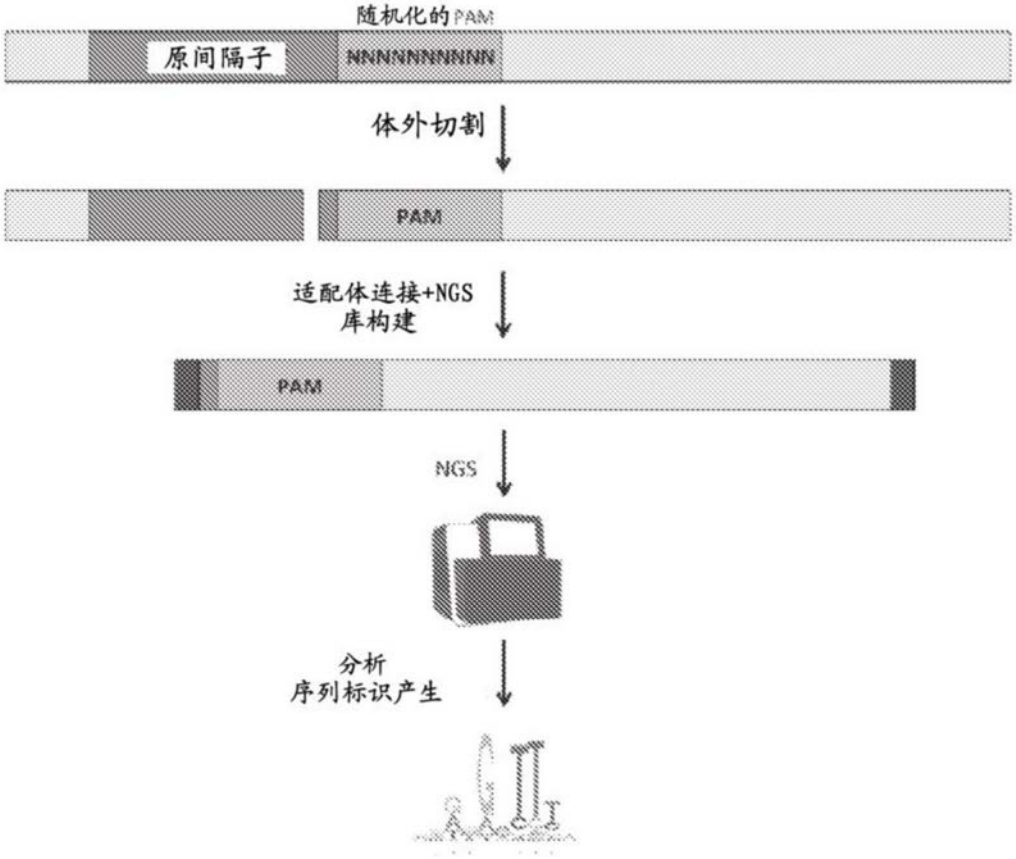

随机化的PAM
原间隔子
NNNNNNNNN
体外切割
PAM
适配体连接+NGS
库构建
PAM
NGS
分析
序列标识产生

图17

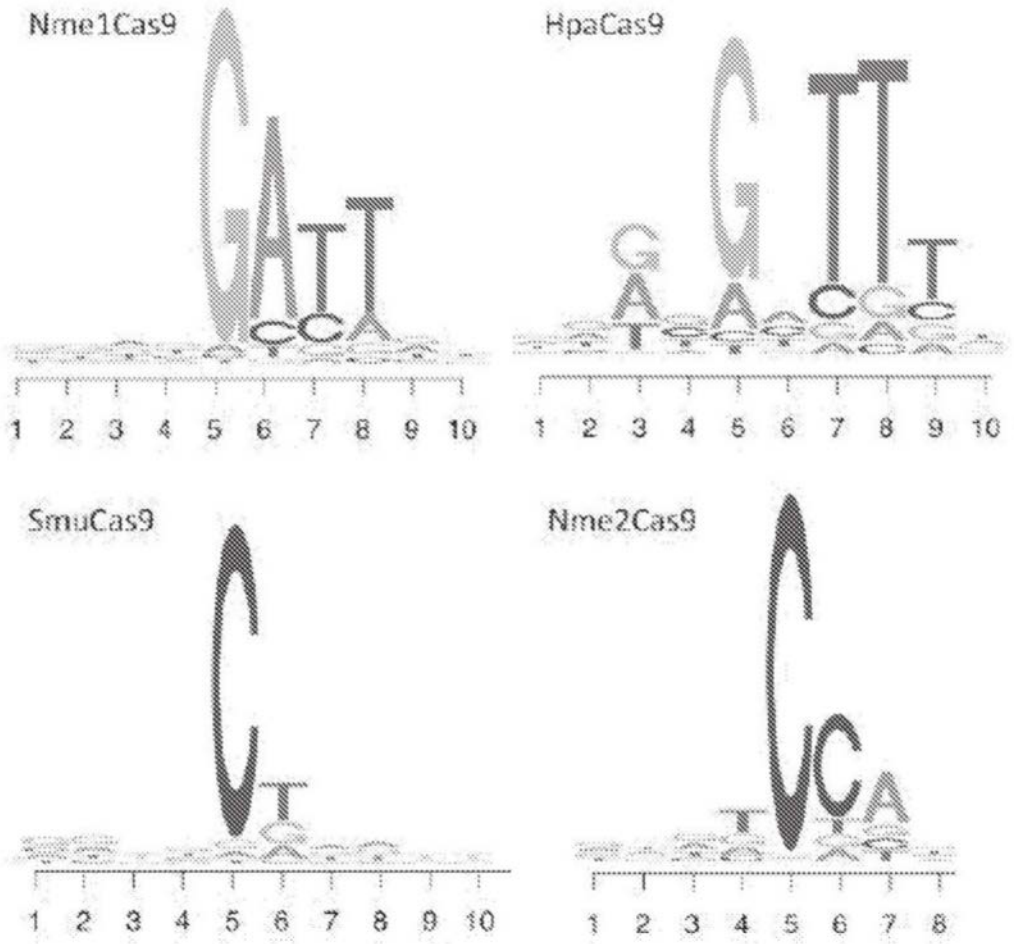

Nme1Cas9
HpaCas9
SmuCas9
Nme2Cas9
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8

图18

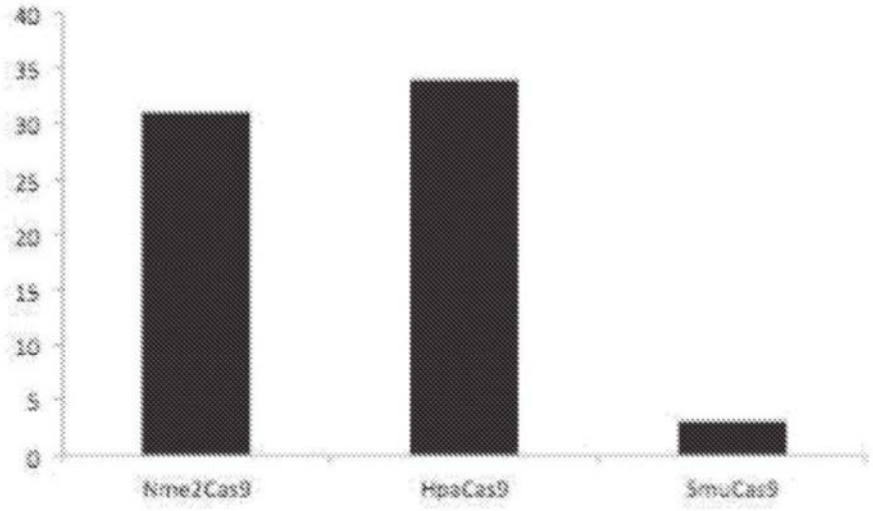

40
35
30
25
20
15
10
5
0
Nme2Cas9
HpaCas9
SmuCas9

图19

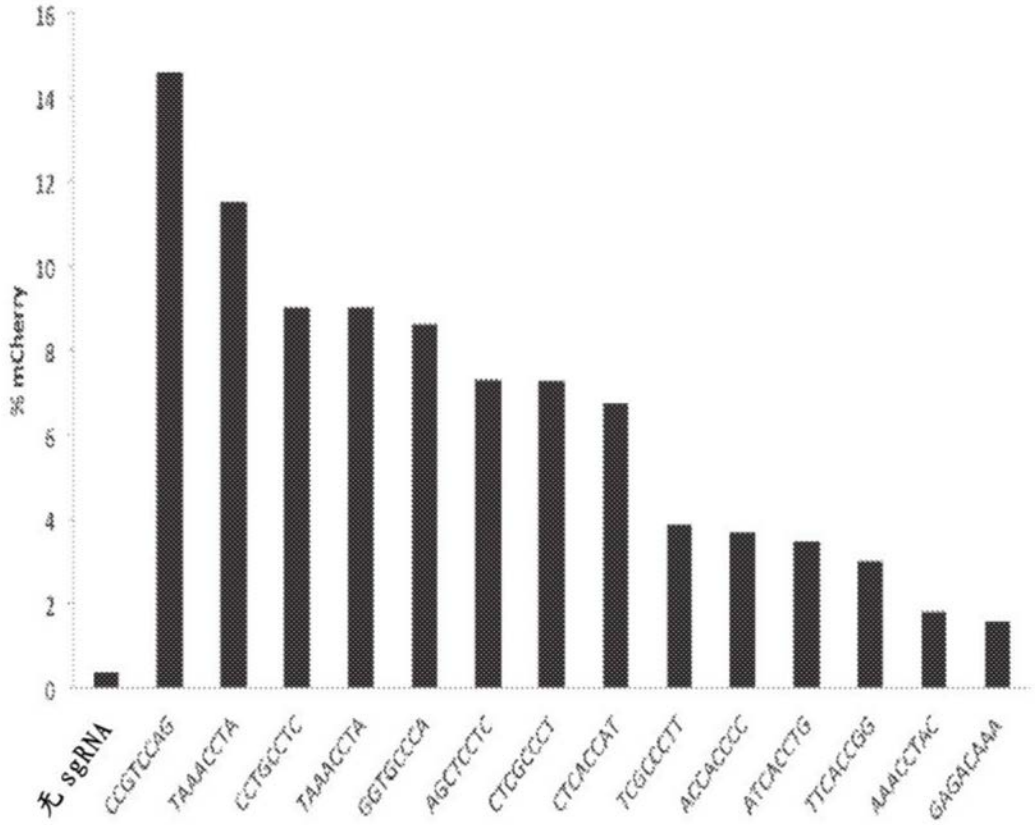

% mCherry
16
14
12
10
8
6
4
2
0
无 sgRNA
CCGTCCAG
TAAACCTA
CCTGCCTC
TAAACCTA
GGTGCCCA
AGCTCCTC
CTCGCCCT
CTCACCAT
TCGCCCTT
ACCACCCC
ATCACCTG
TTCACCGG
AAACCTAC
GAGACAAA

图20

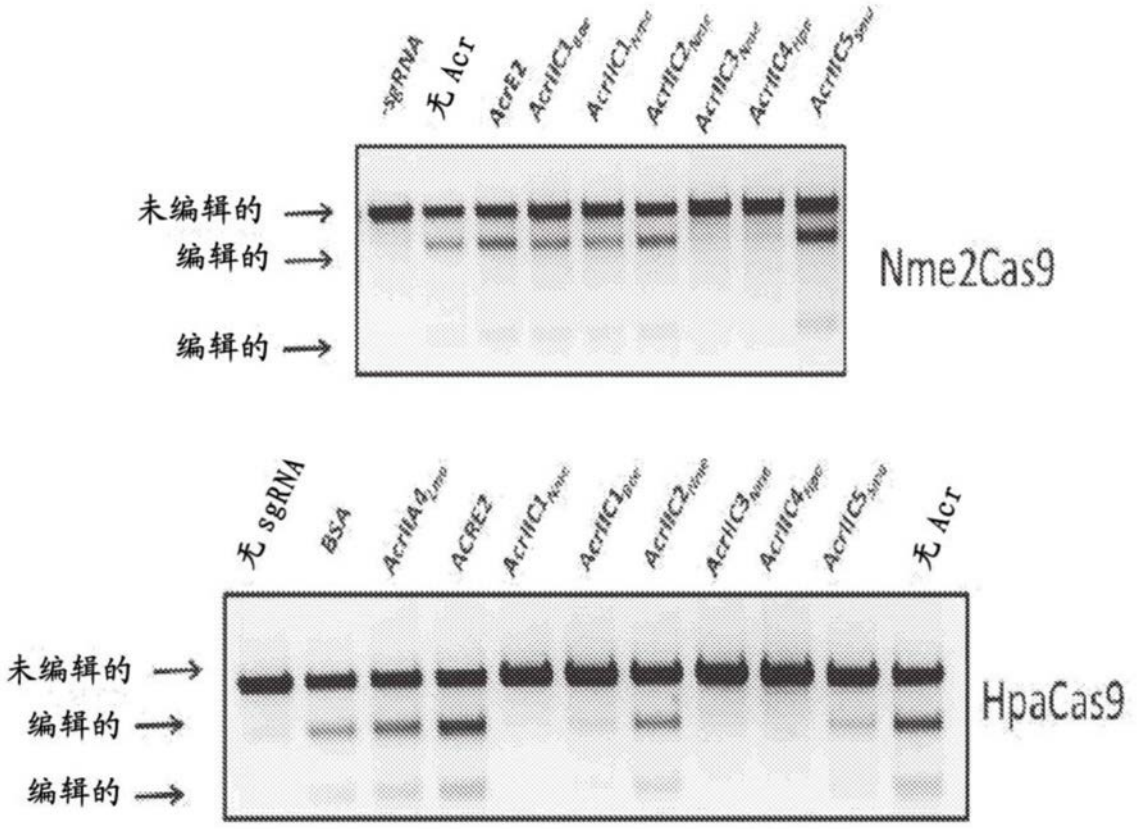

-sgRNA
无 Acr
AcrE2
AcrIIC1
AcrIIC1
AcrIIC2
AcrIIC3
AcrIIC4
AcrIIC5
未编辑的
编辑的
编辑的
Nme2Cas9
无 sgRNA
BSA
AcrIIA4
ACRE2
AcrIIC1
AcrIIC1
AcrIIC2
AcrIIC3
AcrIIC4
AcrIIC5
无 Acr
未编辑的
编辑的
编辑的
HpaCas9

图21

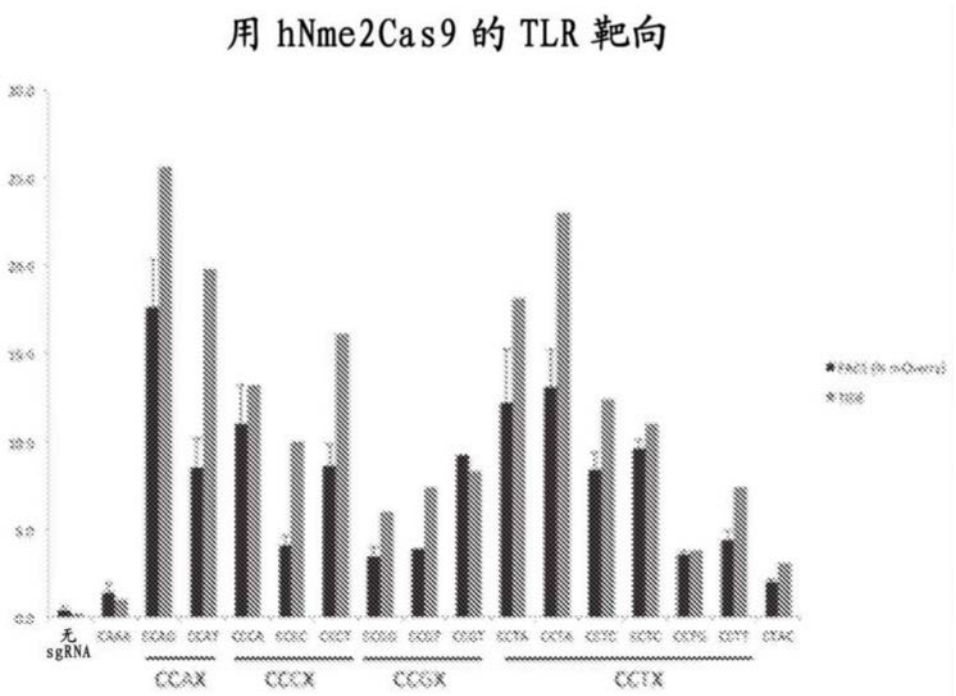

用 hNme2Cas9 的 TLR 靶向
无 sgRNA
CCAX
CCCX
CCGX
CCTX

图22

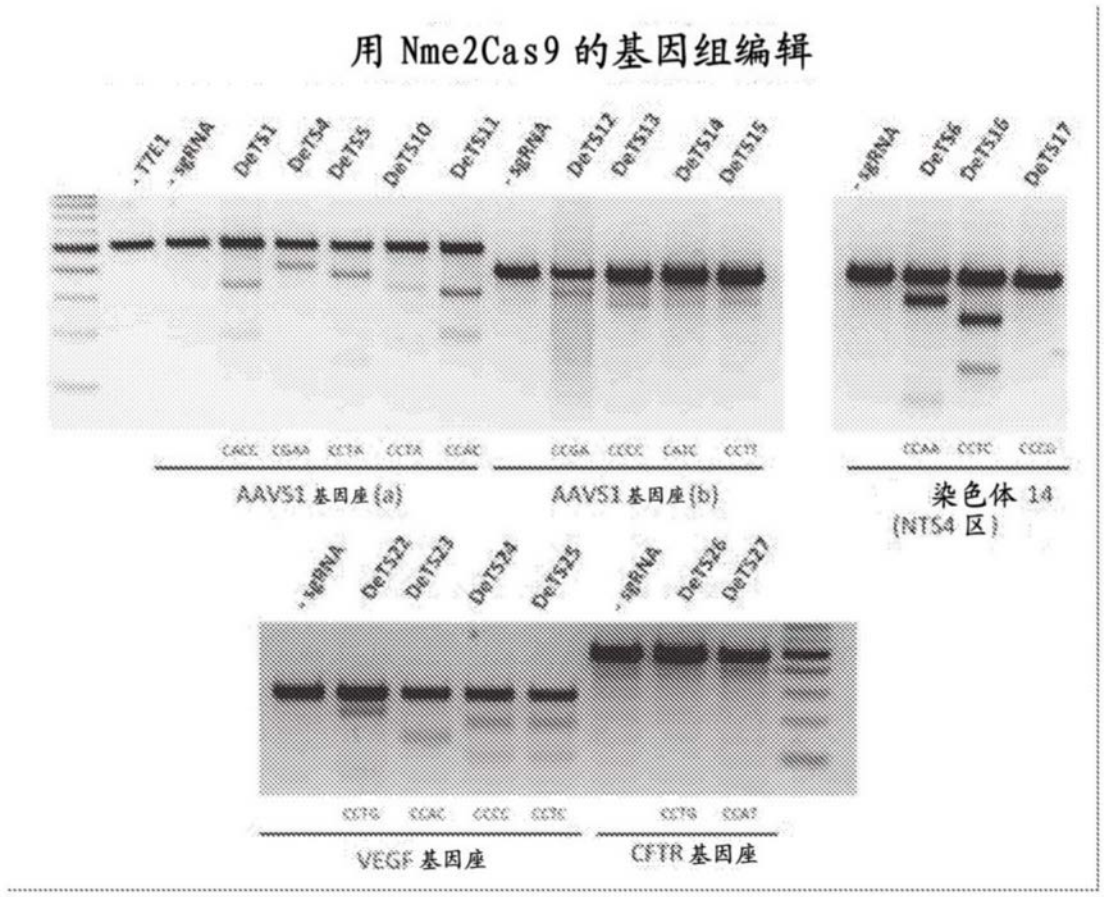

用 Nme2Cas9 的基因组编辑
-TTE1
-sgRNA
DeTS1
DeTS4
DeTS5
DeTS10
DeTS11
-sgRNA
DeTS12
DeTS13
DeTS14
DeTS15
-sgRNA
DeTS6
DeTS16
DeTS17
AAVS1 基因座(a)
AAVS1 基因座(b)
染色体 14
(NTS4 区)
-sgRNA
DeTS22
DeTS23
DeTS24
DeTS25
-sgRNA
DeTS26
DeTS27
VEGF 基因座
CFTR 基因座

图23

ATGGCTGCCTTCAAACCTAATCCAATCAACTACATCCTCGGCCTCGATATCGGCATCGCATCCGTCGGC
TGGGCGATGGTAGAAATTGACGAAGAAGAAAACCCCATCCGCCTGATTGATTTGGGCGTGCGCGTATT
TGAGCGTGCCGAAGTACCGAAAACAGGCGACTCCCTTGCCATGGCAAGGCGTTTGGCGCGCAGTGTCC
GCCGCCTGACCCGCCGTCGCGCCCATCGCCTGCTTCGGGCCCGCCGCCTATTGAAACGCGAAGGCGTA
TTACAAGCCGCTGATTTTGACGAAAACGGCTTGATCAAATCCTTACCGAATACACCATGGCAACTTCG
CGCAGCCGCATTAGACCGCAAACTGACGCCTTTAGAGTGGTCGGCAGTCTTGTTGCATTTAATCAAAC
ACCGCGGTTATTTGTCGCAAAGAAAAAACGAGGGCGAAACTGCCGATAAAGAGCTTGGCGCTTTGCTT
AAAGGCGTGGCCAACAATGCCCATGCCTTACAGACAGGCGATTTCCGCACACCGGCCGAATTGGCTTT
AAATAAATTTGAGAAAGAAAGCGGCCATATCCGCAATCAGCGCGGCGATTATTCGCATACGTTCAGCC
GCAAAGATTTACAGGCGGAGCTGATTTTGCTGTTTGAAAAACAAAAAGAATTTGGCAATCCGCATGTT
TCAGGCGGCCTTAAAGAAGGTATTGAAACCCTACTGATGACGCAACGCCCTGCCCTGTCCGGCGATGC
CGTTCAAAAAATGTTGGGGCATTGCACCTTCGAACCGGCAGAGCCGAAAGCCGCTAAAAACACCTAC
ACAGCCGAACGTTTCATCTGGCTGACCAAGCTGAACAACCTGCGTATTTTAGAGCAAGGCAGCGAGCG
GCCATTGACCGATACCGAACGCGCCACGCTTATGGACGAGCCATACAGAAAATCCAAACTGACTTACG
CACAAGCCCGTAAGCTGCTGGGTTTAGAAGATACCGCCTTTTTCAAAGGCTTGCGCTATGGTAAAG
ACAATGCCGAAGCCTCAACATTGATGGAAATGAAGGCCTACCATGCCATCAGCCGTGCACTGGAAAA
AGAAGGATTGAAAGATAAAAAATCCCCATTAAACCTTTCTTCCGAATTACAAGATGAAATCGGCACGG
CATTCTCCCTGTTCAAAACCGATGAAGACATTACAGGCCGTCTGAAAGACCGTGTTCAGCCCGAAATC
TTAGAAGCGCTGTTGAAACACATCAGCTTCGATAAGTTCGTCCAAATTTCCTTGAAAGCATTGCGCCG
AATTGTGCCTCTAATGGAGCAAGGCAAACGTTACGATGAAGCCTGCGCCGAAATCTACGGAGACCATT
ACGGCAAGAAAAATACGGAAGAAAAAATTTATCTGCCGCCCATCCCTGCCGACGAGATCCGCAACCC
CGTCGTCTTGCGCGCCTTATCTCAAGCACGTAAGGTCATTAACGGCGTGGTACGCCGTTACGGCTCCCC
AGCTCGTATCCATATTGAAACGGCAAGGGAAGTAGGTAAATCGTTTAAAGACCGCAAAGAAATCGAA
AAACGCCAAGAAGAAAACCGCAAAGACCGGGAAAAAGCCGCCGCCAAATTCCGAGAGTATTTCCCCA
ATTTTGTCGGCGAACCCAAATCAAAAGATATTCTGAAACTGCGCCTGTACGAGCAACAACACGGCAAA
TGCCTGTATTCGGGCAAAGAAATCAACTTAGTCCGTCTGAACGAAAAAGGCTATGTCGAAATCGACCA
TGCCCTGCCGTTTTCGCGCACATGGGACGACAGTTTCAACAATAAAGTGCTGGTATTGGGCAGCGAAA
ACCAAAACAAAGGCAATCAAACCCCTTACGAATACTTCAACGGCAAAGACAACAGCCGCGAATGGCA
GGAATTTAAAGCGCGTGTCGAAACCAGCCGTTTCCCGCGCAGTAAAAAACAACGGATTCTGCTGCAAA
AATTCGATGAAGACGGCTTTAAAGAATGCAATCTGAACGACACGCGCTACGTCAACCGTTTCCTGTGC
CAATTTGTTGCCGACCATATATTGCTGACAGGTAAAGGGAAAAGACGTGTCTTTGCCTCAAACGGACA
AATTACCAATCTGTTGCGCGGCTTTTGGGGATTGCGCAAAGTGCGTGCGGAAAACGACCGCCATCACG
CCTTGGACGCTGTAGTCGTTGCCTGCTCGACCGTTGCCATGCAGCAGAAAATTACCCGTTTTGTACGCT
ATAAAGAGATGAACGCGTTTGACGGTAAAACCATAGACAAAGAAACAGGAAAAGTGCTGCATCAAAA
AACACACTTCCCACAACCTTGGGAATTTTTCGCACAAGAAGTCATGATTCGCGTCTTCGGCAAACCGG
ACGGCAAACCCGAATTCGAAGAAGCCGATACCCCAGAAAAACTGCGCACGTTGCTTGCCGAAAAATT
ATCATCTCGCCCCGAAGCCGTACACGAATACGTTACGCCACTGTTTGTTTCACGCGCGCCCAATCGGA
AGATGAGCGGTGCACATAAAGATACTTTGAGATCTGCTAAACGATTTGTTAAACATAATGAAAAAATT
AGTGTTAAACGAGTATGGTTAACCGAAATCAAGTTGGCCGACCTTGAAAATATGGTTAATTATAAAAA
TGGTAGAGAGATTGAATTATATGAGGCTCTTAAGGCGCGTTTAGAGGCATATGGAGGTAATGCTAAAC
AAGCATTTGACCCTAAGGACAATCCGTTTTATAAAAAGGGAGGACAACTGGTTAAAG
CTGTGAGGGTCGAAAAAACCCAAGAGAGCGGAGTCTTATTAAATAAAAAAAATGCTTATACCATTGC
AGATAATGGAGACATGGTACGTGTTGATGTGTTCTGTAAAGTAGATAAGAAAGGAAAAAATCAGTATT
TTATTGTTCCCATCTATGCTTGGCAGGTTGCTGAAAACATTCTTCCCGATATTGATTGTAAGGGATACC
GGATTGATGATAGCTATACATTCTGTTTTAGCTTGCATAAGTATGATCTGATTGCTTTTCAAAAAGATG
AAAAATCTAAAGTAGAATTCGCCTACTATATCAACTGTGATAGCTCTAATGGACGATTCTATTTAGCTT
GGCATGATAAAGGCTCTAAAGAACAGCAATTCCGTATTAGCACCCAAAATCTTGTATTGATACAAAAA
TACCAAGTTAACGAACTGGGCAAAGAAATCAGACCATGCCGTCTGAAAAAACGC
CCACCTGTCCGTTAA

图24

MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAM
ARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAAL
DRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRT
PAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIE
TLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSER
PLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKA
YHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLK
HISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRN
PVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKA
AAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFS
RTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQ
RILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGF
WGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVL
HQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYV
TPLFVSRAPNRKMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNG
REIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLNKK
NAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTF
CFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQ
KYQVNELGKEIRPCRLKKRPPVR

图25


ATGGCCGCCTTCAAGCCTAACCCAATCAATTACATCCTGGGACTGGACATCGGAATCGCATCCGTGGGATGGGCTATGGTGGAGATCGACGA
GGAGGAGAATCCTATCCGGCTGATCGATCTGGGCGTGAGAGTGTTTGAGAGGGCCGAGGTGCCAAAGACCGGCGATTCTCTGGCTATGGCCC
GGAGACTGGCACGGAGCGTGAGGCGCCTGACACGGAGAAGGGCACACAGGCTGCTGAGGGCACGCCGGCTGCTGAAGAGAGAGGGCGTGCT
GCAGGCAGCAGACTTCGATGAGAATGGCCTGATCAAGAGCCTGCCAAACACCCCCTGGCAGCTGAGAGCAGCCGCCCTGGACAGGAAGCTG
ACACCACTGGAGTGGTCTGCCGTGCTGCTGCACCTGATCAAGCACCGCGGCTACCTGAGCCAGCGGAAGAACGAGGGAGAGACAGCAGACA
AGGAGCTGGGCGCCCTGCTGAAGGGAGTGGCCAACAATGCCCACGCCCTGCAGACCGGCGATTTCAGGACACCTGCCGAGCTGGCCCTGAAT
AAGTTTGAGAAGGAGTCCGGCCACATCAGAAACCAGAGGGGCGACTATAGCCACACCTTCTCCCGCAAGGATCTGCAGGCCGAGCTGATCCT
GCTGTTCGAGAAGCAGAAGGAGTTTGGCAATCCACACGTGAGCGGAGGCCTGAAGGAGGGAATCGAGACCCTGCTGATGACACAGAGGCCT
GCCCTGTCCGGCGACGCAGTGCAGAAGATGCTGGGACACTGCACCTTCGAGCCTGCAGAGCCAAAGGCCGCCAAGAACACCTACACAGCCGA
GCGGTTTATCTGGCTGACAAAGCTGAACAATCTGAGAATCCTGGAGCAGGGATCCGAGAGGCCACTGACCGACACAGAGAGGGCCACCCTGA
TGGATGAGCCTTACCGGAAGTCTAAGCTGACATATGCCCAGGCCAGAAAGCTGCTGGGCCTGGAGGACACCGCCTTCTTTAAGGGCCTGAGA
TACGGCAAGGATAATGCCGAGGCCTCCACACTGATGGAGATGAAGGCCTATCACGCCATCTCTCGCGCCCTGGAGAAGGAGGGCCTGAAGGA
CAAGAAGTCCCCCCTGAACCTGAGCTCCGAGCTGCAGGATGAGATCGGCACCGCCTTCTCTCTGTTTAAGACCGACGAGGATATCACAGGCC
GCCTGAAGGACAGGGTGCAGCCTGAGATCCTGGAGGCCCTGCTGAAGCACATCTCTTTCGATAAGTTTGTGCAGATCAGCCTGAAGGCCCTG
AGAAGGATCGTGCCACTGATGGAGCAGGGCAAGCGGTACGACGAGGCCTGCGCCGAGATCTACGGCGATCACTATGGCAAGAAGAACACAG
AGGAGAAGATCTATCTGCCCCCTATCCCTGCCGACGAGATCAGAAATCCTGTGGTGCTGAGGGCCCTGTCCCAGGCAAGAAAAGTGATCAAC
GGAGTGGTGCGCCGGTACGGATCTCCAGCCCGGATCCACATCGAGACCGCCAGAGAAGTGGGCAAGAGCTTCAAGGACCGGAAGGAGATCG
AGAAGAGACAGGAGGAGAATCGCAAGGATCGGGAGAAGGCCGCCGCCAAGTTTAGGGAGTACTTCCCTAACTTTGTGGGCGAGCCAAAGTC
TAAGGACATCCTGAAGCTGCGCCTGTACGAGCAGCAGCACGGCAAGTGTCTGTATAGCGGCAAGGAGATCAATCTGGTGCCGCTGAACGAGA
AGGGCTATGTGGAGATCGATCACGCCCTGCCTTTCTCCAGAACCTGGGACGATTCTTTTAACAATAAGGTGCTGGTGCTGGGCAGCGAGAACC
AGAATAAGGGCAATCAGACACCATACGAGTATTTCAATGGCAAGGACAACTCCAGGGAGTGGCAGGAGTTCAAGGCCCGCGTGGAGACCTC
TAGATTTCCCAGGAGCAAGAAGCAGCGGATCCTGCTGCAGAAGTTCGACGAGGATGGCTTTAAGGAGTGCAACCTGAATGACACCAGATACG
TGAACCGGTTCCTGTGCCAGTTTGTGGCCGATCACATCCTGCTGACCGGCAAGGGCAAGAGAAGGGTGTTCGCCTCTAATGGCCAGATCACAA
ACCTGCTGAGGGGATTTTGGGGACTGAGGAAGGTGCGGGCAGAGAATGACAGACACCACGCACTGGATGCAGTGGTGGTGGCATGCAGCAC
CGTGGCAATGCAGCAGAAGATCACAAGATTCGTGAGGTATAAGGAGATGAACGCCTTTGACGGCAAGACCATCGATAAGGAGACAGGCAAG
GTGCTGCACCAGAAGACCCACTTCCCCCAGCCTTGGGAGTTCTTTGCCCAGGAAGTGATGATCCGGGTGTTCGGCAAGCCAGACGGCAAGCCT
GAGTTTGAGGAGGCCGATACCCCAGAGAAGCTGAGGACACTGCTGGCAGAGAAGCTGTCTAGCAGGCCAGAGGCAGTGCACGAGTACGTGA
CCCCACTGTTCGTGTCCAGGGCACCCAATCGGAAGATGTCTGGCGCCCACAAGGACACACTGAGAAGCGCCAAGAGGTTTGTGAAGCACAAC
GAGAAGATCTCCGTGAAGAGAGTGTGGCTGACCGAGATCAAGCTGGCCGATCTGGAGAACATGGTGAATTACAAGAACGGCAGGGAGATCG
AGCTGTATGAGGCCCTGAAGGCAAGGCTGGAGGCCTACGGAGGAAATGCCAAGCAGGCCTTCGACCCAAAGGATAACCCCTTTTATAAGAAG
GGAGGACAGCTGGTGAAGGCCGTGCGGGTGGAGAAGACCCAGGAGAGCGGCGTGCTGCTGAATAAGAAGAACGCCTACACAATCGCCGACA
ATGGCGATATGGTGAGAGTGGACGTGTTCTGTAAGGTGGATAAGAAGGGCAAGAATCAGTACTTTATCGTGCCTATCTATGCCTGGCAGGTG
GCCGAGAACATCCTGCCAGACATCGATTGCAAGGGCTACAGAATCGACGATAGCTATACATTCTGTTTTTCCCTGCACAAGTATGACCTGATC
GCCTTCCAGAAGGATGAGAAGTCCAAGGTGGAGTTTGCCTACTATATCAATTGCGACTCCTCTAACGGCAGGTTCTACCTGGCCTGGCACGAT
AAGGGCAGCAAGGAGCAGCAGTTTCGCATCTCCACCCAGAATCTGGTGCTGATCCAGAAGTATCAGGTGAACGAGCTGGGCAAGGAGATCA
GGCCATGTCGGCTGAAGAAGCGCCCACCCGTGCGGGGCACCGGCGGGCCCAAGAAGAAGAGGAAGGTATACCCATACGATGTTCCTGACTAT
GCGGGCTATCCCTATGACGTCCCGGACTATGCAGGATCGTATCCTTATGACGTTCCAGATTACGCTGGATCCGCCGCTCCGGCAGCTAAGAAA
AAGAAACTGGATTTCGAATCCGGATAA

图26

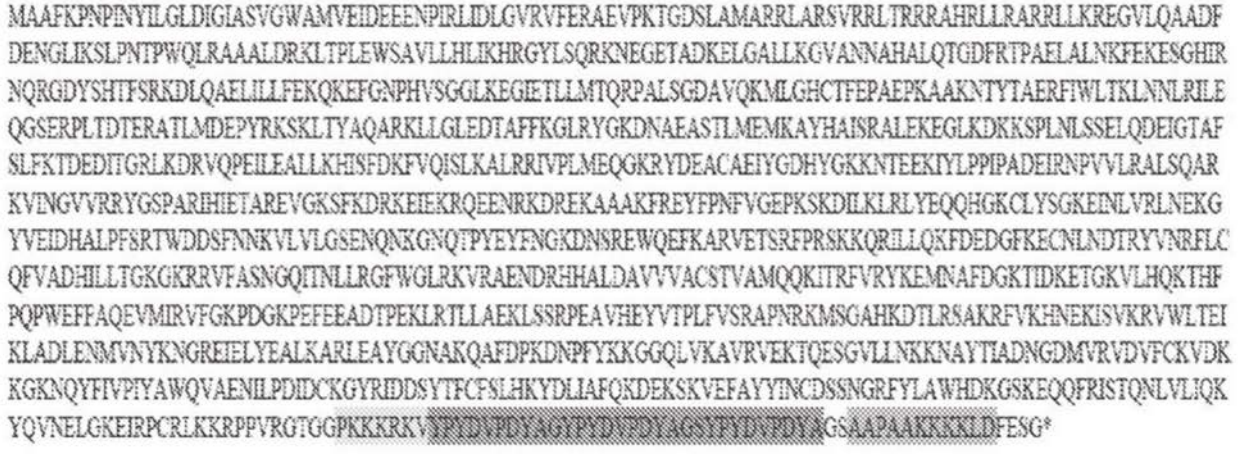

MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADF
DENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIR
NQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILE
QGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAF
SLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQAR
KVINGVVRRYGSPARIHIETARFVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKG
YVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLC
QFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHF
PQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEI
KLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCKVDK
KGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQK
YQVNELGKEIRPCRLKKRPPVRGTGGPKKKRKVYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAGSAAPAAKKKKLDFESG*

图27

MENKNLNYILGLDLGIASVGWAVVEIDEKENPLRLIDVGVRTFERAEVPKTGESLALSRR
LARSARRLTQRRVARLKKAKRLLKSENILLSTDERLPHQVWQLRVEGLDHKLERQEWA
AVLLHLIKHRGYLSQRKNESKSENKELGALLSGVDNNHKLLQQATYRSPAELAVKKFE
VEEGHIRNQQGAYTHTFSRLDLLAEMELLFSRQQHFGNPFASEKLLENLTALLMWQKPA
LSGEAILKMLGKCTFEDEYKAAKNTYSAERFVWITKLNNLRIQENGLERALNDNERLAL
MEQPYDKNRLFYSQVRSILKLSDEAIFKGLRYSGEDKKAIETKAVLMEMKAYHQIRKVL
EGNNLKAEWAELKANPTLLDEIGTAFSLYKTDEDISAYLAGKLSQPVLNALLENLSFDK
FIQLSLKALYKLLPLMQQGLRYDEACREIYGDHYGKKTEENHHFLPQIPADEIRNPVVLR
TLTQARKVINGVVRLYGSPARIHIETGREVGKSYKDRRELEKRQEENRKQRENAIKEFKE
YFPHFAGEPKAKDILKMRLYKQQNAKCLYSGKPIELHRLLEKGYVEVDHALPFSRTWD
DSFNNKVLVLANENQNKGNLTPFEWLDGKHNSERWRAFKALVETSAFPYAKKQRILSQ
KLDEKGFIERNLNDTRYVARFLCNFIADNMHLTGEGKRKVFASNGQITALLRSRWGLAK
SREDNDRHHALDAVVVACSTVAMQQKITRFVRFEAGDVFTGERIDRETGEIIPLHFPTPW
QFFKQEVEIRIFSDNPKLELENRLPDRPQANHEFVQPLFVSRMPTRKMTGQGHMETVKS
AKRLNEGISVIKMPLTKLKLKDLELMVNREREKDLYDTLKARLEAFNDDPAKAFAEPFI
KKGGAIVKSVRVEQIQKSGVLVREGNGVADNASMVRVDVFTKGGKYFLVPIYTWQVA
KGILPNKAATQYKDEEDWEVMDNSATFKFSLHPNDLVKLVTKKKTILGYFNGLNRATG
NIDIKEHDLDKSKGKQGIFEGVGIKLALSFEKYQVDELGKNIRLCKPSKRQPVR*

图28

ATGGAAAAATTTCACTATGTATTGGGTTTGGATTTGGGTATCGCCTCTGTGGGGTGGGCTGCCATTGAA
ATTGACAAGGAAACCGAAACATCAATCGGTTTATTGGATTGCGGTGTCAGAACATTTGAACGTGCAGA
AGTACCCAAAACAGGCGATTCTCTTGCCAAAGCTCGCCGTGAAGCCAGAAGTACTCGCCGTTTAATTC
GCAGACGTTCGCATCGCTTATTACGTTTAAAACGTTTATTGAAACGTGAAATTTTCAGGCAGCCTGAAA
CGTTTAAAGACTTACCAATCAATGCTTGGCAATTGCGTGTTAAAGGCTTGGATAGTCGGTTGAATGAA
TATGAATGGGCGGCCGTTTTATTGCATTTGGTGAAGCATCGCGGTTATTTATCGCAACGCAAAAGCGA
AATGAGCGAAACAGACAGCAAATCTGAAATGGGCAGATTACTGGCAGGTGTGGCGGAAAATCACCAA
CTTTTACAACAAGAACAATATCGTACACCAGCCGAATTAGCACTCAAAAAATTTGTGAAACATTTTCG
CAATAAAGGTGGCGATTATGCACACACTTTCAACCGTTTGGATTTGCAAGCCGAATTGCATTTATTGTT
TCAAAAACAACGTGAATTAGGCAATCCATTCACTTCACCAGAATTGGAACGGCAAGTTGATGATTTGT
TGATGACGCAGCGCAGTGCTTTACAAGGTGATGCGATTTTGAAAATGTTGGGTCATTGTGGGTTTGAA
CCTGAACAATTCAAAGCAGCGAAAAACACATTCAGTGCCGAACGTTTTATTTGGTTGACAAAACTCAA
TAATCTTCGCATTCAAGACCAAGGCAAAGAACGTGCGTTAACTGCCGATGAGCGTACCAAATTGTTGG
ACGAGCCTTATAAAAAAAGTAAATTGACTTACGCACAAGTTCGCAAATTATTAAGCTTGCCTCAAACT
GCTATTTTTAAAGGTTTGCGTTATGATTTGGAACATGACAAAAAAGCAGAAAACAGTACGTTGATGGA
AATGAAATCCTATCACAACATCCGCCAAACATTGGAAAAATCAGGTTTGAAAACAGAATGGCAAAGT
ATTGCCACGCAGCCTGAAATTTTAGATGCAATTGGCACGGCGTTTTCCATTTATAAAACCGATGAAGA
TATTTCGCATGAATTAAAAACGTGCAGGCTGCCTGAAAACGTATTGAATGAATTACTGAAAAACATCA
ATTTTGATGGATTCATTCAATTATCGTTGACTGCATTACGCAAAATTTTGCCCTTGATGGAACAAGGCT
ACCGTTATGATGAAGCGTGTACCCAAATTTACGGTAATCATCATTCAGGCAGCTTGCAACAAGAATCA
AAGCAATTTTTGCCACATATTCCGATTGATGATGTCCGAAATCCTGTGGTGTTCCGTACTTTGACCCAA
GCAAGAAAAGTGGTGAATGCGATTATTCGTCGGTATGGTTCGCCAGCTCGTGTGCATATTGAGATGGC
GCGTGAATTGGGTAAGTCTAAATCAGACCGTGACCGAATTGAAAAACAACAACAAAAAAATAAAAAA
GAACGTGAAAACGCAGTCGCCAAATTCAAAGAAGATTTCCCTGATTTTGTGGGCGAACCCAGAGGGA
AAGATATTTTGAAAATGCGTTTGTATGAACAACAACACGGCAAATGTTTGTATTCGGGTCATGATATT
GATATTAATCGATTGAATGAAAAAGGTTATGTTGAAATTGACCATGCCCTGCCATTTTCACGGACTTGG
GATGATAGTCAAAATAATAAAGTTTTGGTACTTGGCAGCGAAAACCAAAATAAACGCAATCAAACGC
CTGATGAATATTTGGACGGTGCAAACAATAGCCAACGTTGGCTTGAATTTCAAGCGCGTGTACAAACT
TGTCATTTTTCTTACGGTAAAAAACAACGCATTCAATTAGCCAAATTAGACGATGAAACCGAAAAAGG
ATTTTTAGAACGCAATCTAAATGATACGCGGTATATTGCTCGTTTTATGTGTCAATTTGTCCAAGAAAA
TTTATATTTGACAGGTAAAGGAAAACGTCTTGTTTTTGCATCAAACGGCGGAATGACCGCAACATTGA
GAAATTTATGGGGTTTGAGAAAAGTCCGTGAAGACAACGACCGCCATCATGCTCTTGATGCGATTGTG
GTGGCGTGTTCCACTGCTTCTATGCAACAAAAAATAACCAAAGCATTTCAACGCCATGAAAGCATTGA
ATATGTGGATACCGAAACGGGCGAAGTAAAATTTCGTATTCCACAGCCTTGGGATTTTTTCCGTCAGG
AAGTGATGATTCGTGTGTTCAGCGACCAACCGTGTGAAGATTTGGTAGAAAAATTGTCGGCTCGTCCC
GAAGCTTTGCATGACAACGTAACGCCTTTATTTGTCTCGCGTGCACCAAATCGCAAAATGTCGGGGCA
AGGGCATTTGGAAACCATCAAATCTGCAAAAAGGCTGTCTGAAGAAAACAGTATGGTTAAAAAACCA
TTAACCACATTGAAATTAAAAGATATTCCAGAAATCGTAGGCTACCCGAGTCGTGAACCTCAATTGTA
TGCCGCATTGAAAACACGTTTAGAAACGCATGATGATGACCCAATTAAAGCCTTTGCCAAACCCTTTT
ACAAACCCAATAAAAATGGTGAATTGGGGGCGTTGGTTCGATCGGTGCGTGTGAAAGGTGTACAAAA
TACGGGTGTAATGGTTCATGATGGCAAAGGCATTGCCGATAATGCCACAATGGTTCGTGTTGATGTCT
ATACCAAAGCGGGCAAAAATTACCTTGTTCCTGTGTATGTTTGGCAGGTGGCTCAAGGAATTTTGCCA
AATCGGGCGGTTACTTCTGGCAAAAGTGAAGCAGATTGGGATTTAATTGATGAAAGTTTTGAATTTAA
ATTTTCGCTGTCTCGTGGGGATTTAGTGGAAATGATTAGCAATAAAGGAAGAATTTTTGGTTATTACAA
TGGGTTAGATCGTGCAAATGGAAGTATTGGGATTCGTGAACATGATTTGGAAAAGTCCAAAGGAAAA
GATGGTGTTCATCGTGTTGGCGTGAAAACCGCCACCGCATTCAACAAATACCACGTTGACCCACTTGG
TAAAGAAATTCATCGGTGTTCATCTGAACCACGCCCCACATTAAAAATCAAATCCAAGAAATAA

图29

MEKFHYVLGLDLGIASVGWAAIEIDKETETSIGLLDCGVRTFERAEVPKTGDSLAKARRE
ARSTRRLIRRRSHRLLRLKRLLKREIFRQPETFKDLPINAWQLRVKGLDSRLNEYEWAAV
LLHLVKHRGYLSQRKSEMSETDSKSEMGRLLAGVAENHQLLQQEQYRTPAELALKKFV
KHFRNKGGDYAHTFNRLDLQAELHLLFQKQRELGNPFTSPELERQVDDLLMTQRSALQ
GDAILKMLGHCGFEPEQFKAAKNTFSAERFIWLTKLNNLRIQDQGKERALTADERTKLL
DEPYKKSKLTYAQVRKLLSLPQTAIFKGLRYDLEHDKKAENSTLMEMKSYHNIRQTLEK
SGLKTEWQSIATQPEILDAIGTAFSIYKTDEDISHELKTCRLPENVLNELLKNINFDGFIQL
SLTALRKILPLMEQGYRYDEACTQIYGNHHSGSLQQESKQFLPHIPIDDVRNPVVFRTLT
QARKVVNAIIRRYGSPARVHIEMARELGKSKSDRDRIEKQQQKNKKERENAVAKFKED
FPDFVGEPRGKDILKMRLYEQQHGKCLYSGHDIDINRLNEKGYVEIDHALPFSRTWDDS
QNNKVLVLGSENQNKRNQTPDEYLDGANNSQRWLEFQARVQTCHFSYGKKQRIQLAK
LDDETEKGFLERNLNDTRYIARFMCQFVQENLYLTGKGKRLVFASNGGMTATLRNLWG
LRKVREDNDRHHALDAIVVACSTASMQQKITKAFQRHESIEYVDTETGEVKFRIPQPWD
FFRQEVMIRVFSDQPCEDLVEKLSARPEALHDNVTPLFVSRAPNRKMSGQGHLETIKSA
KRLSEENSMVKKPLTTLKLKDIPEIVGYPSREPQLYAALKTRLETHDDDPIKAFAKPFYK
PNKNGELGALVRSVRVKGVQNTGVMVHDGKGIADNATMVRVDVYTKAGKNYLVPVY
VWQVAQGILPNRAVTSGKSEADWDLIDESFEFKFSLSRGDLVEMISNKGRIFGYYNGLD
RANGSIGIREHDLEKSKGKDGVHRVGVKTATAFNKYHVDPLGKEIHRCSSEPRPTLKIKS
KK*

图30


ATGGAGAAGTTCCACTACGTGCTGGGACTGGATCTGGGAATCGCAAGCGTGGGATGGGCAGCAATCGAGATCGATAAGGAGACCGAGACAT
CCATCGGCCTGCTGGACTGCGGCGTGAGGACCTTTGAGAGGGCAGAGGTGCCTAAGACAGGCGACAGCCTGGCAAAGGCAAGGAGAGAGGC
AAGGTCTACAAGGCGCCTGATCCGGAGAAGGAGCCACAGGCTGCTGCGGCTGAAGAGACTGCTGAAGCGGGAGATCTTTAGACAGCCAGAG
ACCTTCAAGGATCTGCCCATCAACGCATGGCAGCTGAGGGTGAAGGGACTGGACTCTCGGCTGAATGAGTACGAGTGGGCAGCCGTGCTGCT
GCACCTGGTGAAGCACAGGGGCTATCTGAGCCAGCGCAAGTCCGAGATGTCTGAGACCGACTCTAAGAGCGAGATGGGCAGGCTGCTGGCA
GGAGTGGCCGAGAACCACCAGCTGCTGCAGCAGGAGCAGTACAGGACCCCAGCAGAGCTGGCCCTGAAGAAGTTTGTGAAGCACTTCCGCA
ACAAGGGCGGCGATTATGCCCACACATTCAATAGGCTGGACCTGCAGGCAGAGCTGCACCTGCTGTTTCAGAAGCAGAGAGAGCTGGGCAAC
CCCTTCACCTCTCCTGAGCTGGAGCGCCAGGTGGACGATCTGCTGATGACACAGCGGAGCGCCCTGCAGGGCGATGCAATCCTGAAGATGCT
GGGCCACTGTGGCTTTGAGCCTGAGCAGTTCAAGGCCGCCAAGAATACCTTTAGCGCCGAGAGATTCATCTGGCTGACAAAGCTGAACAATC
TGAGGATCCAGGACCAGGGCAAGGAGAGAGCCCTGACCGCCGATGAGAGGACAAAGCTGCTGGACGAGCCTTACAAGAAGTCTAAGCTGAC
CTATGCCCAGGTGAGGAAGCTGCTGAGCCTGCCTCAGACAGCCATCTTCAAGGGCCTGCGCTACGATCTGGAGCACGACAAGAAGGCCGAGA
ACTCTACCCTGATGGAGATGAAGAGCTATCACAATATCCGGCAGACACTGGAGAAGTCCGGCCTGAAGACCGAGTGGCAGTCTATCGCCACA
CAGCCAGAGATCCTGGACGCAATCGGAACCGCCTTTTCCATCTACAAGACAGATGAGGACATCTCTCACGAGCTGAAGACCTGCAGACTGCC
TGAGAACGTGCTGAATGAGCTGCTGAAGAACATCAATTTTGATGGCTTCATCCAGCTGAGCCTGACCGCCCTGCGCAAGATCCTGCCACTGAT
GGAGCAGGGCTACCGGTATGACGAGGCCTGTACACAGATCTACGGCAACCACCACTCCGGCTCTCTGCAGCAGGAGTCCAAGCAGTTTCTGC
CTCACATCCCAATCGACGATGTGCGGAACCCAGTGGTGTTCAGAACCCTGACACAGGCCAGGAAGGTGGTGAATGCCATCATCCGCCGGTAT
GGATCTCCAGCAAGGGTGCACATCGAGATGGCAAGGGAGCTGGGCAAGAGCAAGTCCGATAGAGACAGGATCGAGAAGCAGCAGCAGAAG
AACAAGAAGGAGAGGGAGAATGCCGTGGCCAAGTTCAAGGAGGATTTTCCAGACTTCGTGGGAGAGCCTAGGGGCAAGGATATCCTGAAGA
TGCGGCTGTACGAGCAGCAGCACGGCAAGTGCCTGTATTCCGGCCACGATATCGACATCAACCGGCTGAATGAGAAGGGCTACGTGGAGATC
GACCACGCCCTGCCTTTTAGCAGAACCTGGGACGATTCCCAGAACAATAAGGTGCTGGTGCTGGGCAGCGAGAACCAGAATAAGCGCAATCA
GACACCAGATGAGTACCTGGACGGCGCCAACAATTCCCAGAGATGGCTGGAGTTTCAGGCCAGGGTGCAGACCTGCCACTTCTCTTATGGCA
AGAAGCAGAGGATCCAGCTGGCCAAGCTGGACGATGAGACCGAGAAGGGCTTCCTGGAGCGCAACCTGAATGATACAAGGTACATCGCCCG
GTTCATGTGCCAGTTCGTGCAGGAGAACCTGTATCTGACCGGCAAGGGCAAGCGCCTGGTGTTTGCCTCCAACGGCGGCATGACCGCCACACT
GCGGAATCTGTGGGGCCTGAGGAAGGTGCGCGAGGATAATGACAGACACCACGCACTGGACGCAATCGTGGTGGCATGCAGCACCGCATCC
ATGCAGCAGAAGATCACAAAGGCCTTTCAGCGGCACGAGAGCATCGAGTATGTGGATACCGAGACAGGCGAGGTGAAGTTCAGAATCCCCC
AGCCTTGGGACTTCTTTCGCCAGGAAGTGATGATCCGGGTGTTTTCCGATCAGCCATGTGAGGACCTGGTGGAGAAGCTGTCTGCCAGGCCAG
AGGCCCTGCACGATAACGTGACCCCTCTGTTCGTGAGCAGGGCACCAAATAGAAAGATGTCCGGCCAGGGCCACCTGGAGACAATCAAGTCC
GCCAAGCGCCTGTCCGAGGAGAACTCTATGGTGAAGAAGCCCCTGACCACACTGAAGCTGAAGGACATCCCTGAGATCGTGGGCTACCCATC
TAGAGAGCCCCAGCTGTATGCCGCCCTGAAGACCAGGCTGGAGACACACGACGATGACCCAATCAAGGCCTTTGCCAAGCCTTTCTACAAGC
CAAACAAGAATGGAGAGCTGGGCGCCCTGGTGCGGTCCGTGAGAGTGAAGGGCGTGCAGAACACAGGCGTGATGGTGCACGATGGCAAGGG
CATCGCCGACAATGCCACAATGGTGCGGGTGGACGTGTATACAAAGGCCGGCAAGAACTACCTGGTGCCCGTGTACGTGTGGCAGGTGGCAC
AGGGAATCCTGCCAAATAGAGCCGTGACCTCTGGCAAGAGCGAGGCCGATTGGGACCTGATCGATGAGAGCTTCGAGTTTAAGTTCTCTCTG
AGCAGAGGCGACCTGGTGGAGATGATCTCCAACAAGGGCAGGATCTTCGGCTACTATAACGGCCTGGATAGAGCCAATGGCAGCATCGGCAT
CAGGGAGCACGATCTGGAGAAGTCCAAGGGCAAGGACGGAGTGCACAGGGTGGGAGTGAAGACCGCAACAGCCTTTAATAAGTACCACGTG
GACCCCCTGGGCAAGGAGATCCACAGATGTAGCTCCGAGCCAAGGCCCACCCTGAAGATCAAGAGCAAGAAGGGCACCGGCGGG
CCCAAGAAGAAGAGGAAGGTA
GGATCCGCC
TTCGAATCCGGATAA

图31

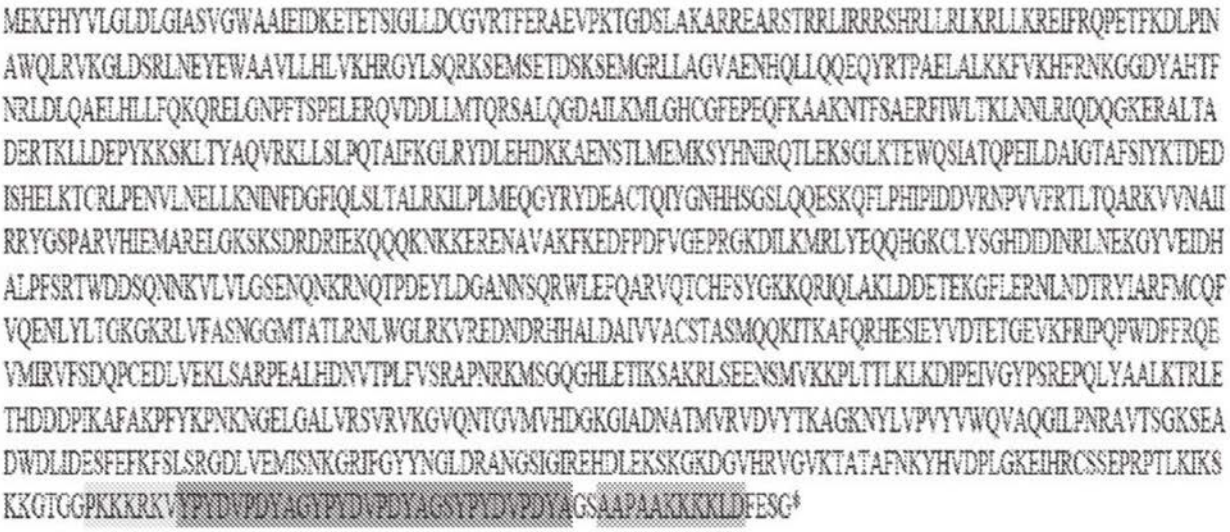

MEKFHYVLGLDLGIASVGWAAIEIDKETETSIGLLDCGVRTFERAEVPKTGDSLAKARREARSTRRLIRRRSHRLLRLKRLLKREIFRQPETFKDLPIN
AWQLRVKGLDSRLNEYEWAAVLLHLVKHRGYLSQRKSEMSETDSKSEMGRLLAGVAENHQLLQQEQYRTPAELALKKFVKHFRNKGGDYAHTF
NRLDLQAELHLLFQKQRELGNPFTSPELERQVDDLLMTQRSALQGDAILKMLGHCGFEPEQFKAAKNTFSAERFIWLTKLNNLRIQDQGKERALTA
DERTKLLDEPYKKSKLTYAQVRKLLSLPQTAIFKGLRYDLEHDKKAENSTLMEMKSYHNIRQTLEKSGLKTEWQSIATQPEILDAIGTAFSIYKTDED
ISHELKTCRLPENVLNELLKNINFDGFIQLSLTALRKILPLMEQGYRYDEACTQIYGNHHSGSLQQESKQFLPHIPIDDVRNPVVFRTLTQARKVVNAII
RRYGSPARVHIEMARELGKSKSDRDRIEKQQQKNKKERENAVAKFKEDFPDFVGEPRGKDILKMRLYEQQHGKCLYSGHDIDINRLNEKGYVEIDH
ALPFSRTWDDSQNNKVLVLGSENQNKRNQTPDEYLDGANNSQRWLEFQARVQTCHFSYGKKQRIQLAKLDDETEKGFLERNLNDTRYIARFMCQF
VQENLYLTGKGKRLVFASNGGMTATLRNLWGLRKVREDNDRHHALDAIVVACSTASMQQKITKAFQRHESIEYVDTETGEVKFRIPQPWDFFRQE
VMIRVFSDQPCEDLVEKLSARPEALHDNVTPLFVSRAPNRKMSGQGHLETIKSAKRLSEENSMVKKPLTTLKLKDIPEIVGYPSREPQLYAALKTRLE
THDDDPIKAFAKPFYKPNKNGELGALVRSVRVKGVQNTGVMVHDGKGIADNATMVRVDVYTKAGKNYLVPVYVWQVAQGILPNRAVTSGKSEA
DWDLIDESFEFKFSLSRGDLVEMISNKGRIFGYYNGLDRANGSIGIREHDLEKSKGKDGVHRVGVKTATAFNKYHVDPLGKEIHRCSSEPRPTLKIKS
KKGTGGPKKKRKV
YPYDVPDYAGYPYDVPDYAGSYPYDVPDYA
GSAAPAAKKKKLD
FESG*

图32



图33

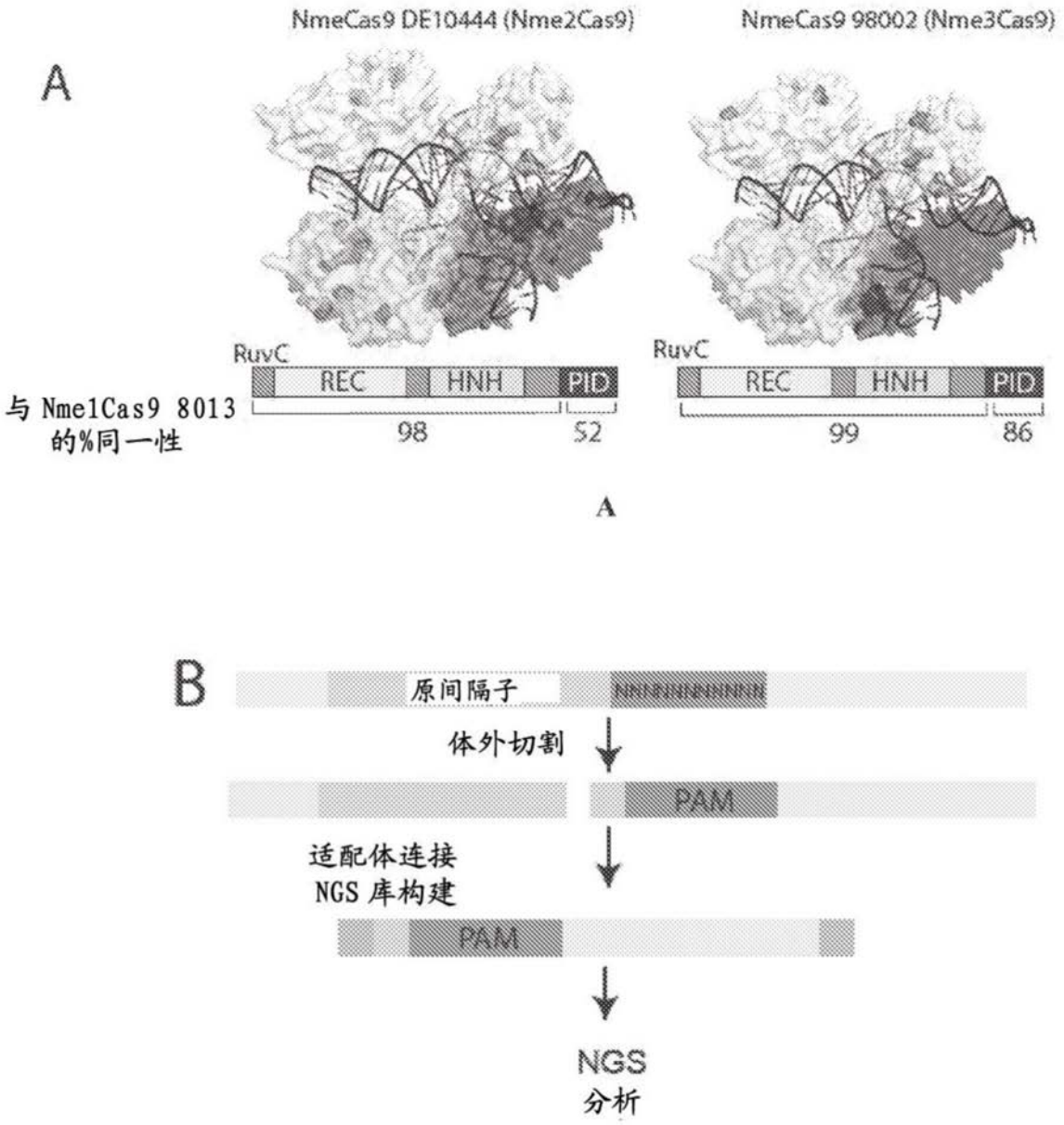

A
NmeCas9 DE10444 (Nme2Cas9)
NmeCas9 98002 (Nme3Cas9)
RuvC
REC
HNH
PID
98
52
99
86
与 Nme1Cas9 8013 的%同一性
A
B
原间隔子
体外切割
PAM
适配体连接
NGS 库构建
PAM
NGS
分析

图34

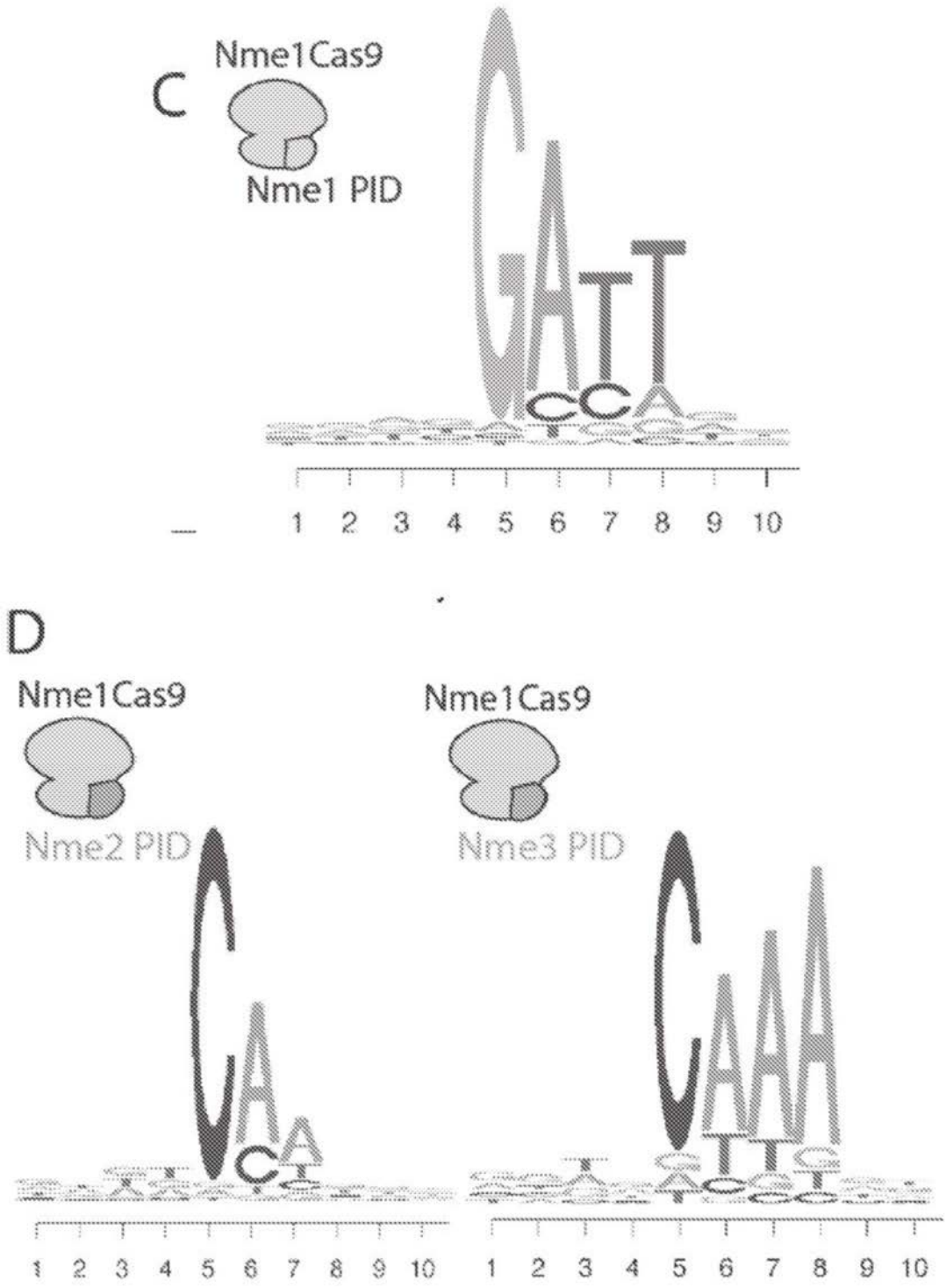

C
Nme1Cas9
Nme1 PID
1 2 3 4 5 6 7 8 9 10
D
Nme1Cas9
Nme2 PID
Nme1Cas9
Nme3 PID
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10

图34续

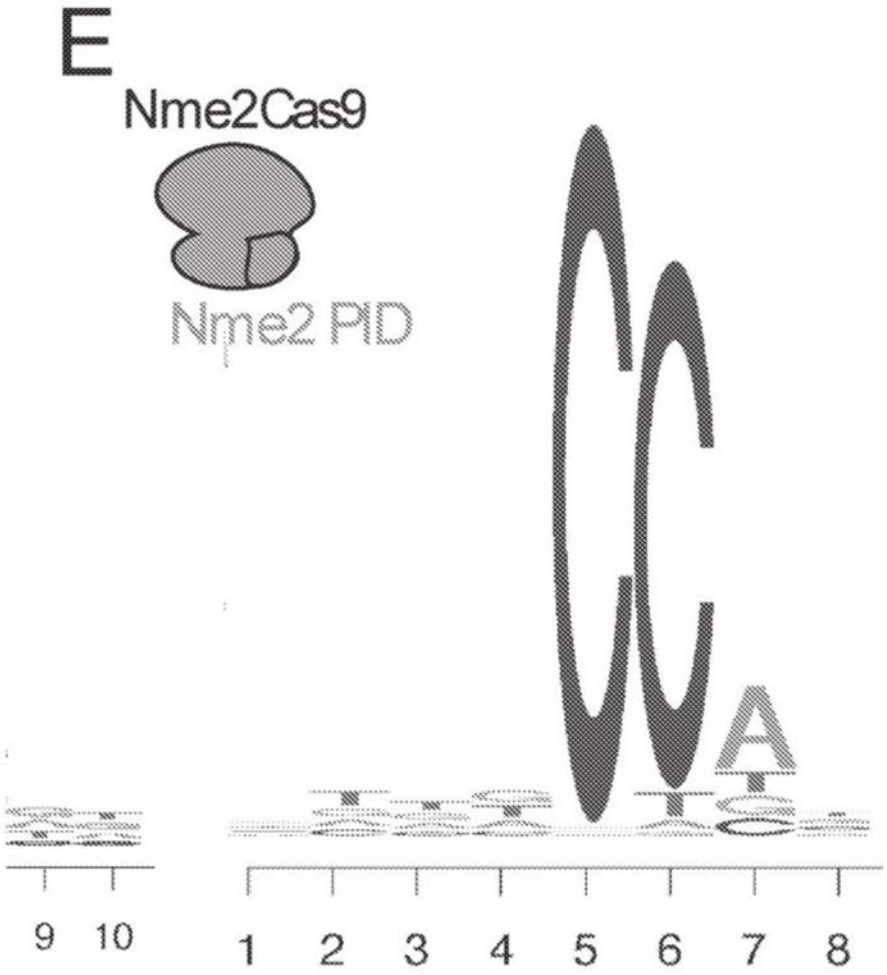


图34续

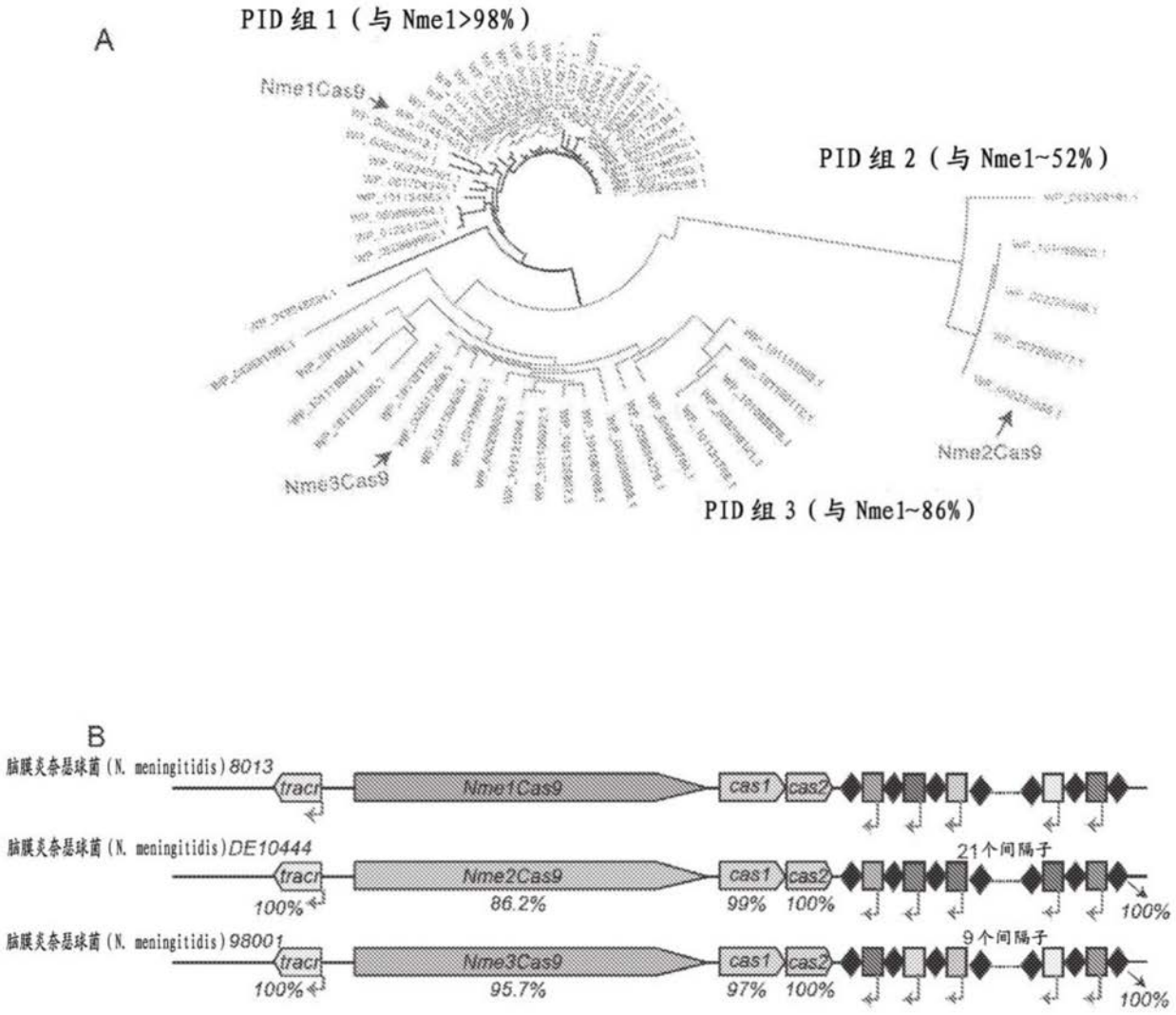

A
PID 组 1（与 Nme1>98%）
Nme1Cas9
PID 组 2（与 Nme1~52%）
Nme2Cas9
Nme3Cas9
PID 组 3（与 Nme1~86%）
B
脑膜炎奈瑟球菌(N. meningitidis)8013
tracr
Nme1Cas9
cas1
cas2
脑膜炎奈瑟球菌(N. meningitidis)DE10444
tracr
Nme2Cas9
cas1
cas2
21个间隔子
100%
86.2%
99%
100%
100%
脑膜炎奈瑟球菌(N. meningitidis)98001
tracr
Nme3Cas9
cas1
cas2
9个间隔子
100%
95.7%
97%
100%
100%

图35

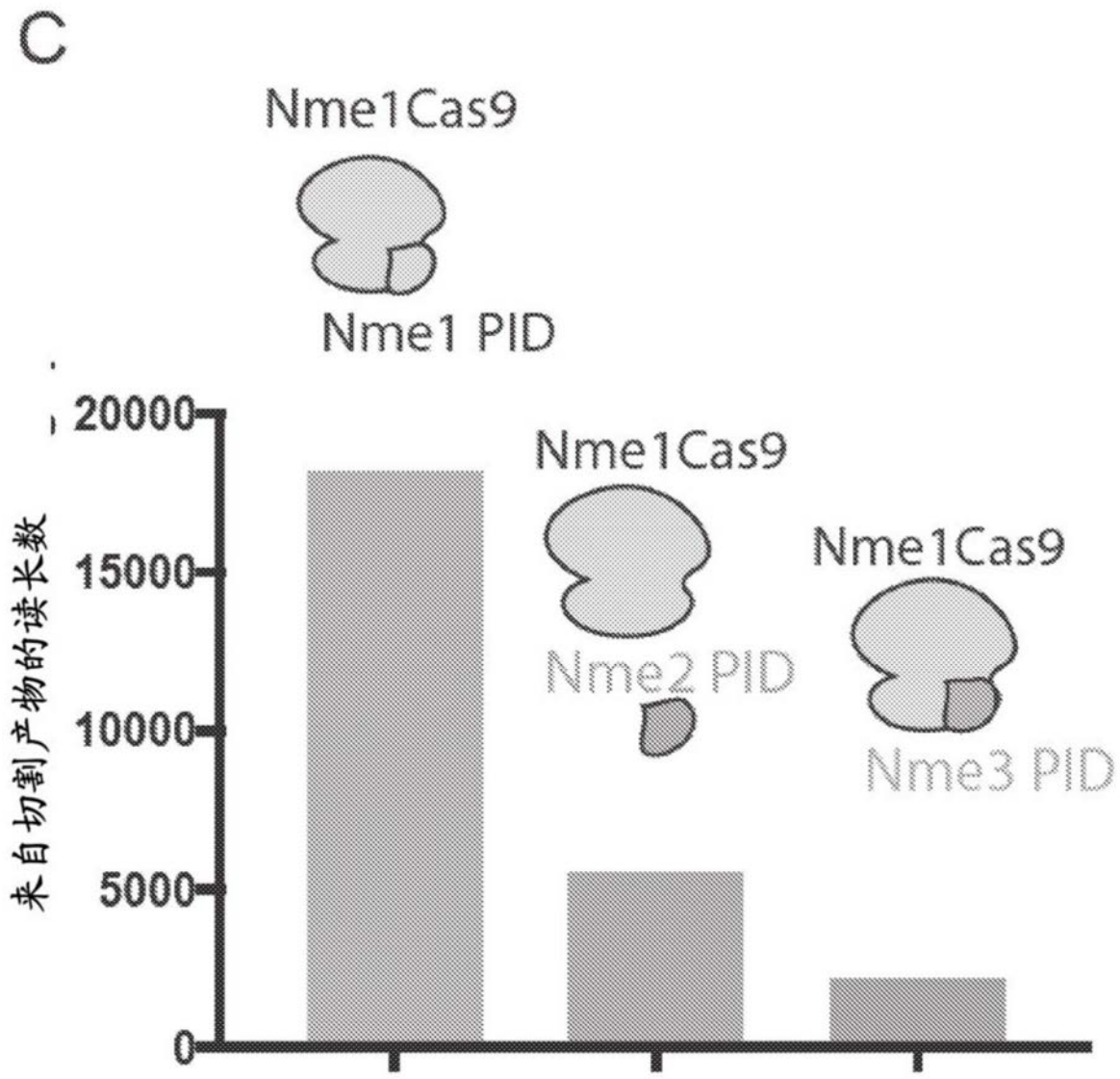

C
Nme1Cas9
Nme1 PID
来自切割产物的读长数
20000
15000
10000
5000
0
Nme1Cas9
Nme2 PID
Nme1Cas9
Nme3 PID

图35续

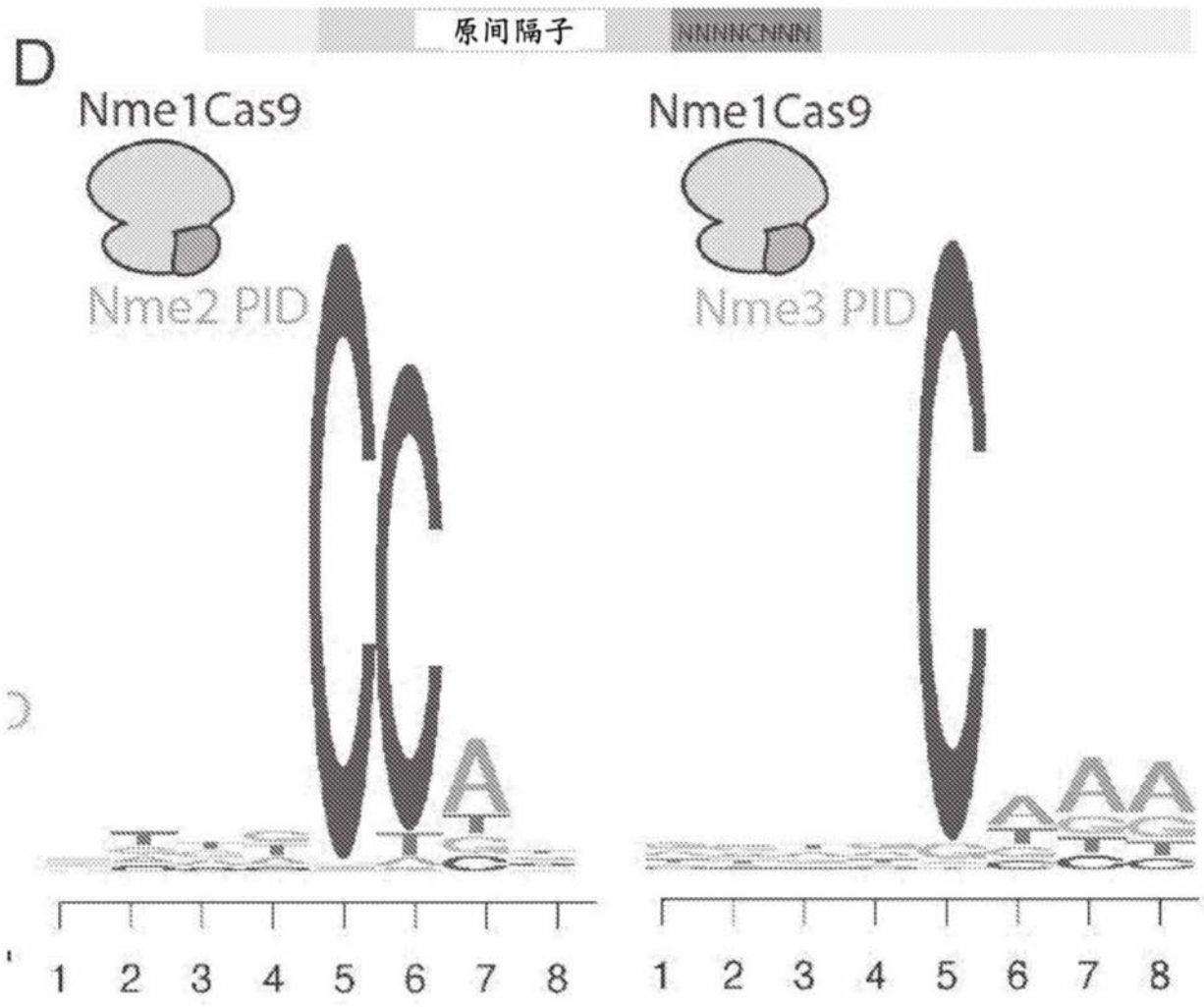

原间隔子
NNNNCNNN
D
Nme1Cas9
Nme2 PID
Nme1Cas9
Nme3 PID
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8

图35续

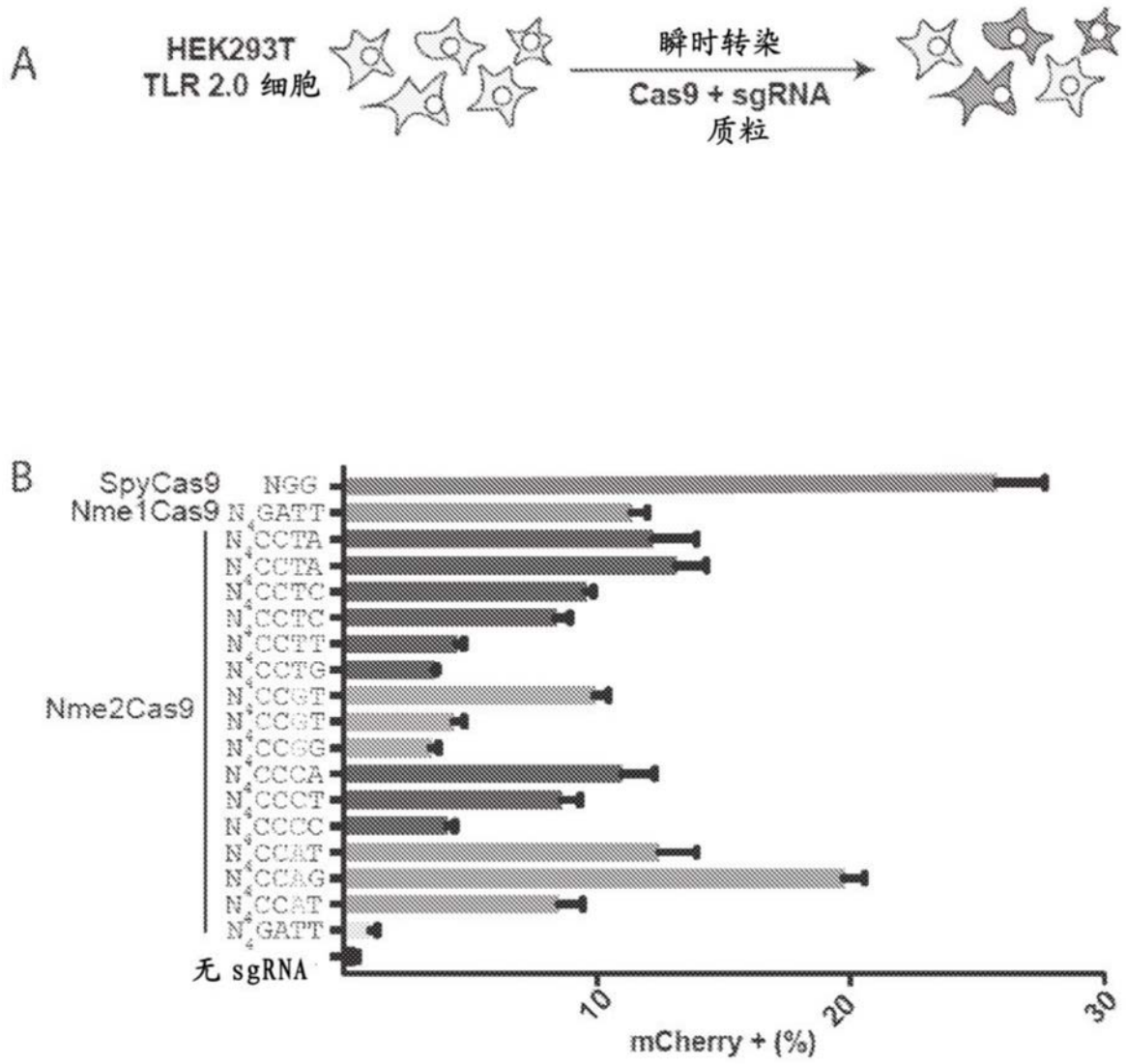
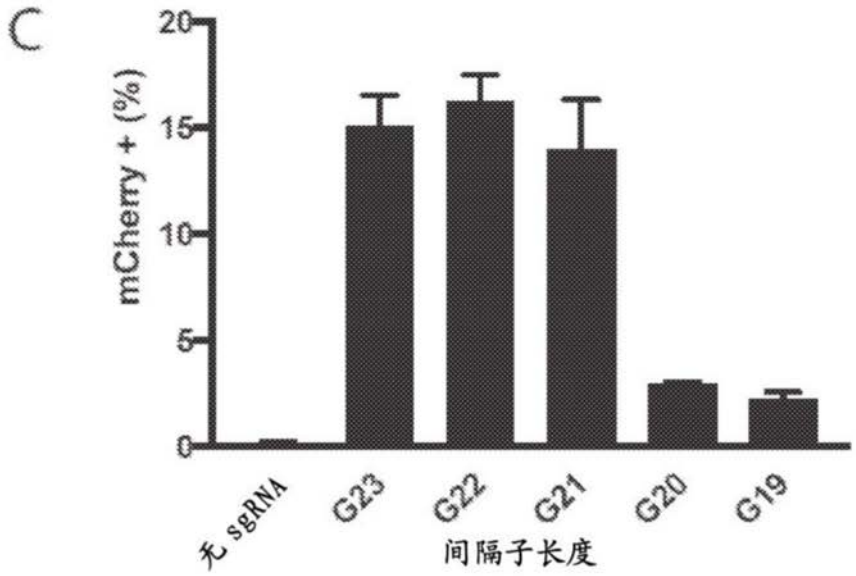

A
HEK293T
TLR 2.0 细胞
瞬时转染
Cas9 + sgRNA
质粒
B
SpyCas9
Nme1Cas9
Nme2Cas9
NGG
N4GATT
N4CCTA
N4CCTA
N4CCTC
N4CCTC
N4CCTT
N4CCTG
N4CCGT
N4CCGT
N4CCGG
N4CCCA
N4CCCT
N4CCCC
N4CCAT
N4CCAG
N4CCAT
N4GATT
无 sgRNA
10
20
30
mCherry + (%)

图36


C
mCherry + (%)
20
15
10
5
0
无 sgRNA
G23
G22
G21
G20
G19
间隔子长度

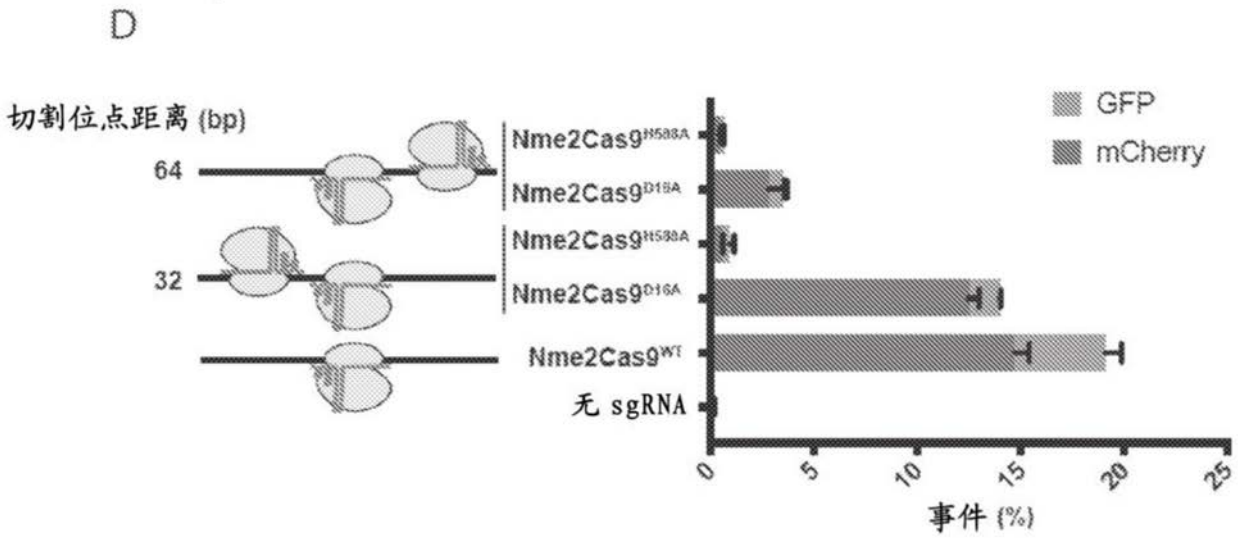

D
切割位点距离 (bp)
64
32
Nme2Cas9H588A
Nme2Cas9D16A
Nme2Cas9H588A
Nme2Cas9D16A
Nme2Cas9WT
无 sgRNA
GFP
mCherry
0
5
10
15
20
25
事件 (%)

图36续

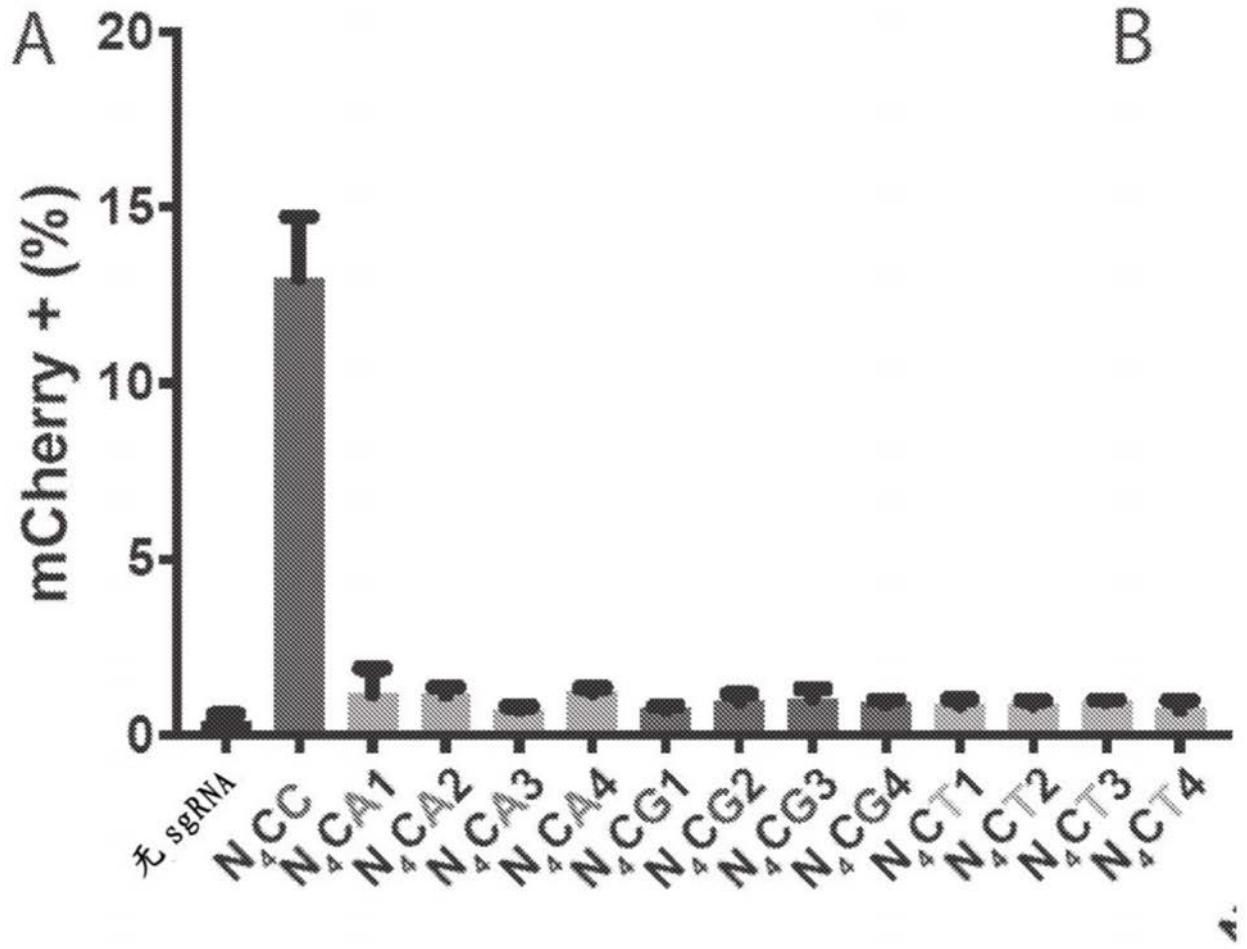

A
B
20
15
10
5
0
mCherry + (%)
无 sgRNA
NCC
NCA1
NCA2
NCA3
NCA4
NCG1
NCG2
NCG3
NCG4
NCT1
NCT2
NCT3
NCT4

图37

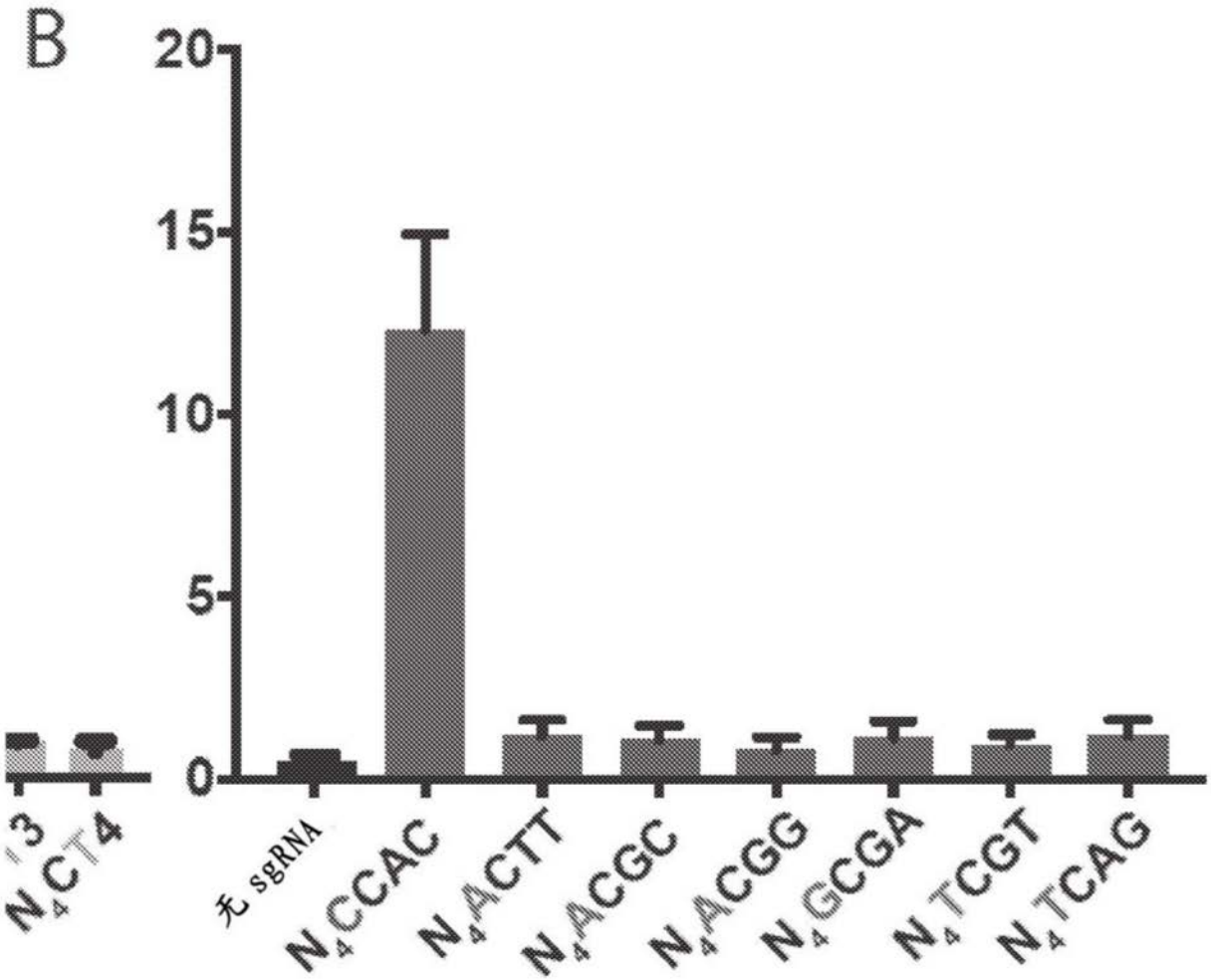
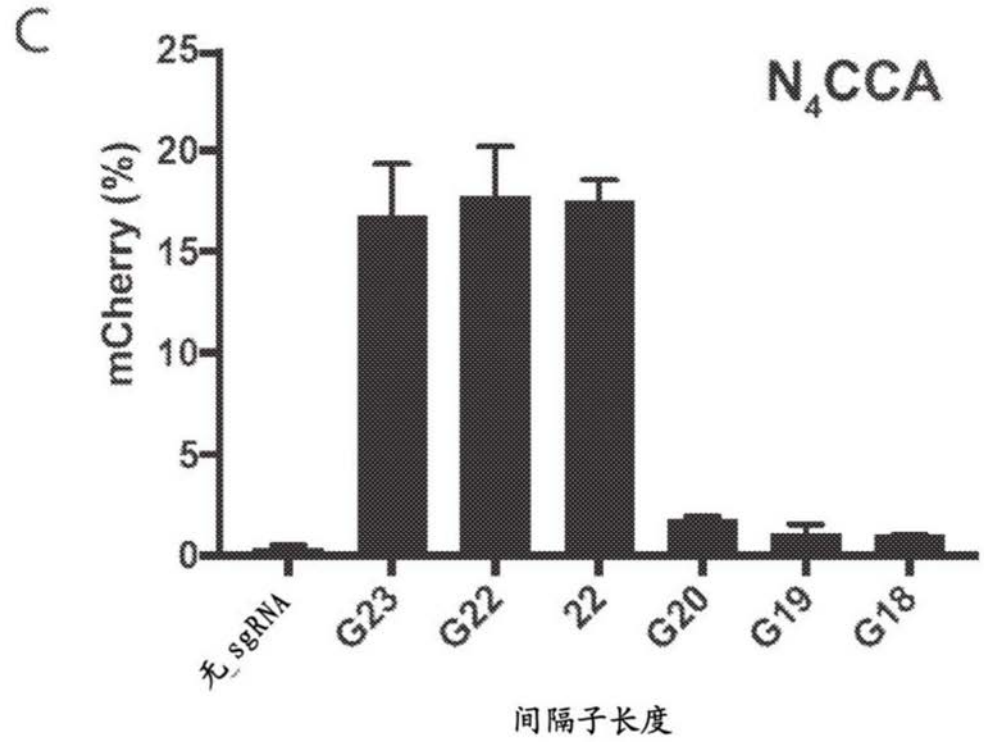

B
20
15
10
5
0
无 sgRNA
N4CCAC
N4ACTT
N4ACGC
N4ACGG
N4GCGA
N4TCGT
N4TCAG

图37续


C
N4CCA
mCherry (%)
25
20
15
10
5
0
无_sgRNA
G23
G22
22
G20
G19
G18
间隔子长度

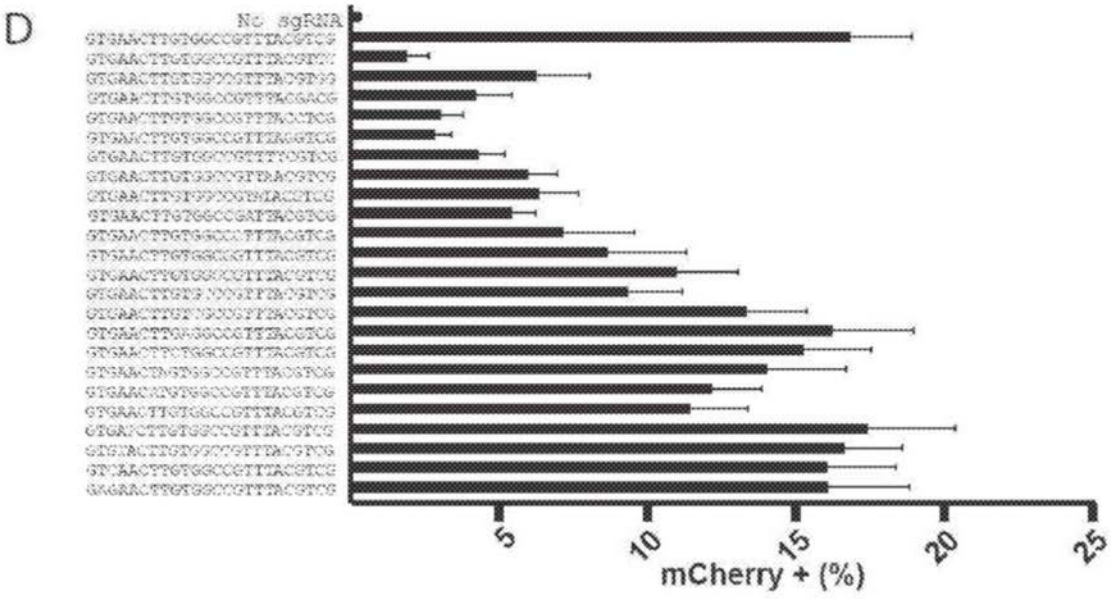

D
No sgRNA
5
10
15
20
25
mCherry + (%)

图37续

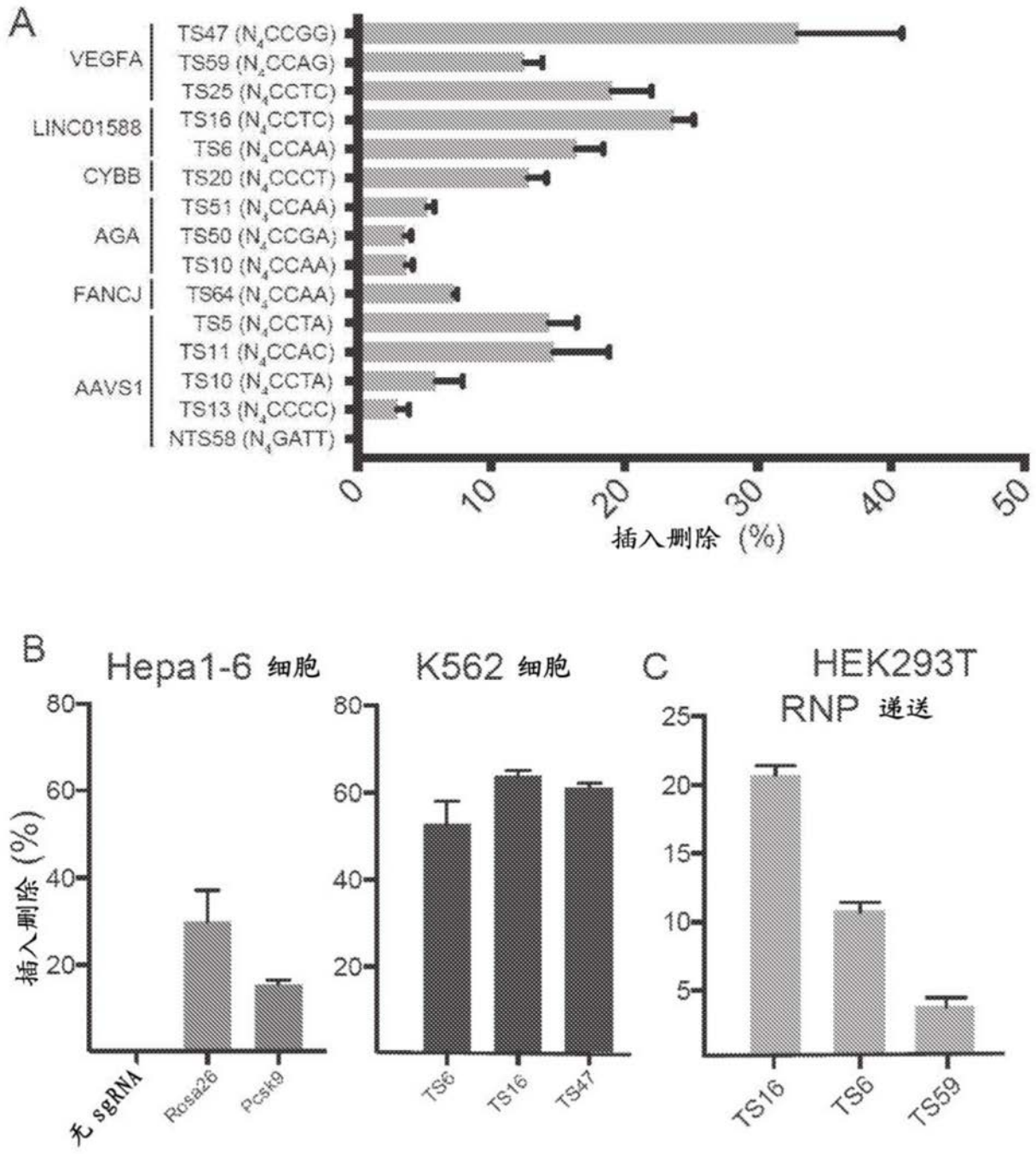

A
VEGFA
LINC01588
CYBB
AGA
FANCJ
AAVS1
TS47 (N4CCGG)
TS59 (N4CCAG)
TS25 (N4CCTC)
TS16 (N4CCTC)
TS6 (N4CCAA)
TS20 (N4CCCT)
TS51 (N3CCAA)
TS50 (N4CCGA)
TS10 (N4CCAA)
TS64 (N4CCAA)
TS5 (N4CCTA)
TS11 (N4CCAC)
TS10 (N4CCTA)
TS13 (N4CCCC)
NTS58 (N4GATT)
0 10 20 30 40 50
插入删除 (%)
B
Hepa1-6 细胞
K562 细胞
插入删除 (%)
80 60 40 20
无 sgRNA
Rosa26
Pcsk9
TS6
TS16
TS47
C
HEK293T
RNP 递送
25 20 15 10 5
TS16
TS6
TS59

图38

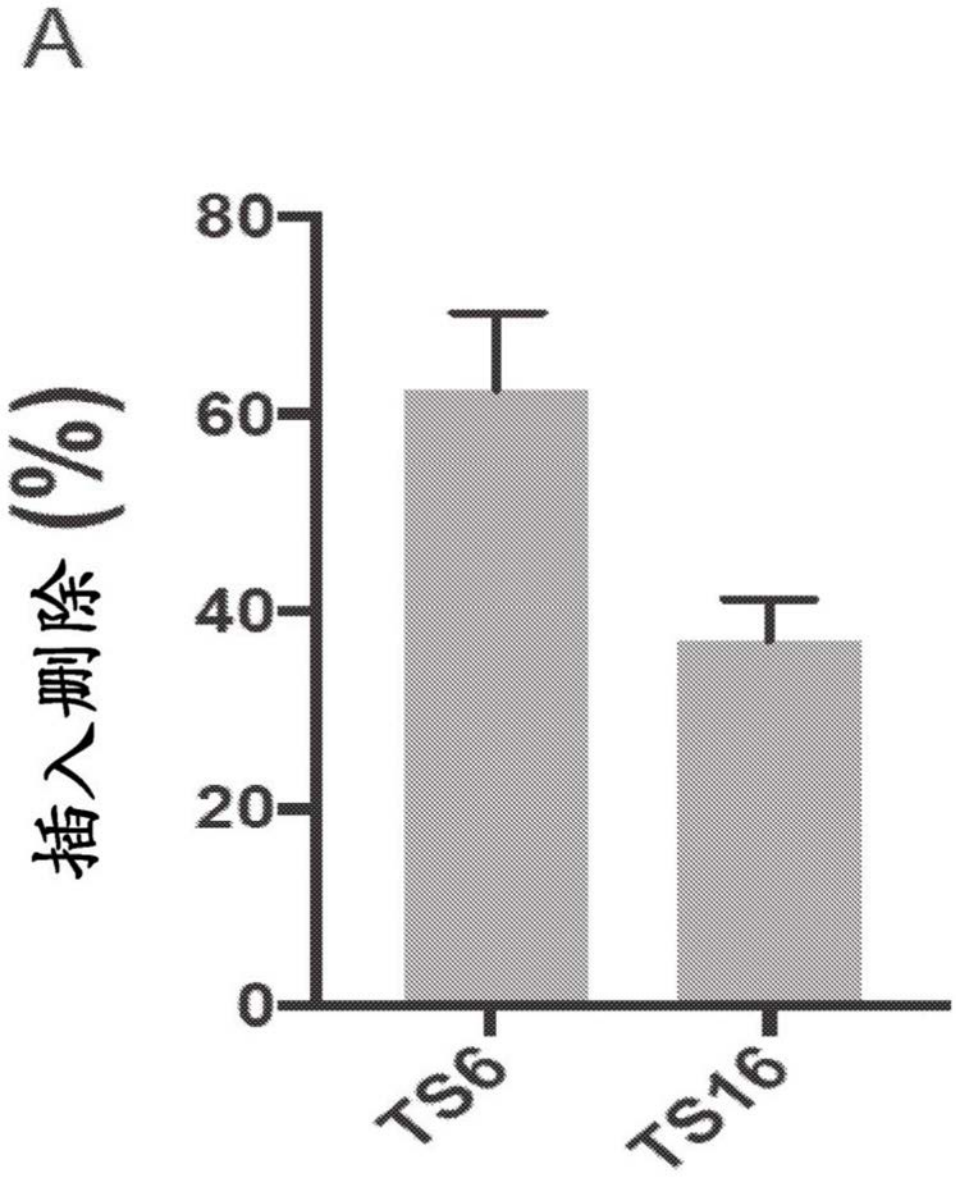


图39

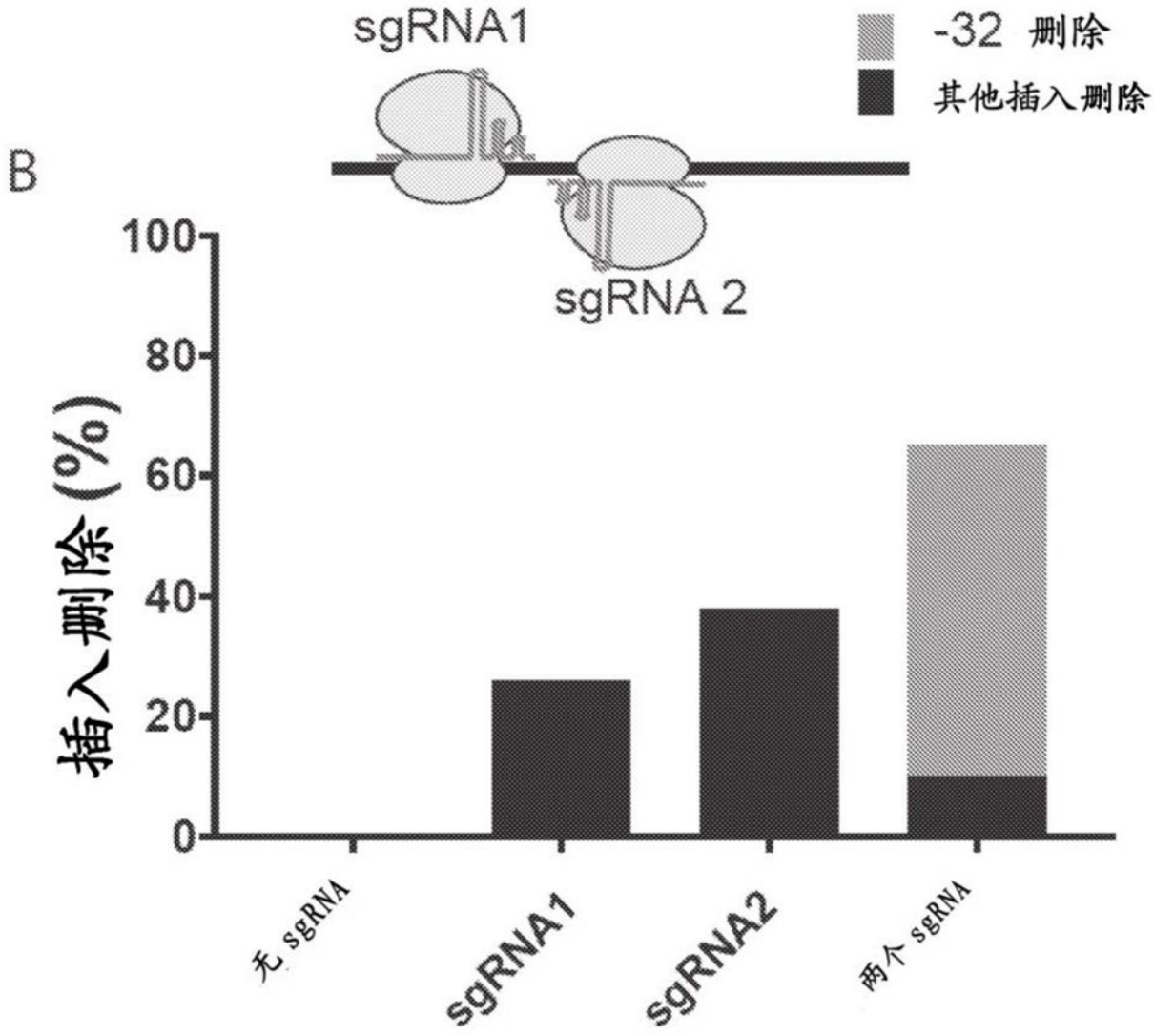

B
sgRNA1
sgRNA 2
-32 删除
其他插入删除
插入删除 (%)
100
80
60
40
20
0
无 sgRNA
sgRNA1
sgRNA2
两个 sgRNA

图39续

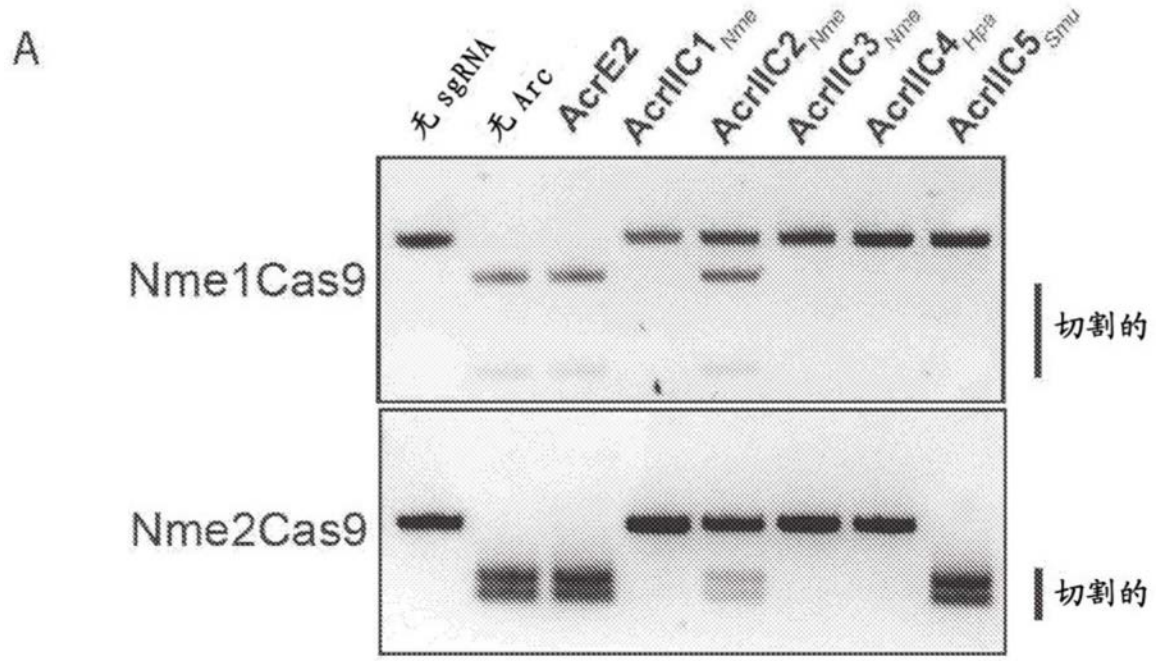

A
无 sgRNA
无 Arc
AcrE2
AcrIIC1 Nme
AcrIIC2 Nme
AcrIIC3 Nme
AcrIIC4 Hpa
AcrIIC5 Smu
Nme1Cas9
切割的
Nme2Cas9
切割的

图40

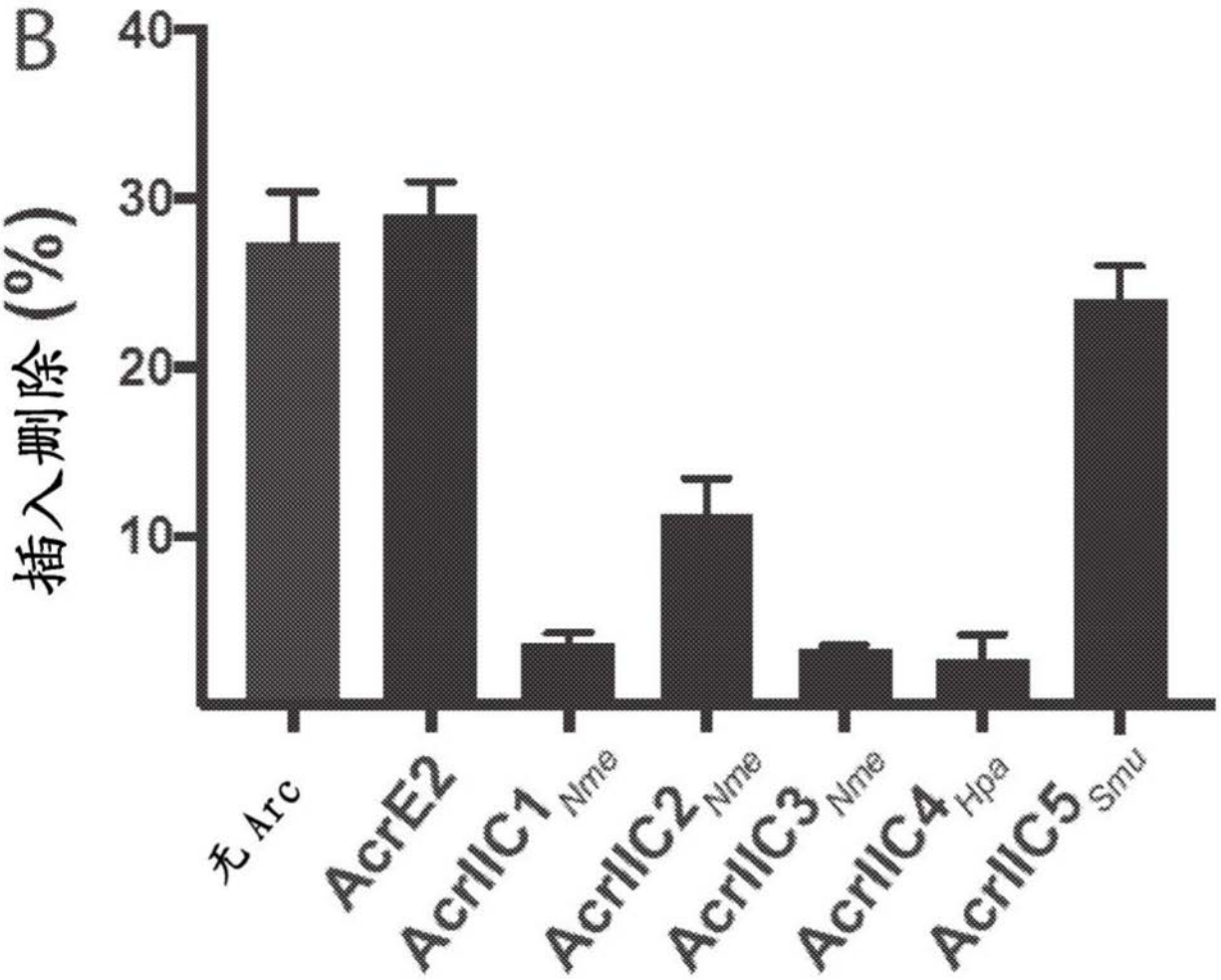

B
插入删除 (%)
40
30
20
10
无 Arc
AcrE2
AcrIIC1 Nme
AcrIIC2 Nme
AcrIIC3 Nme
AcrIIC4 Hpa
AcrIIC5 Smu

图40续

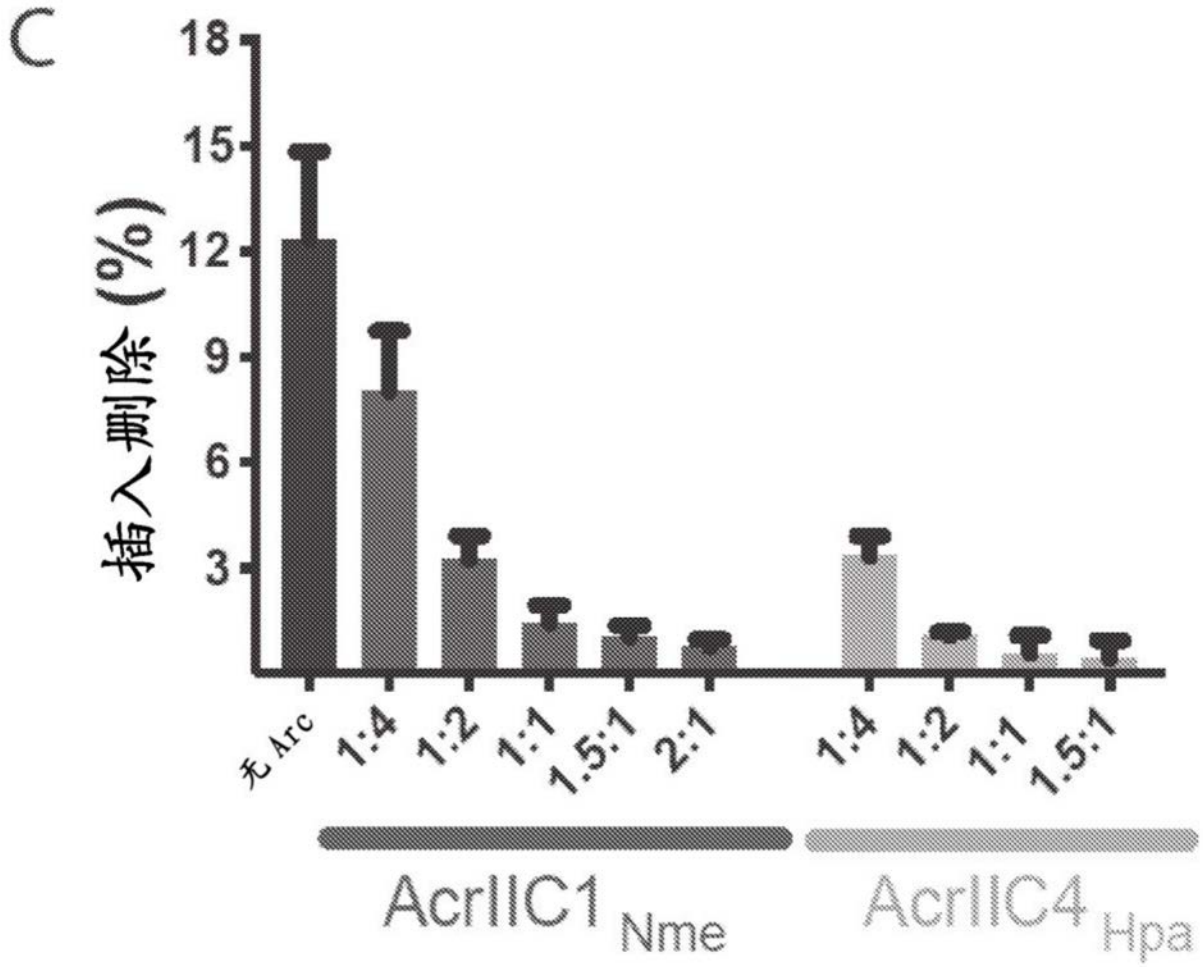

C
插入删除 (%)
18
15
12
9
6
3
无 Arc
1:4
1:2
1:1
1.5:1
2:1
1:4
1:2
1:1
1.5:1
AcrIIC1 Nme
AcrIIC4 Hpa

图40续

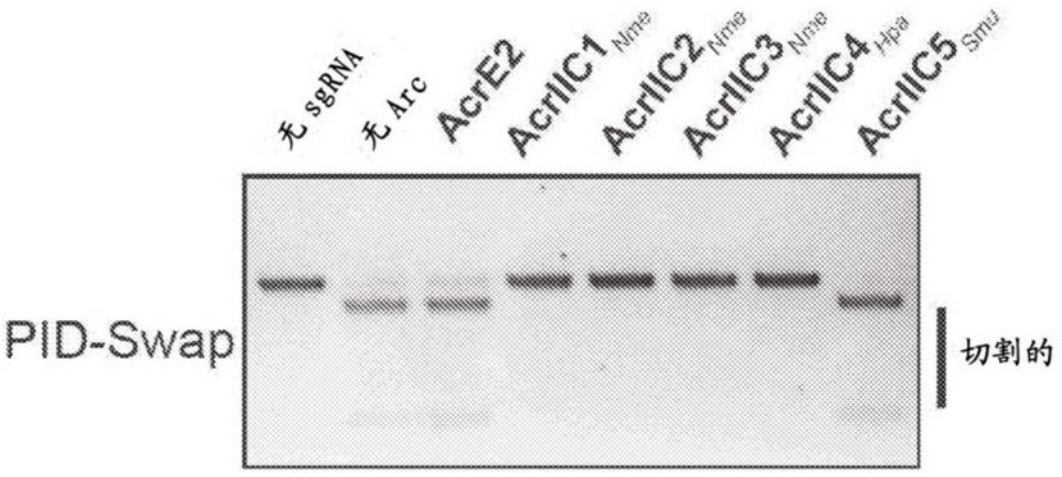

无 sgRNA
无 Arc
AcrE2
AcrIIC1 Nme
AcrIIC2 Nme
AcrIIC3 Nme
AcrIIC4 Hpa
AcrIIC5 Smu
PID-Swap
切割的

图41

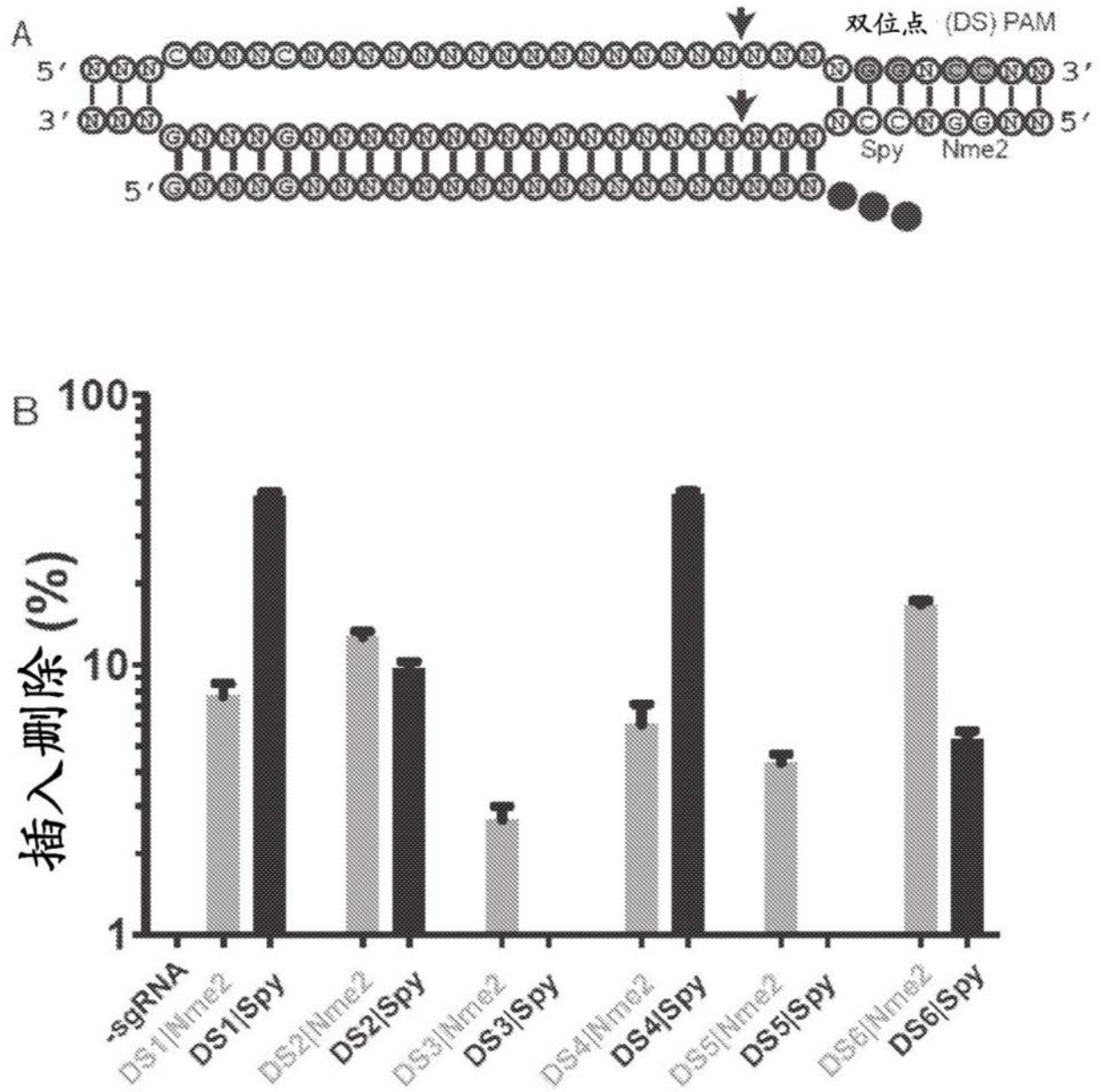

A
双位点 (DS) PAM
5′
3′
Spy
Nme2
B
插入删除 (%)
100
10
1
-sgRNA
DS1|Nme2
DS1|Spy
DS2|Nme2
DS2|Spy
DS3|Nme2
DS3|Spy
DS4|Nme2
DS4|Spy
DS5|Nme2
DS5|Spy
DS6|Nme2
DS6|Spy

图42

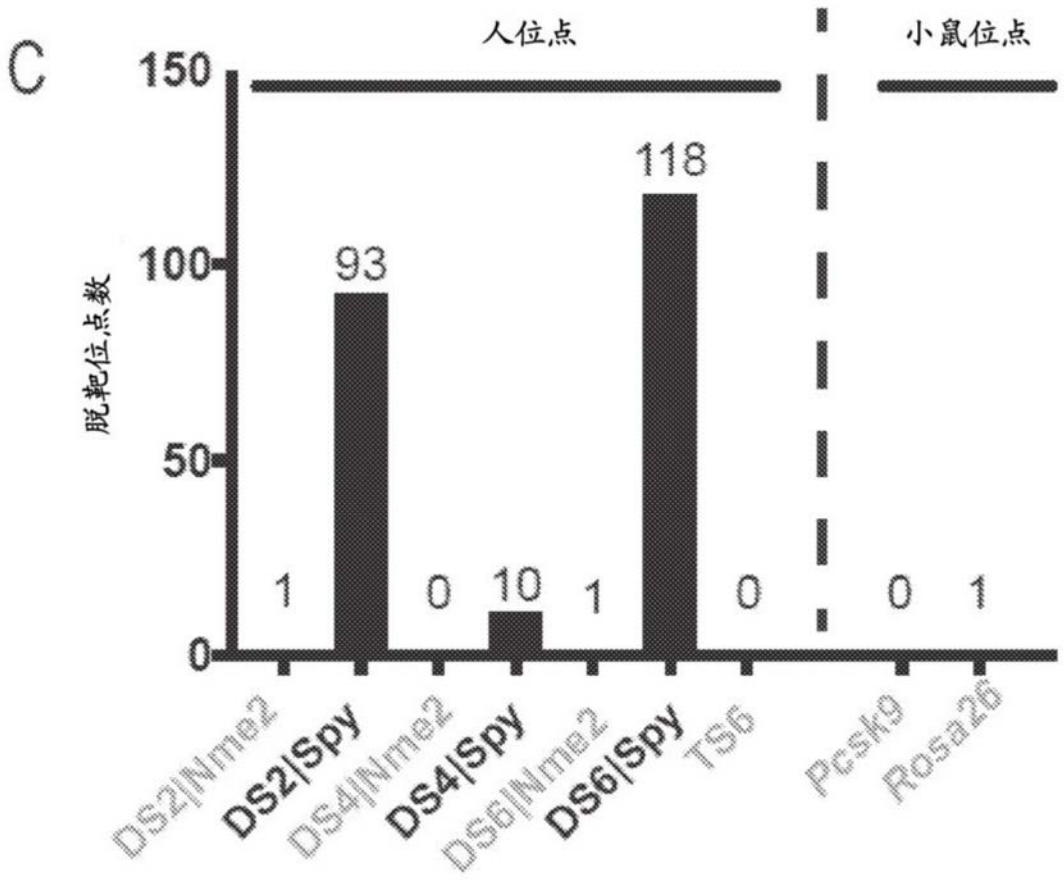

C
人位点
小鼠位点
脱靶位点数
150
100
50
0
1
93
0
10
1
118
0
0
1
DS2|Nme2
DS2|Spy
DS4|Nme2
DS4|Spy
DS6|Nme2
DS6|Spy
TS6
Pcsk9
Rosa26

图42续

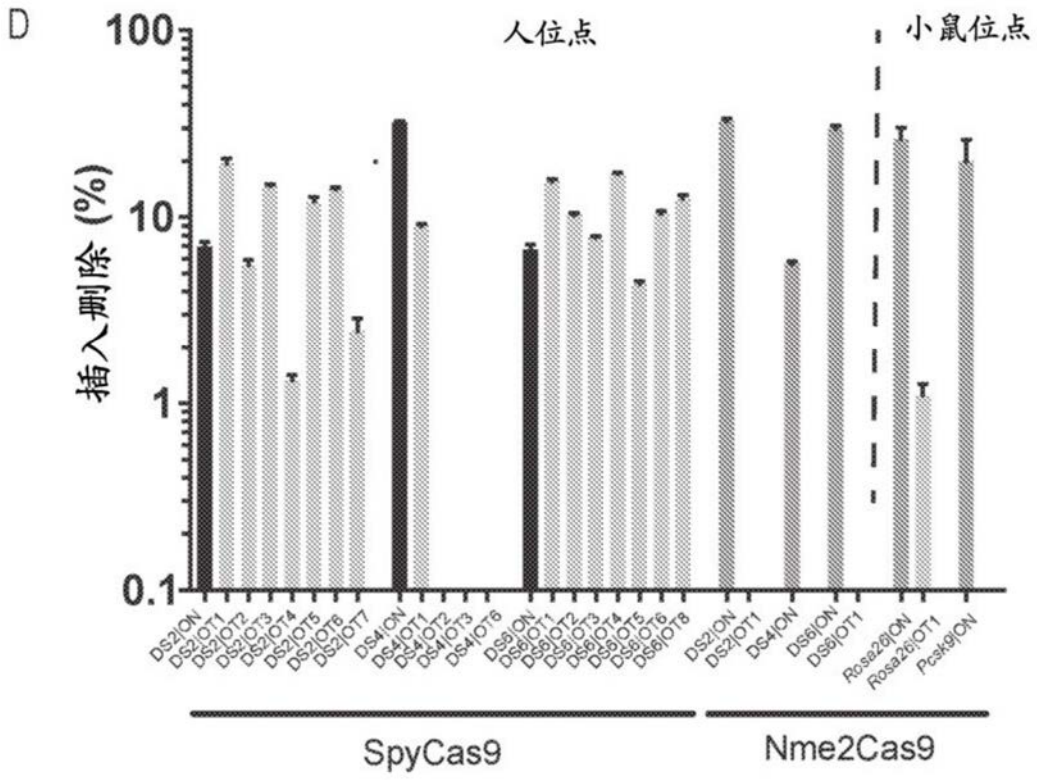

D
100
10
1
0.1
插入删除 (%)
人位点
小鼠位点
DS2|ON
DS2|OT1
DS2|OT2
DS2|OT3
DS2|OT4
DS2|OT5
DS2|OT6
DS2|OT7
DS4|ON
DS4|OT1
DS4|OT2
DS4|OT3
DS4|OT6
DS6|ON
DS6|OT1
DS6|OT2
DS6|OT3
DS6|OT4
DS6|OT5
DS6|OT6
DS6|OT8
DS2|ON
DS2|OT1
DS4|ON
DS6|ON
DS6|OT1
Rosa26|ON
Rosa26|OT1
Pcsk9|ON
SpyCas9
Nme2Cas9

图42续

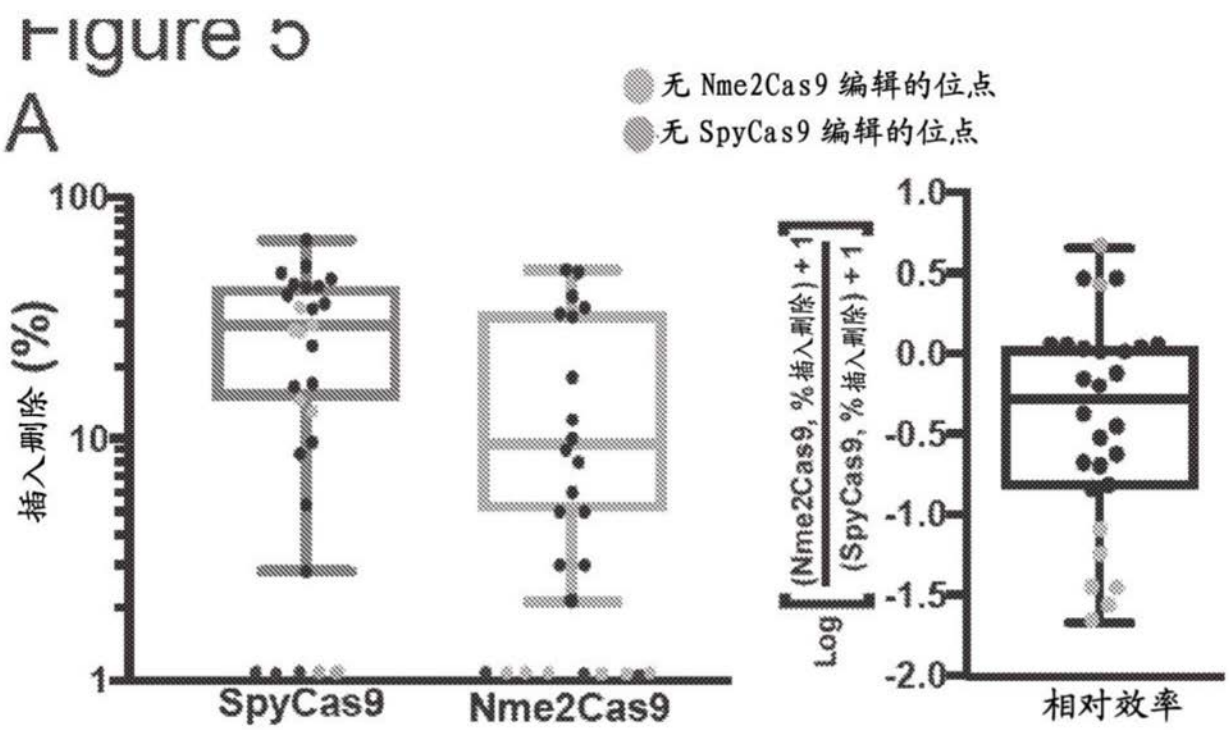


图42E

Rosa26|On : TGAGGACCGCCCTGGGCCTGGGAGAATCCCT
Rosa26|OT1 : GAAGGACCACCCTAGGCCTGGGAGACTCCCT

图42F

图42续

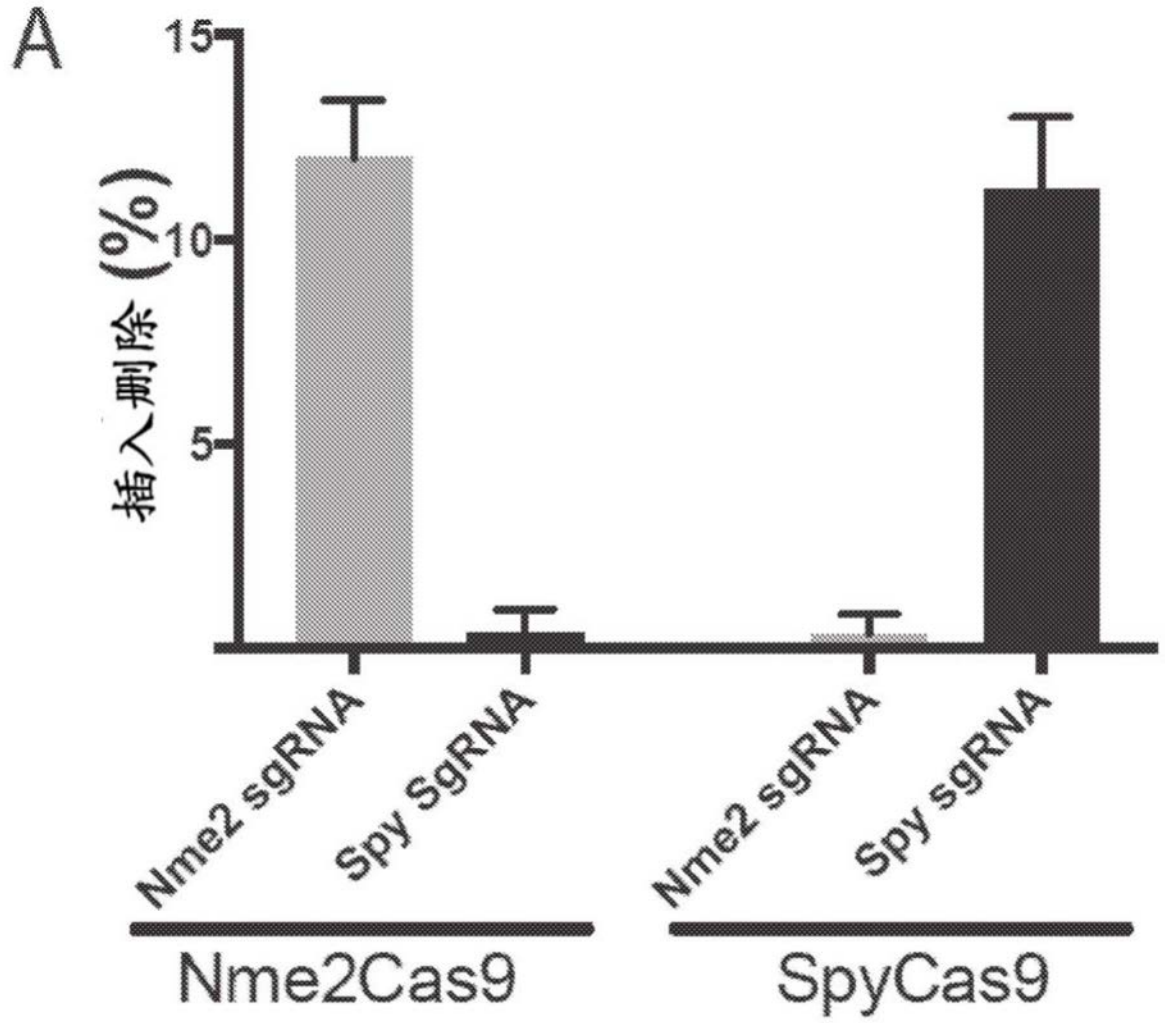

A
15
10
5
插入删除 (%)
Nme2 sgRNA
Spy SgRNA
Nme2 sgRNA
Spy sgRNA
Nme2Cas9
SpyCas9

图43

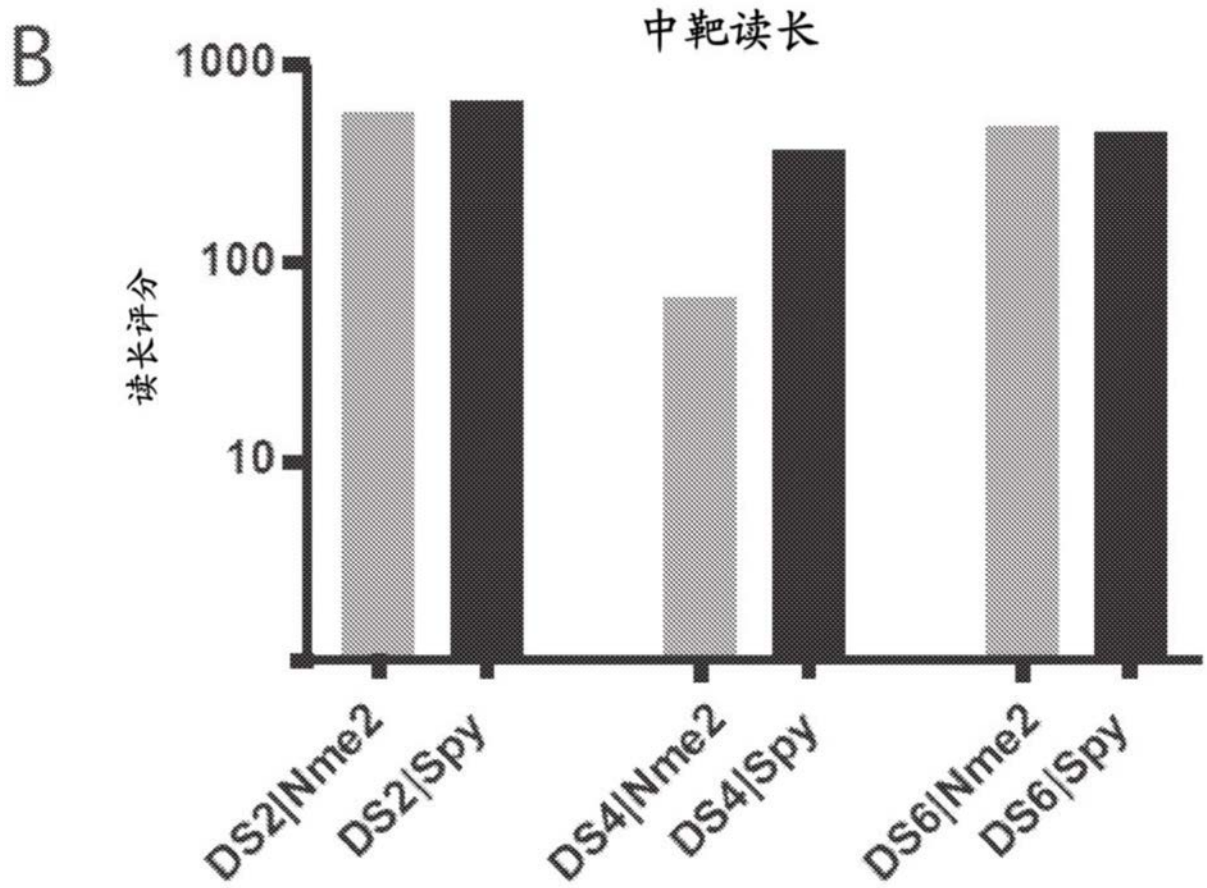

B
中靶读长
1000
100
10
读长评分
DS2|Nme2
DS2|Spy
DS4|Nme2
DS4|Spy
DS6|Nme2
DS6|Spy

图43续

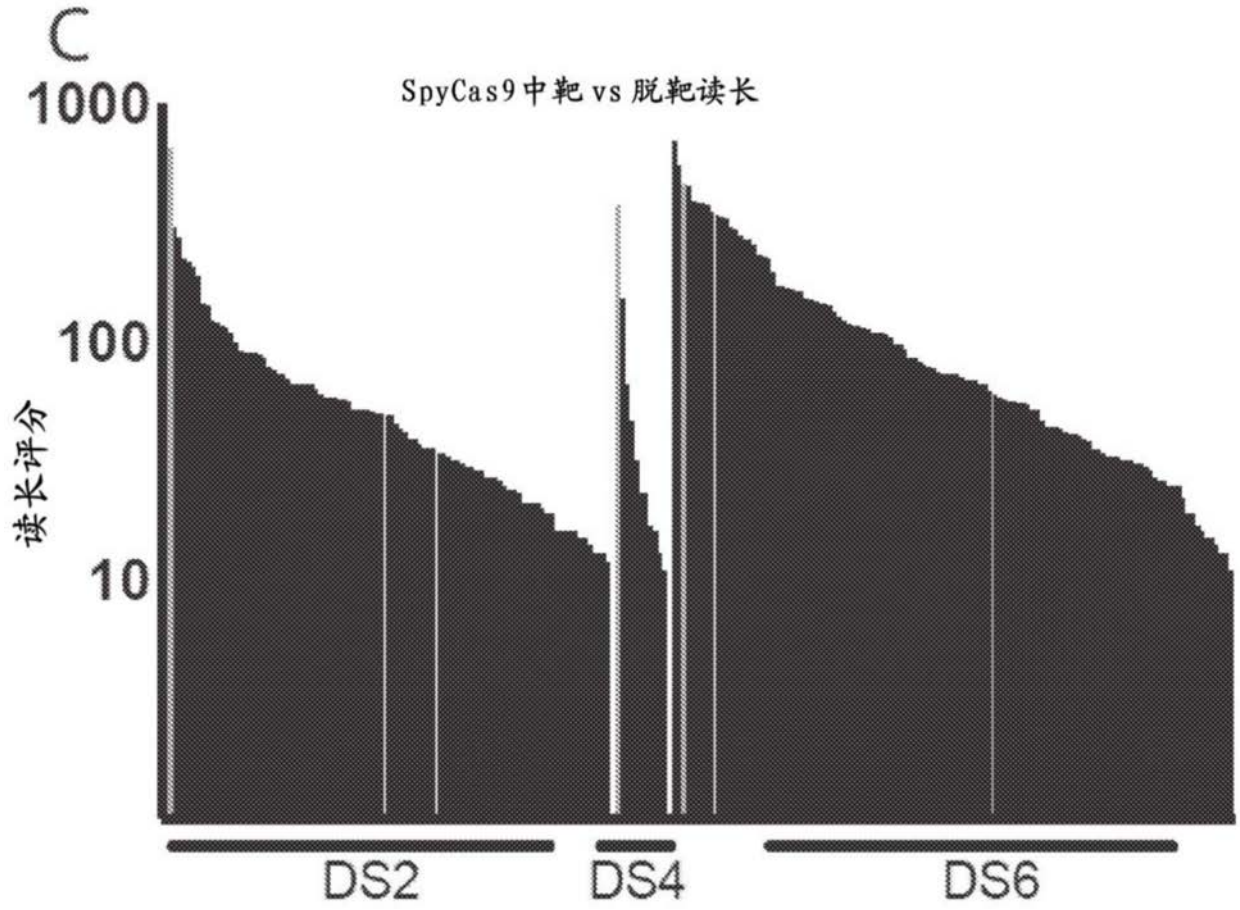


图43续

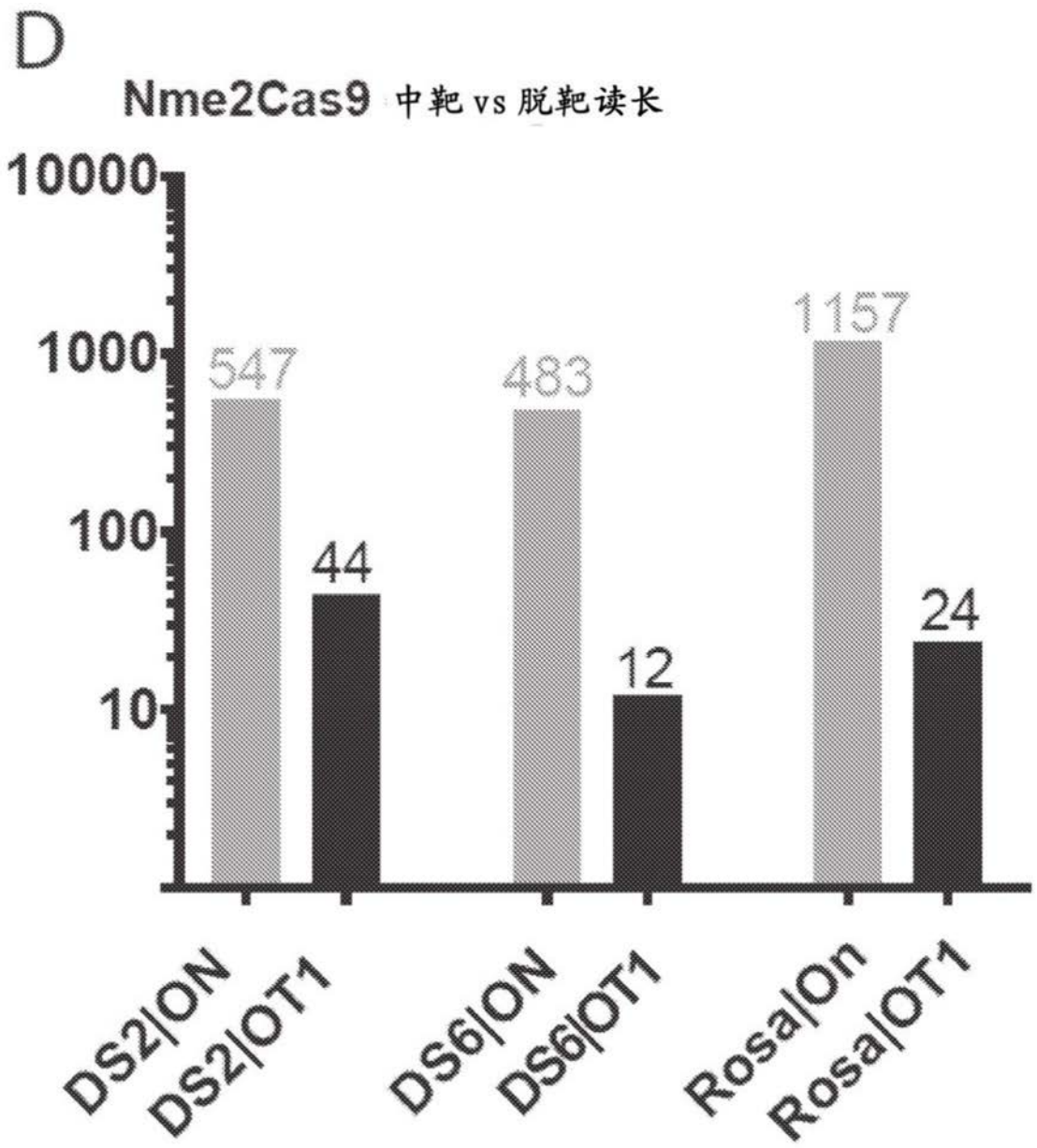

D
Nme2Cas9 中靶 vs 脱靶读长
10000
1000
100
10
547
44
483
12
1157
24
DS2|ON
DS2|OT1
DS6|ON
DS6|OT1
Rosa|On
Rosa|OT1

图43续

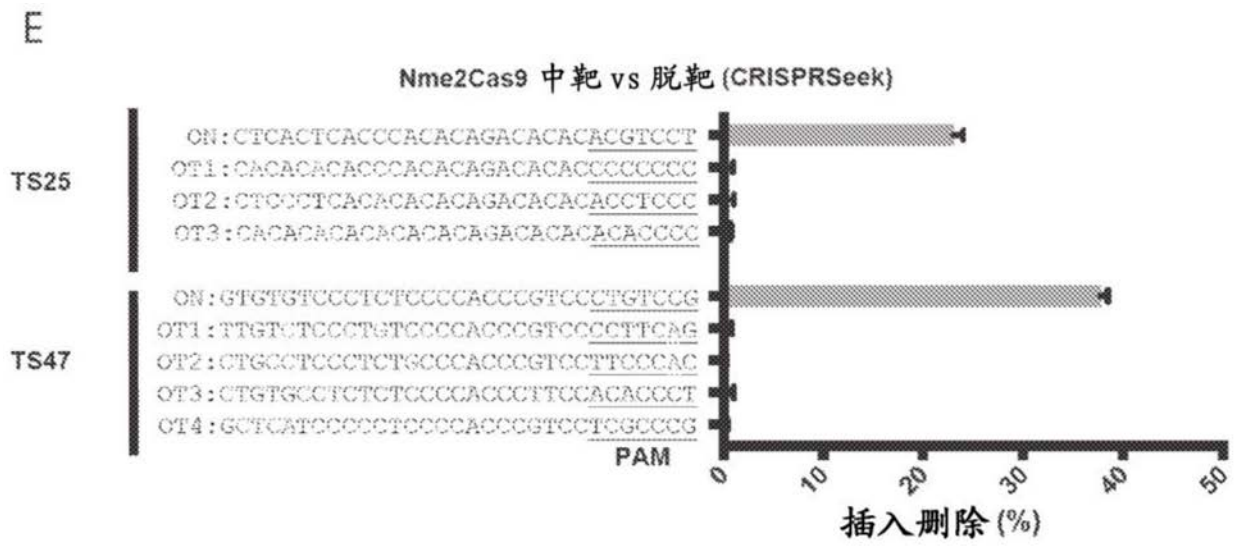
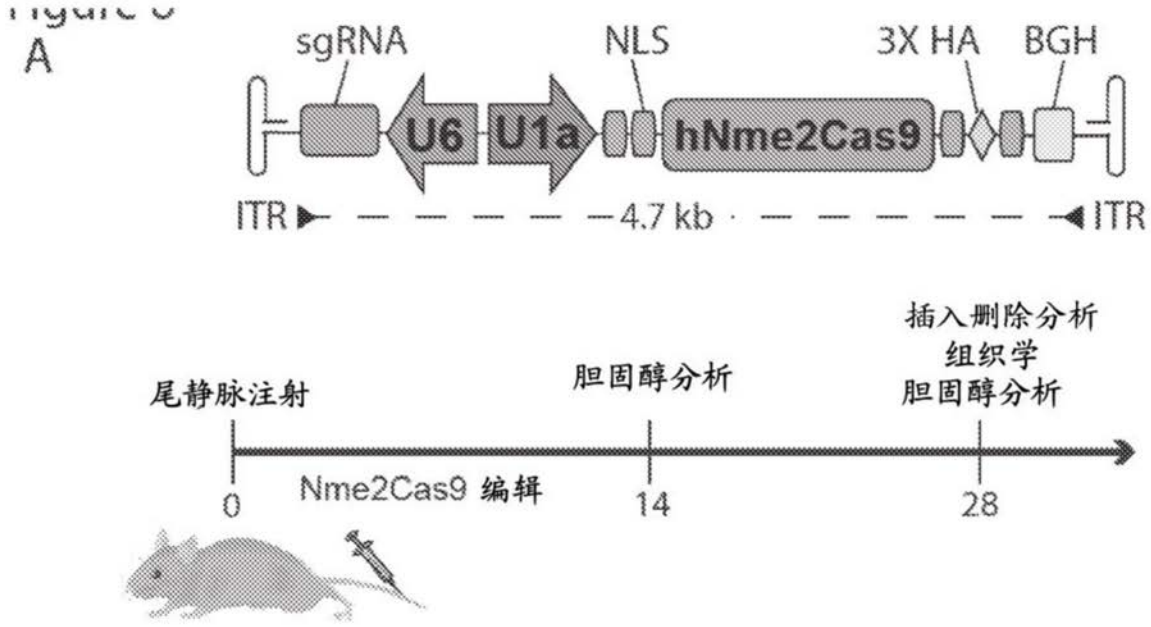

E
Nme2Cas9 中靶 vs 脱靶 (CRISPRSeek)
TS25
ON:CTCACTCACCCACACAGACACACACGTCCT
OT1:CACACACACCCACACAGACACACCCCCCCC
OT2:CTCCCTCACACACACAGACACACACCTCCC
OT3:CACACACACACACACAGACACACACACACCCC
TS47
ON:GTGTGTCCCTCTCCCCACCCGTCCCTGTCCG
OT1:TTGTCTCCCTGTCCCCACCCGTCCCCTTCAG
OT2:CTGCCTCCCTCTGCCCACCCGTCCTTCCCAC
OT3:CTGTGCCTCTCTCCCCACCCTTCCACACCCT
OT4:GCTCATCCCCCTCCCCACCCGTCCTCGCCCG
PAM
0
10
20
30
40
50
插入删除 (%)

图43续


A
sgRNA
NLS
3X HA
BGH
U6
U1a
hNme2Cas9
ITR
4.7 kb
ITR
插入删除分析
组织学
胆固醇分析
尾静脉注射
胆固醇分析
0
Nme2Cas9 编辑
14
28

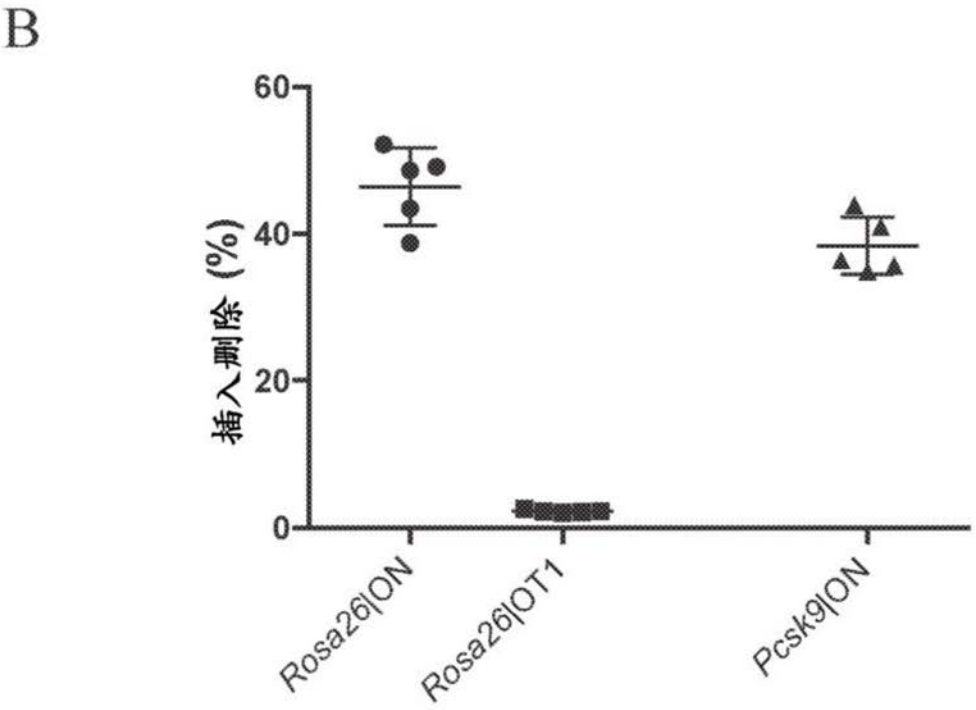
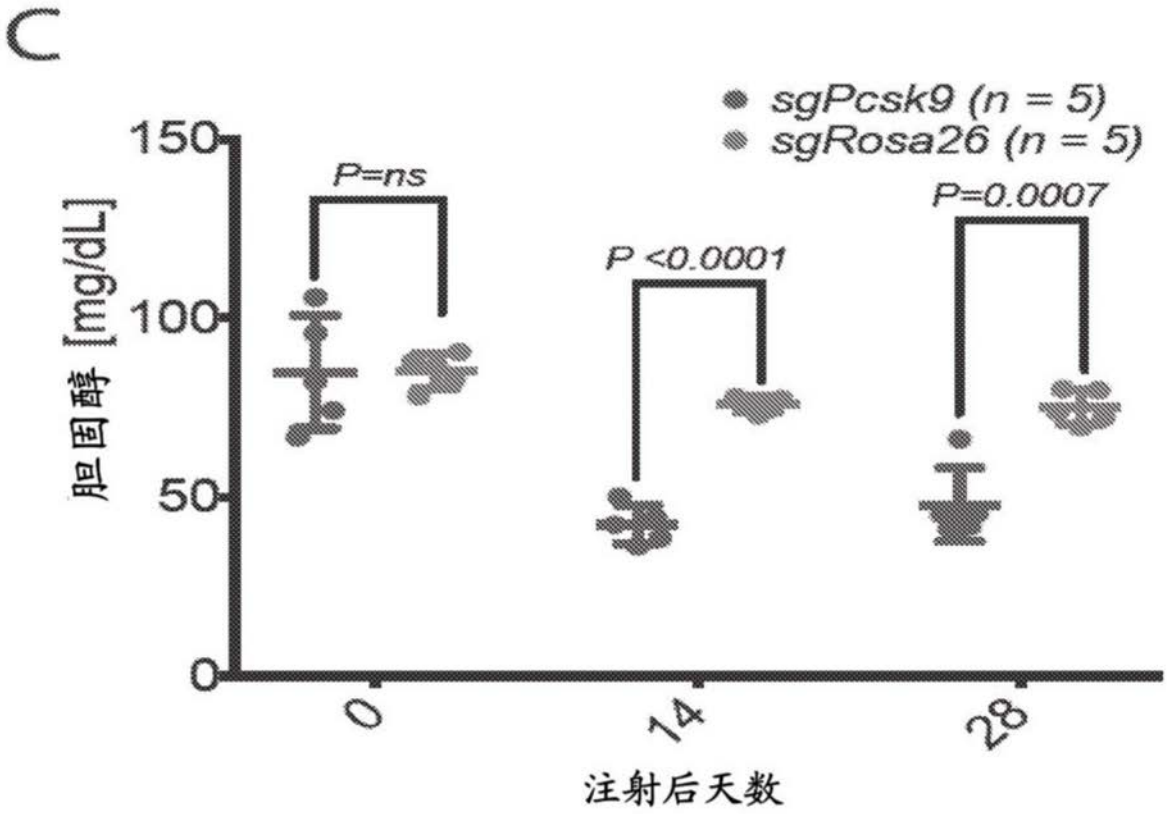

B
60
40
20
0
插入删除 (%)
Rosa26|ON
Rosa26|OT1
Pcsk9|ON

图44


C
sgPcsk9 (n = 5)
sgRosa26 (n = 5)
P=ns
P <0.0001
P=0.0007
胆固醇 [mg/dL]
150
100
50
0
0
14
28
注射后天数

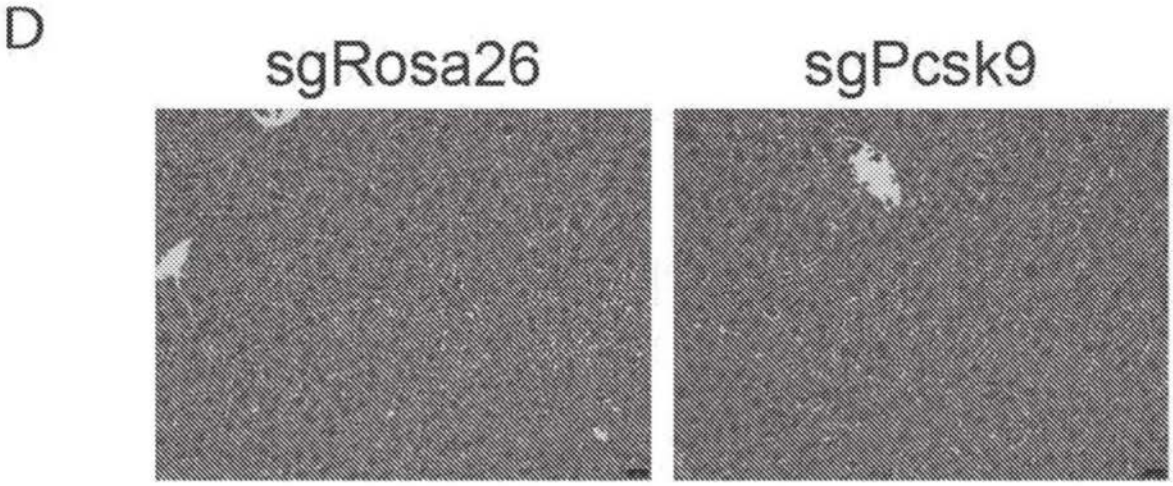

D
sgRosa26
sgPcsk9

图44续

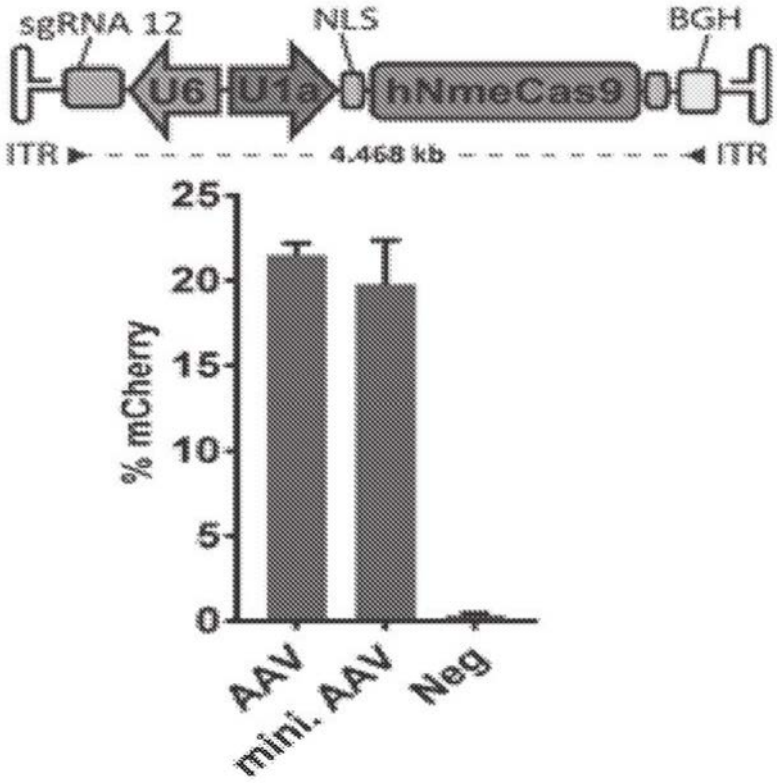

sgRNA 12
NLS
BGH
U6
U1a
hNmeCas9
ITR
4.468 kb
ITR
25
20
15
10
5
0
% mCherry
AAV
mini. AAV
Neg

图45

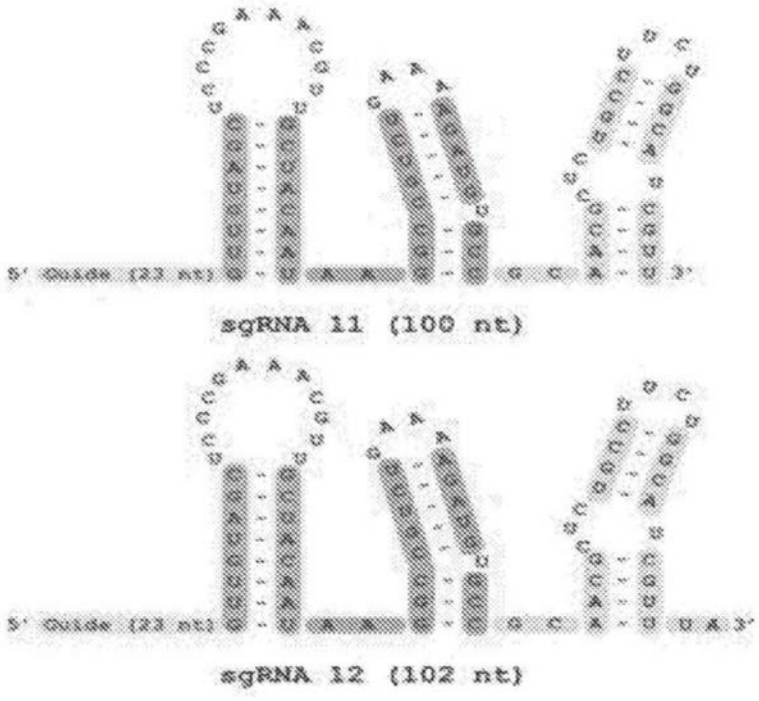

5' Guide (23 nt)
sgRNA 11 (100 nt)
5' Guide (23 nt)
sgRNA 12 (102 nt)

图46

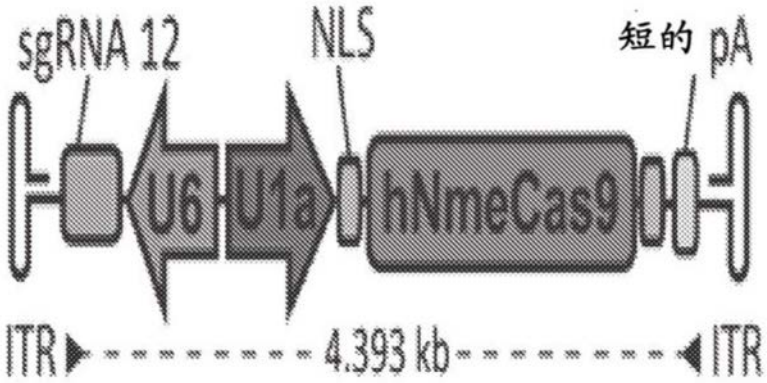


图47

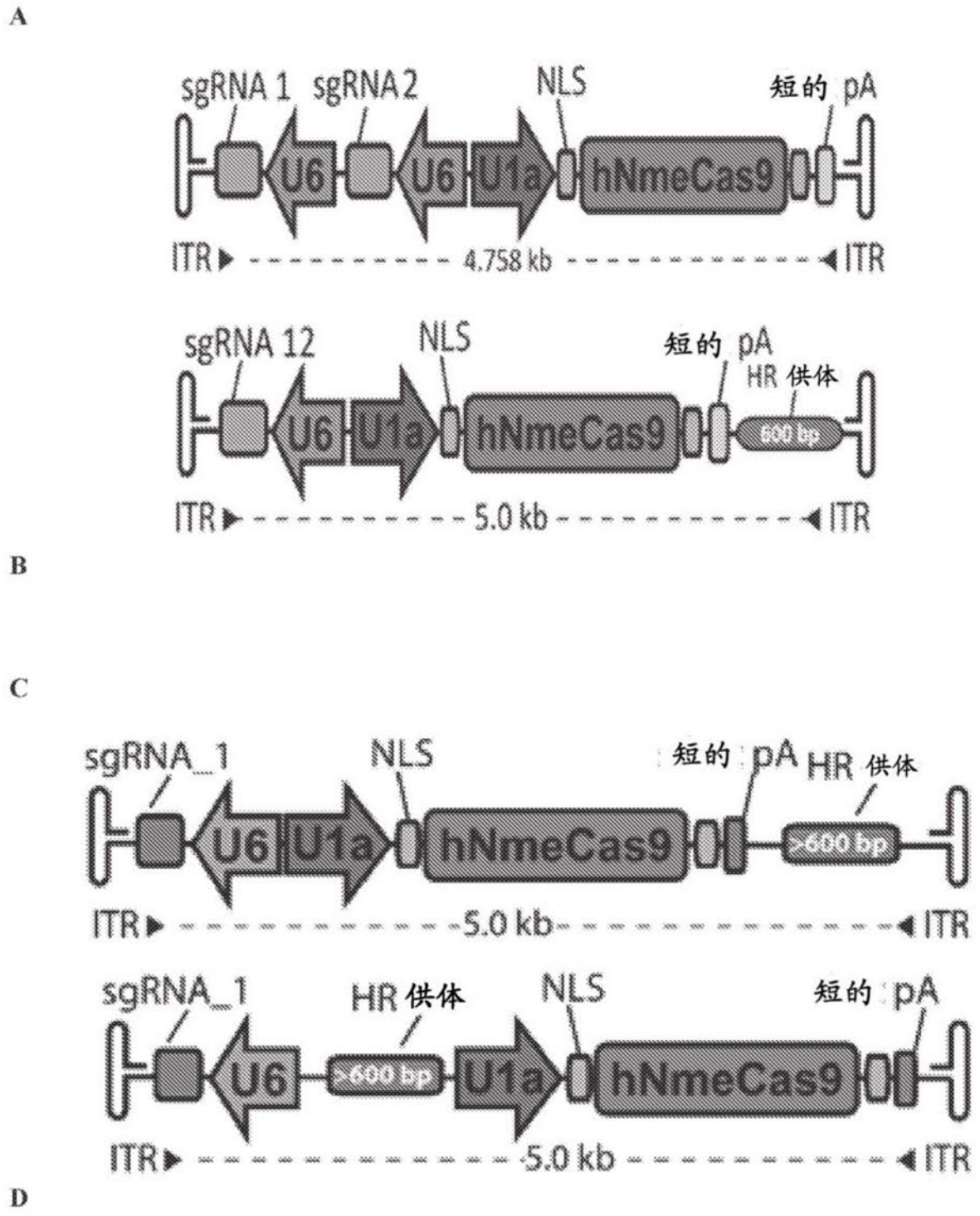
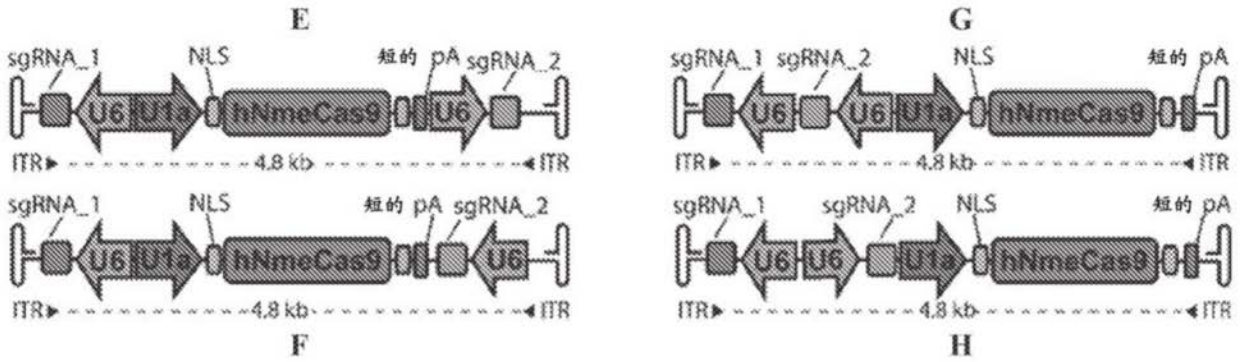

A
sgRNA 1
sgRNA2
NLS
短的 pA
U6
U6
U1a
hNmeCas9
ITR
4.758 kb
ITR
sgRNA 12
NLS
短的 pA
HR 供体
U6
U1a
hNmeCas9
600 bp
ITR
5.0 kb
ITR
B
C
sgRNA_1
NLS
短的 pA
HR 供体
U6
U1a
hNmeCas9
>600 bp
ITR
5.0 kb
ITR
sgRNA_1
HR 供体
NLS
短的 pA
U6
>600 bp
U1a
hNmeCas9
ITR
5.0 kb
ITR
D

图48


E
sgRNA_1
NLS
短的 pA
sgRNA_2
U6
U1a
hNmeCas9
ITR
4.8 kb
F
G
H


I

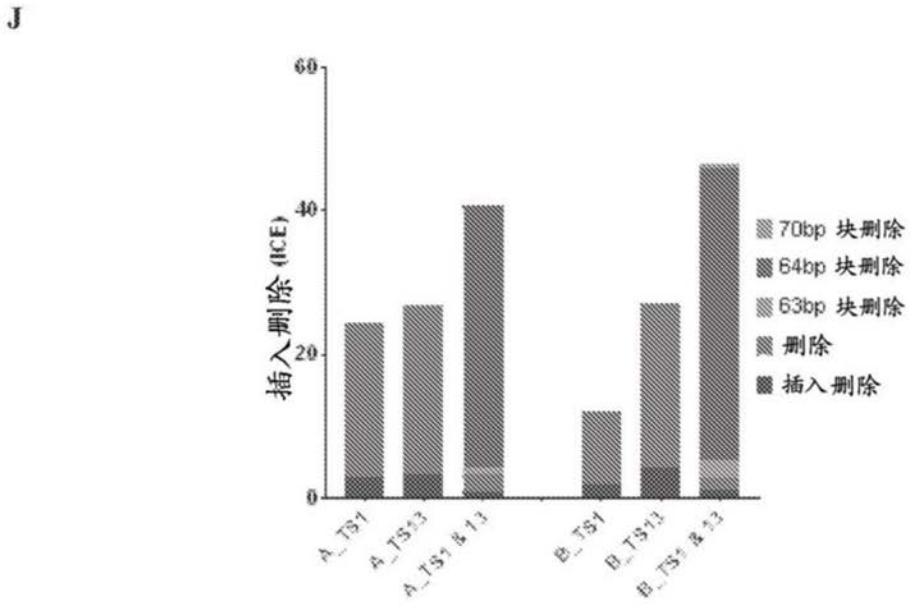

J
插入删除 (ICE)
60
40
20
0
70bp 块删除
64bp 块删除
63bp 块删除
删除
插入删除
A_TS1
A_TS13
A_TS1 & 13
B_TS1
B_TS13
B_TS1 & 13

图48续

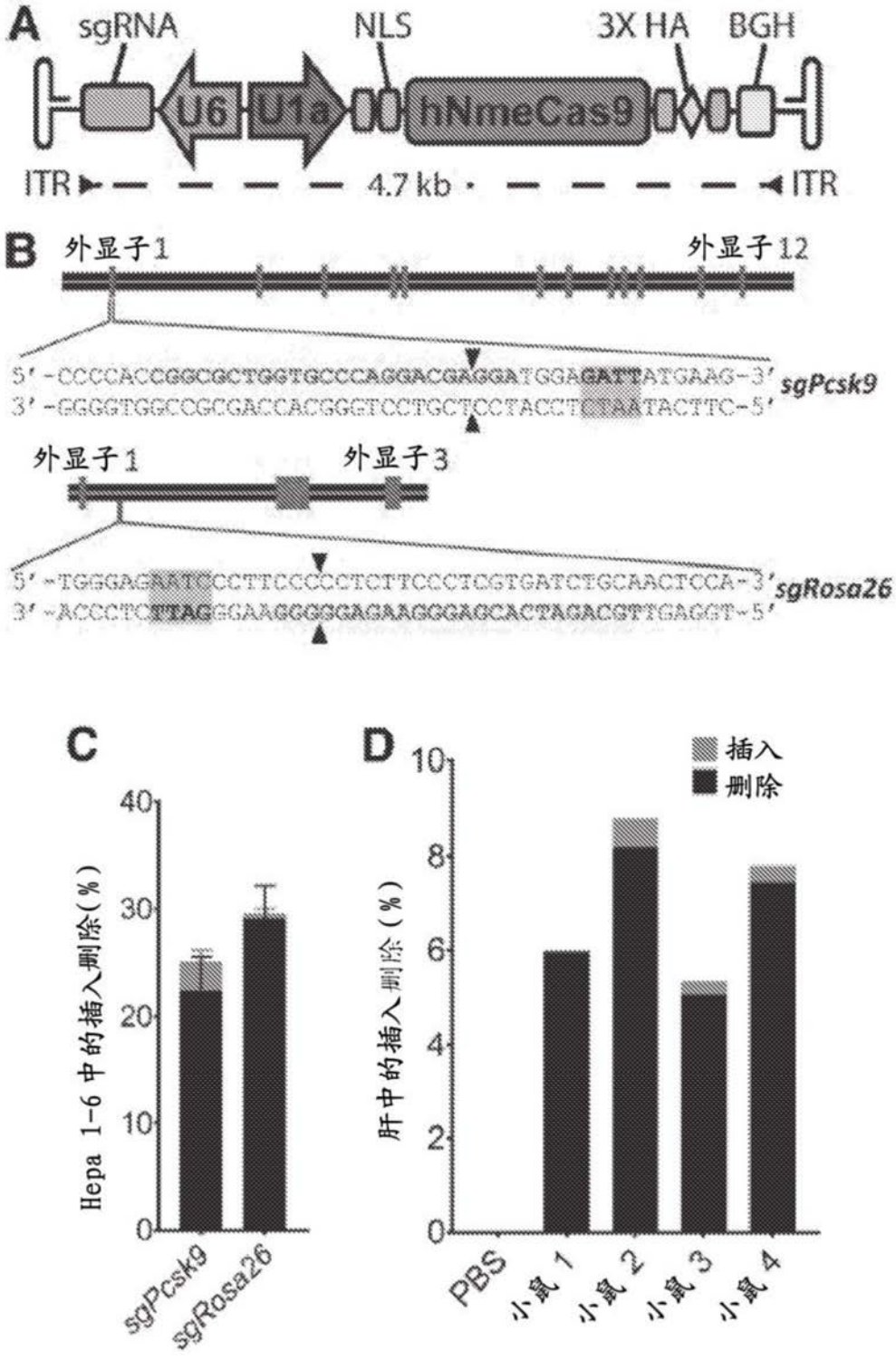

A
sgRNA
NLS
3X HA
BGH
U6
U1a
hNmeCas9
ITR
4.7 kb
ITR
B
外显子1
外显子12
5′-CCCCACCGGCGCTGGTGCCCAGGACGAGGATGGAGATTATGAAG-3′
3′-GGGGTGGCCGCGACCACGGGTCCTGCTCCTACCTCTAATACTTC-5′
sgPcsk9
外显子1
外显子3
5′-TGGGAGAATCCCTTCCCCCTCTTCCCTCGTGATCTGCAACTCCA-3′
3′-ACCCTCTTAGGGAAGGGGGAGAAGGGAGCACTAGACGTTGAGGT-5′
sgRosa26
C
Hepa 1-6 中的插入删除(%)
40
30
20
10
0
sgPcsk9
sgRosa26
D
插入
删除
肝中的插入删除(%)
10
8
6
4
2
0
PBS
小鼠 1
小鼠 2
小鼠 3
小鼠 4

图49

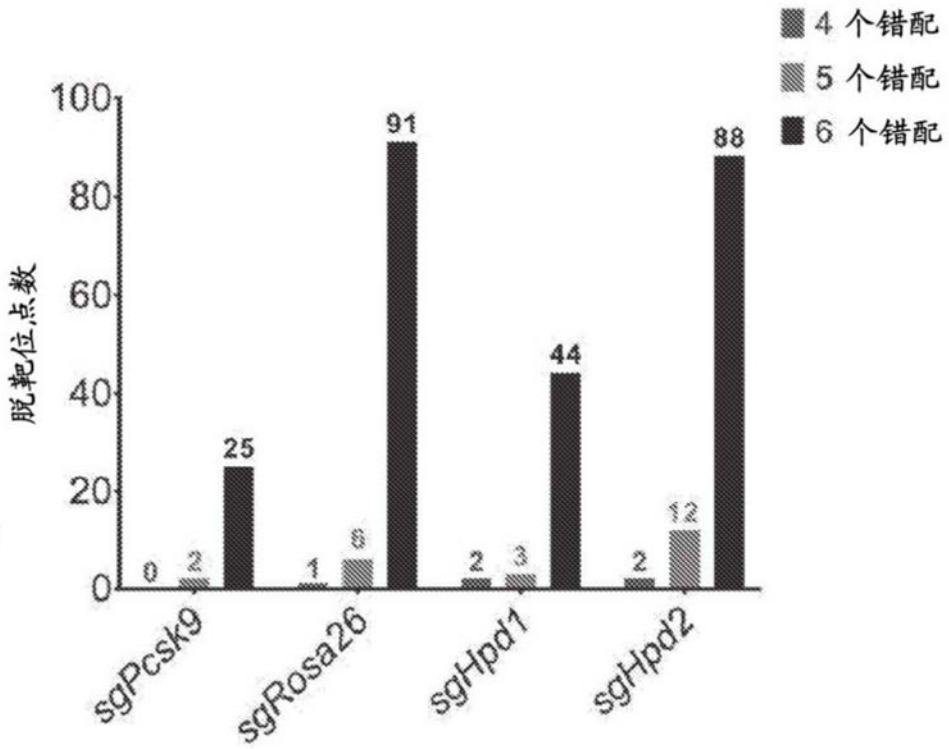

4 个错配
5 个错配
6 个错配
脱靶位点数
100
80
60
40
20
0
0
2
25
1
6
91
2
3
44
2
12
88
sgPcsk9
sgRosa26
sgHpd1
sgHpd2

图50

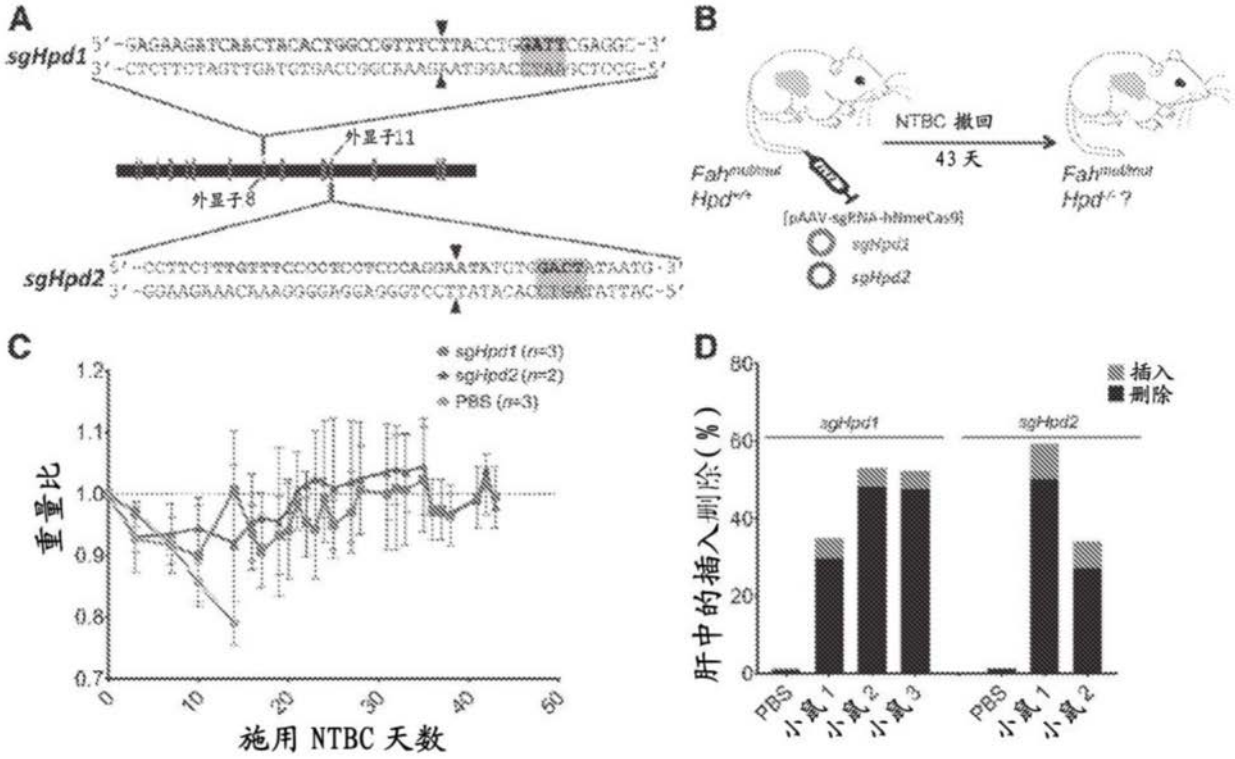

A
sgHpd1
外显子11
外显子8
sgHpd2
B
NTBC 撤回
43 天
[pAAV-sgRNA-hNmeCas9]
sgHpd1
sgHpd2
C
sgHpd1 (n=3)
sgHpd2 (n=2)
PBS (n=3)
重量比
施用 NTBC 天数
D
插入
删除
sgHpd1
sgHpd2
肝中的插入/删除(%)
PBS
小鼠 1
小鼠 2
小鼠 3

图51

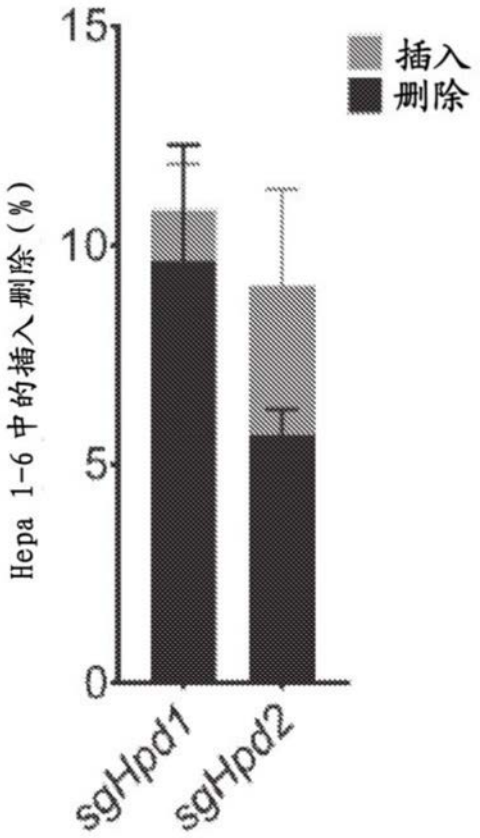


图52

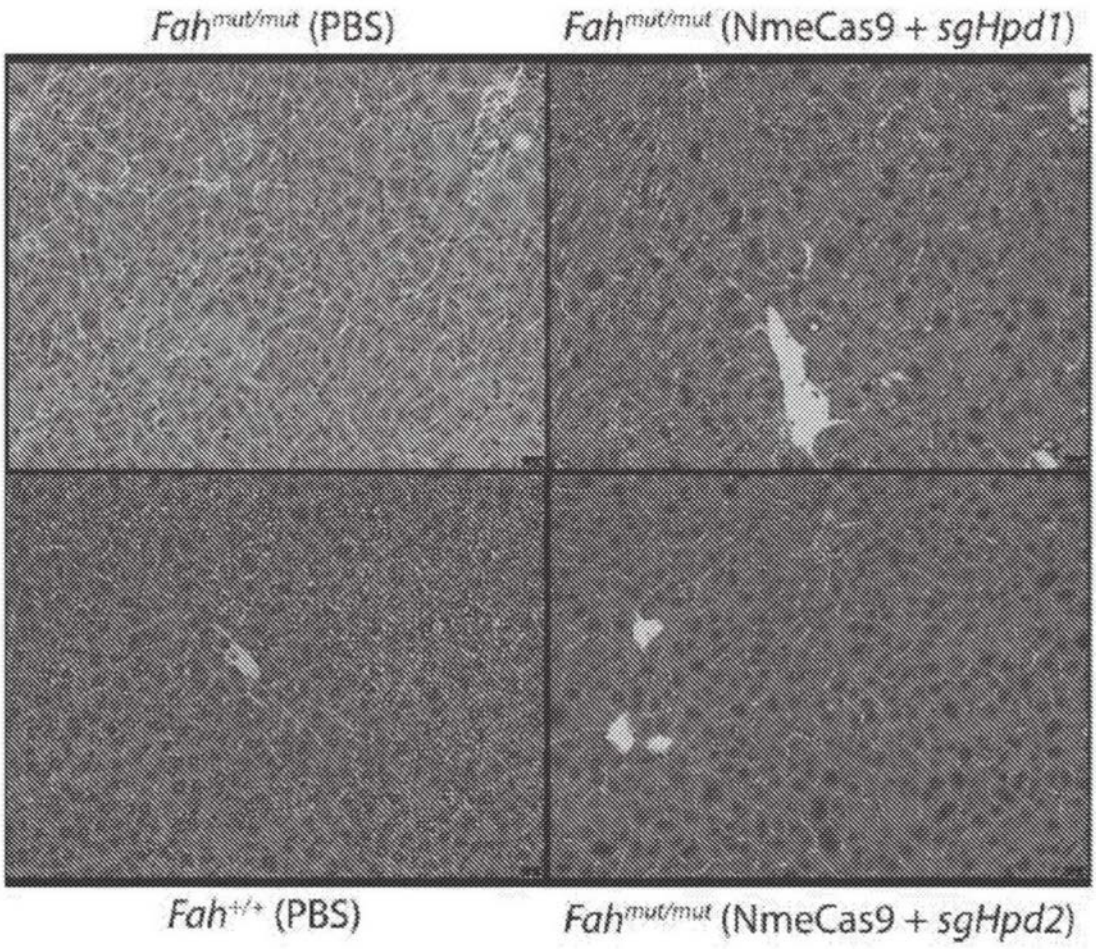

Fah^mut/mut (PBS)
Fah^mut/mut (NmeCas9 + sgHpd1)
Fah^+/+ (PBS)
Fah^mut/mut (NmeCas9 + sgHpd2)

图53

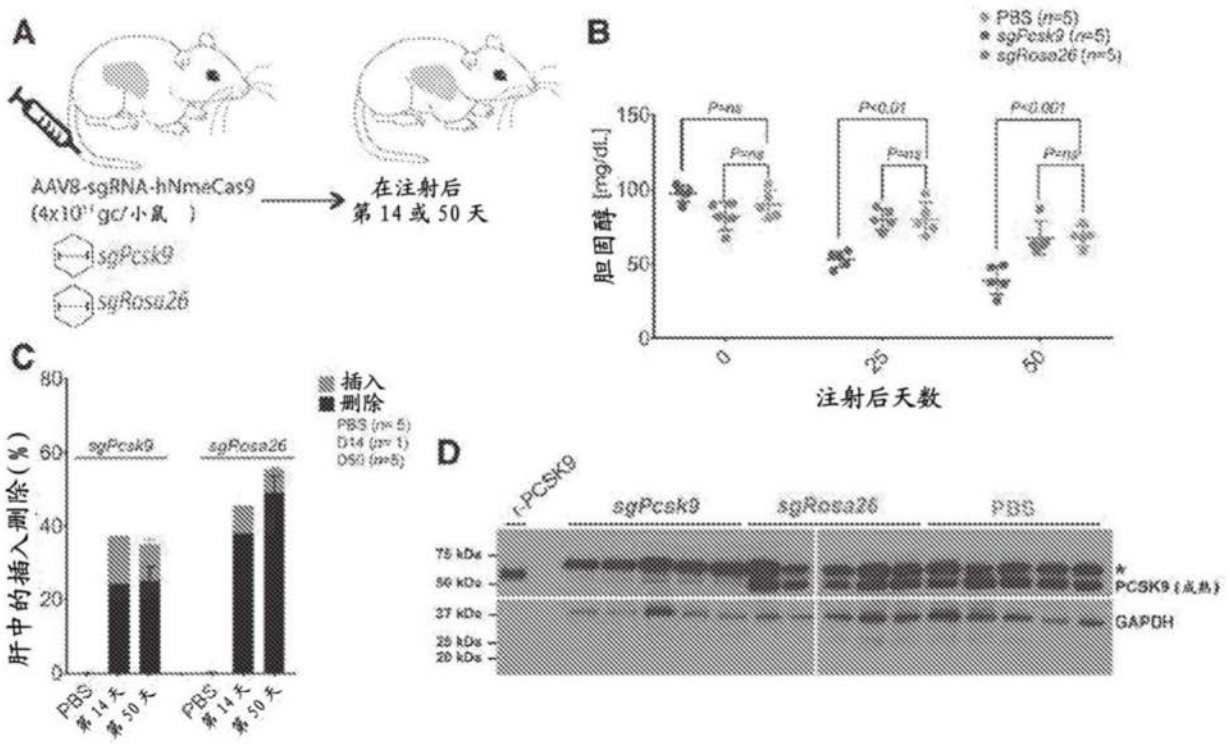
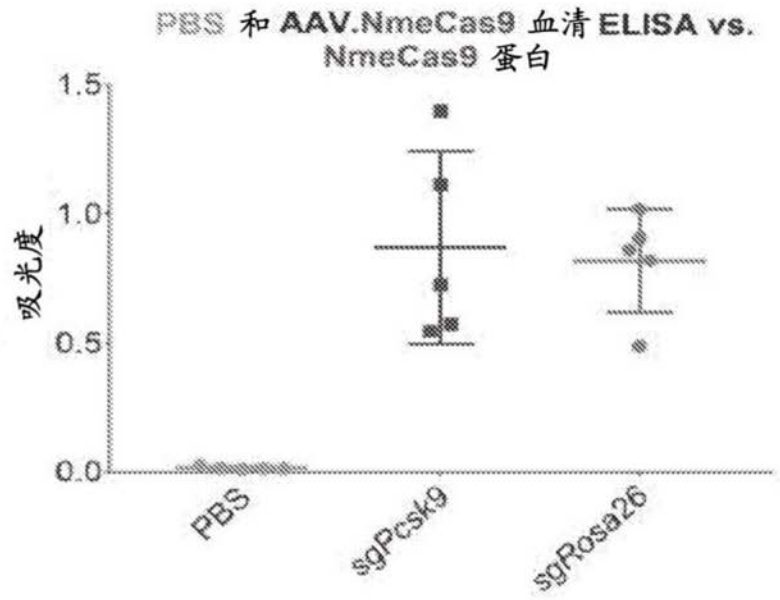

A
AAV8-sgRNA-hNmeCas9
(4x10^11 gc/小鼠)
sgPcsk9
sgRosa26
在注射后
第 14 或 50 天
B
PBS (n=5)
sgPcsk9 (n=5)
sgRosa26 (n=5)
胆固醇 [mg/dL]
注射后天数
C
插入
删除
肝中的插入/删除(%)
sgPcsk9
sgRosa26
PBS
第 14 天
第 50 天
D
r-PCSK9
sgPcsk9
sgRosa26
PBS
75 kDa
50 kDa
37 kDa
25 kDa
20 kDa
PCSK9
GAPDH

图54

A


PBS 和 AAV.NmeCas9 血清 ELISA vs.
NmeCas9 蛋白
吸光度
1.5
1.0
0.5
0.0
PBS
sgPcsk9
sgRosa26

B

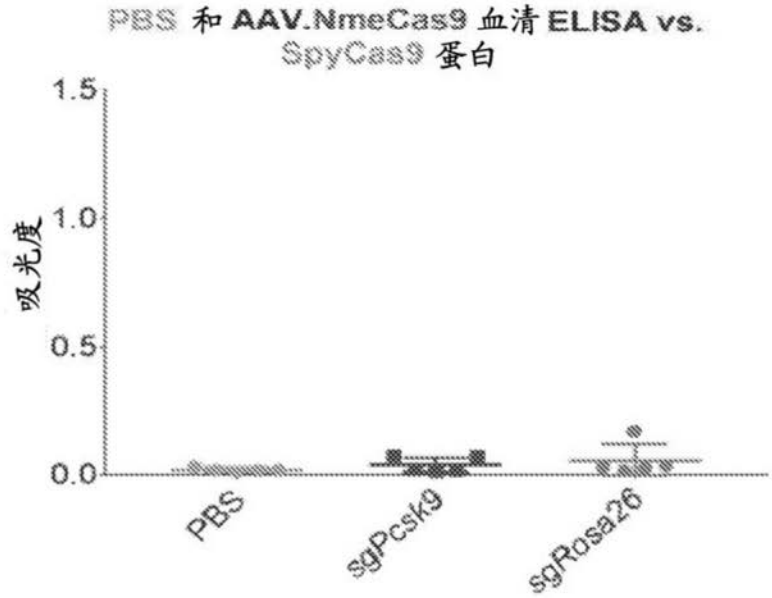

PBS 和 AAV.NmeCas9 血清 ELISA vs.
SpyCas9 蛋白
吸光度
1.5
1.0
0.5
0.0
PBS
sgPcsk9
sgRosa26

图55

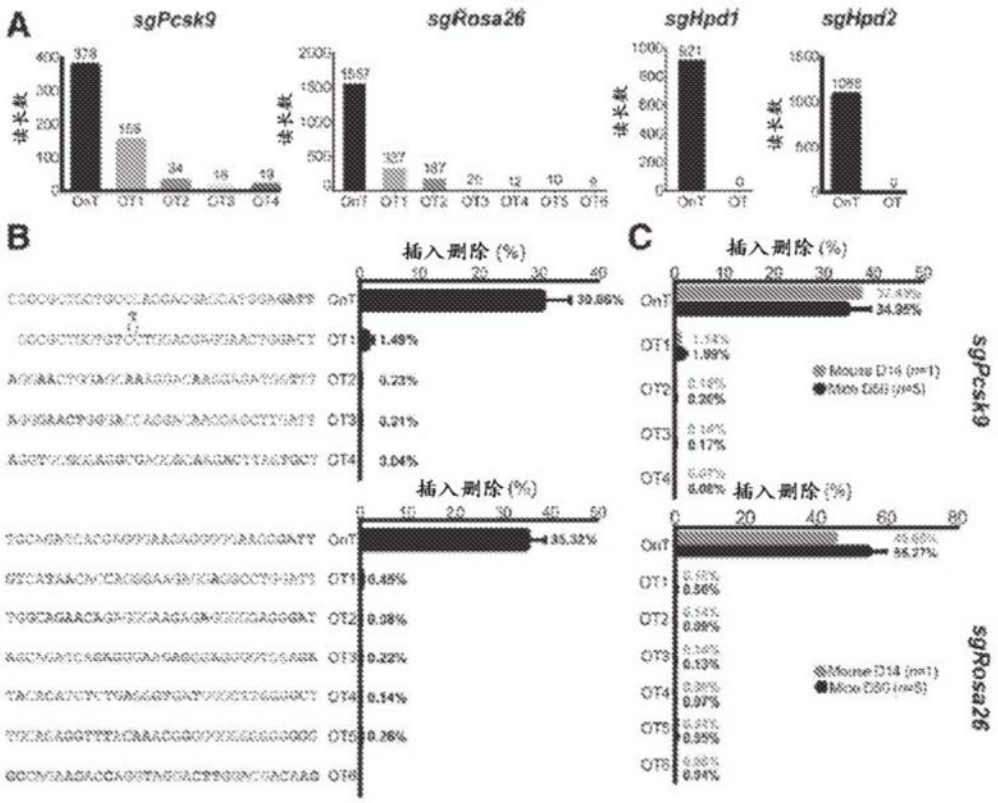
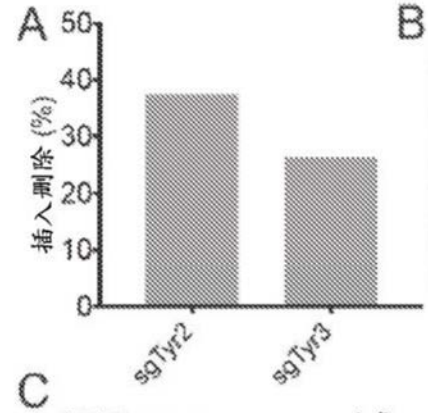

A
sgPcsk9
sgRosa26
sgHpd1
sgHpd2
读长数
B
插入删除 (%)
C
插入删除 (%)
sgPcsk9
sgRosa26

图56



B

小鼠ID	AAV 剂量	插入删除(%)	毛色
9-1	3 x 10^9 GCs	100	白色
9-2	3 x 10^9 GCs	100	白色
9-3	3 x 10^9 GCs	84	杂色及次形态
9-4	3 x 10^9 GCs	98	杂色及次形态
9-5	3 x 10^9 GCs	100	杂色及次形态
8-1	3 x 10^8 GCs	100	杂色及次形态
8-2	3 x 10^8 GCs	98	杂色及次形态

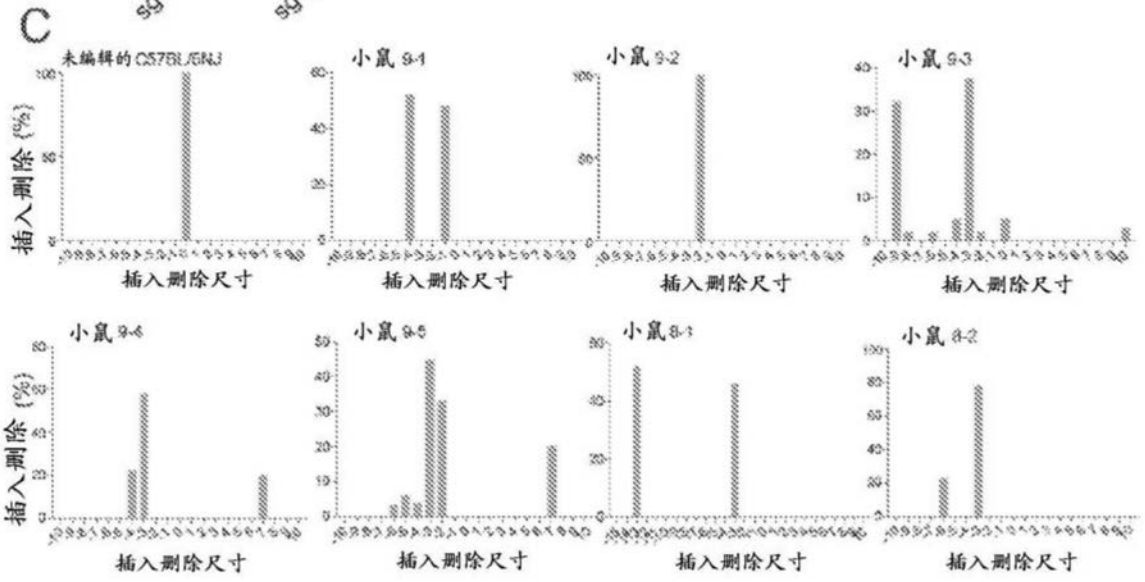
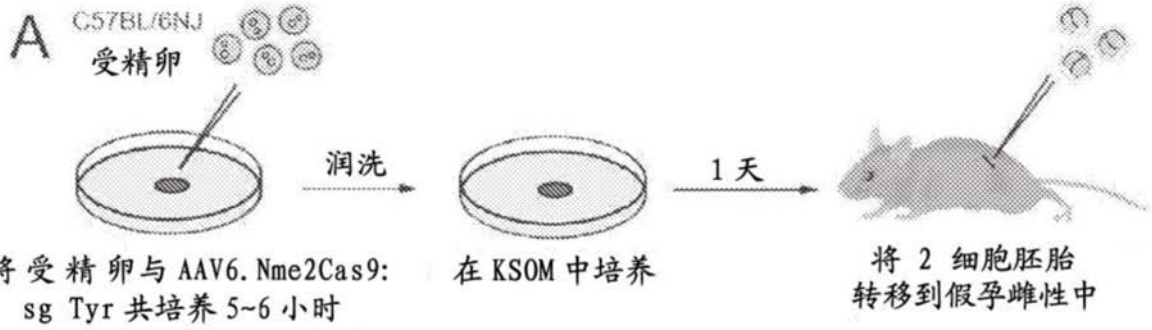
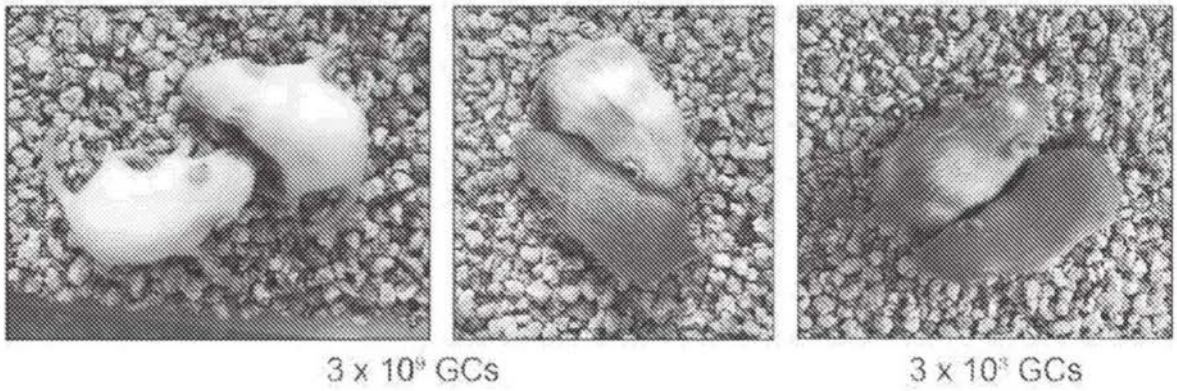


图57

A



B



C

AAV 剂量	转移的胚胎数	恢复的幼崽数	Tyr-编辑的幼崽（白色或绒毛丝鼠）数	Tyr 编辑频率(%)
3 x 10^9 GCs	50	5	5 (3)	100
3 x 10^8 GCs	28	2*	2 (1)	100

图58

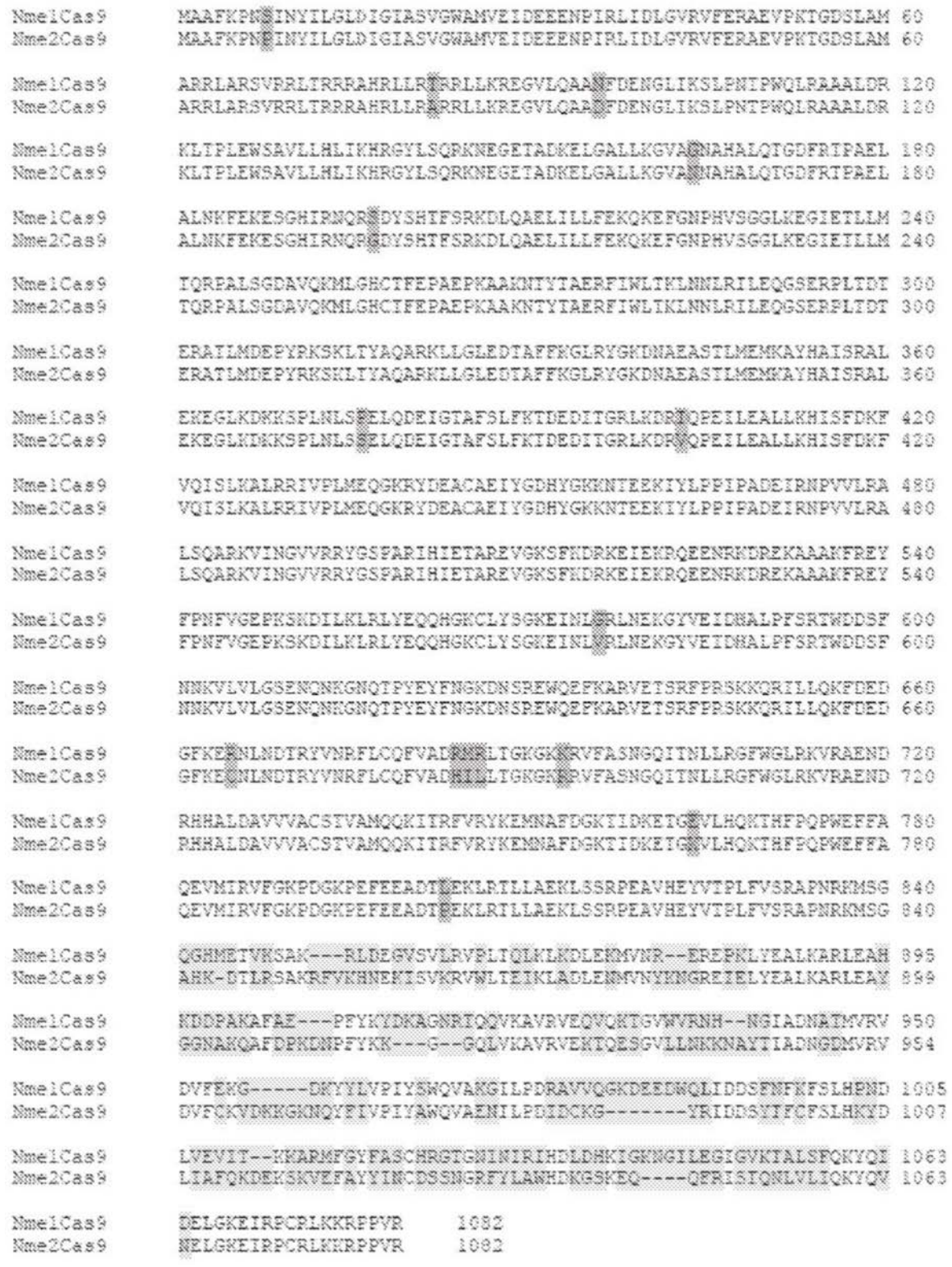

Nme1Cas9 MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAM 60
Nme2Cas9 MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAM 60
Nme1Cas9 ARRLARSVPRLTRRRAHRLLRTRRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAALDR 120
Nme2Cas9 ARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAALDR 120
Nme1Cas9 KLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAEL 180
Nme2Cas9 KLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAEL 180
Nme1Cas9 ALNKFEKESGHIRNQRSDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLM 240
Nme2Cas9 ALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLM 240
Nme1Cas9 TQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDT 300
Nme2Cas9 TQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDT 300
Nme1Cas9 ERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRAL 360
Nme2Cas9 ERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRAL 360
Nme1Cas9 EKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKF 420
Nme2Cas9 EKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKF 420
Nme1Cas9 VQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRA 480
Nme2Cas9 VQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRA 480
Nme1Cas9 LSQARKVINGVVRRYGSPARIHIETARDVGKSFKDRKEIEKRQEENRKDREKAAAKFREY 540
Nme2Cas9 LSQARKVINGVVRRYGSPARIHIETARDVGKSFKDRKEIEKRQEENRKDREKAAAKFREY 540
Nme1Cas9 FPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSF 600
Nme2Cas9 FPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLARLNEKGYVEIDHALPFSRTWDDSF 600
Nme1Cas9 NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDED 660
Nme2Cas9 NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDED 660
Nme1Cas9 GFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFWGLRKVRAEND 720
Nme2Cas9 GFKEQNLNDIRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAEND 720
Nme1Cas9 RHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWEFFA 780
Nme2Cas9 PHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWEFFA 780
Nme1Cas9 QEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSG 840
Nme2Cas9 QEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSG 840
Nme1Cas9 QGHMETVKSAK---RLDEGVSVLRVPLTQLKLKDLEKMVNR--EREPKLYEALKARLEAH 895
Nme2Cas9 AHK-DTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAY 899
Nme1Cas9 KDDPAKAFAE---PFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNH--NGIADNATMVRV 950
Nme2Cas9 GGNAKQAFDPKDNPFYKK---G--GQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRV 954
Nme1Cas9 DVFEKG-----DKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFNFKFSLHPND 1005
Nme2Cas9 DVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKG-------YRIDDSYTFCFSLHKYD 1007
Nme1Cas9 LVEVIT--KKARMFGYFASCHRGTGNIKIRIHDLDHKIGKNGILEGIGVKTALSFQKYQI 1063
Nme2Cas9 LIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQ----QFRISTQNLVLIQKYQV 1063
Nme1Cas9 DELGKEIRPCRLKKRPPVR 1082
Nme2Cas9 NELGKEIRPCRLKKRPPVR 1082

图59



图60

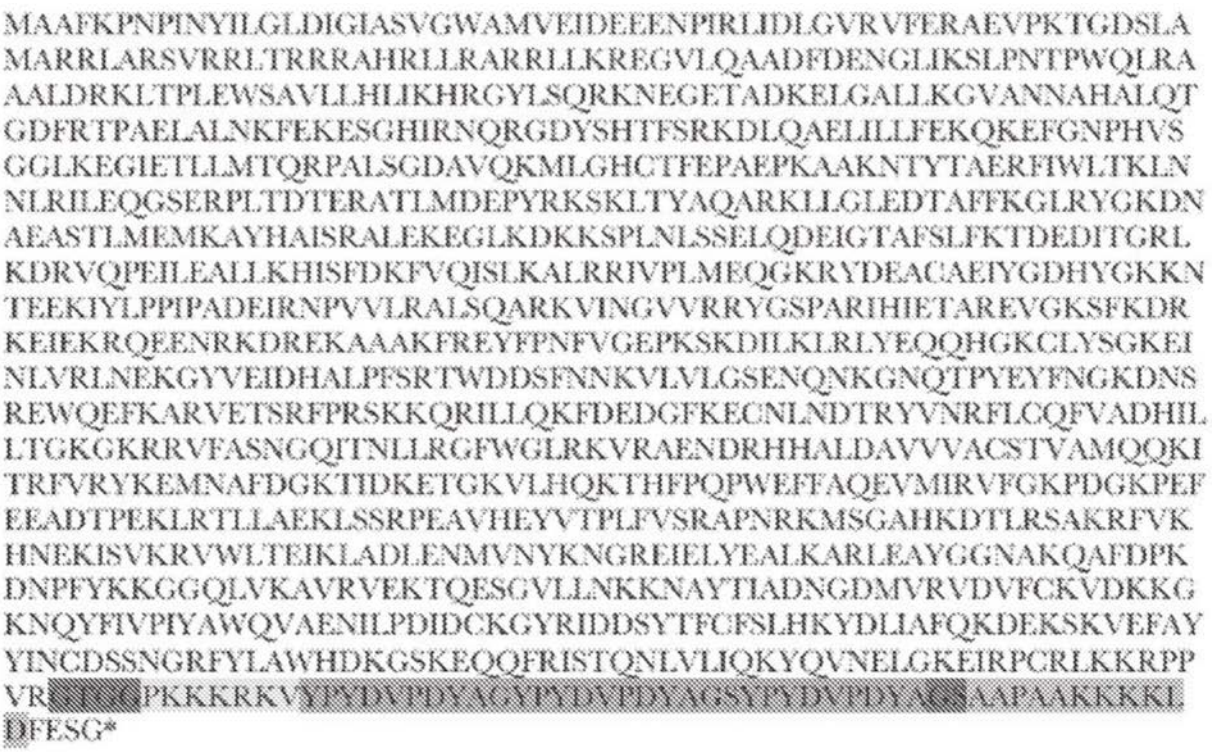

MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLA
MARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRA
AALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQT
GDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVS
GGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLN
NLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDN
AEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRL
KDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKN
TEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETARDVGKSFKDR
KEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEI
NLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNS
REWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHIL
LTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKI
TRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEF
EEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSAKRFVK
HNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPK
DNPFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKG
KNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFAY
YINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPP
VRSTGGPKKKRKVYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAGSAAPAAKKKKL
DFESG*

图61

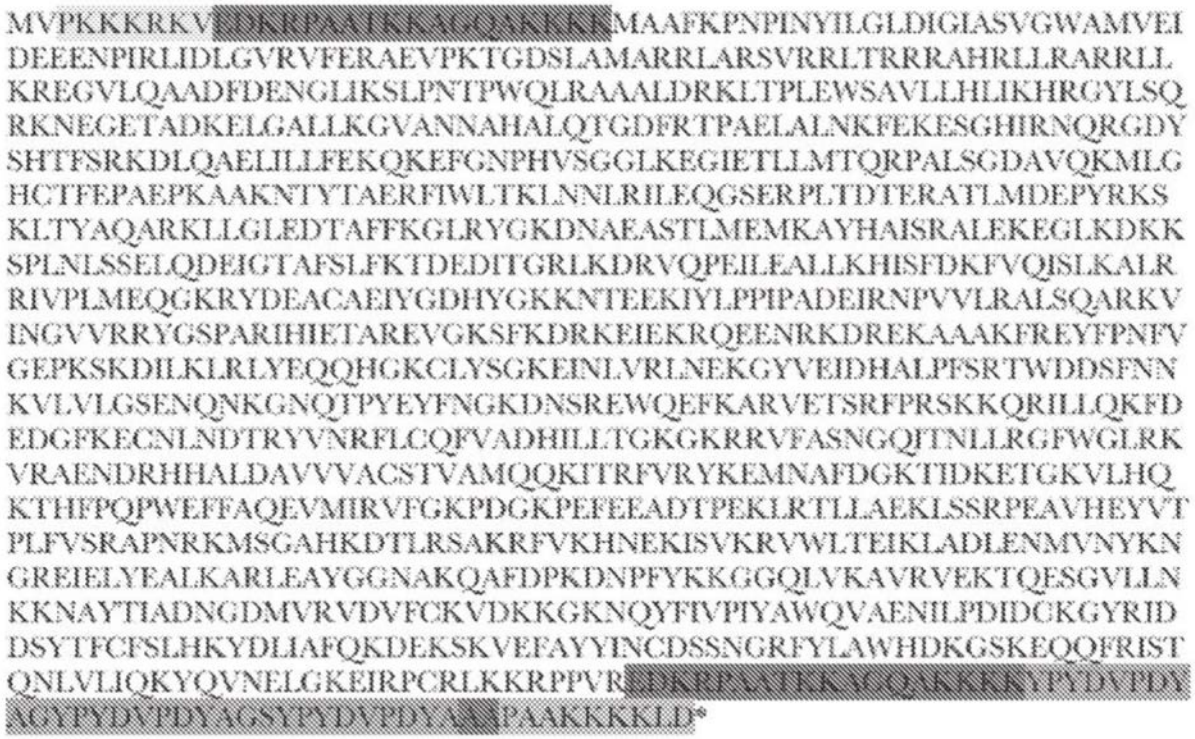

MVPKKKRKVEDKRPAATKKAGQAKKKKMAAFKPNPINYILGLDIGIASVGWAMVEI
DEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLL
KREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQ
RKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDY
SHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLG
HCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKS
KLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKK
SPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALR
RIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKV
INGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFV
GEPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNN
KVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD
EDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRK
VRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQ
KTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVT
PLFVSRAPNRKMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKN
GREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLN
KKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRID
DSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRIST
QNLVLIQKYQVNELGKEIRPCRLKKRPPVRSDKRPAATKKAGQAKKKKYPYDVPDY
AGYPYDVPDYAGSYPYDVPDYAGSPAAKKKKLD*

图62


PKKKRKV[redacted]MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAE
VPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLP
NTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVA
NNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQK
EFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAER
FIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFK
GLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKT
DEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYG
DHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREV
GKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGK
CLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEY
FNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLC
QFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACS
TVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVF
GKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDT
LRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGN
AKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVF
CKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDE
KSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRP
CRLKKRPPVR[redacted]GGGSGGGGSGGGGS[redacted]PAAKKKKLD[redacted]GGSKRPAATKKAGQAKKKK
*

图63

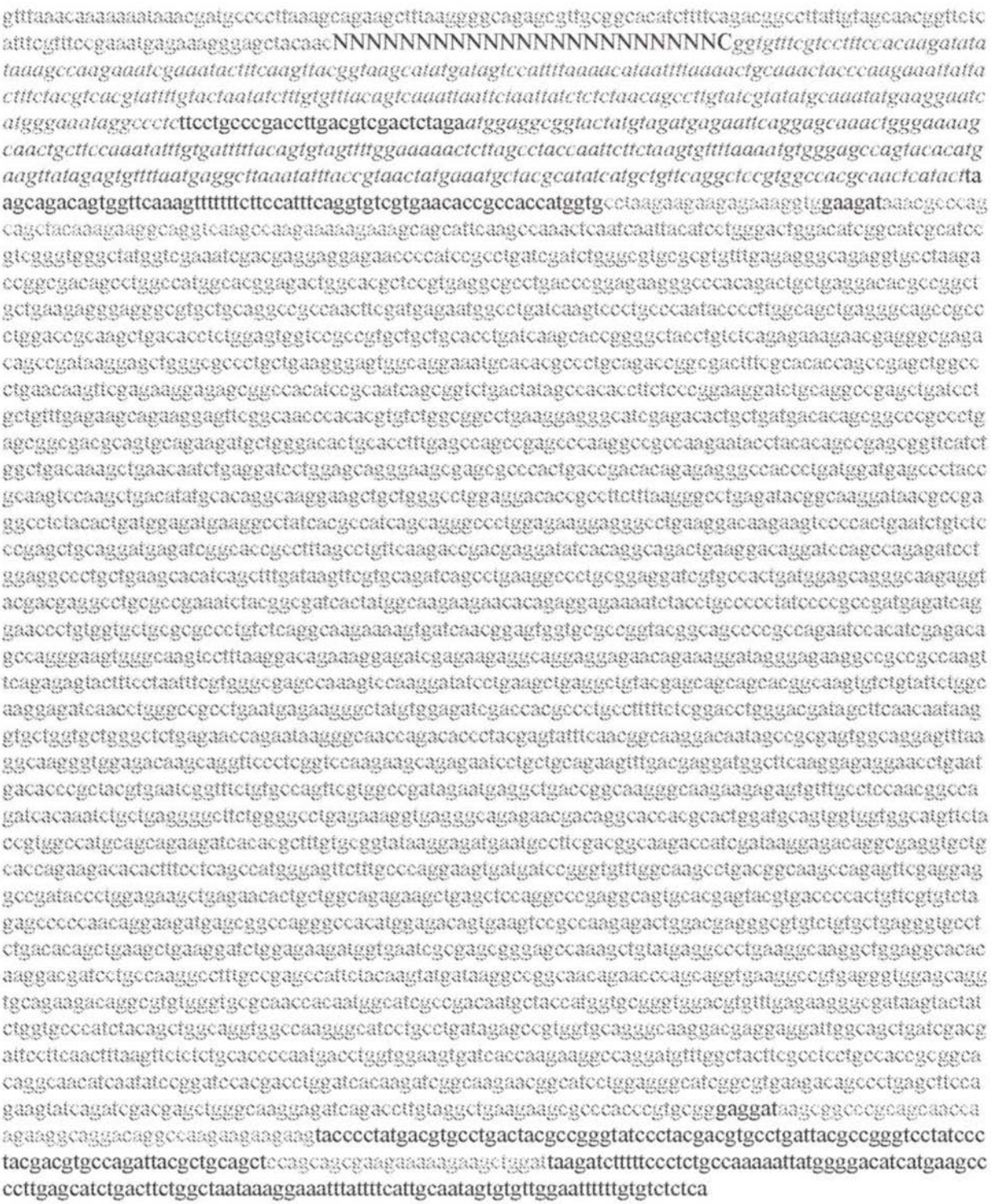

NNNNNNNNNNNNNNNNNNNNNNC
ttcctgcccgaccttgacgtcgactctaga
gaagat
tacccctatgacgtgcctgactacgccgggtatccctacgacgtgcctgattacgccgggtcctatccc
tacgacgtgccagattacgctgcagct
taagatcttttccctctgccaaaaattatggggacatcatgaagcc
ccttgagcatctgacttctggctaataaaggaaatttatttttcattgcaatagtgtgttggaatttttgtgtctctca

图64